

2021-04

Binary Outcomes and Linear Interactions

Vincent Boucher
Yann Bramoullé

Avril / April 2021

Centre de recherche sur les risques
les enjeux économiques et les politiques publiques

www.crrep.ca



ABSTRACT

We revisit and rehabilitate linear models of interactions in binary outcomes. Building on Heckman and MaCurdy (1985), we characterize when these models are statistically well defined. We then characterize and assess their game-theoretic microfoundations. We show that linear models of interactions admit sensible microfoundations under incomplete information and independence, but unconventional ones under complete information. We propose two simple estimators and revisit the empirical analyses of teenage smoking and peer effects of Lee, Li, and Lin (2014) and of entry into airline markets of Ciliberto and Tamer (2009). Our reanalyses highlight the main advantages of the linear framework and suggest that the estimations in these two studies suffer from endogeneity problems.

Keywords: Binary Outcomes, Linear Probability Model, Peer Effects, Econometrics of Games.

Vincent Boucher: Laval University, CRREP, CREATE.

Yann Bramoullé: Aix-Marseille University.

We thank Habiba Djebbari, Arthur Lewbel, participants at seminars and conferences for their helpful comments, and Aureo de Paula for invaluable feedback. This work was supported by the French National Research Agency Grant ANR-17-EURE-0020. This research uses data from Add Health, a program directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by Grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is given to Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain Add Health data files is available on the Add Health website <http://www.cpc.unc.edu/addhealth>). No direct support was received from Grant P01-HD31921 for this research. Replication codes are available at <https://github.com/vincentboucherecn/linearbinaryinteractions>. We used Hlavac (2018) for some tables.

1 Introduction

In many contexts, researchers are interested in estimating interactions in agents' decisions and outcomes. Teenagers' smoking may depend on whether their friends smoke. A firm's entry into a market may depend on the entry of its competitors. Social and strategic interactions likely play a key role in many important issues, including health, academic achievement, public good provision, consumption and imperfect competition.¹

Obtaining credible causal estimates of social and strategic interactions, however, requires addressing some formidable econometric challenges. With binary outcomes, simultaneity in the behavior of interacting agents may notably yield multiple equilibria, see [Brock and Durlauf \(2001\)](#) and [Tamer \(2003\)](#). Addressing multiplicity is a central objective of the econometrics of games, surveyed in [Bajari, Hong, and Nekipelov \(2013\)](#) and [De Paula \(2013\)](#). In the past 30 years, researchers have made great progress on this issue and have developed econometric frameworks that can, in principle, be used to analyze models with multiple equilibria. The ability to account for multiplicity, however, often comes at a significant cost in terms of practical implementation. At this stage, estimating interactions in binary outcomes under multiplicity may be computationally demanding, may require massive amounts of data and generally does not permit controls for unobserved heterogeneity with standard fixed effects.

To address these concerns, we revisit linear models of interactions in binary outcomes. [Heckman and MaCurdy \(1985\)](#) showed that binary outcomes are compatible with classical econometric models where an agent's outcome depends linearly on others' outcomes. This insight, however, was discarded by the recent literature on the econometrics of games. None of the articles and surveys in our references cite [Heckman and MaCurdy \(1985\)](#) or estimate a linear model of interactions. This neglect is perhaps due to a belief that these models are incompatible with the game-theoretic microfoundations analyzed in the literature. We show that this belief is unfounded. Under incomplete information and independence, in particular, we show that the linear model of interactions corresponds to the unique interior Bayes-Nash

¹See, e.g., [Fadlon and Nielsen \(2019\)](#) on health, [Sacerdote \(2011\)](#) on academic achievement, [Foster and Rosenzweig \(1995\)](#) on public good provision, [Kuhn et al. \(2011\)](#) on consumption, and [Berry \(1992\)](#) on imperfect competition.

equilibrium of a game with linear utilities and iid, uniformly distributed preference shocks. The uniform distribution may be unusual but, in our view, it is no more arbitrary than the standard logistic or normal distributions.² Moreover, it is, arguably, more convenient, given the well-known advantages of linear models. Estimation of linear models is straightforward, they have minimal data requirements, and these models and their estimators can easily handle fixed effects. We thus believe that linear models of interactions in binary outcomes have a legitimate place in the toolkit of applied researchers interested in such interactions.

We consider a general model of linear interactions in binary outcomes. The model notably nests linear-in-means models of peer effects in networks (Bramoullé, Djebbari, and Fortin, 2009) and linear models of entry games (Jovanovic, 1989). We develop our analysis in several stages. We first build on Heckman and MaCurdy (1985) and characterize when this model is statistically well defined (Theorem 1). We show how to extend the stochastic structure of the linear probability model to account for outcome interactions. The model thus inherits well-known properties of linear interaction models with continuous outcomes. It generically has a unique solution, and identification is characterized by standard rank conditions.

We then provide the first analysis of the model’s game-theoretic microfoundations, adopting standard assumptions from the literature. In our main result, we characterize these microfoundations under incomplete information, independence, and linear utilities (Theorem 2). We show that the linear model of interactions corresponds to the unique interior Bayes-Nash equilibrium of the game for all possible parameters if and only if preference shocks are iid, uniformly distributed. Furthermore, this is the unique Bayes-Nash equilibrium under moderate interactions. We also characterize microfoundations under complete information and linear utilities (Proposition 1). We show that the linear model of interactions can be rationalized as a Nash equilibrium of the game, but only under unconventional assumptions on preference shocks. In particular, preference shocks that are independent must have non-convex support.³ Overall, we characterize and assess the microfoundations of the linear model of interactions, arguing that these microfoundations are reasonable under incomplete

²Heckman and Snyder Jr (1997) show that with a single decision maker, the classical linear probability model can be microfounded through additive random utilities and uniformly distributed preference shocks. We extend this result to a game-theoretic context with outcome interactions.

³The linear model of interactions then corresponds to the unique Nash equilibrium robust to increases in shock dispersion (Proposition 2).

information, but unconventional under complete information.

Finally, we propose two simple consistent estimators and analyze real data using our proposed linear framework. One estimator is a classical Two-Stage Least Squares (2SLS), the other is a Nonlinear Least Squares (NLS).⁴ To illustrate how this differs from existing approaches, we revisit two studies: [Lee, Li, and Lin \(2014\)](#) on peer effects in teenage smoking and [Ciliberto and Tamer \(2009\)](#) on entry into airline markets. We reanalyze the same data as in the original studies and assess the robustness of the original results. These reanalyses highlight the main advantages of the linear framework: ease of implementation, availability of overidentification tests, and ability to handle fixed effects. In contrast, existing nonlinear frameworks are generally computationally demanding, may lack overidentification tests, and generally cannot handle large sets of fixed effects.

In the case of [Lee, Li, and Lin \(2014\)](#), we are able to include fixed effects at the school-grade level, a natural feature missing from the original analysis. With or without these fixed effects, linear estimates of endogenous peer effects are positive and significant, as observed in the original study. The joint validity of the instruments is, however, strongly rejected by overidentification tests. For [Ciliberto and Tamer \(2009\)](#), we are able to include airline fixed effects, also absent from the original analysis. Results from our reanalysis are qualitatively different from the original results. In a linear framework, estimates of strategic interactions between airlines are generally positive and significant, whereas they are negative and significant in [Ciliberto and Tamer \(2009\)](#).

Absent a proper means of testing one specification versus another, we can only speculate on the causes behind these differences.⁵ As in many studies in the econometrics of games, the first step of [Ciliberto and Tamer \(2009\)](#)'s estimation method is to obtain nonparametric estimates of conditional choice probabilities. These estimates capture how the probabilities of

⁴We discuss the estimators' properties and compare their small-sample performances through Monte Carlo simulations in Section 4. The NLS appears to be more efficient. Including fixed effects in a NLS estimation may be problematic, however, due to the incidental parameter problem. By contrast, eliminating group-level unobservables through within-group deviations is standard in 2SLS estimations, and hence the 2SLS may be preferred, in practice, for most applications.

⁵[Ciliberto and Tamer \(2009\)](#) develop and implement a partial identification approach. Although no direct specification tests are available, measures of goodness of fit are quantitatively similar across the two specifications, with a slight advantage to the linear models. While [Ciliberto and Tamer \(2009\)](#) consider a game of complete information, incomplete information may be a more appropriate assumption in a context of imperfect competition.

all possible market outcomes depend on all covariates. The estimation of conditional choice probabilities suffers from a well-known curse of dimensionality in practice, see [Andrews, Berry, and Barwick \(2004\)](#). This problem appears to be severe in [Ciliberto and Tamer \(2009\)](#)'s application, as researchers must obtain nonparametric estimates of 63 functions of 20 variables with only 2,742 observations. In contrast, linear model estimation does not require estimates of conditional choice probabilities. More generally, linear estimations are not affected by a curse of dimensionality, and we suspect this is key to explaining the differing results. In addition, the joint validity of the instruments is also strongly rejected by overidentification tests. This suggests that both the original analyses of [Lee, Li, and Lin \(2014\)](#) and [Ciliberto and Tamer \(2009\)](#) and our reanalyses suffer from endogeneity problems.

Our analysis contributes to a large and still-expanding literature on peer effects.⁶ Early studies focused on group interactions. The population is therefore partitioned into groups; agents interact with all other members of their group and with no members of another group. [Brock and Durlauf \(2001\)](#) were the first to propose a microfounded econometric framework to analyze peer effects on binary outcomes. They consider a setup of incomplete information under group interactions. They show that the model has a unique equilibrium under moderate interactions and multiple equilibria under strong interactions. [Soetevent and Kooreman \(2007\)](#) analyze peer effects on binary outcomes under complete information and group interactions. They find that the game typically has a large number of equilibria. They propose a simulated maximum likelihood estimator based on the assumption that all Nash equilibria are equally likely. [Nakajima \(2007\)](#) also analyzes peer effects on binary outcomes under complete information and group interactions. He considers a stochastic Markov process where agents sequentially and myopically play a best response. He assumes that the likelihood function is equal to the steady-state distribution of this process. Recent studies consider more complex network interactions.⁷ [Li and Zhao \(2016\)](#) adapt partial identification approaches under complete information to the analysis of peer effects in networks and binary outcomes.

⁶See, for instance, [Manski \(2000\)](#), [Angrist \(2014\)](#), and [Kline and Tamer \(Forthcoming\)](#).

⁷About 10 years ago, four studies independently understood that the reflection problem ([Manski \(1993\)](#)) is naturally solved by network interactions ([Bramoullé, Djebbari, and Fortin \(2009\)](#), [De Giorgi, Pellizzari, and Redaelli \(2010\)](#), [Lin \(2010\)](#), and [Laschever \(2013\)](#)). Since then, the literature on peer effects in networks has rapidly grown and extended in many directions, see [Boucher and Fortin \(2015\)](#), [De Paula \(2017\)](#), and [Bramoullé, Djebbari, and Fortin \(2020\)](#). Relatively few studies, however, analyze peer effects in networks and binary outcomes.

Lee, Li, and Lin (2014) extend the incomplete information framework of Brock and Durlauf (2001) to networks. They show that uniqueness holds under moderate interactions and propose an iterative simulated maximum likelihood estimator based on a subroutine that repeatedly computes the solution of a high-dimensional nonlinear fixed-point system. All these studies develop nonlinear frameworks to analyze peer effects on binary outcomes.

In contrast, we show that linear models of peer effects, traditionally used to study continuous outcomes (Manski (1993)), can also be used for binary outcomes. These models maintain key properties when applied to binary outcomes and, in particular, the identification results and insights of Bramoullé, Djebbari, and Fortin (2009), exploiting holes in the network structure to solve the reflection problem.⁸ We characterize and assess the game-theoretic microfoundations of these models. We revisit the empirical analysis of peer effects and teenage smoking of Lee, Li, and Lin (2014). We obtain similar estimates of endogenous peer effects through a much simpler estimation procedure. In addition, we are able to control for school-grade fixed effects and verify whether the network-based instruments pass overidentification tests, two features absent from the original study.

Our analysis contributes, more generally, to the literature on the econometrics of games. Since the early work of Jovanovic (1989) and Bjorn and Vuong (1997), researchers have made great progress on the empirical analysis of models with multiple equilibria. Applied researchers who wish to estimate interactions in binary outcomes under multiplicity can, notably, specify a flexible selection mechanism dependent on estimated parameters (Bajari, Hong, and Ryan, 2010), assume that the same equilibrium is played across games under incomplete information (Aguirregabiria and Mira, 2007), or adopt a partial identification approach under complete information (Ciliberto and Tamer, 2009). Preference shocks are generally assumed to be either logistically or normally distributed. Different assumptions yield different nonlinear econometric frameworks; a common first step is often to obtain flexible estimates of conditional choice probabilities. Despite such major methodological progress, however, two features may limit the usefulness of these approaches for applied work. First,

⁸The linear framework can also, of course, be applied to analyze peer effects in binary outcomes with group interactions, under appropriate identification conditions. For instance, Soetevent and Kooreman (2007) and Nakajima (2007) assume that an agent's outcome does not depend directly on their peers' characteristics (no contextual peer effects). We can estimate linear interaction models of their data under the same assumption.

and depending on the context, the data requirements may be massive and unrealistic. We argue that this applies to the application in [Ciliberto and Tamer \(2009\)](#). Second, in nonlinear frameworks, introducing unobserved heterogeneity is generally impractical or unfeasible. Due to the incidental parameter problem, nonlinear estimators usually cannot handle fixed effects, whose numbers grow at the same rate as sample size. While they could, in principle, account for a finite number of fixed effects, their introduction further intensifies data requirements.⁹

We show that these limitations are not inevitable, as they do not apply to linear models of interactions in binary outcomes. We provide the first systematic analysis of the game-theoretic microfoundations of these models. We find that they admit reasonable microfoundations under incomplete information and independence, but rather unconventional microfoundations under complete information. The assumption of incomplete information seems particularly appropriate to analyze imperfect competition, since firms generally have a strong incentive not to divulge their private information on benefits, costs and market operations. We revisit the empirical analysis of entry into airline markets of [Ciliberto and Tamer \(2009\)](#), and find that the original analysis likely suffers from endogeneity problems.

Finally, our analysis contributes to the literature on the econometrics of discrete variables and on linear probability models in particular. Researchers hold diverse views on the use of linear models to analyze binary outcomes. Following [Angrist and Pischke \(2008\)](#), applied economists seeking reduced-form causal estimates generally estimate linear probability models. [Heckman and Snyder Jr \(1997\)](#) provide microfoundations for linear probability models with a single decision-maker. A main contribution of our analysis is to characterize and assess the game-theoretic microfoundations of linear probability models with outcome interactions. Given the many advantages of a linear framework, we believe that it could become a natural benchmark in empirical studies of binary outcome interactions.

The paper proceeds as follows. We present the econometric framework in Section 2 and analyze its microfoundations in Section 3. We propose estimators and discuss their properties in Section 4. We revisit existing studies in Section 5 and conclude in Section 6.

