SOME INTERPRETATION OF THE LINEAR-IN-MEANS MODEL OF SOCIAL INTERACTIONS

BRENDAN KLINE AND ELIE TAMER

NORTHWESTERN UNIVERSITY

MANUSCRIPT

Abstract. The linear-in-means model is often used in applied work to empirically study the role of social interactions and peer effects. We document the subtle relationship between the parameters of the linear-in-means model and the parameters relevant for policy analysis, and study the interpretations of the model under two different scenarios. First, we show that without further assumptions on the model the direct analogs of standard policy relevant parameters are either undefined or are complicated functions not only of the parameters of the linear-in-means model but also the parameters of the distribution of the unobservables. This complicates the interpretation of the results. Second, and as in the literature on simultaneous equations, we show that it is possible to interpret the parameters of the linear-in-means model under additional assumptions on the social interaction, mainly that this interaction is a result of a particular *economic game*. These assumptions that the game is built on rule out economically relevant models. We illustrate this using examples of social interactions in educational achievement. We conclude that care should be taken when estimating and especially when interpreting coefficients from linear in means models.

Keywords: social interactions, peer effects, linear-in-means model, equilibrium, educational achievement

---

Date: Original Date: June 2011.

$*$ Department of Economics, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208, USA. Support from the National Science Foundation is greatly appreciated. Any errors are ours. Part of the results here were presented (and published) as a discussion at the AEA meetings in Chicago - 2012 as a discussion of Goldsmith-Pinkham and Imbens (2013) paper on networks which is prepared for the JBES. This article can be be reproduced for research and teaching purposes. Please cite accordingly.

1. Introduction

Models of social interactions allow a "social interaction effect" or "peer effect" through which the outcome of any given individual is related to the outcomes, choices, treatments, and/or characteristics of the other individuals in that individual's reference group. The standard model used in applications is the linear-in-means model. There are a number of possible specifications of this model, and in this paper, we focus on the specification that the outcome y_{ig} for individual i in group g is given by

$$y_{ig} = x_{ig}\beta + z_g\gamma + \frac{\phi}{N_g - 1} \sum_{j=1, j\neq i}^{N_g} y_{jg} + \epsilon_{ig},$$

where N_g is the number of individuals in group g. The observed exogenous covariates x_{ig} are individual specific, while the observed exogenous covariates z_g are common to all individuals in the same group. All exogenous covariates for group g are collected in w_g, and it is assumed that $E(\epsilon_g|w_g) = 0$. We define w_g to be an $N_g \times (K + L)$ matrix where K is the dimension of x_{ig} and L is the dimension of z_g, and where the ith row of w_g contains x_{ig} in the first K columns and z_g in the last L columns. The parameter ϕ gives the "social interaction effect" of the average outcome in the reference group on an individual's own outcome; this effect is the key methodological departure of this model from a typical model in econometrics.

The identification of the linear-in-means model has been previously addressed by Manski (1993), Graham and Hahn (2005), Lee (2007), Graham (2008), Bramoullé, Djebbari, and Fortin (2009), Davezies, D'Haultfoeuille, and Fougère (2009), De Giorgi, Pellizzari, and Redaelli (2010) and Blume, Brock, Durlauf, and Ioannides (2011). A related model for binary outcomes is studied by Brock and Durlauf (2001). Manski (1993) establishes the lack of identification in the general linear-in-means model[1], but subsequent papers provide additional conditions on the model under which the parameters are point identified. We consequently assume point identification[2] [3] in order

[1]Manski considers a variant of the model we study in that he assumes that outcomes depend on the *population average* in one's group.

[2]In our specification of the model, for the parameters to be point identified, it is enough for each individual to have an individual specific covariate that is excluded from the equations explaining the outcomes of each other individual in the group, in addition to the standard full rank condition. This result is in analogy to results on identification in simultaneous equations models.

[3]In order to focus on the interpretation of the model, we also rule out "contextual effects" by which y_{ig} also depends on $\frac{1}{N_g-1}\sum_{j=1,j\neq i}^{N_g} x_{jg}$.

to focus on the interpretation of the model. A recent review of identification results can be found in Blume, Brock, Durlauf, and Ioannides (2011).

This model is applicable to a variety of settings of interest, in the same sense that the ordinary linear model is applicable to a variety of settings of interest. For example, the outcome y_{ig} might be the test score of student i, in which case the social interaction is the commonly studied "peer effect" in educational outcomes. The extensive application of the linear-in-means model to education is reviewed in Sacerdote (2011). More generally, this model is used to make statements like: an individual whose peers have some characteristic is "caused" or "induced" to also tend to have that characteristic.

This paper studies the interpretation of the "social interaction" parameter ϕ in particular, but also the interpretation of the linear-in-means model in general. In the ordinary linear model

$$y_i = x_i\beta + \epsilon_i$$

where it is assumed that $E(\epsilon_i|x_i) = 0$, the interpretation of the parameter β is standard. This model implies that $E(y_i|x_i) = x_i\beta$, so that β gives the marginal effect of x on the average outcome of y given x (See Goldberger (1991) for more on the interpretation of the linear model). Under some assumptions, the linear model justifies the use of estimates of β to make statements like: a change in x tends on average to "cause" a response in y. Note that if instead $E(\epsilon_i|x_i) = f(x_i)$ and $f(\cdot)$ is not identically zero, then $E(y_i|x_i) = x_i\beta + f(x_i)$, and this is not identically equal to $x_i\beta$. Then the marginal effect of x on the average outcome of y given x is $\beta + f'(x_i)$, assuming $f(\cdot)$ is differentiable for simplicity, so β is not the "effect" of x on y. Also, under the alternative assumption that $Q_\alpha(\epsilon_i|x_i) = 0$, β has the interpretation as the marginal effect of x_i on the α quantile of the distribution of $y_i|x_i$. It is also possible to interpret β in this model as the *ceteris paribus* effect of x_i on y_i holding fixed ϵ_i, but without a structural functional form assumption on ϵ_i as a function of x_i this interpretation may be dubious[4].

[4] For example, in the case of multiplicative heteroskedasticity, where $\epsilon_i = \sigma(x_i)\nu_i$ and $\nu_i|\epsilon_i$ satisfies $E(\nu_i|x_i) = 0$, the "*ceteris paribus* effect" (but not the average marginal effect) of x_i on y_i depends also on the unobservable ϵ_i since in this model the unobservable necessarily depends on the observable, and so the *ceteris paribus* interpretation is not possible.

Applications of the linear-in-means model tend to give a similar average marginal effect interpretation to the "social interaction" parameter ϕ. For example, in a recent review of the application of the linear-in-means model to education, Sacerdote (2011) reports the "effect of a 1.0 move in average peer score" from a variety of papers. Bramoullé, Djebbari, and Fortin (2009, p. 50) interprets the estimate in the context of recreational activity participation among high school students as implying that an "increase in the mean recreational activities index of student's friends induces him to increase his recreational activities index [...]." And Gaviria and Raphael (2001, p. 262) interpret the estimate in the context of drug use as implying that "moving a teenager from a school where none of his classmates use drugs to one where half use drugs would increase the probability that she will use drugs by approximately thirteen percentage points." These seem to be representative uses of the linear-in-means model.

