ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Treatment Effects With Heterogeneous Externalities

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Availability:
This version is available at: https://hdl.handle.net/11585/715949 since: 2022-04-12

Published:

DOI: http://doi.org/10.1080/07350015.2019.1592755

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

28 April 2024

This is the final peer-reviewed accepted manuscript of:

Arduini, T., Patacchini, E., & Rainone, E. (2020). Treatment effects with heterogeneous externalities. Journal of Business & Economic Statistics, 38(4), 826-838.

The final published version is available online at:

https://doi.org/10.1080/07350015.2019.1592755

Rights / License:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/)

When citing, please refer to the published version.

Treatment Effects with Heterogeneous Externalities

Tiziano Arduini* Eleonora Patacchini† Edoardo Rainone‡

Abstract

This paper proposes a new method for estimating heterogeneous externalities in policy analysis when social interactions take the linear-in-means form. We establish that the parameters of interest can be identified and consistently estimated using specific functions of the share of the eligible population. We also study the finite sample performance of the proposed estimators using Monte Carlo simulations. The method is illustrated using data on the PROGRESA program. We find that more than 50 percent of the effects of the program on schooling attendance are due to externalities, which are heterogeneous within and between poor and nonpoor households.

Keywords: Program evaluation, Two-Stage Least Squares, Indirect Treatment Effect.

JEL classification: C13, C21, D62

*University of Bologna, E-mail: tiziano.arduini@unibo.it.

†Corresponding author. Cornell University, EIEF, CEPR and IZA. E-mail: ep454@cornell.edu

‡Bank of Italy, E-mail: edoardo.rainone@bancaditalia.it. The views expressed here do not necessarily reflect those of the Bank of Italy.

1 Introduction

Conditional cash transfers, or similar welfare programs, generate indirect effects on untreated subjects in the presence of social interactions. Examining housing mobility studies in which households in poor areas are financed to relocate to better neighborhoods, Sobel (2006) shows that ignoring indirect effects leads to entirely wrong conclusions about the effectiveness of the program. In recent years, an applied literature has considered experimental designs and experimental sources of variation that allow the assumption of no interference between research subjects to be relaxed (see, e.g. Angelucci and De Giorgi, 2009; Barrera-Osorio et al., 2011). This literature has considered the presence of externalities in treatment evaluation, but has focused on the aggregate effect (i.e. the mean impact of the program over untreated subjects). Indirect effects, however, may be heterogeneous across population subgroups.

The analysis of heterogeneous indirect effects is complicated because of simultaneity issues arising from social interactions. Progress in this respect requires either improving the design of randomized control trials (Baird et al., 2018) or a methodological infrastructure, which is not available in the existing literature on policy evaluation.

Our paper aims to fill this gap, and to do so in a way that can be of immediate applicability for the practitioner. Our working tools are a linear-in-means model with a group-level structure (Lee, 2007) and a partial-population experiment (Moffitt et al., 2001). In this protocol, groups are assigned to treatment or control, and a subset of individuals are offered treatment within clusters assigned to treatment according to certain rules. We introduce heterogeneity both in exogenous and endogenous spillover effects.

Our main result demonstrates that all parameters of interest can be identified from experimental variation in the size of eligible individuals across groups. Our methodology exploits the statistical properties of a spatial autoregressive model, which embeds a recursive formulation of the endogenous regressors (representing externalities) over space. The reduced-form model provides exclusion restrictions stemming from higher order effects which are not included in the structural model. The fact that this property of the spatial autoregressive model can be exploited for the identification of peer effects was first noted by Calvó-Armengol et al. (2009) and Bramoullé et al. (2009) when network data are used. Its use is now pervasive in applied work in the social sciences, under the assumption that characteristics of friends of friends are valid instruments for the en-

dogenous effect (peer effects). The contribution of this paper is to bring this methodology to the treatment evaluation literature and to note that, in this context, we do not need network data: identification can be achieved by excluding the individual i from his or her peer group and using variations in the shares of treated individuals across groups in partial population experiments. Importantly, the randomization procedure guarantees that the shares of treated peers are random.