⁹[Aguirregabiria and Mira \(2019\)](#) show that identification may hold under incomplete information, multiplicity, and unobserved heterogeneity when the unobservables have finite support. This is a potentially promising result. Its empirical applicability has not yet been demonstrated, however, and the assumption of finite support represents, in any case, a significant restriction.

2 Econometric framework

A researcher has data on n agents and analyzes interactions affecting a binary outcome. Let $y_i \in \{0, 1\}$ denote agent i 's outcome. Let $\mathbf{y} \in \{0, 1\}^n$ denote the vector of outcomes and \mathbf{x} a matrix containing all observables. Throughout our analysis, we find it useful to distinguish between two types of stochastic terms. By convention, *errors* are defined directly from the data generating process. In particular, define *reduced-form errors* as $\nu_i = y_i - \mathbb{E}(y_i|\mathbf{x})$, such that $\mathbb{E}(\nu_i|\mathbf{x}) = 0$. By contrast, *preference shocks* refer to stochastic terms that appear in underlying microfoundations.

As is well known, the binary nature of the outcome imposes strong restrictions on the data generating process. In particular, $\mathbb{P}(y_i = 1|\mathbf{x}) = \mathbb{E}(y_i|\mathbf{x})$, and the reduced-form error ν_i is a binary, Bernoulli stochastic variable: $\nu_i = 1 - \mathbb{E}(y_i|\mathbf{x})$ with probability $\mathbb{E}(y_i|\mathbf{x})$ and $-\mathbb{E}(y_i|\mathbf{x})$ with probability $1 - \mathbb{E}(y_i|\mathbf{x})$.¹⁰ Reduced-form errors are always binary, even when preference shocks are continuously distributed.

In our analysis, we consider the following general model of linear interactions

$$y_i = f_i(\mathbf{x}, \boldsymbol{\theta}) + \sum_j \beta_{ij} y_j + \varepsilon_i, \quad (1)$$

under the exogeneity assumption, $\mathbb{E}(\varepsilon_i|\mathbf{x}) = 0$. Outcome y_i is affected by observables through function f_i and parameters $\boldsymbol{\theta}$ and by others' outcomes through linear interactions $\sum_j \beta_{ij} y_j$. Let $\boldsymbol{\beta}$ denote the interaction matrix, where $\beta_{ii} = 0$, and β_{ij} can potentially have any sign.

Whether there exists an error structure such that equation (1) holds with binary outcomes is not immediate. The interaction term $\sum_j \beta_{ij} y_j$ can take up to 2^{n-1} values and partly determines y_i , which can take only 2 values. In this Section, we clarify the conditions under which this model is statistically well defined. We analyze underlying microfoundations in the next Section.

Model (1) nests two important cases of interactions in binary outcomes: *peer effects* and *entry games*. First, consider the benchmark linear-in-means model of peer effects in networks, see [Bramoullé, Djebbari, and Fortin \(2009\)](#). For each agent i , the researcher

¹⁰This further implies that $\mathbb{V}(y_i|\mathbf{x}) = \mathbb{E}(y_i|\mathbf{x})(1 - \mathbb{E}(y_i|\mathbf{x}))$. The conditional variance and, more generally, higher moments of the conditional outcome distribution do not contain extra information with respect to the conditional expectation.

observes characteristics \mathbf{x}_i and set of peers N_i . Peer relationships form a binary directed network. Let $d_i = |N_i|$ denote i 's degree, i.e., the number of peers of i . Assume that no agent is isolated, $d_i > 0$.¹¹ Define \mathbf{G} as the linear-in-means matrix of interactions: $g_{ij} = \frac{1}{d_i}$ if $j \in N_i$ and 0 otherwise. The linear-in-means model of peer effects in networks can be written as

$$y_i = \alpha + \mathbf{x}_i \boldsymbol{\gamma} + \sum_j g_{ij} \mathbf{x}_j \boldsymbol{\delta} + \beta \sum_j g_{ij} y_j + \varepsilon_i, \quad (2)$$

under the assumption that $\mathbb{E}(\varepsilon_i | \mathbf{x}, \mathbf{G}) = 0$. In this model, outcomes can be affected by individual characteristics (individual effects, $\boldsymbol{\gamma}$), peers' characteristics (contextual peer effects, $\boldsymbol{\delta}$), and peers' outcomes (endogenous peer effects, β). Model (2) is a case of model (1) with $\boldsymbol{\theta} = (\alpha, \boldsymbol{\gamma}, \boldsymbol{\delta})$, $f_i(\mathbf{x}, \mathbf{G}, \boldsymbol{\theta}) = \alpha + \mathbf{x}_i \boldsymbol{\gamma} + \sum_j g_{ij} \mathbf{x}_j \boldsymbol{\delta}$, and the interaction matrix $\boldsymbol{\beta} = \beta \mathbf{G}$. In this model, the structure of the interactions \mathbf{G} is known but not their extent, β . The assumption $\mathbb{E}(\varepsilon_i | \mathbf{x}, \mathbf{G}) = 0$ means that characteristics and the network are strictly exogenous and the problem of correlated effects has been solved.¹² This framework has generally been applied to study continuous outcomes. We show below that it is also compatible with binary outcomes.

Our second main application is entry games. These games were introduced to study competition between a small number of firms in a large number of markets. Firm i 's decision to enter market m may depend on characteristics of the firm and the market and on the entry decisions of its competitors. In the literature, researchers generally consider nonlinear models of entry games, e.g., [Ciliberto and Tamer \(2009\)](#). In contrast, we consider the following linear model. Let $y_{im} \in \{0, 1\}$ denote the entry of firm i into market m . Then,

$$y_{im} = \alpha + \mathbf{x}_{im} \boldsymbol{\gamma} + \mathbf{z}_m \boldsymbol{\lambda} + \sum_j \beta_{ij} y_{jm} + \varepsilon_{im}, \quad (3)$$

under the assumption that $\mathbb{E}(\varepsilon_{im} | \mathbf{x}, \mathbf{z}) = 0$. Firm i 's entry depends on firm-market characteristics \mathbf{x}_{im} , on market characteristics \mathbf{z}_m , and on other firms' entries $\sum_j \beta_{ij} y_{jm}$. Observe that model (3) is a case of model (1) applied to firm-market observations with $\boldsymbol{\theta} = (\alpha, \boldsymbol{\gamma}, \boldsymbol{\lambda})$,

¹¹The model can easily be extended to account for isolated individuals.

¹²[Bramoullé, Djebbari, and Fortin \(2020\)](#) show that exogeneity of either a characteristic or the network can be sufficient to identify peer effects in model (2).

$f_{im}(\mathbf{x}_{im}, \mathbf{z}_m, \boldsymbol{\theta}) = \alpha + \mathbf{x}_{im}\boldsymbol{\gamma} + \mathbf{z}_m\boldsymbol{\lambda}$, and under the assumption that the interaction matrix $\boldsymbol{\beta}$ is constant across markets.

We now characterize when binary outcomes are statistically compatible with linear interactions, following arguments in Heckman and MaCurdy (1985). In what follows, the notation \mathbf{x} refers to a matrix containing all observables, including the network in a peer-effect application and market characteristics in an entry game. Note that equation (1) defines a fixed-point system in the outcome profile \mathbf{y} . In matrix notations,

$$\mathbf{y} = \mathbf{f} + \boldsymbol{\beta}\mathbf{y} + \boldsymbol{\varepsilon}.$$

We assume that the matrix $\mathbf{I} - \boldsymbol{\beta}$ is invertible. This holds generically and implies that this system has a unique solution. The reduced-form of model (1), expressing outcomes \mathbf{y} as a function of observables, parameters and errors, is equal to

$$\mathbf{y} = (\mathbf{I} - \boldsymbol{\beta})^{-1}\mathbf{f} + (\mathbf{I} - \boldsymbol{\beta})^{-1}\boldsymbol{\varepsilon}.$$

Let $P_i = [(\mathbf{I} - \boldsymbol{\beta})^{-1}\mathbf{f}]_i$ and $\nu_i = [(\mathbf{I} - \boldsymbol{\beta})^{-1}\boldsymbol{\varepsilon}]_i$. Here, $P_i = \mathbb{E}(y_i|\mathbf{x}) = \mathbb{P}(y_i = 1|\mathbf{x})$ is the conditional expected outcome and hence must lie between 0 and 1. Then, $\nu_i = y_i - \mathbb{E}(y_i|\mathbf{x})$ is the reduced-form error of the data generating process. We have $y_i = P_i + \nu_i$, and $y_i = 0$ and $\nu_i = -P_i$ with probability $1 - P_i$, while $y_i = 1$ and $\nu_i = 1 - P_i$ with probability P_i . Then, $\boldsymbol{\varepsilon} = (\mathbf{I} - \boldsymbol{\beta})\boldsymbol{\nu}$, leading to our first result.

Theorem 1. *Assume that $\mathbf{I} - \boldsymbol{\beta}$ is invertible and that $\forall i, P_i = [(\mathbf{I} - \boldsymbol{\beta})^{-1}\mathbf{f}]_i \in [0, 1]$. Outcomes in the unique solution to model (1) are binary, $y_i \in \{0, 1\}$, if and only if*

$$\varepsilon_i = \nu_i - \sum_j \beta_{ij}\nu_j,$$

where $\nu_i = -P_i$ with probability $1 - P_i$ and $1 - P_i$ with probability P_i .

Theorem 1 shows how to extend the stochastic structure of the linear probability model to account for outcome interactions. In the standard case without interactions, $\boldsymbol{\beta} = \mathbf{0}$ and $\varepsilon_i = \nu_i = -f_i$ with probability $1 - f_i$ and $1 - f_i$ with probability f_i . Under interactions, by

contrast, the ε_i 's are specific linear functions of the ν_i 's. For instance in model (2) of peer effects in networks, ε_i generally takes 2^{d_i+1} values. Further, ε_i and ε_j are generally correlated if i and j are peers or peers of peers, and even when the ν_i 's are uncorrelated. Thus, errors have a specific discrete structure, which depends on the network of interactions.

Theorem 1 demonstrates that writing Equation (1) under $\mathbb{E}(\varepsilon_i|\mathbf{x}) = 0$ is equivalent to assuming $P(y_i = 1|\mathbf{x}) = [(\mathbf{I} - \boldsymbol{\beta})^{-1}\mathbf{f}]_i$. As for any model with binary outcomes, we can then rewrite the model using additive errors. These additive errors generally have no structural interpretation. We analyze the microfoundations of model (1) in the next Section, and show that preference shocks in underlying microfoundations can have very different properties. In particular preference shocks can be iid and uniformly distributed over an interval - see Theorem 2 - even though the additive errors in the econometric model are discrete and correlated.

Probabilities must of course lie between 0 and 1. This is guaranteed in model (1) with binary outcomes when for any i , $[(\mathbf{I} - \boldsymbol{\beta})^{-1}\mathbf{f}]_i \in [0, 1]$. This condition depends both on interactions $\boldsymbol{\beta}$ and expected outcomes in the absence of interactions \mathbf{f} . In any application we can easily compute the proportion of observations for which the estimated probability lies between 0 and 1. As with the standard linear probability model, this provides a simple measure of whether the estimated model is appropriate. We report these proportions in our estimations in Section 5.

A key property of the linear framework is that if $\mathbf{I} - \boldsymbol{\beta}$ is invertible, there is a unique solution to the fixed-point system defined by model (1). In other words, the econometric model is both coherent and complete, see Tamer (2003) and Lewbel (2019). Furthermore, identification in a linear framework follows from well-known results. For the linear-in-means model of peer effects in networks, the identification results of Bramoullé, Djebbari, and Fortin (2009) apply when outcomes are binary under the assumptions of Theorem 1. This holds because their analysis does not impose restrictions on the nature of the outcome or on the error terms, other than the exogeneity assumption $\mathbb{E}(\varepsilon_i|\mathbf{x}, \mathbf{G}) = 0$.

Corollary 1. *(Bramoullé, Djebbari, and Fortin, 2009) Consider the linear-in-means model of peer effects in networks with binary outcomes and under the assumptions of Theorem 1. Assume $\boldsymbol{\delta} + \beta\boldsymbol{\gamma} \neq \mathbf{0}$. The model is identified if the matrices \mathbf{I} , \mathbf{G} , and \mathbf{G}^2 are linearly*

independent.

Identification notably holds when the network’s diameter is greater than or equal to 2, or under group interactions when group sizes differ.

With entry games, model (3) is generally identified when the entry of firm i is affected by some firm-market characteristic x_{im} that does not directly affect the entry of other firms, a standard assumption in the literature. The entry of firm j can thus be instrumented by x_{jm} in equation (3), see e.g., [Bajari, Hong, and Nekipelov \(2013\)](#). More generally, model (1) defines simultaneous linear equations in outcomes. When functions f_i are also linear, classical rank conditions for identification apply, see e.g., [Wooldridge \(2010, Section 9\)](#).

Overall, [Theorem 1](#) demonstrates that, from a statistical and econometric point of view, interactions in binary outcomes can be analyzed with a linear model. In some contexts, the outcome of interest is also a choice, and this raises the question of the microfoundations behind model (1).

3 Microfoundations

We now analyze the microfoundations of model (1), under the assumption that outcome $y_i \in \{0, 1\}$ is a choice of agent i . We consider games of incomplete or complete information under the assumptions of [Theorem 1](#) that $\mathbf{I} - \boldsymbol{\beta}$ is invertible and $P_i = [(\mathbf{I} - \boldsymbol{\beta})^{-1}\mathbf{f}]_i \in [0, 1]$.

We adopt standard assumptions of the literature on the econometrics of games. We consider a classical additive random utility framework ([McFadden, 1974](#)). Agent i derives utility $v_i(y_i, \mathbf{y}_{-i})$ from playing y_i when other agents play \mathbf{y}_{-i} . Utility $v_i(y_i, \mathbf{y}_{-i})$ is the sum of deterministic utility $u_i(y_i, \mathbf{y}_{-i})$ and preference shock $e_i(y_i)$. Let $\Delta u_i(\mathbf{y}_{-i}) = u_i(1, \mathbf{y}_{-i}) - u_i(0, \mathbf{y}_{-i})$ denote the relative deterministic utility of playing 1, and let $e_i = e_i(1) - e_i(0)$ denote the relative preference shock in the utility of playing 1. Under incomplete information, the deterministic utilities of all agents and the distribution of preference shocks are common knowledge. Agent i observes the realization of her own shock e_i but not the realization of others’ shocks \mathbf{e}_{-i} . Under complete information, the deterministic utilities and preference shocks of all agents are common knowledge.

Existing studies of interactions in binary outcomes generally assume that relative utility

is linear in others' actions:

$$\Delta u_i(\mathbf{y}_{-i}) = f_i - \frac{1}{2} + \sum_j \beta_{ij} y_j, \quad (4)$$

and we maintain this assumption in our main results below.¹³

3.1 Incomplete information

Under incomplete information, agent i observes her preference shock e_i but not others' shocks \mathbf{e}_{-i} . A strategy of agent i is a function of own shock $y_i(e_i) \in \{0, 1\}$. Outcomes \mathbf{y} are assumed to form a Bayes-Nash equilibrium of the game for any realization of preference shocks \mathbf{e} . Conditional on others' strategies, agent i plays a best response and her beliefs on others' play are consistent with their strategies.