However, this interpretation is *not necessarily* justified under the assumptions of the linear-in-means model. We discuss this in section 2 and derive a basic result that shows that without further assumptions, deriving policy relevant parameters in the linear-in-means model is complicated and non-standard. We then provide an alternative interpretation of the parameters in section 3 which provides a justification for the more "standard" partial effect interpretation. This relies on treating the above models as generated by a particular economic game in which the outcome equation above becomes a best response function from that game. We illustrate our results in the specific context of educational achievement in section 4 where we show that the class of games that lead to the above formulation might not cover some economically relevant models. Finally, we conclude with some views about the econometric modeling of social interactions in section 5.

2. Policy relevant parameters in the linear-in-means model

Following the same approach as with the ordinary linear model, conditioning on all of the observed exogenous variables, the assumptions of the linear-in-means model imply that $E(y_{ig}|w_g) = x_{ig}\beta + z_g\gamma + \frac{\phi}{N_g-1}\sum_{j=1,j\neq i}^{N_g} E(y_{jg}|w_g)$. However, this is not sufficient to interpret ϕ in the same way as β is interpreted in the ordinary linear model. This is because the outcomes $y_{\cdot g}$ are endogenous and simultaneously determined in the linear-in-means model. Rather, this shows that ϕ is the marginal effect of $\frac{1}{N_g-1}\sum_{j=1,j\neq i}^{N_g} E(y_{jg}|w_g)$ on $E(y_{ig}|w_g)$. Thus, ϕ is the "average marginal effect" of the group average outcome at the average outcomes in a group with observables w_g.

So, if w_g changes in such a way that $\frac{1}{N_g-1}\sum_{j=1,j\neq i}^{N_g} E(y_{jg}|w_g)$ increases by one unit, then $E(y_{ig}|w_g)$ is caused to change by ϕ units as a "partial" effect through the social interaction process. But, since $y_{\cdot g}$ are endogenous, this approach does not provide a way to interpret this coefficient in terms of a policy experiment that manipulates some part of the environment, because the effect of w_g on $E(y_{jg}|w_g)$ is part of the model. If the object of interest is the effect of w_g on average outcomes, then as shown below the appropriate object of interest is a known function of β and ϕ. Thus, this approach does not seem to provide a useful interpretation of ϕ.

An alternative approach is to consider also conditioning on the outcomes of individuals other than i, as in $E(y_{ig}|w_g, y_{-i,g}) = x_{ig}\beta + z_g\gamma + \frac{\phi}{N_g-1}\sum_{j=1,j\neq i}^{N_g} y_{jg} + E(\epsilon_{ig}|w_g, y_{-i,g})$. In this expression, although y_{jg} now appears unconditionally as part of the "social interaction" effect term that depends on ϕ, potentially providing a justification of the interpretation of ϕ as the "effect" of others' outcomes, it also appears in the conditioning set on the expectation of ϵ_{ig}. This is to be expected because $y_{-i,g}$ is endogenous. The assumption that $E(\epsilon_g|w_g) = 0$ does not necessarily imply that $E(\epsilon_{ig}|w_g, y_{-i,g}) = 0$. Indeed, unless the unobservables ϵ_{ig} are distributed in a very precise way within a group conditional on w_g, especially related to the correlation of unobservables, then the condition that $E(\epsilon_{ig}|w_g, y_{-i,g}) = 0$ fails. But, potentially surprisingly, at least with multivariate normally distributed unobservables, that condition is *not* the independence of the unobservables across individuals in the same group.

We obtain the following theorem on the average effect of $y_{-i,g}$ on y_{ig} under the assumption of multivariate normally distributed unobservables. The next Theorems take as given that (β, γ, ϕ) are "known," fixes a group g and hence interprets these parameters within the model for members of group g.

Theorem 2.1. *Suppose the model is* $y_{ig} = x_{ig}\beta + z_g\gamma + \frac{\phi}{N_g-1}\sum_{j=1,j\neq i}^{N_g} y_{jg} + \epsilon_{ig}$ *as before. Further, suppose that* ϵ_g *are distributed according to a multivariate normal distribution with mean* 0*, unit variances, and covariance* $\rho \in (-1, 1)$*, independently of* w_g*. Suppose that* $\phi \neq 1 - N_g$ *and* $\phi \neq 1$*. Then,* $E(y_{ig}|w_g, y_{-i,g})$ *is linear in* $y_{j,g}$ *for* $j \neq i$*, and has slope equal to*

$$\eta = \frac{\phi(2(N_g-1) - \phi(N_g-2)) + (N_g-1)\rho(N_g-1+\phi^2)}{(N_g-1)((N_g-1)(1+(N_g-2)\rho+2\phi\rho)+\phi^2)}.$$

Equivalently, the slope on $\frac{1}{N_g-1}\sum_{j\neq i} y_{jg}$ *is* $\zeta = (N_g-1)\eta$*. In general,* $\zeta \neq \phi$*.*

Proof. Let Σ be the covariance matrix for the unobservables and let Υ be the $N_g \times N_g$ matrix with 1 on the diagonal and $\frac{-\phi}{N_g-1}$ on the off diagonal. Then, $y_g = \Upsilon^{-1}(w_g\tau + \epsilon_g)$. Consequently, $y_g|w_g$ is distributed as multivariate normal with mean $\Upsilon^{-1}w_g\tau$ and covariance $S = \Upsilon^{-1}\Sigma\Upsilon^{-1}$. Then by standard facts about the multivariate normal distribution, $y_{ig}|w_g, y_{-i,g}$ is normally distributed with mean $[\Upsilon^{-1}w_g\tau]_i - [S]_{i,-i}([S]_{-i,-i})^{-1}[\Upsilon^{-1}w_g\tau]_{-i} + [S]_{i,-i}([S]_{-i,-i})^{-1}y_{-i,g}$. Then, write $\Upsilon = u_1 I_{N_g} + u_2 1_{N_g\times N_g}$ and $\Sigma = s_1 I_{N_g} + s_2 1_{N_g\times N_g}$ where $u_1 = 1 + \frac{\phi}{N_g-1}$ and $u_2 = -\frac{\phi}{N_g-1}$ and $s_1 = 1-\rho$ and $s_2 = \rho$. Then, $\Upsilon^{-1} = \frac{1}{u_1}I_{N_g} - \frac{u_2}{u_1(u_1+N_g u_2)}1_{N_g\times N_g}$, so $S = \Upsilon^{-1}\Sigma\Upsilon^{-1} = k_1 I_{N_g} + k_2 1_{N_g\times N_g}$, where $k_1 = \frac{s_1}{u_1^2}$ and $k_2 = \left(\frac{s_2}{u_1} - \frac{u_2 s_1 + N_g u_2 s_2}{u_1(u_1+N_g u_2)}\right)\frac{1}{u_1} - \frac{s_1}{u_1}\frac{u_2}{u_1(u_1+N_g u_2)} - \left(\frac{s_2}{u_1} - \frac{u_2 s_1 + N_g u_2 s_2}{u_1(u_1+N_g u_2)}\right)\frac{u_2 N_g}{u_1(u_1+N_g u_2)}$. This implies that $[S]_{i,-i} = k_2 1_{1\times(N_g-1)}$ and $[S]_{-i,-i} = k_1 I_{N_g-1} + k_2 1_{(N_g-1)\times(N_g-1)}$. Consequently, $([S]_{-i,-i})^{-1} = \frac{1}{k_1}I_{N_g-1} - \frac{k_2}{k_1(k_1+(N_g-1)k_2)}1_{(N_g-1)\times(N_g-1)}$. Therefore, $[S]_{i,-i}([S]_{-i,-i})^{-1} = \frac{k_2}{k_1+k_2(N_g-1)}1_{1\times(N_g-1)}$. This gives the claimed expression for η after substitution back in terms of the model primitives. $\square$