We also show how the parameters can be consistently estimated and we study the finite sample performance of the proposed estimators using Monte Carlo simulations. With the parameter estimates in hand, we then analytically derive causal estimands (the average treatment effect and the indirect treatment effect) in presence of heterogeneous externalities.

Our approach comes at the cost of specifying a parametric model. The main limitation in the use of parametric methods for estimating average treatment effects is the sensitivity of the results to misspecification of the conditional mean function. When a linear regression model is used to predict outcomes and the averages of the covariates in the treated and control samples are very different, the results can be sensitive to minor changes in the specification (Imbens and Woolridge, 2009). The commonly used nonparametric approach in the program evaluation literature is the potential outcome approach. The adoption of this framework in the presence of spillovers presents several challenges. In particular, when allowing for spillovers, the number of combinations of potential outcomes explodes with the number of interacting units and strong assumptions need to be made. A commonly used assumption in the existing studies in statistics imposes that spillover effects do not depend on the identity of each treated neighbor, so that interactions are anonymous (Sobel, 2006; Hudgens and Halloran, 2008; Tchetgen and VanderWeele, 2012; Liu and Hudgens, 2014; Rigdon and Hudgens, 2015, among others). Recently, in the context of social networks and using observational data, Forastiere et al. (2016) assume bounded interactions, specifically no interactions above distance one (i.e interferences are allowed only between direct friends). The problem is exacerbated when allowing for heterogeneous externalities. The statistical framework reduces quickly into an extremely large number of cells for which we need enough treated and untreated units. The presence of a parametric model allows us to instead make inference on empty cells. This comes at the cost of assuming constant functional form and parameters' values, even on a domain that is not used for the estimation of the structural parameters. We assume a linear structure of the model, which is the most commonly used parametric specification in applied work. Under this model specification, we can also decompose the estimands into spillovers due to

the direct effect of the treatment on peers and the indirect effect due to variations in the peers' behavior.

Our methodology is illustrated using data from one of the most-studied programs of poverty alleviation: the Mexican conditional cash transfer program PROGRESA (Programa de Educacion, Salud, y Alimentacion). We focus our analysis on the effects of the program on schooling decisions. We find that more than 50 percent of the effects are due to externalities, which are highly heterogeneous within and between household types. Most notably, the indirect effects on ineligible households are not entirely due to spillovers from eligible households. On the contrary, these treatment externalities are small. The aggregate effect of the treatment on nonpoor households is the result of a large social multiplier within nonpoor households. A 10 percentage point increase in school enrollment of eligible students is associated with a 4 percentage point increase in ineligible students' school attendance, whereas this effect more than doubles (8.8 percentage points) for a 10 percentage point increase in school attendance of ineligible students. Subgroup variation in impacts in PROGRESA has been documented by Djebbari and Smith (2008) and Lee and Shaikh (2014). Our study contributes to understanding the mechanisms driving such heterogeneity by showing that externalities differ within and across subgroups. The paper is organized as follows. Section 2 introduces our structural model. Section 3 shows how the model parameters are identified and estimated. Section 4 derives and decomposes the estimands that are of interest for policy purposes. Section 5 is devoted to the application of our framework in the case of PROGRESA. In Section 6, we discuss the implications of our analysis for the design of experiments and identify a variety of contexts in which our methodology can be applied. Section 6 concludes.