In our second main result, we characterize the incomplete information microfoundations of model (1) under independence, a standard assumption. Denote by F_i the CDF of $-e_i$ and assume that $e_i \perp \mathbf{e}_{-i}, \mathbf{x}$. When beliefs are consistent, i 's belief that j plays 1 is equal to the probability that j plays 1. Under independence, $\mathbb{P}(y_j = 1 | e_i, \mathbf{x}) = \mathbb{P}(y_j = 1 | \mathbf{x})$, and let $p_j = \mathbb{P}(y_j = 1 | \mathbf{x})$. Under linear utilities (4), agent i 's expected relative utility of playing 1 is equal to

$$\mathbb{E}u_i(1|e_i) - \mathbb{E}u_i(0|e_i) + e_i = f_i - \frac{1}{2} + \sum_j \beta_{ij} p_j + e_i,$$

and agent i plays 1 when her expected utility is positive. This is equivalent to $-e_i < f_i - \frac{1}{2} + \sum_j \beta_{ij} p_j$. Since $p_i = \mathbb{P}(y_i = 1 | \mathbf{x})$, a Bayes-Nash equilibrium under independence is characterized by the following system of equations:¹⁴

$$p_i = F_i\left(f_i - \frac{1}{2} + \sum_j \beta_{ij} p_j\right). \quad (5)$$

¹³Note that linear relative utility (4) is consistent with quadratic utility $u_i(y_i, \mathbf{y}_{-i}) = y_i f_i - \frac{1}{2} y_i^2 + \sum_j \beta_{ij} y_i y_j$. This quadratic utility has been extensively studied in network games with continuous actions, see [Ballester, Calvó-Armengol, and Zenou \(2006\)](#) and [Bramoullé, Kranton, and D'amours \(2014\)](#), and has been proposed as a microfoundation of the econometric model of peer effects with continuous outcomes, see e.g., [Davezies, d'Haultfoeuille, and Fougère \(2009\)](#). This provides a common game-theoretic framework for binary and continuous action games.

¹⁴Equation (1) in [Lee, Li, and Lin \(2014\)](#) is a particular case of equation (5) in a model of peer effects and when shocks are identically distributed.

This is, in general, a nonlinear fixed point system which can have multiple solutions. In our next result, we show that the linear model of interactions corresponds to the unique interior Bayes-Nash equilibrium of the incomplete information game with linear utilities when preference shocks are iid and uniformly distributed. This is also the unique Bayes-Nash equilibrium when interactions are not too large in magnitude. We also derive a necessary condition. Say that β and \mathbf{f} take all possible values when $(\mathbf{I} - \beta)^{-1}\mathbf{f}$ takes all possible values in $[0, 1]^n$. We show that when preference shocks are independent between agents and independent of observables, the linear model of interactions corresponds to a Bayes-Nash equilibrium of the game for all possible values of β and \mathbf{f} only if preference shocks are iid uniformly distributed.

Theorem 2. *Assume that reduced-form errors ν_i are uncorrelated. Consider the game of incomplete information with linear relative utilities (4) and preference shocks e_i . (Sufficiency): If the e_i 's are iid and uniformly distributed over $[-\frac{1}{2}, \frac{1}{2}]$, then the unique solution to the linear model of interactions (1) is the unique interior Bayes-Nash equilibrium of the game. If in addition $\forall i, \sum_j |\beta_{ij}| < 1$, this is the unique Bayes-Nash equilibrium. (Necessity): Assume that $e_i \perp \mathbf{e}_{-i}, \mathbf{x}$. If the unique solution to the linear model of interactions (1) is a Bayes-Nash equilibrium of the game for all possible values of β and \mathbf{f} , then the e_i 's are iid and uniformly distributed over $[-\frac{1}{2}, \frac{1}{2}]$.*

Our proof in Appendix relies on a well-known property, i.e., that a uniform distribution is the only distribution with a linear CDF over its convex support.¹⁵ The uniqueness condition comes from a classical contraction mapping argument and we show in Appendix that uniqueness holds when $\|\beta\| < 1$ for any submultiplicative matrix norm $\|\cdot\|$.¹⁶ It is well-known that uniqueness holds under incomplete information when interactions are moderate, see e.g., Brock and Durlauf (2001) and Lee, Li, and Lin (2014). However, existing applications typically assume that preference shocks have full support, thereby leading to nonlinear econometric models.

We also show in Appendix that the game can have multiple Bayes-Nash equilibria when the uniqueness condition is not satisfied. Under multiplicity, the estimation of the linear

¹⁵Precisely, $F_i(e) = \min(\max(e + \frac{1}{2}, 0), 1)$ when e_i is uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]$.

¹⁶A matrix norm is submultiplicative if $\|\mathbf{MN}\| \leq \|\mathbf{M}\|\|\mathbf{N}\|, \forall \mathbf{M}, \mathbf{N}$.

model of interactions (1) is valid under an equilibrium selection assumption, namely that the data is generated by the unique interior equilibrium of the incomplete information game. This contrasts with the estimator proposed in Lee, Li, and Lin (2014), which is only valid if there exists a unique equilibrium. In any case, the uniqueness condition of Theorem 2 is easy to verify for estimated parameters, and we will see that it holds in the empirical applications in Section 5.

Theorem 2 illustrates key differences between stochastic shocks on underlying utilities and errors in the econometric model. Here, preference shocks are iid - hence independent of observables - and continuously distributed over an interval. By contrast and as shown in Theorem 1, errors in model (1) depend on observables, have a discrete support, and display a specific correlation pattern.

In our view, Theorem 2 demonstrates that models of linear interactions in binary outcomes admit reasonable game-theoretic microfoundations. The random utility framework and the independence assumption are standard in the literature, e.g., De Paula (2013). The only unusual assumption is that preference shocks are uniformly distributed over an interval. We see no valid reason to reject this assumption *a priori*. On the contrary, we would argue that a uniform distribution may be more convenient than the usual logistic and normal distribution, since it gives rise to a linear model of interactions.

3.2 Complete information

Under complete information, outcomes \mathbf{y} are assumed to form a Nash equilibrium of the game for any realization of preference shocks \mathbf{e} . This means that for every agent i , if $y_i = 1$ then $\Delta u_i(\mathbf{y}_{-i}) + e_i \geq 0$, whereas if $y_i = 0$ then $\Delta u_i(\mathbf{y}_{-i}) + e_i \leq 0$. Let $\mathbb{1}(Y_i \geq 0) = 1$ if $Y_i > 0$ and 0 if $Y_i < 0$. Therefore, $y_i = \mathbb{1}(\Delta u_i(\mathbf{y}_{-i}) + e_i \geq 0)$ and the system of equilibrium conditions is equivalent to a latent variable model with interactions. Under linear utility (4), this yields

$$y_i = \mathbb{1}\left(f_i(\mathbf{x}, \boldsymbol{\theta}) - \frac{1}{2} + \sum_j \beta_{ij} y_j + e_i \geq 0\right). \quad (6)$$

An important early finding is that multiple equilibria necessarily appear in games of complete information with linear utility when interactions are positive and preference shocks have full

support over \mathbb{R} , see [Tamer \(2003\)](#) and [De Paula \(2013\)](#). Multiplicity appears when preference shocks take intermediate values.¹⁷

In our next result, we characterize the complete information microfoundations of model (1). We obtain a necessary and sufficient condition on preference shocks under which the linear model of interactions corresponds to a Nash equilibrium of the complete information game with linear utilities and a sufficient condition under which this is the unique Nash equilibrium in dominant strategies.

Proposition 1. *Consider the game of complete information with linear relative utilities (4) and preference shocks e_i . The unique solution to the linear model of interactions (1) is a Nash equilibrium of the game if and only if $\nu_i > 0 \Rightarrow e_i \geq \varepsilon_i - \frac{1}{2}$ and $\nu_i < 0 \Rightarrow e_i \leq \varepsilon_i + \frac{1}{2}$. It is the unique Nash equilibrium in dominant strategies if $\nu_i > 0 \Rightarrow e_i > \varepsilon_i - \frac{1}{2} + \sum_j |\beta_{ij}|$ and $\nu_i < 0 \Rightarrow e_i < \varepsilon_i + \frac{1}{2} - \sum_j |\beta_{ij}|$.*

Proposition 1 describes how preference shocks must be related to the errors of the econometric model. Intuitively, preference shocks must be large enough in situations where the agent plays 1 and small enough in situations where the agent plays 0.

Proposition 1 implies that the linear model of interactions (1) is a particular case of the latent variable model with interactions (6). When preference shocks satisfy the first set of inequalities described in Proposition 1, the linear model corresponds to one of the possible solutions of the system of equations defined by model (6). When preference shocks satisfy the second set of inequalities, the linear model corresponds to the unique solution and hence the two models are formally equivalent.

Proposition 1 also shows that very different kinds of preference shock are consistent with the econometric model. Preference shocks can notably be independent and continuously distributed. To see why, consider a situation where the ν_i 's are uncorrelated. Assume that the e_i 's are independent and that e_i is continuously distributed with no probability on $[-L_i, M_i]$, an overall probability of $1 - P_i$ on $] -\infty, -L_i]$, and an overall probability of P_i on $[M_i, \infty[$. Further assume $L_i, M_i \geq \frac{1}{2} + 2 \sum_j |\beta_{ij}|$. This captures, for instance, situations where the

¹⁷To see why, consider the linear relative utility (4). Note that $v_i(1, \mathbf{0}) - v_i(0, \mathbf{0}) = f_i - \frac{1}{2} + e_i$. Then, $(0, 0, \dots, 0)$ is an equilibrium iff $\forall i, e_i \leq -(f_i - \frac{1}{2})$. Similarly, $v_i(1, \mathbf{1}) - v_i(0, \mathbf{1}) = f_i - \frac{1}{2} + \sum_j \beta_{ij} + e_i$, and $(1, 1, \dots, 1)$ is an equilibrium iff $\forall i, e_i \geq -(f_i - \frac{1}{2}) - \sum_j \beta_{ij}$. Therefore, both $\mathbf{0}$ and $\mathbf{1}$ are Nash equilibria iff $\forall i, -(f_i - \frac{1}{2}) - \sum_j \beta_{ij} \leq e_i \leq -(f_i - \frac{1}{2})$.

relative cost of playing 1 is bimodal, either quite high or quite low. Proposition 1 shows that in this case, the unique solution to the linear model of interactions is the unique Nash equilibrium of the corresponding game. As with incomplete information, preference shocks under complete information can thus have very different properties from errors in the econometric model.

More generally, the linear model of interactions corresponds to the unique Nash equilibrium in dominant strategies when preference shocks are sufficiently dispersed: sufficiently high when high and sufficiently low when low. Preference shocks do not therefore take intermediate values, bypassing the multiplicity domain. This implies that even in the presence of multiple Nash equilibria, the linear model of interactions becomes the unique equilibrium following specific changes in preference shocks. The linear model of interactions is thus, in a sense, the only robust Nash equilibrium. We now develop these arguments formally. In our next result, we show that this reasoning holds for any deterministic utility and preference shocks.

Proposition 2. *Suppose that the unique solution to model (1) is a Nash equilibrium of the game of complete information with deterministic utilities U_i and preference shocks e_i . Consider preference shocks e'_i where $e'_i = e_i + M_i$ if $\nu_i > 0$ and $e_i - L_i$ if $\nu_i < 0$ and $L_i, M_i \geq 0$. Then, the unique solution to model (1) remains a Nash equilibrium of the game of complete information with deterministic utilities U_i and preference shocks e'_i , and it is the unique Nash equilibrium in dominant strategies when L_i and M_i are sufficiently large.*

Next, say that a Nash equilibrium \mathbf{y} for preference shocks \mathbf{e} is *robust to increases in shock dispersion* when \mathbf{y} remains an equilibrium for preference shocks e'_i where $e'_i = e_i + M_i$ if $\nu_i > 0$ and $e_i - L_i$ if $\nu_i < 0$ and $L_i, M_i \geq 0$.¹⁸

Corollary 2. *Suppose that the unique solution to model (1) is a Nash equilibrium of the game of complete information with deterministic utilities U_i and preference shocks e_i . Then, this is the unique equilibrium robust to increases in shock dispersion.*

¹⁸An extensive game-theoretic literature on robust equilibria proposes various definitions of robustness. Key, as here, is the idea that an equilibrium is robust when it remains an equilibrium following perturbations of the underlying game (e.g., Trembling Hand perfection, Selten (1975)).

Even in the presence of multiple equilibria, therefore, the estimation of the linear model of interactions (1) is valid under the assumption that the data is generated by the unique robust Nash equilibrium. Overall, the results in this Subsection show that the linear model of interactions can be microfounded as an outcome of a game of complete information, and they clarify the conditions under which this can be done.

3.3 Discussion

To conclude, it should be pointed out that our results on incomplete and complete information are quite different in nature. We think that the microfoundations under incomplete information described in Theorem 2 are *a priori* reasonable. The main difference with respect to the usual approaches described in the literature is the assumption that preference shocks are uniformly distributed. This assumption is strong - but no stronger, in our view, than the usual assumptions of a logistic or normal distribution.

By contrast, the *a priori* appeal of microfoundations under complete information is less clear. In particular, assuming that reduced-form errors are uncorrelated, a direct consequence of Proposition 1 is the following impossibility result. The linear model of interactions *cannot* be obtained as a Nash equilibrium of a game of complete information with linear utilities when preference shocks are independent and continuously distributed over a convex support.¹⁹ Thus, either preference shocks must be correlated or the support of their distribution must be non-convex.

Our results also confirm that microfoundations are generally not identifiable without additional assumptions. Combining Theorem 2 and Proposition 1 shows that the same data generating process can be obtained as a unique Bayes-Nash equilibrium of a game of incomplete information with preference shocks that are iid uniformly distributed over an interval or as a unique Nash equilibrium of a game of complete information with preference shocks that are iid and continuously distributed over a non-convex support. Thus, even holding deterministic utilities constant, the nature of the preference shocks and of the information available to the agents cannot be identified from the data.

¹⁹To see why, note that $\varepsilon_i - \frac{1}{2}$ when $\nu_i > 0$ and $\varepsilon_i + \frac{1}{2}$ when $\nu_i < 0$ are both equal to $\frac{1}{2} - P_i - \sum_j \beta_{ij} \nu_j$. If the support of e_i is convex, this value lies in this support. By Proposition 1, e_i then depends on ν_j if $\beta_{ij} \neq 0$. When $\beta \neq \mathbf{0}$, therefore, underlying preference shocks with convex support can not be independent.

4 Estimators

In this Section, we propose two simple estimators to analyze data with binary outcomes generated by model (1). We consider a many-groups asymptotic framework with independent groups of bounded size. This corresponds to a many-network asymptotic framework in a peer effect setting²⁰ or to a many-market asymptotic framework for entry games. The number of groups thus goes to infinity with sample size, and consistency and the asymptotic normality of extremum estimators are guaranteed under standard assumptions, see, e.g., [Cameron and Trivedi \(2005\)](#), Section 5.3.²¹ We explore the small sample properties of our proposed estimators through Monte Carlo simulations and apply these estimators to real data in Section 5.