Also, note that the fact that $E(y_{ig}|w_g, y_{-i,g})$ is linear in $y_{-i,g}$ depends on the extra assumption that ϵ_g are distributed according to a multivariate normal distribution, and is not implied by the linear-in-means model alone. Some other distributions share this property (e.g., the multivariate student t distribution as in Kotz and Nadarajah (2004)), while others do not (e.g., the skewed multivariate normal distribution of Azzalini and Valle (1996) or the bivariate exponential distributions of Gumbel (1960)).

Similarly, the parameter $\tau = [\beta' \ \gamma']'$ does not directly give the effect of w_g on $E(y_g|w_g)$. Rather, it holds that $E(y_g|w_g) = \Upsilon^{-1}w_g\tau$, so that the effect of w_g on $E(y_g|w_g)$ depends both on τ and ϕ. This leads us to the following result:

Theorem 2.2. *Suppose the model is $y_{ig} = x_{ig}\beta + z_g\gamma + \frac{\phi}{N_g-1}\sum_{j=1,j\neq i}^{N_g} y_{jg} + \epsilon_{ig}$ with $E(\epsilon_g|w_g) = 0$. Suppose that $\phi \neq 1 - N_g$ and $\phi \neq 1$. Then, the effects of the various elements of w_g on $E(y_{ig}|w_g)$ are the following:*

1. *The effect of $x_{ig,s}$ is $\beta_s\lambda$.*
2. *The effect of $x_{jg,s}$ is $\beta_s\psi$, for $j \neq i$.*
3. *The effect of $z_{g,s}$ is $\tau_s\omega$,*

where $\lambda = \frac{1-N_g+\phi(N_g-2)}{(\phi-1)(N_g-1+\phi)}$, $\psi = \frac{-\phi}{(\phi-1)(N_g-1+\phi)}$, and $\omega = \frac{1}{1-\phi}$.

Proof. Let Υ be the $N_g \times N_g$ matrix with 1 on the diagonal and $\frac{-\phi}{N_g-1}$ on the off diagonal. Then, $y_g = \Upsilon^{-1}(w_g\tau + \epsilon_g)$. Consequently, $E(y_g|w_g) = \Upsilon^{-1}w_g\tau$.

Let C be the $N_g \times (N_g+1)$ matrix which is the I_{N_g} matrix in the first N_g rows and columns and $1_{N_g\times 1}$ in the last column. And let W_g be the $(N_g+1)\times(K+L)$

matrix where K is the dimension of the individual specific observables x_{ig} and L is the dimension of the dimension of the group specific observables z_g. Then let W_g be such that the first K columns of row i contain x_{ig}, and the last L columns of row $N_g + 1$ contain z_g. Consequently, $CW_g = w_g$, so $E(y_g|W_g) = \Upsilon^{-1}CW_g\tau$, and so $E(y_{ig}|W_g) = e_i\Upsilon^{-1}CW_g\tau$ where e_i is the unit row vector with a 1 in the ith column and 0s everywhere else. Consequently, $\frac{\partial E(y_{ig}|W_g)}{\partial W_g} = \tau e_i \Upsilon^{-1}C$. Using the same notation as in the proof of theorem 2.1, it holds that $\Upsilon^{-1} = \frac{1}{u_1}I_{N_g} - \frac{u_2}{u_1(u_1+N_g u_2)}1_{N_g \times N_g}$. After some algebra, $\tau e_i \Upsilon^{-1}C$ is the $(K + L) \times (N_g + 1)$ matrix where the ith column is $\tau\lambda$, all of the first N_g columns but the ith column is $\tau\psi$, and the $N_g + 1$st column is $\tau\omega$, where $\lambda = \frac{1}{u_1} - \frac{u_2}{u_1(u_1+N_g u_2)} = \frac{1-N_g+\phi(N_g-2)}{(\phi-1)(N_g-1+\phi)}$, $\psi = \frac{-u_2}{u_1(u_1+N_g u_2)} = \frac{-\phi}{(\phi-1)(N_g-1+\phi)}$, and $\omega = \frac{1}{u_1} - \frac{N_g u_2}{u_1(u_1+N_g u_2)} = \frac{1}{1-\phi}$. □

These expressions are equivalently also the *ceteris paribus* effects of the respective observables, holding the rest of the observables and unobservables fixed, and allowing the social interaction to re-equilibrate according to the linear-in-means model.

This theorem implies that for some comparisons (e.g., comparing the effect of $x_{ig,s}$ and $x_{ig,t}$) the relative effect on $E(y_{ig}|w_g)$ can be determined as the ratio of the parameters β and τ, while for others (e.g., comparing the effect of $x_{ig,s}$ with $x_{jg,s}$) the ratio of the parameters β and τ does not give the relative effects. Also, since $\omega = \lambda + (N_g - 1)\psi$, the effect of a marginal change in $z_{g,s}$ has the same effect as the same marginal change in everyone in the group's $x_{\cdot g,s}$, as long as the corresponding parameter values on $z_{g,s}$ equals that on $x_{\cdot g,s}$. This is expected since increasing a group specific observable affects everyone directly, which is equivalent to increasing everyone's individual specific observables. Also, under the conditions of corollary 2.1, when the corresponding parameter values are the same, the effect of a group specific observable is greater in magnitude than the effect of an individual's own individual specific observable when $\phi > 0$ and is lesser in magnitude when $\phi < 0$. This implies the necessity to explicitly specify which observables are individual specific and which are group specific, and to treat these two sets of observables in distinct ways in the analysis.

This theorem also implies that the direction of the effect of components of w_g depends not only on the parameters τ, but also, in the case of the effect of an individual specific observable of a peer, also on ϕ. Thus, we are led to the following result:

Corollary 2.1. *Under the same conditions, as long as $-1 < \phi < 1$ and $N_g > 2$, it holds that:*

(1) *The sign of the effect of* $x_{ig,s}$ *is the same as the sign of* β_s.
(2) *The sign of the effect of* $x_{jg,s}$ *for* $j \neq i$ *is the same as the sign of* β_s *times the sign of* ϕ.
(3) *The sign of the effect of* $z_{g,s}$ *is the same as the sign of* τ_s.