2 The Model

We observe a population with n agents distributed into $\bar{c}$ groups. The population is indexed by i. Groups are indexed by c with numerosity m_c. If externalities are constrained to be the same between and within groups, the outcome of interest y_{ic} is given by

$$y_{ic} = \phi \bar{y}_{-ic} + \delta t_{ic} + x_{ic}\beta + \bar{x}_{-ic}\gamma + \epsilon_{ic}, \tag{1}$$

where $\bar{y}_{-ic}$ indicates the mean outcome of the group (excluding individual i), x_{ic} and $\bar{x}_{-ic}$ denote vectors of individual and means of group characteristics (excluding individual i) respectively, and ϵ_{ic} is an error with mean zero and variance σ^2 for all i and c. In model (1), ϕ represents *the endogenous effects*, in which an agent's choice/outcome may depend on that of his/her peers, and γ represents the contextual effects, in which an agent's choice/outcome may depend on the exogenous characteristics of his/her peers, β captures the effects of individual characteristics. t_{ic} is a dummy variable indicating whether unit i in group c is treated and δ represents its effect (the direct effect of the treatment). In this model, we assume that a subset E of agents is eligible for treatment. Groups of eligible agents are randomly assigned to treatment following a randomized control trial. Let us suppose without loss of generality that all agents eligible for the treatment in treated groups are treated. The complement N is composed of ineligible agents. Denote e_c and n_c as the cardinalities of E and N in group c. If we allow externalities to be heterogeneous, treatment spillover effects can be studied using the model:

$$y_{ic}^E = \phi^E \bar{y}_{-ic}^E + \phi^{EN} \bar{y}_{ic}^N + \delta t_{ic} + x_{ic}^E \beta^E + \bar{x}_{-ic}^E \gamma^E + \bar{x}_{ic}^N \gamma^{EN} + \epsilon_{ic}^E, \tag{2}$$

$$y_{ic}^N = \phi^N \bar{y}_{-ic}^N + \phi^{NE} \bar{y}_{ic}^E + x_{ic}^N \beta^N + \bar{x}_{-ic}^N \gamma^N + \bar{x}_{ic}^E \gamma^{NE} + \epsilon_{ic}^N, \tag{3}$$

where $\bar{y}_{-ic}^E = \frac{e_c-1}{m_c-1} \frac{\sum_{j \in E, j \neq i} y_{jc}^E}{e_c-1}$ and $\bar{y}_{ic}^N = \frac{n_c}{m_c-1} \frac{\sum_{j \in N} y_{jc}^N}{n_c}$ with $\bar{y}_{-ic}^E + \bar{y}_{ic}^N = \bar{y}_{-ic}$ and $\bar{y}_{-ic}^N = \frac{n_c-1}{m_c-1} \frac{\sum_{j \in N, j \neq i} y_{jc}^N}{n_c-1}$ and $\bar{y}_{ic}^E = \frac{e_c}{m_c-1} \frac{\sum_{j \in E} y_{jc}^E}{e_c}$ with $\bar{y}_{-ic}^N + \bar{y}_{ic}^E = \bar{y}_{-ic}$, $\bar{x}_{-ic}^E = \frac{e_c-1}{m_c-1} \frac{\sum_{j \in E, j \neq i} x_{jc}^E}{e_c-1}$ and $\bar{x}_{ic}^N = \frac{n_c}{m_c-1} \frac{\sum_{j \in N} x_{jc}^N}{n_c}$ with $\bar{x}_{-ic}^E + \bar{x}_{ic}^N = \bar{x}_{-ic}$ and $\bar{x}_{-ic}^N = \frac{n_c-1}{m_c-1} \frac{\sum_{j \in N} x_{jc}^N}{n_c-1}$ and $\bar{x}_{ic}^E = \frac{e_c}{m_c-1} \frac{\sum_{j \in E} x_{jc}^E}{e_c}$ with $\bar{x}_{-ic}^N + \bar{x}_{ic}^E = \bar{x}_{-ic}$. ϕ^E and ϕ^N capture the *within-group externalities,* while ϕ^{EN} and ϕ^{NE} capture the *between-group externalities.* γ^E and γ^N are the *within-group contextual effects*, while γ^{EN} and γ^{NE} represent the*between-group contextual effects.*

3 Identification and Estimation

3.1 Identification

The following proposition establishes the conditions under which the parameters in model (2)-(3) are identified. Here, identification means that a consistent estimator of the parameters in equations (2) and (3) exists.