We consider the following variant of model (1). We assume that \mathbf{f} is linear in observables and that the interaction matrix $\boldsymbol{\beta}$ depends linearly on a fixed number of parameters to estimate. Formally, $\mathbf{f} = \mathbf{X}\boldsymbol{\theta}$ and $\boldsymbol{\beta} = \sum_{k=1}^K \beta_k \mathbf{G}_k$, where K is finite and independent of sample size. This yields

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \sum_{k=1}^K \beta_k \mathbf{G}_k \mathbf{y} + \boldsymbol{\varepsilon}. \quad (7)$$

While not essential, the linearity of \mathbf{f} is sufficient for most applications and facilitates exposition. The assumption on the interaction matrix ensures that the number of parameters to estimate does not grow with sample size²² and allows us to provide a common framework for our two applications. Denote by $\boldsymbol{\beta}_{(K)} = (\beta_1, \dots, \beta_K)$ and similarly for $\mathbf{G}_{(K)}$.

In the benchmark linear-in-means model of peer effects in networks, there is only one interaction parameter to estimate. In that model, $K = 1$ and $\boldsymbol{\beta} = \beta \mathbf{G}$, where $\mathbf{G}_{ij} = 1/d_i$ if i and j are linked and 0 otherwise. Extended versions of the model with heterogeneous peer effects, as in [Nakajima \(2007\)](#), [Soetevent and Kooreman \(2007\)](#), and [Dieye and Fortin \(2017\)](#), are also cases of model (7). For instance, when men and women can be differentially

²⁰Population is then partitioned into groups, and agents can only be affected by others in their own group. The overall network is composed of disjoint subnetworks with a block diagonal interaction matrix.

²¹The analysis of estimators' properties in a single-network asymptotic framework is an active area of research, see e.g., [Lee \(2004\)](#), [Menzel \(2016\)](#), and [Leung \(2016\)](#).

²²[Peng \(2019\)](#) proposes a penalized regression strategy that depends on the weaker assumption that $K \leq \frac{c\sqrt{n}}{\ln n}$ for some c as sample size n goes to infinity. However, his analysis depends on the errors being independent and sub-Gaussian.

affected by their male and female peers, there are $K = 4$ interaction parameters to estimate. In that case, $\boldsymbol{\beta} = \beta_{FF}\mathbf{G}_{FF} + \beta_{FM}\mathbf{G}_{FM} + \beta_{MM}\mathbf{G}_{MM} + \beta_{MF}\mathbf{G}_{MF}$, where, for example, \mathbf{G}_{FM} models the structure of interactions between female individuals and male peers.

In the linear model of entry games, there are $K = n(n - 1)$ interaction parameters to estimate, where n is the number of firms competing across markets. Here, $\boldsymbol{\beta} = \sum_{i,j} \beta_{ij}\mathbf{G}_{ij}$, where \mathbf{G}_{ij} has 1 at entry (i, j) and 0 elsewhere. In the next Section, we estimate a version of this model where the entry of a firm has a common impact on the entry of other firms, i.e., $\beta_{ij} = \beta_j$. In that version, there are $K = n$ parameters to estimate, and $\boldsymbol{\beta} = \sum_{j=1}^N \beta_j\mathbf{G}_j$, where $\mathbf{G}_j = \sum_i \mathbf{G}_{ij}$ has 1 in its j th column and 0 elsewhere.

We assume that the model to be estimated is identified, see Section 2. The linear-in-means model of peer effects is identified under conditions described in Corollary 1. Similar conditions hold when peer effects are heterogeneous, see e.g., Dieye and Fortin (2017). In the linear model of entry games, identification holds under the exclusion restrictions that some characteristics of firm j that affect its profit do not affect the profit of firm i .

Our first proposed estimator is a Two-Stage Least Squares (2SLS) estimator, building on the 2SLS estimation strategies proposed by Kelejian and Prucha (1998) and Bramoullé, Djebbari, and Fortin (2009). Since $\mathbb{E}(\boldsymbol{\varepsilon}|\mathbf{X}, \mathbf{G}_{(K)}) = 0$, we have

$$\mathbb{E}(\mathbf{G}_k\mathbf{y}|\mathbf{X}, \mathbf{G}_{(K)}) = \mathbf{G}_k\mathbf{X}\boldsymbol{\theta} + \sum_{l=1}^K \beta_l\mathbf{G}_k\mathbf{G}_l(\mathbf{I} - \boldsymbol{\beta})^{-1}\mathbf{X}\boldsymbol{\theta}.$$

In particular, variables in $\mathbf{G}_k\mathbf{X}$ that are not already in \mathbf{X} provide natural instruments for $\mathbf{G}_k\mathbf{y}$ in equation (7).

In the linear-in-means model (2), we have $\mathbf{X} = [\mathbf{1}, \mathbf{x}, \mathbf{G}\mathbf{x}]$ and hence $\mathbf{G}\mathbf{X} = [\mathbf{G}\mathbf{1}, \mathbf{G}\mathbf{x}, \mathbf{G}^2\mathbf{x}]$, so $\mathbf{G}^2\mathbf{x}$ can be used as an instrument for $\mathbf{G}\mathbf{y}$. This instrument is valid under the conditions described in Corollary 1. Intuitively, characteristics of peers of peers who are not peers affect individual outcome only through their impact on peers' outcomes, see Bramoullé, Djebbari, and Fortin (2009). In the linear model of entry games (3), for any market m , we have $\mathbf{X}_m = [\mathbf{1}, \mathbf{x}_m, \mathbf{z}_m]$ and, hence, $\mathbf{G}_{ij}\mathbf{X}_m = [\mathbf{G}_{ij}\mathbf{1}, \mathbf{G}_{ij}\mathbf{x}_m, \mathbf{G}_{ij}\mathbf{z}_m]$. Here, $[\mathbf{G}_{ij}\mathbf{y}]_i = y_j$ can be instrumented by $[\mathbf{G}_{ij}\mathbf{x}_m]_i = \mathbf{x}_{jm}$. The impact of the entry of firm j on i 's entry can be instrumented by the characteristics of firm j .

The validity of this IV strategy relies on the exogeneity condition $\mathbb{E}(\varepsilon_i|\mathbf{X}, \mathbf{G}_{(K)}) = 0$ but not on the specific structure of the errors. This strategy is thus valid when errors have the discrete structure uncovered in Theorem 1. Error structure may, of course, matter for inference. In particular, errors in the linear-in-means model (2) with binary outcomes are heteroscedastic and correlated among peers and peers of peers, as shown in Section 2. In a many-groups asymptotic framework, we propose to use group-level cluster-robust standard errors for inference since they allow for arbitrary within-group correlations, see [Cameron and Miller \(2015\)](#).²³

Our second proposed estimator is a Nonlinear Least Squares (NLS) estimator, exploiting the structure of the reduced-form equations. Recall that $y_i = P_i + \nu_i$ with $\mathbb{E}(\nu_i|\mathbf{x}) = 0$, and $\mathbf{P}(\boldsymbol{\beta}_{(K)}, \boldsymbol{\theta}) = (\mathbf{I} - \sum_{k=1}^K \beta_k \mathbf{G}_k)^{-1} \mathbf{X} \boldsymbol{\theta}$. The model can thus be consistently estimated by the following Nonlinear Least Squares estimator:

$$(\hat{\boldsymbol{\beta}}_{(K)}, \hat{\boldsymbol{\theta}}) = \arg \min_{\boldsymbol{\beta}_{(K)}, \boldsymbol{\theta}} [\mathbf{y} - \mathbf{P}(\boldsymbol{\beta}_{(K)}, \boldsymbol{\theta})]' [\mathbf{y} - \mathbf{P}(\boldsymbol{\beta}_{(K)}, \boldsymbol{\theta})]. \quad (8)$$

Conveniently, the estimator in (8) can be concentrated around $\boldsymbol{\beta}_{(K)}$. Indeed, taking the first-order conditions with respect to $\boldsymbol{\theta}$, conditional on $\boldsymbol{\beta}_{(K)}$, we obtain

$$\hat{\boldsymbol{\theta}}(\boldsymbol{\beta}_{(K)}) = [\mathbf{Z}'(\boldsymbol{\beta}_{(K)}) \mathbf{Z}(\boldsymbol{\beta}_{(K)})]^{-1} \mathbf{Z}'(\boldsymbol{\beta}_{(K)}) \mathbf{y},$$

where $\mathbf{Z}(\boldsymbol{\beta}_{(K)}) = (\mathbf{I} - \sum_{k=1}^K \beta_k \mathbf{G}_k)^{-1} \mathbf{X}$. Substituting in the objective function (8), we obtain the concentrated NLS estimator:

$$\hat{\boldsymbol{\beta}}_{(K)} = \arg \min_{\boldsymbol{\beta}_{(K)}} \mathbf{y}' [\mathbf{I} - \mathbf{Z}(\boldsymbol{\beta}_{(K)}) [\mathbf{Z}'(\boldsymbol{\beta}_{(K)}) \mathbf{Z}(\boldsymbol{\beta}_{(K)})]^{-1} \mathbf{Z}'(\boldsymbol{\beta}_{(K)})] \mathbf{y}. \quad (9)$$

While the objective function in (9) may not be convex, numerical optimization of the concentrated NLS is relatively straightforward when K is small.

To analyze interactions in binary outcomes, the 2SLS and NLS estimators are natural and

²³In other asymptotic frameworks, spatial Heteroscedastic and Autocorrelation Consistent (HAC) variance estimators could be appropriate (e.g., [Conley \(1999\)](#), [Kelejian and Prucha \(2007\)](#), [Leung \(2019\)](#)). Cluster-robust standard errors can still yield valid inferences in some situations where groups are not independent, see [Bester, Conley, and Hansen \(2011\)](#).

easy to implement with standard statistical software. There are, of course, other estimators that can be used to estimate model (7) and its variants. Following [Bramoullé, Djebbari, and Fortin \(2009\)](#), many studies proposed alternative strategies to estimate model (2) with continuous outcomes. Some of these strategies can be applied, or extended, to binary outcomes. Notably, these include moment-based estimators (e.g., [Kelejian and Prucha \(1998\)](#) and [Lee and Liu \(2010\)](#)) and different ways to compute instruments (e.g., [Kelejian and Piras \(2014\)](#)). In applied studies, researchers often instrument average peers' outcome by the average characteristics among a subset of peers of peers who are not peers, see [Nicoletti, Salvanes, and Tominey \(2018\)](#) and [De Giorgi, Frederiksen, and Pistaferri \(2020\)](#). In contrast, [Theorem 1](#) indicates that quasi-maximum likelihood approaches based on independence and homoscedasticity, as in [Lee, Liu, and Lin \(2010\)](#), cannot be used with binary outcomes.

Well-known estimators developed to estimate standard linear probability models can also be extended to account for social and strategic interactions. For instance, if the ν_i 's are uncorrelated, Maximum Likelihood and (feasible) Weighted Nonlinear Least Squares provide valid alternatives. While these estimators are more efficient, however, their practical implementation raises empirical and computational issues that will likely be aggravated by interactions.²⁴ In the absence of interactions, these practical considerations and the fact that actual efficiency gains appear to be small have led researchers to focus on Ordinary Least Squares, see Section 3.4.1 in [Angrist and Pischke \(2008\)](#). Our proposed 2SLS and NLS estimators provide natural counterparts of Ordinary Least Squares in a setup with interactions.

We next compare the small-sample performances of our two proposed two estimators using Monte Carlo simulations based on linear-in-means model (2). We let $x_i \sim U[0, 1]$. Note that setting $\alpha, \beta, \gamma, \delta > 0$, with $\alpha + \beta + \gamma + \delta < 1$, ensures that $P_i \in [0, 1]$ when $x_i \in [0, 1]$. We assume that the population is partitioned into $M > 0$ groups of size $N > 0$ such that $g_{ij} = 0$ whenever i and j belong to different groups. For any two agents i and j belonging to the same group, we let $g_{ij} = 1$ with a probability $p = 0.1$. The overall network is thus composed of $M = 500$ disjoint instances of Erdős-Renyi subnetworks connecting N agents.

²⁴Assuming error independence, [Amemiya \(1977\)](#) shows that weighted least squares is as efficient as the MLE. Moreover, since the MLE is maximized using a numerical algorithm, all proposed parameters have to be such that $P_i(\theta) \in (0, 1)$. A similar issue arises for the implementation of feasible weighted least squares: predicted probabilities used to weight the estimator have to fall between 0 and 1.

Results for two sets of parameters (high and low β), $N = 30$, $M = 100$ and 300 —corresponding to $NM = 3,000$ and $9,000$ observations—are presented in Table 1. Both estimators display moderate small-sample bias. Bias and estimates’ dispersion tend to be smaller for the NLS and when β is high. In the Appendix, we report results for the same parameters, $N = 20$ and 50 , $M = 500$ —corresponding to $10,000$ and $25,000$ observations—in Table 5. Bias is very low in all scenarios. Estimates’ dispersion is also low, and dispersion is lower for the NLS estimator and for larger groups. Overall, these simulations show that the 2SLS and NLS estimators perform well in small samples of artificial data where binary outcomes are subject to endogenous peer effects. The NLS appears to outperform the 2SLS, especially when β is high.

Nonetheless, the 2SLS estimator offers a major attractive feature for empirical applications: it can easily handle group fixed effects. Formally, suppose that α varies across groups $r = 1, \dots, M$. We have

$$\mathbf{y}_r = \alpha_r \mathbf{1} + \mathbf{X}_r \boldsymbol{\theta} + \sum_{k=1}^K \beta_k \mathbf{G}_{k,r} \mathbf{y}_r + \boldsymbol{\varepsilon}_r.$$

Since group size is bounded, the number of groups—and hence the number of parameters α_r to estimate—goes to infinity at the same rate as sample size. This is known as the *incidental parameter problem* and can notably yield inconsistent estimates of $\boldsymbol{\theta}$ and $\boldsymbol{\beta}_{(K)}$, see Lancaster (2000) for a review.

For linear models, however, a standard workaround is to rewrite the model in deviation from the group average (see Cameron and Trivedi (2005), Section 21.6). Let $\mathbf{J}_r = \mathbf{I}_r - \mathbf{1}_r \mathbf{1}'_r$. We obtain

$$\mathbf{J}_r \mathbf{y}_r = \mathbf{J}_r \mathbf{X}_r \boldsymbol{\theta} + \sum_{k=1}^K \beta_k \mathbf{J}_r \mathbf{G}_{k,r} \mathbf{y}_r + \mathbf{J}_r \boldsymbol{\varepsilon}_r,$$

which does not depend on α_r . The 2SLS strategy can then easily be adapted to estimate the model in deviation.²⁵ This issue may be critical in practice, in contexts where common unobservables may generate spurious correlations in outcomes.

²⁵Identification may of course be affected by the presence of fixed effects. Bramoullé, Djebbari, and Fortin (2009) derive identification conditions in variants of model (2) in the presence of group fixed effects.