Proof. This follows by inspecting the signs of λ, ψ, and ω. Since $\phi < 1$, $\phi - 1$ is negative, and since $\phi > -1$ and $N_g > 2$, $N_g - 1 + \phi$ is positive. Thus, the denominators of λ and ψ are negative. The numerator of λ is increasing in ϕ since $N_g - 2 > 0$, and equals 0 when $\phi = \frac{N_g-1}{N_g-2} > 1$, so is negative since $\phi < 1$. Thus, λ is positive. Since the denominator of ψ is negative, the sign of ψ is the same as the sign of ϕ. And, it is obvious by inspection that the sign of ω is positive since $\phi < 1$. □

This means that the parameters β and τ alone are sufficient to determine the direction of the effect of an individual's own individual specific observables, and the group specific observables.

However, the parameters β and τ alone can be misleading about the magnitude of the effects. This theorem implies that the effect of $x_{ig,s}$ on $E(y_{ig}|w_g)$ is equal to β_s if and only if $\phi = 0$, and otherwise has an effect that is greater in magnitude. Similarly, the effect of z_g on $E(y_{ig}|w_g)$ is equal to τ_s if and only if $\phi = 0$, and otherwise has an effect that is greater in magnitude. Thus, we have the following result:

Corollary 2.2. *Under the same conditions, as long as* $-1 < \phi < 1$ *and* $N_g > 2$, *it holds that:*

(1) *The effect of* $x_{ig,s}$ *is larger in magnitude than (the same as)* β_s *if and only if* $\phi \neq 0 (= 0)$.
(2) *The effect of* $z_{g,s}$ *is larger in magnitude than (the same as)* τ_s *if and only if* $\phi \neq 0 (= 0)$.

Also, as $\phi \nearrow 1$, *all of these effects approach* ∞ *in magnitude.*

Proof. The first claims follow by comparing λ and ω to 1. The parameter λ equals 1 if and only if $\phi = 0$. The derivative of λ with respect to ϕ is $\phi\frac{(2(\phi-1)-N_g(\phi-2))}{(\phi-1)^2(N_g-1+\phi)^2}$. The denominator is positive, and the numerator has the same sign as ϕ since $N_g > 2$. Thus, λ decreases to 1 as ϕ increases for $\phi < 0$ and increases from 1 as ϕ increases for $\phi > 0$. And, it is obvious by inspection that the parameter ω is equal to 1 if and only if $\phi = 0$, and otherwise is greater than 1, since $\phi < 1$.

The second claim follows by observing that λ, ψ, and ω all tend to $+\infty$ as $\phi \nearrow 1$. □

This result shows that the parameters β and τ necessarily understate the magnitude of these average marginal effects except in the special case of $\phi = 0$, and indeed may understate the effects by an arbitrarily large ratio, depending on the value of ϕ.

2.1. **Implications of these results to the application of the linear-in-means model.** Both Theorem 2.1 and Theorem 2.2 show that if these average marginal effects, or in the case of w_g these *ceteris paribus* effects, are the objects of interest then applications of the linear-in-means model should report these functions of the parameters of the model rather than the parameters of the model directly. This is similar to the standard that when estimating discrete choice models (e.g., a binary probit model) the estimated marginal effects are reported, not the parameters of the model directly. This requires also the reporting of a different standard error to account for the different statistic being reported, which might be accomplished by the delta method.

Also, without further restrictions, the linear in means model above might not be *coherent* in general, i.e., the exogenous variables (both observed and unobserved) might not generate a well defined joint distribution of the outcomes without further restrictions. These restrictions are similar to whether the predicted joint distribution of the observables factors out. See Arnold, Castillo, and Sarabia (1999) for more on this.

In short, these results imply that the interpretation of the linear-in-means model, and the parameter ϕ in particular, is complicated and non-standard. The fundamental reason for this, as has been seen in the above, is that *all* of the $y_{\cdot g}$ are endogenous in the linear-in-means model, and it is this *simultaneity* that causes the interpretation of coefficients directly to be problematic: A "change" in the outcomes of the rest of the group, according to the linear-in-means model, affects the treated individual, but then the change in that individual's outcome feeds back onto the "change" in the outcomes of the rest of the group. Consequently, it requires some work to interpret the model as providing the "effect" of the outcomes of the rest of the group. Similarly, because of the social interaction effect, and hence the underlying simultaneity, the effect of the observables w_g is not simply the parameters β and τ. The next section provides conditions under which the linear in means models inherits a structural interpretation by treating it as a model of an economic game with the corresponding response functions as given above.

3. The linear-in-means game and the interpretation of the parameters as a best response

We propose in this paper that if the object of interest is in some sense of the "effect" of a policy experiment that manipulates the average outcome of an individual's group, perhaps motivated by the experiment of moving that individual to a new group, then it is useful to interpret the linear-in-means model as corresponding to a game. This will allow the interpretation of ϕ as a "best response" based on that game.

In particular, suppose that group g is playing the following complete information game. Each individual in group g has a publicly known type θ_{ig}, and the utility of individual i of type θ_{ig} is $u(y_{ig}, y_{-i,g}; \theta_{ig}) = \theta_{ig} y_{ig} - \frac{1}{2} y_{ig}^2 + \frac{\phi}{N_g - 1} y_i \sum_{j=1, j \neq i}^{N_g} y_{jg}$. This is strategically equivalent to the game with utilities $u(y_{ig}, y_{-i,g}; \theta_{ig}) = \theta_{ig} y_{ig} - \frac{1}{2}(y_{ig} - \frac{2\phi}{N_g - 1} \sum_{j=1, j \neq i}^{N_g} y_{jg})^2$. Consequently, this game can be interpreted to be a game in which individuals have a preference to conform their own outcome to $\frac{2\phi}{N_g - 1} \sum_{j=1, j \neq i}^{N_g} y_{jg}$ under square loss. Also, individuals have a "private" utility of θ_{ig} per unit of their outcome y_{ig}. We suppose that individuals choose their outcome y_{ig} and that the outcome of this game of choosing $y_{\cdot g}$ is described by Nash equilibrium. This game generates the linear-in-means econometric model.