Let us assume the model (2)-(3) represents a social equilibrium, so that the reduced form can be derived.

Proposition 1 *Under a partial-population experiment, the parameters of model (2)-(3) are identified if and only if the share of eligible agents varies across groups.*

The proof is given in the Appendix. The result can be easily extended to models with any finite number of groups. The intuition is as follows. In the presence of social spillovers, the partial-population experiment generates exogenous variations of a variety of combinations of nonlinear functions of the share of the eligible population. If such shares vary across groups, the richness in the combinations can be used to identify the social spillover effects, even if they are assumed to be heterogeneous.

3.2 IV estimator

The spatial econometrics tradition builds instruments for the group average from its expected value conditional on the exogenous variables (see Kelejian and Prucha, 1998; Kelejian and Prucha, 1999; Lee, 2003). We follow this approach. Let $T_c = \{t_{ic}\}_{i \in E}$. From the reduced form model (equations (33)-(34) in the Appendix), we can write the conditional expected values of the endogenous variables as functions of the treatment vector. The expected values of the endogenous terms in equation (2) conditional on the treatment are:

$$TIV_{Ec} = E(\bar{Y}^E_{-c}|T_c) = R^{\infty}_{Ec} T_c \psi^*, \tag{4}$$

$$TIV_{ENc} = E(\bar{Y}^N_{c}|T_c) = R^{\infty}_{ENc} T_c \zeta^*, \tag{5}$$

while for equation (3) they are:

$$TIV_{Nc} = E(\bar{Y}^N_{-c}|T_c) = R^{\infty}_{Nc} T_c \mu^*, \tag{6}$$

$$TIV_{NEc} = E(\bar{Y}^E_{c}|T_c) = R^{\infty}_{NEc} T_c \iota^*, \tag{7}$$

where μ^*,ψ^*, ι^* and μ^* are convolutions of parameters from the structural equations. The vectors $\bar{Y}^E_{-c} = \{\bar{y}^E_{-ic}\}_{i \in E}$, $\bar{Y}^N_{-c} = \{\bar{y}^N_{-ic}\}_{i \in N}$, $\bar{Y}^E_{c} = \{\bar{y}^E_{ic}\}_{i \in N}$ and $\bar{Y}^N_{c} = \{\bar{y}^N_{ic}\}_{i \in E}$ contain the endogenous

terms. R^{∞} are infinite sets of matrices multiplying the treatment vector, as defined in the proof of Proposition 1.

These vectors can be used as instruments for the endogenous terms. In fact, they are the expected values of our endogenous variables conditional on the treatment, and thus they are correlated with the endogenous variables but not with the error term since the treatment is administered at random. As shown in the proof of Proposition 1, they are linear combinations of powers of the shares of eligible agents in the groups, which can be approximated as follows:

$$TIV_{Ec} \approxeq Q_E^{\infty}\psi = \sum_{\substack{v=1 \\ r>0}}^{\infty}\sum_{\substack{r=0 \\ if s>0}}^{\infty}\sum_{s=0}^{\infty} \frac{(e_c-1)^v(e_cn_c)^s(n_c-1)^r}{(m_c-1)^{2s+r+v}} 1_{e_c,1}\psi_{vrs}, \tag{8}$$

$$TIV_{ENc} \approxeq Q_{EN}^{\infty}\zeta = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{q=0}^{\infty} \frac{n_ce_c(e_c-1)^q(e_cn_c)^s(n_c-1)^r}{(m_c-1)^{2s+2+r+q}} 1_{e_c,1}\zeta_{rsq}. \tag{9}$$