Table 1: Monte Carlo Simulations – Number of Groups

High β					
$M = 100$			$M = 300$		
Parameters	2SLS	NLS	Parameters	2SLS	NLS
$\alpha = 0.1$	0.098 (0.031)	0.098 (0.029)	$\alpha = 0.1$	0.099 (0.017)	0.100 (0.015)
$\beta = 0.7$	0.712 (0.139)	0.705 (0.086)	$\beta = 0.7$	0.704 (0.077)	0.700 (0.045)
$\gamma = 0.05$	0.053 (0.042)	0.053 (0.041)	$\gamma = 0.05$	0.050 (0.024)	0.050 (0.023)
$\delta = 0.1$	0.088 (0.126)	0.096 (0.079)	$\delta = 0.1$	0.096 (0.070)	0.101 (0.042)
Low β					
$M = 100$			$M = 300$		
Parameters	2SLS	NLS	Parameters	2SLS	NLS
$\alpha = 0.1$	0.099 (0.027)	0.098 (0.027)	$\alpha = 0.1$	0.099 (0.015)	0.099 (0.015)
$\beta = 0.25$	0.284 (0.242)	0.262 (0.188)	$\beta = 0.25$	0.263 (0.134)	0.254 (0.104)
$\gamma = 0.2$	0.199 (0.040)	0.201 (0.040)	$\gamma = 0.2$	0.199 (0.024)	0.200 (0.023)
$\delta = 0.3$	0.081 (0.130)	0.095 (0.102)	$\delta = 0.3$	0.093 (0.070)	0.098 (0.056)

Note: For each simulation, M networks are generated among the $N = 30$ individuals using iid Bernoulli trials with a probability $p = 0.1$. Thus, the expected number of links for each individual is $0.1(N - 1)$. Values represent the average (standard deviation) of the 1000 simulations.

5 Applications

We now apply the linear framework to real data. To highlight how it differs from existing approaches, we revisit two studies: [Lee, Li, and Lin \(2014\)](#) on peer effects in teenage smoking and [Ciliberto and Tamer \(2009\)](#) on entry into airline markets. We reanalyze the same data used in the original studies.

5.1 Peer effects in teenage smoking

We first revisit the analysis of peer effects in teenage smoking by [Lee, Li, and Lin \(2014\)](#).²⁶ This study is based on data from the National Longitudinal Survey of Adolescent to Adult Health, or Add Health, which provides rich information on the outcomes, behaviors, and characteristics of middle and high school students in the US. The data notably include detailed information on self-reported friendship relationships, widely used to analyze peer effects in networks. For the sake of comparison, we focus on the same sample, outcomes, characteristics, and networks as [Lee, Li, and Lin \(2014\)](#).

The data come from Wave I of the In-School Add Health survey, collected from 1994 to 1995. The sample contains information on the smoking behavior of 74,783 students in 127 schools. [Lee, Li, and Lin \(2014\)](#) classify a student as a non-smoker if they declare never having smoked or having smoked only once or twice in the past twelve months. A student’s peers are his or her self-reported friends in the same school and grade. There are 532 school-grade groups, and hence the overall network is composed of 532 disjoint subnetworks. Summary statistics are presented in Table 6 of Appendix 7.1. The proportion of smokers among the students is 23%.

To analyze peer effects on binary outcomes, [Lee, Li, and Lin \(2014\)](#) develop an incomplete information framework, extending [Brock and Durlauf \(2001\)](#) to networks. In their framework, conditional probabilities $p_i = \mathbb{P}(y_i = 1|\mathbf{x})$ satisfy the following fixed-point equation (see equation (1) in [Lee, Li, and Lin \(2014\)](#)):

$$p_i = F\left(f_i - \frac{1}{2} + \beta \sum_j g_{ij} p_j\right), \quad (10)$$

²⁶We are grateful to the authors for providing the replication codes.

Since this is a particular case of equation (5), their econometric framework can be micro-founded as follows, see also Liu (2019). Outcomes \mathbf{y} correspond to a Bayes-Nash equilibrium of the game with incomplete information, linear utilities (4), iid preference shocks e_i and where $F(\cdot)$ is the CDF of $-e_i$. By contrast, model (2) yields $p_i = f_i + \beta \sum_j g_{ij} p_j$. Theorem 2 provides the conditions under which model (2) is a case of this framework: reduced-form errors must be uncorrelated, interactions must be moderate, and preference shocks must be uniformly distributed. Lee, Li, and Lin (2014) consider a logit framework in their empirical analysis. They assume that the probability that student i smokes tobacco is equal to

$$p_i = \frac{\exp(\alpha + \mathbf{x}_i \boldsymbol{\gamma} + \sum_j g_{ij} \mathbf{x}_j \boldsymbol{\delta} + \beta \sum_j g_{ij} p_j)}{1 + \exp(\alpha + \mathbf{x}_i \boldsymbol{\gamma} + \sum_j g_{ij} \mathbf{x}_j \boldsymbol{\delta} + \beta \sum_j g_{ij} p_j)}. \quad (11)$$

They propose to estimate the model via an iterative simulated maximum likelihood. Each iteration has two steps: solving for p_i 's in the nonlinear fixed-point equation (11), conditional on parameter values, and then re-estimating parameters through (simulated) maximum likelihood, conditional on these p_i 's. These two steps are repeated until convergence. Their preferred specification includes contextual and endogenous peer effects, fixed effects at the school level, and random effects at the school-grade level. They find evidence of statistically significant, positive endogenous peer effects, with estimates of β ranging from 0.598 to 0.665.

Their approach has two drawbacks. First, it is computationally demanding and involves a series of relatively high-dimensional nonlinear optimizations and fixed-point computations. This is likely to limit the method's applicability to other data and may make estimation unfeasible for larger data sets. Moreover, their estimator relies on a contraction property that is only valid for moderate interactions, when there exists a unique equilibrium. Second, and as discussed by Lee, Li, and Lin (2014) in Section IV.B, the model cannot be estimated in deviations. This complicates the inclusion of group fixed effects, a main means of controlling for correlated effects. Simply including group dummies may be computationally unfeasible and may bias the estimates due to the incidental parameter problem.

In contrast, these drawbacks do not arise with the 2SLS estimation of the linear model (2). Fixed effects can be eliminated by taking deviations from the group average. Moreover, efficient computation of 2SLS estimates is pre-programmed in standard statistical software and can be computed quickly even for massive data sets. We therefore reanalyze the same

Table 2: Peer Effects on Smoking

	NLS		2SLS		2SLS		2SLS	
Endogenous effect	0.545	(0.038)	0.608	(0.062)	0.588	(0.058)	0.568	(0.058)
Individual effects								
Constant	-1.043	(0.089)	-0.950	(0.141)	-	(-)	-	(-)
Age	0.159	(0.012)	0.147	(0.020)	0.131	(0.018)	0.102	(0.027)
Age ² /10	-0.044	(0.004)	-0.040	(0.007)	-0.034	(0.006)	-0.024	(0.009)
Years in school	0.001	(0.002)	0.001	(0.002)	-0.001	(0.002)	0.000	(0.002)
Male	0.005	(0.004)	0.007	(0.005)	0.006	(0.005)	0.005	(0.005)
Black	-0.172	(0.006)	-0.157	(0.007)	-0.141	(0.007)	-0.142	(0.007)
Asian	-0.080	(0.007)	-0.070	(0.008)	-0.056	(0.008)	-0.059	(0.008)
Hispanic	-0.080	(0.006)	-0.071	(0.009)	-0.036	(0.007)	-0.036	(0.007)
Other race	0.029	(0.007)	0.022	(0.007)	0.028	(0.007)	0.027	(0.007)
Live with both parents	-0.046	(0.004)	-0.039	(0.004)	-0.042	(0.004)	-0.042	(0.004)
Sports club	-0.040	(0.003)	-0.036	(0.004)	-0.039	(0.004)	-0.040	(0.004)
Mom education less than high school	0.010	(0.006)	0.007	(0.006)	0.010	(0.006)	0.009	(0.006)
Mom education more than high school	-0.012	(0.004)	-0.008	(0.004)	-0.007	(0.004)	-0.007	(0.004)
Mom education missing	-0.030	(0.005)	-0.026	(0.004)	-0.021	(0.004)	-0.021	(0.004)
Mom job is professional	0.022	(0.004)	0.019	(0.005)	0.018	(0.005)	0.019	(0.005)
Mom other jobs	0.024	(0.004)	0.021	(0.004)	0.021	(0.004)	0.021	(0.004)
Mom on welfare	0.027	(0.017)	0.025	(0.016)	0.027	(0.016)	0.024	(0.016)
Mom job is missing	0.014	(0.006)	0.008	(0.006)	0.009	(0.006)	0.009	(0.006)
Contextual effects								
Age	-0.007	(0.002)	-0.010	(0.003)	-0.009	(0.003)	-0.010	(0.003)
Age ² /10	-0.002	(0.001)	-0.001	(0.002)	-0.001	(0.002)	0.000	(0.002)
Years in school	-0.003	(0.002)	-0.003	(0.002)	-0.005	(0.002)	-0.004	(0.003)
Male	-0.006	(0.005)	-0.015	(0.006)	-0.018	(0.006)	-0.017	(0.006)
Black	0.062	(0.010)	0.065	(0.014)	0.062	(0.014)	0.059	(0.014)
Asian	0.001	(0.010)	0.000	(0.012)	0.012	(0.013)	0.010	(0.013)
Hispanic	-0.010	(0.009)	-0.004	(0.011)	0.026	(0.011)	0.024	(0.011)
Other race	0.028	(0.011)	0.016	(0.013)	0.022	(0.014)	0.024	(0.014)
Live with both parents	-0.033	(0.007)	-0.022	(0.009)	-0.027	(0.008)	-0.028	(0.008)
Sports club	-0.004	(0.005)	-0.003	(0.007)	-0.009	(0.007)	-0.010	(0.007)
Mom education less than high school	0.009	(0.009)	0.011	(0.011)	0.011	(0.011)	0.011	(0.011)
Mom education more than high school	-0.028	(0.006)	-0.020	(0.008)	-0.018	(0.008)	-0.018	(0.008)
Mom education missing	0.000	(0.009)	0.017	(0.011)	0.023	(0.011)	0.023	(0.011)
Mom job is professional	0.011	(0.007)	0.001	(0.009)	0.003	(0.009)	0.004	(0.009)
Mom other jobs	0.008	(0.006)	-0.001	(0.007)	0.002	(0.008)	0.004	(0.008)
Mom on welfare	-0.025	(0.026)	0.005	(0.031)	0.015	(0.031)	0.010	(0.031)
Mom job is missing	0.031	(0.010)	0.010	(0.011)	0.012	(0.011)	0.013	(0.011)
School fixed effects					X			
School-grade fixed effects					X			
Weak instruments					39.376	46.510	46.750	
Sargan					44.449	57.190	60.431	
Fraction predicted in [0, 1]	0.973		0.971		0.960			

Note: Estimated coefficients and associated standard errors (in parenthesis). Estimation of linear model (2). Outcome is smoking. Summary statistics are presented in Table 6 of Appendix 7.1. The number of observations is 74,783, the number of schools is 127, and the number of school-grades is 532. Standard errors are clustered at the grade-school level for the 2SLS estimators and are heteroscedastic-robust for the NLS estimator. Instruments for 2SLS estimations are generated using second-degree friends: $\mathbf{G}^2\mathbf{X}$. The weak instrument tests are based on first-stage F-tests. The test statistic under the null hypothesis that all instruments are weak follows a non-central χ^2 distribution (see Stock and Yogo (2005)). The null hypothesis for all specifications is rejected at a confidence level $< 1\%$. The null hypothesis of the Sargan test is that all instruments are exogenous. The test statistic follows a χ^2 distribution under the null hypothesis. The null hypothesis is rejected at $< 1\%$ for all specifications.

data assuming that model (2) holds.

Our estimation results are presented in Table 2. We consider specifications without fixed effects (NLS estimates in Column 1 and 2SLS estimates in Column 2), with school fixed effects (2SLS estimates in Column 3), and with school-grade fixed effects (2SLS estimates in Column 4). Our estimates of the endogenous peer effects are remarkably similar to those obtained in Lee, Li, and Lin (2014): 0.568 in our preferred specification, compared to 0.666 in theirs. Standard errors are of similar magnitude, and this coefficient is very precisely estimated. Incorporating school-grade fixed effects only slightly decreases the estimate of the endogenous peer effect. The proportion of observations having a predicted probability between 0 and 1 lies between 96% and 97%.²⁷ Furthermore, in all specifications, the estimated coefficient satisfies the uniqueness condition of Theorem 2.²⁸

The characteristics of the students and their peers also affect smoking behavior. Another well-known advantage of a linear formulation is that the marginal impact of a characteristic on the outcome is simply equal to the characteristic’s estimated coefficient.²⁹ For instance, being Black rather than White is associated with a 0.14 decrease in likelihood of being a smoker. Students living with both parents and those with a high school-educated mother are less likely to smoke. Interestingly, these beneficial effects appear to spill over to students’ friends: a student with friends who live with both parents or with a high school-educated mother is also less likely to be a smoker.

Finally, we see that the joint validity of the instruments is rejected by overidentification tests. This is perhaps not surprising since there are 17 instruments here for one endogenous variable. Even though the instruments, jointly, appear to be strong, some instruments are likely to be weak. How best to estimate IV regressions and test overidentification in the presence of many weak instruments is an active area of research, see e.g., Davidson and MacKinnon (2015), Carrasco and Tchuente (2016), and Tchuente (2019).

²⁷Note that this proportion cannot be computed when the model is estimated in deviations.

²⁸In linear-in-means model (2), this condition is equivalent to $|\beta| < 1$.

²⁹In Lee, Li, and Lin (2014), this corresponds to their “naive” estimation of the marginal effects. Our estimates and theirs have similar signs - although their estimated marginal effects are generally larger in absolute value.

5.2 Entry into airline markets

In this Section, we revisit the analysis of entry into airline markets by [Ciliberto and Tamer \(2009\)](#). For the sake of comparison, we analyze the same sample, variables, and data as [Ciliberto and Tamer \(2009\)](#). The main data come from the 2001 Airline Origin and Destination Survey, a 10% sample of tickets collected by the US Department of Transportation. A market is defined as the trip between two airports, irrespective of intermediate transfer points and the direction of the flight. The sample includes 2,742 markets. Six firms are assumed to compete across all markets: American (AA), Delta (DL), United (UA), Southwest (WN), and two “composite” firms: Medium Airlines (MA) and Low-Cost Carriers (LCC). Each firm i is either present in or absent from market m , $y_{im} \in \{0, 1\}$. The data include 10 variables assumed to be exogenous: 8 market-level variables z_m and 2 firm-market-level variables, x_{im} : “airport presence” and “cost”. We present summary statistics of these variables in [Table 7](#) of [Appendix 7.1](#).