Theorem 3.1. *Suppose that the outcome of group g is a Nash equilibrium of the complete information game in which $u(y_{ig}, y_{-i,g}; \theta_{ig}) = \theta_{ig} y_{ig} - \frac{1}{2} y_{ig}^2 + \frac{\phi}{N_g - 1} y_i \sum_{j=1, j \neq i}^{N_g} y_{jg}$. Suppose that $\theta_{ig} = x_{ig}\beta + z_g\gamma + \epsilon_{ig}$ where x_{ig} and z_g are observed by the econometrician, but ϵ_{ig} is unobserved and $E(\epsilon_g | w_g) = 0$. Suppose that $\phi \neq 1 - N_g$ and $\phi \neq 1$. Then, the data satisfY the linear-in-means econometric model: $y_{ig} = x_{ig}\beta + z_g\gamma + \frac{\phi}{N_g - 1} \sum_{j=1, j \neq i}^{N_g} y_{jg} + \epsilon_{ig}$ where $E(\epsilon_g | w_g) = 0$. Also, the best response functions in this game are $\upsilon(t_{-ig}; \theta_{ig}) = x_{ig}\beta + z_g\gamma + \epsilon_{ig} + \frac{\phi}{N_g - 1} \sum_{j=1, j \neq i}^{N_g} t_{jg}$, where the notation $t_{\cdot g}$ is used to indicate the best response is a function of* any *possible outcomes of the others, not just the equilibrium outcomes y_{-ig} that are observed in the data. And consequently the average best response function is $E(\upsilon(t_{-i,g}; \theta_{ig}) | w_g) = x_{ig}\beta + z_g\gamma + \frac{\phi}{N_g - 1} \sum_{j=1, j \neq i}^{N_g} t_{jg}$.*

Proof. Any Nash equilibrium of this game must be in pure strategies, because the first order condition of the expected utility function taking as given the strategies of the others is $\theta_{ig} - y_{ig} + \frac{\phi}{N_g - 1} \sum_{j=1, j \neq i}^{N_g} E(y_{jg}) = 0$. The second order condition is -1, so that there is a unique optimal y_{ig} in response to any mixed strategies of the others. Consequently, for any mixed strategy of individuals $-i$, individual i's unique optimizing choice is $y_{ig} = \theta_{ig} + \frac{\phi}{N_g-1}\sum_{j=1,j\neq i}^{N_g} E(y_{jg})$. Consequently, in Nash equilibrium each individual uses a pure strategy.

The best responses in this game are $\upsilon(t_{-ig};\theta_{ig}) = \theta_{ig} + \frac{\phi}{N_g-1}\sum_{j=1,j\neq i}^{N_g} t_{jg}$, where the notation $t_{\cdot g}$ is used to indicate the best response is a function of *any* possible outcomes of the others, not just the equilibrium outcomes y_{-ig} that are observed in the data. Since $\phi \neq 1 - N_g$ and $\phi \neq 1$ the game has a unique equilibrium outcome. Then, in the unique pure strategy Nash equilibrium outcome of this game, it holds that $y_{ig} = \theta_{ig} + \frac{\phi}{N_g-1}\sum_{j=1,j\neq i}^{N_g} y_{jg}$. If $\theta_{ig} = x_{ig}\beta + z_g\gamma + \epsilon_{ig}$, then $y_{ig} = x_{ig}\beta + z_g\gamma + \epsilon_{ig} + \frac{\phi}{N_g-1}\sum_{j=1,j\neq i}^{N_g} y_{jg}$. Thus, given the parameterization of θ_{ig}, this game generates data according to the linear-in-means model. $\square$

Note that in *the best response function*, the outcomes of the others are *held constant and are non-random*, so that it is not necessary to condition on them as it was in the previous section in the context of the linear-in-means model without an underlying game. Thus, ϕ is the marginal effect of the peers' average outcome on the average best response function. This provides a justification for the interpretation of ϕ as a marginal effect of the peers' average outcome.

The extra assumption of the linear-in-means *game* is necessary for this interpretation. Under only the linear-in-means econometric model, all that the model specifies is that y_g is some simultaneously determined function of the observables w_g and unobservables ϵ_g. That model alone does not provide the way to "predict" what would happen when some subset of y_g are specified "off the equilibrium." But, with the assumption that the linear-in-means *game* is generating the data, and describes the social interaction, then ϕ may be interpreted as a parameter of the best response in that game. We elaborate on this point further in the conclusions.

This allows extrapolating from observations of Nash equilibrium behavior to draw conclusions about responses to manipulating an individual's peers' average outcome. This is necessarily a "non-equilibrium quantity" because it involves specifying some of the values of the outcomes of the social interaction. This specification of an individual's peers' average outcome together with the best response to that average will generically not be a Nash equilibrium. This is a limitation of any approach of "estimating" a peer effect of average outcomes on an individuals outcome, because the outcomes are endogenous and simultaneously determined and cannot necessarily be directly manipulated while maintaining equilibrium.

3.1. **Can other games generate the linear-in-means econometric model?** The game considered above is not the only game generating the linear-in-means econometric model. Alternatively, suppose each individual in group g has a publicly known type θ_{ig}, and that the utility of individual i is $u(y_{ig}, y_{-i,g}; \theta_{ig}) = \theta_{ig} y_{ig} - \frac{y_{ig}^2}{2} - \frac{\eta}{2}(y_{i,g} - \frac{1}{N_g - 1}\sum_{j=1,j\neq i}^{N_g} y_{j,g})^2$. This game can be interpreted to be a game in which individuals wish to conform to $\frac{1}{N_g-1}\sum_{j=1,j\neq i}^{N_g} y_{j,g}$ with "intensity" η. The best response in this game is $\upsilon(t_{-ig}; \theta_{ig}) = \frac{\theta_{ig}}{\eta+1} + \frac{\eta}{\eta+1}\frac{1}{N_g-1}\sum_{j=1,j\neq i}^{N_g} t_{jg}$. Suppose that $\eta \neq \frac{1-N_g}{N_g}$, so that the game has a unique equilibrium outcome. Then, in the unique pure strategy Nash equilibrium using the same functional form assumptions on θ_{ig}, $y_{ig} = \frac{x_{ig}\beta + z_g\gamma + \epsilon_{ig}}{1+\eta} + \frac{\eta}{1+\eta}\frac{1}{N_g-1}\sum_{j=1,j\neq i}^{N_g} y_{jg}$. So, up to a re-parametrization of the model, this game also generates data according to the linear-in-means model, and the rest of the analysis above follows similarly. This parameterization of the game may be better, as it results in preferences for *conformity of outcomes* with an intensity parameter η.

However, there is a sense in which this sort class of game is the only game that can be used to interpret the linear-in-means model, as the following theorem establishes.

Theorem 3.2. *Suppose that individual i with type θ_i has utility function $u_i(y_i, y_{-i}; \theta_i)$, defined for actions on the entire real line. Suppose that $u_i(\cdot)$ is a polynomial in y_i, and in particular is such that maximizing $u_i(\cdot)$ with respect to y_i follows from the usual first and second order conditions, and so also in particular that $\frac{\partial^2 u_i}{\partial y_i^2} < 0$. And suppose that the first order condition has the unique solution $y_i = \theta_i + \frac{\phi}{N_g-1}\sum_{j\neq i} y_j$ over the complex numbers. Then it must be that $u_i(y_i, y_{-i})$ is equal to $\theta_i y_i - \frac{y_i^2}{2} + \frac{\phi}{N_g-1}\sum_{j\neq i} y_j$ up to positive affine transformation, where the constants in the transformation may depend on y_{-i} and θ but not y_i, and so in particular for any y_{-i} individual i has an incentive to conform to $\frac{1}{N_g-1}\sum_{j\neq i} y_j$.*