$$TIV_{Nc} \approxeq Q_N^{\infty}\mu = \sum_{v=1}^{\infty}\sum_{s=0}^{\infty}\sum_{q=0}^{\infty} \frac{(n_c-1)^v e_c(e_c-1)^q(e_cn_c)^s}{(m_c-1)^{2s+1+v+q}} 1_{n_c,1}\mu_{vsq}. \tag{10}$$

$$TIV_{NEc} \approxeq Q_{NE}^{\infty}\iota = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{q=0}^{\infty} \frac{e_c(e_c-1)^q(e_cn_c)^s(n_c-1)^r}{(m_c-1)^{2s+1+r+q}} 1_{n_c,1}\iota_{rsq}, \tag{11}$$

where Q^{∞} are vectors containing infinite combinations of products of powers of the share of eligibles and its complement. $1_{e_c,1}$ and $1_{n_c,1}$ are two $e_c \times 1$ and $n_c \times 1$ vectors of ones. If Proposition 1 holds, the IV vectors necessarily have full rank. To see this, let $TIV_S = (TIV'_{S1}, \ldots, TIV'_{S\bar{c}})'$ for $S = E, N, NE, EN$, $T = (T'_1, \ldots, T'_{\bar{c}})'$, $X_E^* = (X_{E1}^{*\prime}, \ldots, X_{E\bar{c}}^{*\prime})'$ with $X_{Ec}^* = \{x_{ic}^{E*}\}_{i\in E}$ and $x_{ic}^{E*} = (x_{ic}^E, \bar{x}_{-ic}^E, \bar{x}_{ic}^N)$, and $X_N^* = (X_{N1}^{*\prime}, \ldots, X_{N\bar{c}}^{*\prime})'$ with $X_{Nc}^* = \{x_{ic}^{N*}\}_{i\in N}$ and $x_{ic}^{N*} = (x_{ic}^N, \bar{x}_{-ic}^N, \bar{x}_{ic}^E)$. If e_c, n_c and m_c do not vary with c, $W_N = [TIV_N, TIV_{NE}]$ and $W_E = [TIV_E, TIV_{EN}, T]$ do not have full column rank, because TIV_E and TIV_{EN} are linear combinations of T. Given that the treatment is administered at random, $[W_N, X_N^*]$ and $[W_E, X_E^*]$ have full column rank if and only if the share of eligible agents varies across the groups. Intuitively, share variation introduces nonlinearities that allow identification and existence of valid instruments.

It appears from equations (8) - (11) that there are infinite ways to approximate TIV using subsets of Q^{∞}. Observe also that these IV vectors are infeasible since they involve the unknown parameters, μ, ψ, ι and ζ. Following the approach used in the literature on spatial econometrics (Kelejian and Prucha, 1998), we use a subset of the vectors in Q as empirical instruments. In

practice, we use the following empirical IVs (EIV),

$$Q_{Ec} = \left[\frac{e_c-1}{m_c-1}, \frac{(e_c-1)^2}{(m_c-1)^2}, \frac{n_c e_c(e_c-1)}{(m_c-1)^3}, \frac{(e_c-1)e_c n_c(n_c-1)}{(m_c-1)^4}\right] 1_{e_c,1}, \tag{12}$$

$$Q_{ENc} = \left[\frac{n_c e_c}{(m_c-1)^2}, \frac{(n_c e_c)^2}{(m_c-1)^4}, \frac{(n_c-1)e_c n_c}{(m_c-1)^3}, \frac{(n_c-1)^2 e_c n_c}{(m_c-1)^4}\right] 1_{e_c,1}, \tag{13}$$

$$Q_{Nc} = \left[\frac{e_c(n_c-1)}{(m_c-1)^2}, \frac{e_c(n_c-1)(e_c-1)}{(m_c-1)^3}, \frac{e_c(n_c-1)^2}{(m_c-1)^3}, \frac{e_c(n_c-1)^3}{(m_c-1)^4}\right] 1_{n_c,1}, \tag{14}$$

$$Q_{NEc} = \left[\frac{e_c}{m_c-1}, \frac{e_c(e_c-1)}{(m_c-1)^2}, \frac{e_c(e_c-1)^2}{(m_c-1)^3}, \frac{e_c^2(e_c-1)n_c}{(m_c-1)^4}\right] 1_{n_c,1}. \tag{15}$$