[Ciliberto and Tamer \(2009\)](#) develop an econometric framework under the assumptions that firms play a static game of complete information with linear relative utility and that preference shocks are continuous and independent from observables. Multiple equilibria then arise when preference shocks take intermediate values. Introduce conditional choice probabilities, $\mathbb{P}(\mathbf{y}|\mathbf{x}, \mathbf{z})$, as the probabilities of observing entry decisions $\mathbf{y} \in \{0, 1\}^n$ conditional on all market and firm-market observables. The authors derive sharp bounds on $\mathbb{P}(\mathbf{y}|\mathbf{x}, \mathbf{z})$ implied by equilibrium behavior and propose a two-step estimation procedure. As in many other approaches in the literature, the first step involves obtaining a consistent estimate of the 2^n conditional probabilities $\mathbb{P}(\mathbf{y}|\mathbf{x}, \mathbf{z})$. In the second step, parameters $\boldsymbol{\theta}, \boldsymbol{\beta}$ are obtained by minimizing a distance from the identified set, built from this consistent estimate and simulated bounds.

As is common in the literature, [Ciliberto and Tamer \(2009\)](#) take the first step as a given when developing their method. However, this step suffers from a curse of dimensionality in practice,³⁰ since $\mathbb{P}(\mathbf{y}|\mathbf{x}, \mathbf{z})$ is a high-dimensional object. In the airline application, there

³⁰The assumption that $\mathbb{P}(\mathbf{y}|\mathbf{x}, \mathbf{z})$ can be estimated nonparametrically from the data is prevalent in the literature (e.g., [Beresteanu, Molchanov, and Molinari \(2011\)](#), [Chesher and Rosen \(2012\)](#), [Manski and Tamer \(2002\)](#), [Galichon and Henry \(2011\)](#), and [Tamer \(2003\)](#)). See [Andrews, Berry, and Barwick \(2004\)](#) for a discussion of the associated curse of dimensionality.

are $n = 6$ firms and $2^6 = 64$ possible market structures. Conditional choice probabilities are thus composed of 63 functions of 20 observed variables: 8 market variables and $2 * 6 = 12$ firm-market variables for all firms. Some of these variables take continuous values. Obtaining reliable nonparametric estimates of these 63 functions requires massive amounts of data. However, there are on average only $2742/63 \approx 44$ observations available to estimate each function of 20 variables in the airline data.³¹ The usual solution, applied by [Ciliberto and Tamer \(2009\)](#), is to discretize the observable space. Given the curse of dimensionality, however, discretization in this context leads to a severe loss of information.

In contrast, our proposed estimators based on linear model (3) do not require estimation of $\mathbb{P}(\mathbf{y}|\mathbf{x}, \mathbf{z})$ and do not suffer from a curse of dimensionality. We next compare estimates based on models with linear interactions to the original estimates. We consider two specifications: one with homogeneous interactions, $\beta_{ij} = \beta$, corresponding to Column 2 in Table 3 in [Ciliberto and Tamer \(2009\)](#), and one with heterogeneous interactions, $\beta_{ij} = \beta_j$, corresponding to Column 3 in their Table 3. We report the original estimates in Column 1, estimates from a 2SLS estimation of model (3) in Column 2, and estimates from a 2SLS estimation of model (3) with airline fixed effects in Column 3 in our Table 3 for homogeneous interactions and in Column 3 in our Table 4 for heterogeneous interactions.

Estimated interactions are seen to be generally positive and significant under linear formulations, whereas they are negative and significant in [Ciliberto and Tamer \(2009\)](#). The two different approaches thus appear to yield qualitatively different results. Overidentification tests show that the joint validity of the exclusion restrictions is strongly rejected in the absence of airline fixed effects for both specifications. We then assess the effect of controlling for airline fixed effects, absent from the specifications analyzed in [Ciliberto and Tamer \(2009\)](#). In the homogeneous specification, the estimated interaction parameter is lower but remains positive and significant. The joint validity of the instruments is no longer rejected at the 10% level. In the heterogeneous specification, adding airline fixed effects has a strong effect on interaction estimates, an indication that airline unobservables matter. The validity of the instruments is strongly rejected in this more general specification. The impact of air-

³¹To put this into perspective, suppose that there is only one binary firm-market characteristic. The matrix of observables \mathbf{x} can then take $2^6 = 64$ values. Estimating $\mathbb{P}(\mathbf{y}|\mathbf{x})$ may then require the estimation of $63 \times 64 = 4032$ parameters.

Table 3: Market Structure of the Airline Industry: Homogeneous Effects

	Ciliberto and Tamer (2009)	2SLS		2SLS	
Endogenous Effect	[-14.151,-10.581]	0.098	(0.002)	0.080	(0.003)
Individual Effects					
Airport presence	[3.052,5.087]	1.504	(0.013)	1.877	(0.020)
Cost	[-0.714,0.024]	-0.044	(0.006)	-0.022	(0.004)
Market Controls					
Wright	[-20.526,-8.612]	-0.096	(0.011)	-0.105	(0.013)
Dallas	[-6.890,-1.087]	0.035	(0.007)	0.040	(0.009)
Market size	[-0.972,2.247]	0.008	(0.001)	0.009	(0.001)
Market distance	[4.356,7.046]	0.001	(0.007)	0.035	(0.006)
Close airport	[4.022,9.831]	-0.004	(0.011)	-0.020	(0.011)
U.S. center distance	[1.452,3.330]	0.003	(0.005)	-0.024	(0.005)
Per capita income	[0,568,2.623]	0.010	(0.005)	0.010	(0.006)
Income growth rate	[0.370,1.003]	0.002	(0.001)	0.002	(0.001)
Constant	[-13.840,-7.796]	-0.345	(0.018)	-	(-)
Airline fixed effects				X	
Correctly predicted	0.328	0.395		0.433	
Weak instruments		5389.230		5464.494	
Sargan		72.490		0.462	
Fraction predicted in [0, 1]		0.846		0.838	

Note: Estimated coefficients and associated standard errors (in parenthesis). Column (1) is reproduced from Ciliberto and Tamer (2009). Standard errors for columns (2) and (3) are clustered at the market level. Predicted values for columns (2) and (3) give the proportions of markets whose observed structure is equal to the structure of highest likelihood. For all columns, there are 2,742 markets and 6 firms in each market. The weak instrument tests are based on first-stage F-tests. The test statistic under the null hypothesis that all instruments are weak follows a non-central χ^2 distribution (see Stock and Yogo (2005)). The null hypothesis is rejected at $< 1\%$ for both specifications. The null hypothesis of the Sargan test is that all instruments are exogenous. The test statistic follows a χ^2 distribution under the null hypothesis. The null hypothesis is rejected at $< 1\%$ for the specification without fixed effects but is not rejected at 10% for the specification with airline fixed effects.

Table 4: Market Structure of the Airline Industry: Heterogeneous Effects

	Ciliberto and Tamer (2009)	2SLS	2SLS
Endogenous Effect			
Presence of AA	[-10.914,-8.822]	0.188 (0.010)	0.065 (0.008)
Presence of DL	[-10.037,-8.631]	0.250 (0.009)	0.133 (0.006)
Presence of UA	[-10.101,-4.938]	0.075 (0.011)	0.090 (0.010)
Presence of MA	[-11.489,-9.414]	-0.007 (0.008)	0.074 (0.007)
Presence of LCC	[-19.623,-14.578]	-0.055 (0.014)	0.079 (0.012)
Presence of WN	[-12.912,-10.969]	0.063 (0.009)	0.060 (0.007)
Individual Effects			
Airport presence	[11.262,14.296]	1.631 (0.014)	1.894 (0.020)
Cost	[-1.197,-0.333]	-0.048 (0.006)	-0.022 (0.004)
Market Controls			
Wright	[-14.738,-12.556]	-0.034 (0.018)	-0.081 (0.013)
Dallas	[-1.186,0.421]	0.005 (0.014)	0.034 (0.008)
Market size	[0.532,1.245]	0.011 (0.002)	0.010 (0.001)
Market distance	[0.106,1.002]	-0.046 (0.008)	0.030 (0.006)
Close airport	[4.022,9.831]	-0.019 (0.015)	-0.019 (0.011)
U.S. center distance	[1.452,3.330]	0.043 (0.007)	-0.022 (0.005)
Per capita income	[-0.080,1.010]	0.014 (0.008)	0.007 (0.006)
Income growth rate	[0.078,0.360]	-0.005 (0.002)	0.001 (0.001)
Constant	[-1.362,2.431]	-0.401 (0.027)	- (-)
Airline fixed effects			X
Predicted	0.326	0.342	0.437
Weak instruments AA		1,350.680	1,090.872
Weak instruments DL		1,373.260	1,274.265
Weak instruments UA		1,023.390	926.734
Weak instruments MA		1,360.520	896.077
Weak instruments LCC		455.520	424.383
Weak instruments WN		1,521.330	1,474.447
Sargan		119.28	93.923
Fraction predicted in [0, 1]		0.780	0.837

Note: Estimated coefficients and associated standard errors (in parenthesis). Column (1) is reproduced from Ciliberto and Tamer (2009). Standard errors for columns (2) and (3) are clustered at the market level. Predicted values for columns (2) and (3) give the proportions of markets whose observed structure is equal to the structure of highest likelihood. For all columns, there are 2,742 markets and 6 firms in each market. The weak instrument tests are based on first-stage F-tests. The test statistic under the null hypothesis that all instruments are weak follows a non-central χ^2 distribution (see Stock and Yogo (2005)). The null hypothesis is rejected at the < 1% level for both specifications. The null hypothesis of the Sargan test is that all instruments are exogenous. The test statistic follows a χ^2 distribution under the null hypothesis. The null hypothesis is rejected at the < 1% level for both specifications.

line fixed effects on estimates and the results from overidentification tests both suggest that endogeneity is a serious concern in the empirical analysis of [Ciliberto and Tamer \(2009\)](#).

The proportion of observations whose predicted probability is between 0 and 1 ranges from 78% to 85%, depending on the specification. Predicting a probability outside $[0, 1]$ is strongly correlated with the variable “airport presence”, however, and is sensitive to how this variable is measured.³² For example, we can replace this variable by a dummy equal to 0 when airport presence is lower than the median and 1 otherwise, as in [Chen, Christensen, and Tamer \(2018\)](#). When we re-estimate linear specifications with a binary airport presence, the interaction estimates are qualitatively similar, and the proportion of predicted probabilities between 0 and 1 now ranges from 85% to 93%.³³ Note that airport presence is a function of outcomes, and its inclusion can only be justified by making strong separability assumptions, see Footnote 27 in [Ciliberto and Tamer \(2009\)](#).

We then verify that the uniqueness condition of [Theorem 2](#) holds for the estimated interaction parameters in all linear specifications. Under homogeneity, this condition is equivalent to $|\beta| < 1/5 = 0.2$, and here $\hat{\beta} = 0.098$ without airline fixed effects and 0.080 with airline fixed effects. Under heterogeneity, uniqueness holds if $\sum_j |\beta_j| < 1$. From [Table 4](#), we see that $\sum_j |\hat{\beta}_j| = 0.637$ without airline fixed effects and 0.501 with airline fixed effects.³⁴

Endogeneity may notably be caused by market-level unobservables, i.e., unobserved characteristics of the markets that affect all firms’ entry decisions. In their analysis, [Ciliberto and Tamer \(2009\)](#) assume that preference shock e_{im} for firm i in market m with origin o and destination d is the sum of a firm-market shock u_{im} , a market shock u_m , an origin shock u_o and a destination shock u_d : $e_{im} = u_{im} + u_m + u_d + u_o$. They further assume that these shocks are independent from each other and from observables. These preference shocks are then random effects in the usual panel terminology. We can easily introduce random or fixed effects in linear models of interactions in binary outcomes.³⁵ Moreover, 2SLS estimates remain consistent in the presence of random effects. In other words, the estimates from linear models

³²“Airport presence” computes the average proportion of other markets served by a carrier out of its departure and arrival airport, see the Supplementary Appendix of [Ciliberto and Tamer \(2009\)](#).

³³This proportion is equal to 90% in the specification with heterogeneous interactions and airline fixed effects.

³⁴The null hypothesis $\sum_j |\beta_j| \geq 1$ is rejected at a significance level of $< 1\%$ for all specifications.

³⁵We describe in [Appendix](#) how [Theorem 1](#) extends in the presence of random or fixed effects. The supports of these effects are generally constrained to guarantee that probabilities lie between 0 and 1.

of interactions presented in Table 3 and 4 are robust to the type of random effects considered in [Ciliberto and Tamer \(2009\)](#).

Endogeneity is likely induced by fixed effects, i.e., market-level unobservables that are correlated with observables. Are interactions still identified in entry models in the presence of market fixed effects? To our knowledge, this question has not been systematically studied in the literature. One advantage of linear models is that they are well suited to analyze identification. To illustrate, we analyze the identification of linear model (3) with heterogeneous interactions ($\beta_{ij} = \beta_j$) and market fixed effects in Appendix. We show that this model is not identified but that it only has one degree of underidentification. Interaction parameters are identified conditional on some normalization, and their ranking is identified under slight sign restriction. A full exploration of identification in linear and nonlinear entry models with market fixed effects is left for future research.

6 Conclusion

We consider a general model of linear interactions in binary outcomes. Building on [Heckman and MaCurdy \(1985\)](#), we first characterize the conditions under which the model is statistically well defined. Additive errors in the econometric model must have a specific discrete structure, imposed by the binary nature of the outcomes. We then characterize and assess the game-theoretic microfoundations of the model. We notably show that the linear model of interactions admits reasonable microfoundations under incomplete information and independence. Finally, we propose two estimators and revisit the analysis of teenage smoking and peer effects by [Lee, Li, and Lin \(2014\)](#) and that of entry into airline markets by [Ciliberto and Tamer \(2009\)](#). These reanalyses highlight the advantages of the linear framework and suggest that the previous analyses suffer from endogeneity issues. We do not claim, of course, that data with binary outcomes are always best represented by a linear model of interactions. We do claim that linear models provide a useful benchmark that has been unduly discarded by the literature on the econometrics of games. We believe that they have a legitimate place in the toolkit of applied researchers interested in interactions in binary outcomes, and hence should be rehabilitated.

One important question for future research is whether estimates of a linear model of interactions in binary outcomes admit any meaningful interpretation when the data generating process is not linear. For instance, are there natural conditions under which nonlinear models of interactions can be approximated by linear models for moderate interactions? Or under which linear estimates of interactions can be interpreted as some kind of average impact, even when interactions are strong? Another important and difficult challenge is to find ways to account for unconstrained fixed effects in nonlinear frameworks, particularly in games of complete information. Researchers often interpret the presence of bunching and clustering in the data as a sign of multiple equilibria. However, bunching and clustering can also be explained by common shocks. We thus need to better understand what can be identified when both multiplicity and unobserved heterogeneity may matter, and to develop appropriate estimation frameworks.