Proof. Since the first order condition has the unique solution $y_i = \theta_i + \frac{\phi}{N_g-1}\sum_{j\neq i} y_j$ over the complex numbers, by the Fundamental Theorem of Algebra it must be that $u_i'(y_i, y_{-i}; \theta_i) = C_i(y_{-i}, \theta)(y_i - \theta_i - \frac{\phi}{N_g-1}\sum_{j\neq i} y_j)$. Consequently, it must be that $u_i(y_i, y_{-i}; \theta_i) = C_i(y_{-i}, \theta)(\frac{y_i^2}{2} - \theta_i y_i - \frac{\phi}{N_g-1} y_i \sum_{j\neq i} y_j) + D_i(y_{-i}, \theta)$. Since the second derivative of utility is assumed to be negative, it must be that $C_i(y_{-i}, \theta) < 0$. So, the utility function is equal to $\theta_i y_i - \frac{y_i^2}{2} + \frac{\phi}{N_g-1}\sum_{j\neq i} y_j$ up to positive affine transformation, where the constants in the transformation may depend on y_{-i} and θ but not y_i. □

3.2. **Limitations of the linear-in-means game above.** We provided above one model of a game in which the linear in means model is a best response function of the game. However, this economic model is making particular assumptions about the relationship between the data and the underlying interaction. These particular assumptions introduce limitations. We list and discuss three such limitations below. These are made in terms of comments.

Comment 1: The linear in means game we specified above is not the only game that leads to the linear in means model. That other economic games can lead to the same econometric model affects the interpretation of the parameters in that latter model in that for example, these parameters cannot be given the best response interpretation. For example, consider a two-stage interaction where in the first stage individuals choose their peers and in the second stage individuals play the linear-in-means game above with their peers. Suppose that in the first stage individuals choose their peers partly on the basis of their predicted outcome of the second stage. Then, *under the assumption that the data come from an equilibrium in this game*, the data will satisfy the linear-in-means econometric model. However, upon manipulating an individual's peers to have different outcomes, it is plausible that the individual will also re-optimize the choice of peers, which implies that the linear-in-means model does not have a best response interpretation. For example, in an equilibrium of this two-stage interaction each individual may "pay attention to" its peers, having chosen its peers, but upon "assigning" peers to that individual with different outcomes than the equilibrium outcomes, the individual may then choose to ignore its peers. This might present difficulties for using data on equilibrium outcomes from the linear in means game to predict outcomes from a policy question that "assigns" peers.

Comment 2: The assumption that data are realizations from the intersection of best responses implicitly entails the assumption of *pure strategy Nash equilibrium play*. It is well known that Nash equilibrium play often is not a reasonable characterization of actual behavior. Many alternatives to Nash equilibrium play are entertained both in the theoretical game theory literature and also the experimental literature. Among many other examples this includes rationalizability (i.e., Bernheim (1984) and Pearce (1984)), quantal response (i.e., McKelvey and Palfrey (1995)), and models of "bounded reasoning" (i.e., Camerer, Ho, and Chong (2004) and Costa-Gomes and Crawford (2006)). Also, it entails the assumption of complete information. It may be plausible in some settings that individuals only have non-degenerate beliefs over the types of its peers. In the case of incomplete information the analog to the best response is an individual's action as a function of its realized type and its *beliefs* over its peers. Thus, the best response interpretation of a counterfactual involving the manipulation of an individuals' peers is not obvious in the case of incomplete information. Rather, the relevant counterfactual would involve the manipulation of an individual's beliefs about its peers, which is not the same as manipulating its actual peers. Under incomplete information it is possible for a manipulation of an individual's actual peers to have no effect on outcomes, but for a manipulation of an individual's beliefs about its peers to have an effect. We leave a further development of this to future work[5]. In addition, the game above implicitly rules out mixed strategy equilibria since we assume that the best response functions are satisfied in the data. This might be important because in some setups of interest (see Section 4 below), it is plausible to have a model for the interaction in which the only equilibria of the game are ones in mixed strategies (this effectively means that every player is indifferent in playing any of the actions on the support of his equilibrium strategy).

Comment 3: The above game can admit multiple equilibria. This is not without loss of generality even in the above quadratic utility game. For example, if one assumes that outcomes have support on $[0, \infty]$, then it is easy to see that the above game can have multiple equilibria many of which are on the "boundary" where many players decide to play $y_i = 0$, and unique "interior" equilibria are a special case that depend on particular values of the parameters. This problem of multiplicity by itself does not present problems since the best response functions for that game still hold in the data regardless of whether these best responses lead to multiplicity. However, care should be taken in interpreting the predictions of the model under multiplicity.

4. Educational achievement with peer effects

In this section we focus on the applications of models of social interactions to classroom production, where the outcome is educational achievement that is measured, for example, by test scores. Our intention in this section is to highlight the assumptions implicit in the linear-in-means game, and thus the assumptions implicit in the interpretation of the parameters of the linear-in-means econometric model as a best response function as described above.

[5]See Blume, Brock, Durlauf, and Ioannides (2011) who have made contributions in the incomplete information game.

One possible mechanism for the social interaction in the case of educational achievement is competition among the students. Students might be competing for grades assigned based on relative performance (e.g., "grading on a curve"), for permission to take subsequent classes that take only the highest achieving students (e.g., ability tracking), for interviews or jobs that depend on class rank (e.g., the labor market for business, law, medical, and perhaps other students in the United States and similar countries, or the labor market for government jobs in some countries), for letters of recommendations from the instructor, or simply for the psychic benefit of doing better than peers. This is related to the idea that education acts as a "filter" or "screen" for ability (i.e., Arrow (1973)).

This suggests that a plausible economic model for educational achievement in a classroom is that if student i has the k-th highest achievement, perhaps as measured by test score, it gets a reward of V_{ik} where $V_{i1} \geq V_{i2} \geq \cdots \geq 0$.[6] The students must pay a cost of effort to produce educational achievement: the cost of educational achievement b_i for student i is $c_i(b_i)$ where $c_i(\cdot)$ is a cost function. This cost might reflect the psychic cost of learning, the opportunity cost of time in general, or, perhaps especially in developing countries, the opportunity cost of time as it relates to foregone wages or contribution to the household. Let $V_i(b_i, b_{-i})$ be the reward V_{ik} when b_i is the k-th highest achievement. Therefore the payoff to student i from educational achievement b_i when the other students have educational achievements b_{-i} is $V_i(b_i, b_{-i}) - c_i(b_i)$. Economic models like this are known variously as *all-pay auctions* or *contests*.

In general any Nash equilibrium of this interaction must involve randomization, either because there is complete information and students use mixed strategies or because there is incomplete information. The intuition for this result is evident in the case of two students and one reward with asymmetric valuation, with complete information and a shared cost function $c_i(b) = b$. Suppose that student 1 chooses educational achievement $b_1 > 0$ as a pure strategy. Student 2 has two possible "best responses:" either it selects an achievement of 0 and gets payoff 0 or it selects an achievement of $b_1 + \epsilon$ for $\epsilon > 0$ very small and gets a payoff of $V_2 - b_1 - \epsilon$.[7] In the

[6] We can allow the valuations of the rewards to be symmetric or asymmetric across students. The valuations are symmetric if $V_{ik} = V_{jk}$ for all $i \neq j$ and all k, and are asymmetric otherwise. If the valuations are symmetric, the indexing by i is irrelevant. Also, when the valuations are asymmetric, the valuations can either be commonly known or private information.