References

- V. Aguirregabiria and P. Mira. Sequential estimation of dynamic discrete games. *Econometrica*, 75(1):1–53, 2007.
- V. Aguirregabiria and P. Mira. Identification of games of incomplete information with multiple equilibria and unobserved heterogeneity. *Quantitative Economics*, 10(4):1659–1701, 2019.
- T. Amemiya. Some theorems in the linear probability model. *International Economic Review*, 18(3):645–650, 1977.
- D. W. Andrews, S. Berry, and P. J. Barwick. Confidence regions for parameters in discrete games with multiple equilibria, with an application to discount chain store location. *Working Paper*, 2004.
- J. D. Angrist. The perils of peer effects. *Labour Economics*, 30:98–108, 2014.
- J. D. Angrist and J.-S. Pischke. *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton university press, 2008.
- P. Bajari, H. Hong, and S. P. Ryan. Identification and estimation of a discrete game of complete information. *Econometrica*, 78(5):1529–1568, 2010.
- P. Bajari, H. Hong, and D. Nekipelov. Game theory and econometrics: A survey of some recent research. In *Advances in Economics and Econometrics, 10th World Congress*, volume 3, pages 3–52, 2013.
- C. Ballester, A. Calvó-Armengol, and Y. Zenou. Who’s who in networks. wanted: The key player. *Econometrica*, 74(5):1403–1417, 2006.
- A. Beresteanu, I. Molchanov, and F. Molinari. Sharp identification regions in models with convex moment predictions. *Econometrica*, 79(6):1785–1821, 2011.
- S. T. Berry. Estimation of a model of entry in the airline industry. *Econometrica*, 60(4):889–917, 1992.
- C. A. Bester, T. G. Conley, and C. B. Hansen. Inference with dependent data using cluster covariance estimators. *Journal of Econometrics*, 165(2):137–151, 2011.
- P. Bjorn and Q. Vuong. Modèles d’équations simultanées pour variables endogènes fictives: Une formulation par la théorie des jeux avec application à la participation au marché du travail. *L’Actualité économique*, 73(1-2-3):161–205, 1997.
- V. Boucher and B. Fortin. Some challenges in the empirics of the effects of networks. *Oxford Handbook on the Economics of Networks*, Y. Bramoullé, B. Rogers and A. Galeotti (Eds.), Oxford University Press, 2015.
- Y. Bramoullé, H. Djebbari, and B. Fortin. Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55, 2009.

- Y. Bramoullé, R. Kranton, and M. D’amours. Strategic interaction and networks. *American Economic Review*, 104(3):898–930, 2014.
- Y. Bramoullé, H. Djebbari, and B. Fortin. Peer effects in networks: a survey. *Annual Review of Economics*, 12:603–629, 2020.
- W. A. Brock and S. N. Durlauf. Discrete choice with social interactions. *The Review of Economic Studies*, 68(2):235–260, 2001.
- A. C. Cameron and D. L. Miller. A practitioner’s guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372, 2015.
- A. C. Cameron and P. K. Trivedi. *Microeconometrics: methods and applications*. Cambridge university press, 2005.
- M. Carrasco and G. Tchuente. Efficient estimation with many weak instruments using regularization techniques. *Econometric Reviews*, 35(8-10):1609–1637, 2016.
- X. Chen, T. M. Christensen, and E. Tamer. Monte carlo confidence sets for identified sets. *Econometrica*, 86(6):1965–2018, 2018.
- A. Chesher and A. M. Rosen. Simultaneous equations models for discrete outcomes: coherence, completeness, and identification. *Working Paper*, 2012.
- F. Ciliberto and E. Tamer. Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828, 2009.
- T. G. Conley. Gmm estimation with cross sectional dependence. *Journal of Econometrics*, 92(1):1–45, 1999.
- L. Davezies, X. d’Haultfoeuille, and D. Fougère. Identification of peer effects using group size variation. *The Econometrics Journal*, 12(3):397–413, 2009.
- R. Davidson and J. G. MacKinnon. Bootstrap tests for overidentification in linear regression models. *Econometrics*, 3(4):825–863, 2015.
- G. De Giorgi, M. Pellizzari, and S. Redaelli. Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*, 2(2): 241–75, 2010.
- G. De Giorgi, A. Frederiksen, and L. Pistaferri. Consumption network effects. *The Review of Economic Studies*, 87(1):130–163, 2020.
- A. De Paula. Econometric analysis of games with multiple equilibria. *Annual Review of Economics*, 5(1):107–131, 2013.
- A. De Paula. Econometrics of network models. In *Advances in economics and econometrics: Theory and applications, eleventh world congress*, pages 268–323. Cambridge University Press Cambridge, 2017.

- R. Dieye and B. Fortin. Gender peer effects heterogeneity in obesity. *Working Paper*, 2017.
- I. Fadlon and T. H. Nielsen. Family health behaviors. *American Economic Review*, 109(9): 3162–91, 2019.
- A. D. Foster and M. R. Rosenzweig. Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, 103(6):1176–1209, 1995.
- A. Galichon and M. Henry. Set identification in models with multiple equilibria. *The Review of Economic Studies*, 78(4):1264–1298, 2011.
- J. J. Heckman and T. E. MaCurdy. A simultaneous equations linear probability model. *Canadian Journal of Economics*, pages 28–37, 1985.
- J. J. Heckman and J. M. Snyder Jr. Linear probability models of the demand for attributes with an empirical application to estimating the preferences of legislators. *The Rand Journal of Economics*, 28:S142–S189, 1997.
- M. Hlavac. *stargazer: Well-Formatted Regression and Summary Statistics Tables*, 2018. URL <http://CRAN.R-project.org/package=stargazer>. R package version 5.2.2.
- B. Jovanovic. Observable implications of models with multiple equilibria. *Econometrica*, 57(6):1431–1437, 1989.
- H. H. Kelejian and G. Piras. Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes. *Regional Science and Urban Economics*, 46:140–149, 2014.
- H. H. Kelejian and I. R. Prucha. A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *The Journal of Real Estate Finance and Economics*, 17(1):99–121, 1998.
- H. H. Kelejian and I. R. Prucha. Hac estimation in a spatial framework. *Journal of Econometrics*, 140(1):131–154, 2007.
- B. Kline and E. Tamer. Econometric analysis of models with social interactions. *The Econometric Analysis of Network Data (B. Graham and A. De Paula Editors)*, Forthcoming.
- P. Kuhn, P. Kooreman, A. Soetevent, and A. Kapteyn. The effects of lottery prizes on winners and their neighbors: Evidence from the dutch postcode lottery. *American Economic Review*, 101(5):2226–47, 2011.
- T. Lancaster. The incidental parameter problem since 1948. *Journal of Econometrics*, 95(2): 391–413, 2000.
- R. Laschever. The doughboys network: Social interactions and the employment of world war i veterans. *Working Paper*, 2013.

- L.-F. Lee. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, 72(6):1899–1925, 2004.
- L.-f. Lee and X. Liu. Efficient gmm estimation of high order spatial autoregressive models with autoregressive disturbances. *Econometric Theory*, 26(1):187–230, 2010.
- L.-f. Lee, X. Liu, and X. Lin. Specification and estimation of social interaction models with network structures. *The Econometrics Journal*, 13(2):145–176, 2010.
- L.-f. Lee, J. Li, and X. Lin. Binary choice models with social network under heterogeneous rational expectations. *Review of Economics and Statistics*, 96(3):402–417, 2014.
- M. P. Leung. Treatment and spillover effects under network interference. *Review of Economics and Statistics*, 102(2):1–42, 2016.
- M. P. Leung. Inference in models of discrete choice with social interactions using network data. *Working Paper*, 2019.
- A. Lewbel. The identification zoo: Meanings of identification in econometrics. *Journal of Economic Literature*, 57(4):835–903, 2019.
- T. Li and L. Zhao. A partial identification subnetwork approach to discrete games in large networks: An application to quantifying peer effects. *Working Paper*, 2016.
- X. Lin. Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables. *Journal of Labor Economics*, 28(4):825–860, 2010.
- X. Liu. Simultaneous equations with binary outcomes and social interactions. *Econometric Reviews*, 38(8):921–937, 2019.
- C. F. Manski. Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3):531–542, 1993.
- C. F. Manski. Economic analysis of social interactions. *Journal of Economic Perspectives*, 14(3):115–136, 2000.
- C. F. Manski and E. Tamer. Inference on regressions with interval data on a regressor or outcome. *Econometrica*, 70(2):519–546, 2002.
- D. McFadden. Conditional logit analysis of qualitative choice behavior. *in Frontiers in Econometrics, Chapter 4*, pages 105–142, 1974.
- K. Menzel. Inference for games with many players. *The Review of Economic Studies*, 83(1):306–337, 2016.
- R. Nakajima. Measuring peer effects on youth smoking behaviour. *The Review of Economic Studies*, 74(3):897–935, 2007.
- C. Nicoletti, K. G. Salvanes, and E. Tominey. The family peer effect on mothers’ labor supply. *American Economic Journal: Applied Economics*, 10(3):206–34, 2018.

- S. Peng. Heterogeneous endogenous effects in networks. *Working Paper*, 2019.
- B. Sacerdote. Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the Economics of Education*, volume 3, pages 249–277. Elsevier, 2011.
- R. Selten. Reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55, 1975.
- A. R. Soetevent and P. Kooreman. A discrete-choice model with social interactions: with an application to high school teen behavior. *Journal of Applied Econometrics*, 22(3):599–624, 2007.
- J. H. Stock and M. Yogo. Testing for weak instruments in linear iv regression. In *Identification and inference for econometric models: essays in honor of Thomas Rothenberg*, pages 80–108. Cambridge University Press, 2005.
- E. Tamer. Incomplete simultaneous discrete response model with multiple equilibria. *The Review of Economic Studies*, 70(1):147–165, 2003.
- G. Tchuente. Weak identification and estimation of social interaction models. *Working Paper*, 2019.
- J. M. Wooldridge. *Econometric analysis of cross section and panel data*. MIT press, 2010.

7 Appendix

7.1 Tables

Table 5: Monte Carlo Simulations – Group Sizes

High β					
$N = 20$			$N = 50$		
Parameters	2SLS	NLS	Parameters	2SLS	NLS
$\alpha = 0.1$	0.099 (0.014)	0.099 (0.012)	$\alpha = 0.1$	0.100 (0.015)	0.100 (0.012)
$\beta = 0.7$	0.706 (0.071)	0.701 (0.039)	$\beta = 0.7$	0.700 (0.054)	0.699 (0.034)
$\gamma = 0.05$	0.051 (0.021)	0.051 (0.020)	$\gamma = 0.05$	0.050 (0.015)	0.050 (0.014)
$\delta = 0.1$	0.094 (0.069)	0.099 (0.038)	$\delta = 0.1$	0.101 (0.044)	0.101 (0.029)
Low β					
$N = 20$			$N = 50$		
Parameters	2SLS	NLS	Parameters	2SLS	NLS
$\alpha = 0.1$	0.099 (0.012)	0.099 (0.012)	$\alpha = 0.1$	0.100 (0.013)	0.100 (0.012)
$\beta = 0.25$	0.254 (0.070)	0.252 (0.063)	$\beta = 0.25$	0.252 (0.060)	0.251 (0.053)
$\gamma = 0.2$	0.201 (0.022)	0.201 (0.022)	$\gamma = 0.2$	0.200 (0.016)	0.200 (0.016)
$\delta = 0.3$	0.297 (0.056)	0.299 (0.050)	$\delta = 0.3$	0.297 (0.039)	0.298 (0.036)

Note: For each simulation, $M = 500$ networks are generated using iid Bernoulli trials with a probability $p = 0.1$. Thus, the expected number of links for each individual is $0.1(N - 1)$. Values represent the average (standard deviation) of the 1000 simulations.

Table 6: Summary Statistics: Teenagers' Smoking Decisions

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Age	15.068	1.685	10	14	16	19
Years in school	2.493	1.407	1	1	3	6
Male	0.488	0.500	0	0	1	1
Black	0.183	0.386	0	0	0	1
Asian	0.066	0.248	0	0	0	1
Hisp.	0.139	0.346	0	0	0	1
Other race	0.056	0.230	0	0	0	1
Live with both parents	0.730	0.444	0	0	1	1
Sports club	0.524	0.499	0	0	1	1
Mom education less than high school	0.101	0.301	0	0	0	1
Mom education more than high school	0.412	0.492	0	0	1	1
Mom education missing	0.107	0.309	0	0	0	1
Mom job is professional	0.262	0.440	0	0	1	1
Mom other jobs	0.358	0.479	0	0	1	1
Mom on welfare	0.009	0.093	0	0	0	1
Mom job is missing	0.090	0.286	0	0	0	1
Smoke	0.231	0.421	0	0	0	1

Note: Summary statistics using the same sample as in [Lee, Li, and Lin \(2014\)](#). The number of observations is 74,783, the number of schools is 127, and the number of school-grades is 532.

Table 7: Summary Statistics: Market Structures in the Airline Industry

Statistic	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Endogenous Variables						
Entry AA	0.426	0.495	0	0	1	1
Entry DL	0.551	0.497	0	0	1	1
Entry UA	0.275	0.447	0	0	1	1
Entry MA	0.548	0.498	0	0	1	1
Entry LCC	0.162	0.369	0	0	0	1
Entry WN	0.247	0.431	0	0	0	1
Firm-Level Variables						
Airport presence AA	0.422	0.167	0.000	0.293	0.548	0.873
Airport presence DL	0.540	0.181	0.000	0.406	0.681	0.987
Airport presence UA	0.265	0.153	0.000	0.143	0.369	0.689
Airport presence MA	0.376	0.135	0.000	0.277	0.459	0.850
Airport presence LCC	0.098	0.077	0.000	0.054	0.127	0.650
Airport presence WN	0.242	0.176	0	0.2	0.4	1
Cost AA	0.736	1.609	0.000	0.016	0.812	27.570
Cost DL	0.420	1.322	0	0.01	0.3	28
Cost UA	0.784	1.476	0.000	0.021	0.933	21.096
Cost MA	0.229	0.615	0.000	0.003	0.191	11.620
Cost Hub LCC	0.043	0.174	0	0	0.01	3
Cost WN	0.303	0.860	0.000	0.001	0.233	16.180
Market-Level Variables						
Market distance	1.085	0.624	0.067	0.602	1.452	2.724
Distance from center	1.571	0.594	0.283	1.138	1.956	3.390
Minimum distance	0.346	0.205	0.102	0.155	0.489	1.505
Income growth rate	4.051	1.478	-0.300	3.050	4.950	10.050
Income per capita	3.240	0.391	1.702	2.965	3.491	4.580
Market size	2.259	1.846	0.310	1.097	2.701	15.236
Wright amendment	0.030	0.169	0	0	0	1
Dallas airport	0.070	0.255	0	0	0	1

7.2 Proofs

7.2.1 Proof of Theorem 2

Suppose that the ν_i 's are uncorrelated and recall that $P_i = [(\mathbf{I} - \boldsymbol{\beta})^{-1}\mathbf{f}]_i = \mathbb{P}(y_i = 1|\mathbf{x})$. Consider preference shocks e_i , that are independent and independent of observables, and let F_i denote the cdf of e_i . A strategy is a function of an agent's shock into their action set, i.e., $y_i(e_i) \in \{0, 1\}$. Payoffs are given by the agent's expected utilities:

$$\mathbb{E}u_i(1|e_i) - \mathbb{E}u_i(0|e_i) = f_i - \frac{1}{2} + e_i + \sum_j \beta_{ij}\mathbb{P}(y_j = 1|e_i).$$

By independence, $\mathbb{P}(y_j = 1|e_i) = \mathbb{P}(y_j = 1)$. Then, agent i with private information e_i chooses $y_i = 1$ iff

$$f_i - \frac{1}{2} + e_i + \sum_j \beta_{ij}\mathbb{P}(y_j = 1) \geq 0.$$

Introduce \bar{e}_i such that

$$f_i - \frac{1}{2} + \bar{e}_i + \sum_j \beta_{ij}\mathbb{P}(y_j = 1) = 0.$$

Then, $y_i = 1 \Leftrightarrow e_i \geq \bar{e}_i$ and $\mathbb{P}(y_i = 1) = \mathbb{P}(e_i \geq \bar{e}_i) = 1 - F_i(\bar{e}_i)$.