[7] Note that student 2 does not technically have a best response to a pure strategy when $b_1 < V_1$. This is not a problem for us, exactly because we are showing that we should expect mixed strategies in equilibrium.

former case student 1 has exerted too much effort and would prefer to deviate to an achievement of $0+\epsilon$, so that it still wins V_1, but at a lower cost. In the latter case student 1 would prefer to deviate either to an achievement of 0, so that it still loses but exerts no effort, or to just out-achieve student 2's achievement of $b_1+\epsilon$. This heuristic suggests that in general that there can be no Nash equilibrium in pure strategies in this sort of model.

Results like this have been proved formally under a variety of specifications: complete information with asymmetric valuations for one reward (i.e., Hillman and Riley (1989), Baye, Kovenock, and De Vries (1993), and Baye, Kovenock, and De Vries (1996)), complete information with asymmetric valuations for one reward under various types of upper bounds on the allowed achievement[8] (i.e., Che and Gale (1998), Kaplan and Wettstein (2006), and Kline (2009)), the equilibrium with complete information with symmetric valuations for many rewards (i.e., Barut and Kovenock (1998)), and the symmetric equilibrium with incomplete information about cost functions with symmetric valuations for many rewards (i.e., Moldovanu and Sela (2001)). A more general study of these types of models is Siegel (2009). We do not claim that any of these models are necessarily completely satisfactory as a model of competition in educational achievement. Rather, we simply make the case that it is a reasonably robust result that randomization is necessary in equilibrium in this type of model.

Thus, this model cannot possibly be "represented" by the linear-in-means game interpretation of the linear-in-means model, because the linear-in-means game entails both complete information and the use of pure strategies. Furthermore, and perhaps a more fundamental problem with the linear-in-means game, the incentives of this game are significantly different from the incentives of the linear-in-means game. The linear-in-means game entails a social interaction where the social interaction is due to a preference for *conformity of outcomes*. This may be plausible in some settings like drug and alcohol use, but is less plausible in the setting of educational achievement. There is not a preference for conformity of outcomes in the interaction based on competition in the all pay contest just discussed above. Sacerdote (2011) dicusses other possible mechanisms for social interaction in the case of educational achievement. Among these possible mechanisms are models where the best student in the peer group matters (as a role model for the rest of the group) and the worst student in

[8]This has been previously studied in the context of political lobbying, where the upper bound is understood to be a law that limits lobbying. In the context of educational achievement this upper bound might reflect the fact that there is an upper bound on "humanly possible" educational achievement, or a maximum possible test score.

the peer group matters (as a distraction for the rest of the group). In neither of these cases, nor the others discussed by Sacerdote (2011), does it appear that the social interaction is due (at least primarily) to a preference for conformity of outcomes.

5. Conclusions

In this paper we have studied the interpretation of the linear-in-means model. First, we have shown that standard policy relevant parameters related to average treatment effects are not directly the parameters of the linear-in-means model, but rather functions of these parameters and possibly also the parameters of the distribution of the unobservables. These results show that a "standard" interpretation of the parameters of the model in analogy to the ordinary linear model are not necessarily justified. Second, we have shown that it is possible to interpret the parameters of the linear-in-means model in almost the "standard" way under the additional assumption of the linear-in-means game. But, third, we have shown that the game that leads to the linear-in-means model places strong conditions on the nature of the social interaction allowed, and may or may not be plausible for any particular social interaction under consideration. In particular the linear-in-means model is shown to make most sense in settings where there is a preference for conformity of outcomes. In addition, the game leading to the linear in means model implicitly assumes Nash behavior in pure strategies. These various results are summarized in table 5, which shows four possible effects of interest in the linear-in-means model and the corresponding result in this paper giving that effect.

The use of the linear-in-means game to interpret the parameters of the linear-in-means model is analogous to the situation with models of supply and demand or other structural simultaneous equations models. In the classical model of supply and demand by linear simultaneous equations the model is

$$q_i = \gamma_s p_i + x_i \beta_s + u_{is}$$

$$p_i = \gamma_d q_i + x_i \beta_d + u_{id}$$

where the first equation is the supply as a function of price and other variables, and the second equation is the willingness to pay as a function of quantity and other variables (e.g., Amemiya (1985, p. 228)). In this model price and quantity are simultaneously

on	Effect of	
	w_g	y_{-ig}
Equilibrium outcomes	$E(y_{ig}\|w_g)$, or *ceteris paribus* equilibrium y_{ig} as a function of w_g discussed in Theorem 2.2	a *ceteris paribus* interpretation of this quantity is not defined, but Theorem 2.1 gives $E(y_{ig}\|w_g, y_{-ig})$
Out of equilibrium responses	given by β and τ interpreted as responses of the linear-in-means game in Theorem 3.1. But effect on equilibrium outcomes may be more interesting.	given by β and τ interpreted as responses of the linear-in-means game in Theorem 3.1

TABLE 1. Possible effects of interest in the linear-in-means model

determined, according to the equilibrium condition that supply equals demand, and so this model alone does not give the response of supply to price. But, in addition to this econometric model is the supply model (alone) from economic theory which implies that a competitive market has a supply function that is well defined at non-equilibrium prices. The condition that each equation of the model has meaning independently from the rest of the simultaneous equations model is referred to as autonomy, and is closely related to the causal interpretation of the parameters of the model. See for example Goldberger (1989) who discusses this issue[9].

In the supply and demand model the economic model is that in addition to the econometric model observed at equilibrium values of quantity and price, it also holds that

$$q_i(t) = \gamma_s t + x_i\beta_s + u_{is}$$

$$p_i(t) = \gamma_d t + x_i\beta_d + u_{id}$$

because of the existence of an economic model that gives supply as a function of any price and willingness to pay as a function of any quantity, not just equilibrium prices and quantities. Then, under this standard assumption, it is possible to interpret γ_s as the response of supply to price even at non-equilibrium prices and to interpret γ_d as the response of willingness to pay even at non-equilibrium quantities. The additional assumption that there is a well-defined supply function and well-defined willingness to pay function independent of the equilibrium supply and demand process is so standard that it need not be stated in applications of this model. But, with the linear-in-means model it is necessary to provide an explicit and rigorous justification for interpreting the parameters of the linear-in-means model as a sort of "response" function, but in this case as a best response function in a particular game rather than as a competitive market supply function or willingness to pay function. In this paper we have shown a way to interpret the linear-in-means model to satisfy this autonomy requirement. But, also we have shown that this does entail additional restrictions on the social interaction.

We conclude with two remarks related to the modeling and interpretation of social interactions.

[9]See also Wooldridge (2002, p. 209-211), which points out that many applications of simultaneous equations models do not satisfy this autonomy requirement, complicating the interpretation of the parameters in those models.

Remark 5.1 (Alternative models of interaction)**.** *There are at least two possible responses to our results about the interpretation of the linear-in-means model.*

First, it is possible to conclude that the behavioral assumptions that are implicit in the linear-in-means game are correct, and if so to interpret the linear-in-means model as usual. This means ruling out a class of plausible economic model, especially when there is incomplete information or the use of mixed strategies, as in the example of competition in educational achievement.

Second, it is possible to use a different model for the social interaction, perhaps based on a game theoretic model that is interaction specific. The most closely related paper in this literature is Kline and Tamer (2010), which studies the partial identification of best responses in binary games, under a variety of behavioral assumptions. Other papers like Bresnahan and Reiss (1991), Tamer (2003), Ciliberto and Tamer (2009), Bajari, Hong, and Ryan (2010), and Bajari, Hong, Krainer, and Nekipelov (2010) study identification of the payoffs of a game, which is stronger than identification of best responses.

Remark 5.2 (Applicability of these conclusions to experimental data)**.** *The results in this paper apply equally to observational data and experimental data. There are obviously many advantages of experiments in areas like development and labor economics. Nevertheless, the conclusions of this paper suggest that, despite these advantages, there is still significant scope for econometric modeling of the data from experiments when there is social interaction. While the experiment helps to resolve endogeneity issues related to the relation between observables and unobservables, econometric modeling helps resolve issues related to the interpretation of the parameters of the model of the social interaction, as discussed in this paper.*

References

Amemiya, T. (1985): *Advanced econometrics*. Harvard Univ Press.

Arnold, B., E. Castillo, and J. Sarabia (1999): *Conditional specification of statistical models*. Springer Verlag.

Arrow, K. (1973): "Higher education as a filter," *Journal of Public Economics*, 2(3), 193–216.

Azzalini, A., and A. Valle (1996): "The multivariate skew-normal distribution," *Biometrika*, 83(4), 715.

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

Bajari, P., H. Hong, and S. P. Ryan (2010): "Identification and estimation of a discrete game of complete information," *Econometrica*, 78(5), 1529–1568.

Barut, Y., and D. Kovenock (1998): "The symmetric multiple prize all-pay auction with complete information," *European Journal of Political Economy*, 14(4), 627–644.

Baye, M., D. Kovenock, and C. De Vries (1993): "Rigging the lobbying process: an application of the all-pay auction," *American Economic Review*, 83(1), 289–294.

——— (1996): "The all-pay auction with complete information," *Economic Theory*, 8(2), 291–305.

Bernheim, B. (1984): "Rationalizable strategic behavior," *Econometrica*, 52(4), 1007–1028.

Blume, L., W. Brock, S. Durlauf, and Y. Ioannides (2011): "Identification of social interactions," *Handbook of Social Economics*, 1, 855–966.

Bramoullé, Y., H. Djebbari, and B. Fortin (2009): "Identification of peer effects through social networks," *Journal of Econometrics*, 150(1), 41–55.

Bresnahan, T., and P. Reiss (1991): "Empirical models of discrete games," *Journal of Econometrics*, 48(1-2), 57–81.

Brock, W. A., and S. N. Durlauf (2001): "Discrete choice with social interactions," *The Review of Economic Studies*, 68(2), 235–260.

Camerer, C., T. Ho, and J. Chong (2004): "A cognitive hierarchy model of games," *Quarterly Journal of Economics*, 119(3), 861–898.

Che, Y., and I. Gale (1998): "Caps on political lobbying," *American Economic Review*, 88(3), 643–651.

Ciliberto, F., and E. Tamer (2009): "Market structure and multiple equilibria in airline markets," *Econometrica*, 77, 1791–1828.

COSTA-GOMES, M., AND V. CRAWFORD (2006): "Cognition and behavior in two-person guessing games: an experimental study," *American Economic Review*, 96(5), 1737–1768.

DAVEZIES, L., X. D'HAULTFOEUILLE, AND D. FOUGÈRE (2009): "Identification of peer effects using group size variation," *Econometrics Journal*, 12(3), 397–413.

DE GIORGI, G., M. PELLIZZARI, AND S. REDAELLI (2010): "Identification of social interactions through partially overlapping peer groups," *American Economic Journal: Applied Economics*, 2(2), 241–275.

GAVIRIA, A., AND S. RAPHAEL (2001): "School-based peer effects and juvenile behavior," *Review of Economics and Statistics*, 83(2), 257–268.

GOLDBERGER, A. (1989): "Interviewed by Nicholas M. Kiefer," *Econometric Theory*, 5(01), 133–160.

——— (1991): *A course in econometrics*. Harvard Univ Pr.

GOLDSMITH-PINKHAM, P., AND G. W. IMBENS (2013): "Social networks and the identification of peer effects," *Journal of Business & Economic Statistics*, 31(3), 253–264.

GRAHAM, B. (2008): "Identifying social interactions through conditional variance restrictions," *Econometrica*, 76(3), 643.

GRAHAM, B., AND J. HAHN (2005): "Identification and estimation of the linear-in-means model of social interactions," *Economics Letters*, 88(1), 1–6.

GUMBEL, E. (1960): "Bivariate exponential distributions," *Journal of the American Statistical Association*, 55(292), 698–707.

HILLMAN, A., AND J. RILEY (1989): "Politically contestable rents and transfers," *Economics & Politics*, 1(1), 17–39.

KAPLAN, T., AND D. WETTSTEIN (2006): "Caps on political lobbying: Comment," *American Economic Review*, pp. 1351–1354.

KLINE, B. (2009): "A restriction on lobbyist donations," *Economics Letters*, 104(3), 129–132.

KLINE, B., AND E. TAMER (2010): "Bounds for best response functions in binary games," forthcoming in *Journal of Econometrics*.

KOTZ, S., AND S. NADARAJAH (2004): *Multivariate t distributions and their applications*. Cambridge Univ Pr.

LEE, L. (2007): "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects," *Journal of Econometrics*, 140(2), 333–374.

Manski, C. (1993): "Identification of endogenous social effects: the reflection problem," *The Review of Economic Studies*, 60(3), 531–542.

McKelvey, R., and T. Palfrey (1995): "Quantal response equilibria for normal form games," *Games and Economic Behavior*, 10(1), 6–38.

Moldovanu, B., and A. Sela (2001): "The optimal allocation of prizes in contests," *American Economic Review*, 91(3), 542–558.

Pearce, D. (1984): "Rationalizable strategic behavior and the problem of perfection," *Econometrica*, 52(4), 1029–1050.

Sacerdote, B. (2011): "Peer effects in education: How might they work, how big are they and how much do we know thus far?," *Handbook of the Economics of Education*, 3, 249–277.

Siegel, R. (2009): "All-pay contests," *Econometrica*, 77(1), 71–92.

Tamer, E. (2003): "Incomplete simultaneous discrete response model with multiple equilibria," *Review of Economic Studies*, 70(1), 147–165.

Wooldridge, J. (2002): *Econometric analysis of cross section and panel data*. The MIT press.