(1) Assume, first, that the profile $y_i = 1$ with probability P_i and 0 with probability $1 - P_i$ is a Bayes-Nash equilibrium, where $P_i = [(\mathbf{I} - \boldsymbol{\beta})^{-1}\mathbf{f}]_i$ for every possible \mathbf{f} and $\boldsymbol{\beta}$. This means that $P_i = \mathbb{P}(y_i = 1) = 1 - F_i(\bar{e}_i)$. Since $P_i = f_i + \sum_j \beta_{ij}P_j$, $1 - F_i(\bar{e}_i) = f_i + \sum_j \beta_{ij}P_j$. By the definition of \bar{e}_i , $f_i + \sum_j \beta_{ij}P_j = \frac{1}{2} - \bar{e}_i$ and hence

$$F_i(\bar{e}_i) = \bar{e}_i + \frac{1}{2}.$$

As \mathbf{f} and $\boldsymbol{\beta}$ take all possible values, \bar{e}_i takes all values in $[-\frac{1}{2}, \frac{1}{2}]$. This shows that F_i is the cdf of the uniform distribution on $[-\frac{1}{2}, \frac{1}{2}]$.

(2) Conversely, assume that $\forall i, e_i$ is uniformly distributed on $[-\frac{1}{2}, \frac{1}{2}]$. Then, $F_i(e) = \min\{\max\{e + \frac{1}{2}, 0\}, 1\}$. This implies that $\mathbb{P}(y_i = 1) = \min\{\max\{0, (1/2 - \bar{e}_i)\}, 1\}$. Next, consider a situation where for every i , $\bar{e}_i \in [-\frac{1}{2}, \frac{1}{2}]$. This corresponds to the following fixed-point equation:

$$\bar{\mathbf{e}} = -\mathbf{f} + \frac{1}{2}\mathbf{1} - \frac{1}{2}\boldsymbol{\beta}\mathbf{1} + \boldsymbol{\beta}\bar{\mathbf{e}},$$

or

$$\bar{\mathbf{e}} = -(\mathbf{I} - \boldsymbol{\beta})^{-1}[\mathbf{f} - \frac{1}{2}\mathbf{1} + \frac{1}{2}\boldsymbol{\beta}\mathbf{1}],$$

and this yields

$$\bar{\mathbf{e}} = -(\mathbf{I} - \boldsymbol{\beta})^{-1}\mathbf{f} + \frac{1}{2}\mathbf{1}.$$

Since $P_i = [(\mathbf{I} - \boldsymbol{\beta})^{-1}\mathbf{f}]_i \in [0, 1]$, we indeed have $\bar{e}_i \in [-1/2, 1/2]$ and hence

$$\mathbb{P}(y_i = 1) = \mathbb{P}(e_i \geq \bar{e}_i) = \frac{1}{2} + [(\mathbf{I} - \boldsymbol{\beta})^{-1}\mathbf{f}]_i - \frac{1}{2} = P_i,$$

and hence the stochastic profile $y_i = 1$ with probability P_i and 0 with probability $1 - P_i$ is a Bayes-Nash equilibrium.

Can the game have other Bayes-Nash equilibria? Any Bayes-Nash equilibrium corresponds to the following fixed point equation:

$$\bar{e}_i = -f_i + \frac{1}{2} - \sum_j \beta_{ij} \min\{\max\{0, (1/2 - \bar{e}_j)\}, 1\}.$$

Rewrite the fixed-point problem in matrix form: $\mathbf{T}(\bar{\mathbf{e}}) = -\mathbf{f} + \frac{1}{2}\mathbf{1} - \beta\mathbf{h}(\bar{\mathbf{e}})$, where $\mathbf{h}(\bar{\mathbf{e}})_i = \min\{\max\{0, (1/2 - \bar{e}_i)\}, 1\}$. We have

$$\|\mathbf{T}(\bar{\mathbf{e}}) - \mathbf{T}(\tilde{\mathbf{e}})\| = \|\beta\mathbf{h}(\tilde{\mathbf{e}}) - \beta\mathbf{h}(\bar{\mathbf{e}})\| \leq \|\beta\| \cdot \|\mathbf{h}(\tilde{\mathbf{e}}) - \mathbf{h}(\bar{\mathbf{e}})\| \leq \|\beta\| \cdot \|\tilde{\mathbf{e}} - \bar{\mathbf{e}}\|$$

for any submultiplicative norm $\|\cdot\|$. If $\|\beta\| < 1$, the fixed-point function is a contraction mapping and thus has a unique fixed point. From the argument above, this fixed point is interior.

With $n = 2$, we can easily verify that when $f_1 = f_2 = -1$ and $\beta_{12} = \beta_{21} = 3$, the game has 3 Bayes-Nash equilibria: the one corresponding to model (1), $\mathbb{P}(y_1 = 1) = P_1 = \frac{1}{2}$ and $\mathbb{P}(y_2 = 1) = P_2 = \frac{1}{2}$, as well as two others, $\mathbb{P}(y_1 = 1) = \mathbb{P}(y_2 = 1) = 0$ and $\mathbb{P}(y_1 = 1) = \mathbb{P}(y_2 = 1) = 1$. QED.

7.2.2 Proof of Proposition 1

Consider preference shocks e_i . We have $v_i(1, \mathbf{y}_{-i}) - v_i(0, \mathbf{y}_{-i}) = f_i - \frac{1}{2} + \sum_j \beta_{ij} y_j + e_i$ and $y_i^* = f_i + \sum_j \beta_{ij} y_j^* + \varepsilon_i$. Therefore, $v_i(1, \mathbf{y}_{-i}^*) - v_i(0, \mathbf{y}_{-i}^*) = y_i^* - \frac{1}{2} + e_i - \varepsilon_i$. If $\nu_i > 0$, then $y_i^* = 1$ and $v_i(1, \mathbf{y}_{-i}^*) - v_i(0, \mathbf{y}_{-i}^*) \geq 0$ iff $e_i \geq \varepsilon_i - \frac{1}{2}$. If $\nu_i < 0$, then $y_i^* = 0$ and $v_i(1, \mathbf{y}_{-i}^*) - v_i(0, \mathbf{y}_{-i}^*) \leq 0$ iff $e_i \leq \varepsilon_i + \frac{1}{2}$. By Theorem 1, $\varepsilon_i = \nu_i - \sum_j \beta_{ij} \nu_j$ and $\nu_i \in \{-P_i, 1 - P_i\}$. Substituting in the inequalities yields the first part of the result, characterizing preference shocks for which \mathbf{y}^* is a Nash equilibrium.

Next, derive a sufficient condition for uniqueness in dominant strategies. If $\nu_i > 0$, then $y_i^* = 1 = f_i + \sum_j \beta_{ij} y_j^* + \varepsilon_i$. This implies that $v_i(1, \mathbf{y}_{-i}) - v_i(0, \mathbf{y}_{-i}) = \frac{1}{2} + \sum_j \beta_{ij} (y_j - y_j^*) + e_i - \varepsilon_i$. Note that $\beta_{ij} (y_j - y_j^*) \geq -|\beta_{ij}|$. Therefore, $v_i(1, \mathbf{y}_{-i}) - v_i(0, \mathbf{y}_{-i}) > 0$ if $-\sum_j |\beta_{ij}| + e_i - \varepsilon_i + \frac{1}{2} > 0$. Here, $y_i = 1$ is a dominant strategy for agent i if $e_i > \varepsilon_i - \frac{1}{2} + \sum_j |\beta_{ij}|$.

If $\nu_i < 0$, then $y_i^* = 0 = f_i + \sum_j \beta_{ij} y_j^* + \varepsilon_i$. This implies that $v_i(1, \mathbf{y}_{-i}) - v_i(0, \mathbf{y}_{-i}) = \sum_j \beta_{ij} (y_j - y_j^*) + e_i - \varepsilon_i - \frac{1}{2}$. Note that $\beta_{ij} (y_j - y_j^*) \leq |\beta_{ij}|$. This means that $v_i(1, \mathbf{y}_{-i}) - v_i(0, \mathbf{y}_{-i}) < 0$ if $\sum_j |\beta_{ij}| + e_i - \varepsilon_i - \frac{1}{2} < 0$. This shows that $y_i = 0$ is a dominant strategy for agent i if $e_i < \varepsilon_i + \frac{1}{2} - \sum_j |\beta_{ij}|$. QED.

7.2.3 Proof of Proposition 2

First, show that \mathbf{y}^* is still a Nash equilibrium for preference shocks e'_i . Consider a realization of errors ν and \mathbf{y}^* the unique solution to equation (1). Consider i such that $\nu_i > 0$ and $y_i^* = 1$. Since \mathbf{y}^* is a Nash equilibrium for shocks e_i , $u_i(1, \mathbf{y}_{-i}^*) - u_i(0, \mathbf{y}_{-i}^*) + e_i \geq 0$. Since $e'_i = e_i + L_i$ and $L_i \geq 0$, $u_i(1, \mathbf{y}_{-i}^*) - u_i(0, \mathbf{y}_{-i}^*) + e'_i \geq 0$. Similarly, if $\nu_i < 0$ and $y_i^* = 0$,

$u_i(1, \mathbf{y}_{-i}^*) - u_i(0, \mathbf{y}_{-i}^*) + e_i \leq 0$. Since $e'_i = e_i - M_i$ and $M_i \geq 0$, then $u_i(1, \mathbf{y}_{-i}^*) - u_i(0, \mathbf{y}_{-i}^*) + e_i \leq 0$. And hence \mathbf{y}^* is a Nash equilibrium for shocks e'_i .

Next, assume that

$$L_i > -e_i - \min_{\mathbf{y}_{-i}} u_i(1, \mathbf{y}_{-i}) - u_i(0, \mathbf{y}_{-i}),$$

$$M_i > e_i + \max_{\mathbf{y}_{-i}} u_i(1, \mathbf{y}_{-i}) - u_i(0, \mathbf{y}_{-i}).$$

The right-hand sides of these inequalities are well defined because \mathbf{y}_{-i} takes a finite number of values.

From the first inequality, we have for every \mathbf{y}_{-i} , $L_i > -u_i(1, \mathbf{y}_{-i}) + u_i(0, \mathbf{y}_{-i}) - e_i$. If $\nu_i > 0$, then $e'_i = e_i + L_i$ and $u_i(1, \mathbf{y}_{-i}) - u_i(0, \mathbf{y}_{-i}) + e'_i > 0$. Playing 1 is a dominant strategy for agent i . From the second inequality, we have for every \mathbf{y}_{-i} , $M_i > u_i(1, \mathbf{y}_{-i}) - u_i(0, \mathbf{y}_{-i}) + e_i$. If $\nu_i < 0$, then $e'_i = e_i - M_i$ and hence $u_i(1, \mathbf{y}_{-i}) - u_i(0, \mathbf{y}_{-i}) + e'_i < 0$. Playing 0 is a dominant strategy for player i . QED.

7.2.4 Partial identification in linear entry games with market fixed effects

Consider the following variant of model (3) with market fixed effects:

$$y_{im} = \alpha_m + \mathbf{x}_{im}\boldsymbol{\gamma} + \sum_{j \neq i} \beta_j y_{jm} + \varepsilon_{im}.$$

The effect of market-level characteristics is of course not identified here, as these characteristics are absorbed in the market fixed effects. These fixed effects must be eliminated. Consider the model in deviation with respect to y_{1m} :

$$y_{im} - y_{1m} = (\mathbf{x}_{im} - \mathbf{x}_{1m})\boldsymbol{\gamma} + \beta_1 y_{1m} - \beta_i y_{im} + \varepsilon_{im} - \varepsilon_{1m}.$$

If $\beta_i \neq -1$, this is equivalent to

$$y_{im} - y_{1m} = \frac{1}{1 + \beta_i} (\mathbf{x}_{im} - \mathbf{x}_{1m})\boldsymbol{\gamma} + \frac{\beta_1 - \beta_i}{1 + \beta_i} y_{1m} + \frac{1}{1 + \beta_i} (\varepsilon_{im} - \varepsilon_{1m}).$$

This equation can be estimated by instrumenting y_{1m} on the right-hand side by \mathbf{x}_{1m} . This implies that the composite parameters $\boldsymbol{\gamma}/(1 + \beta_i)$ and $b_i = (\beta_1 - \beta_i)/(1 + \beta_i) = (1 + \beta_1)/(1 + \beta_i) - 1$ are identified, and hence the ratios $(1 + \beta_i)/(1 + \beta_j)$ are identified. The β 's are not identified without further assumptions. However, β_i for $i \neq 1$ is identified when β_1 is known (except when $b_i = -1$), which shows that there is one degree of underidentification. The β 's are thus identified conditional on some normalization. Furthermore, the ordering of the β 's is identified under some slight restriction on signs, for instance that $1 + \beta_1 > 0$. QED.

7.2.5 Accounting for random and fixed effects

To illustrate, introduce market fixed effects α_m in the linear entry model: $y_{im} = \alpha_m + f_{im} + \sum_j \beta_{ij} y_{jm} + \varepsilon_{im}$ with $\mathbb{E}(\varepsilon_{im} | \boldsymbol{\alpha}, \mathbf{x}) = 0$. Random effects are a particular case where $\mathbb{E}(\alpha_m | \mathbf{x}) = 0$. Then, Theorem 1 extends when conditioning on observables and fixed effects. Formally, assume that $\mathbf{I} - \boldsymbol{\beta}$ invertible. Let $\mathbf{Q}_m = (\mathbf{I} - \boldsymbol{\beta})^{-1}(\alpha_m \mathbf{1} + \mathbf{f}_m) = \mathbb{E}(\mathbf{y}_m | \boldsymbol{\alpha}, \mathbf{x})$ and assume that $\forall i, m, Q_{im} \in [0, 1]$. Then outcomes in the unique solution to the linear model of

interactions are binary, $y_{im} \in \{0, 1\}$, if and only if $\varepsilon_{im} = \nu_{im} - \sum_j \beta_{ij} \nu_{jm}$, where $\nu_{im} = -Q_{im}$ with probability $1 - Q_{im}$ and $1 - Q_{im}$ with probability Q_{im} . This clarifies when linear models of entry with random or fixed effects are statistically well defined. Note that the assumption $Q_{im} \in [0, 1]$ puts constraints on the support of the α_m 's.