The specific vectors used in the approximation are not relevant in terms of identification, provided that Q_{Ec}, Q_{Nc}, Q_{NEc} and Q_{ENc} have full column rank. The behavior of the estimator with the instruments (12)-(15) in terms of efficiency is shown in the Monte Carlo experiments in Section 3.3. More details about the asymptotic behavior of this type of IV estimator can be found in Lee (2007).

Let $Q_S = (Q'_{S1}, \ldots, Q'_{S\bar{c}})'$ for $S = E, N, NE, EN$, $Y^E = (Y_1^{E\prime}, \ldots, Y_{\bar{c}}^{E\prime})'$, $\bar{Y}^E = (\bar{Y}_{-1}^{E\prime}, \ldots, \bar{Y}_{-\bar{c}}^{E\prime})'$, $\bar{Y}^N = (\bar{Y}_1^{N\prime}, \ldots, \bar{Y}_{\bar{c}}^{N\prime})'$ with $Y_c^E = \{y_{-ic}\}_{i\in E}$, $\bar{Y}_c^E = \{\bar{y}_{-ic}\}_{i\in E}$, $\bar{Y}_c^N = \{\bar{y}_{ic}\}_{i\in N}$ and $Z_E = (\bar{Y}^E, \bar{Y}^N, T, X_E^*)$. The IV estimator for model (2) is

$$\hat{\kappa}^E = (Z'_E P_E Z_E)^{-1} Z'_E P_E Y^E, \tag{16}$$

where $\kappa^E = (\phi^E, \phi^{EN}, \delta, \beta^{E\prime}, \gamma^{E\prime}, \gamma^{EN\prime})'$ and $P_E = \tilde{Q}_E(\tilde{Q}'_E\tilde{Q}_E)^{-1}\tilde{Q}'_E$,
where $\tilde{Q}_E = (Q_E, Q_{EN}, T, X_E^*)$. We can define in the same fashion the estimator for equation (3). Observe that while we split the equations in model (2)-(3) into two groups according to their treatment status, we could have split the model according to other types. Technically, Proposition 1 holds true for any two types of agents and any number of types (up to a finite number). The challenging task is to prove that the share of each specific type of agents varies randomly across groups.

Three features of this methodology are worth noting. First, the exclusion of the contextual effects of the treatment in our model specification (2)-(3) (that is $\frac{1}{m_c-1}\sum_{j\neq i} t_{jc} = \frac{e_c-1}{m_c-1}$) is not a crucial restriction for the identification of the model. The intuition is that the vectors of instruments are vast arrays of (nonlinear) transformations of the share of eligibles. The inclusion of

the contextual effects decreases the number of valid instruments, but the model remains (over-) identified. We formally prove that our model is identified when contextual treatment effects are included in the model specification in Proposition 2 in the Supplementary Appendix S. Second, identification is not stemming from the fact that the externalities are assumed to be heterogeneous. The model is still identified if the peer effects are homogeneous. The intuition is that the effects of the four endogenous variables are not identified because the same instruments are multiplied by different parameters, but rather from the fact that we have different instruments for each of them. As a result, even if the parameter is the same across the four endogenous variables, the model remains identified. The presence of heterogeneous externalities can be tested using a standard Wald test. To test for differences in peer effects in different equations, one would need to pool our model and jointly estimate both equations. Notice that if the hypothesis of homogeneity cannot be rejected and a constrained model is estimated, then inference should take into account the sequential structure of the testing procedure. We formally prove that our model is identified when externalities are constrained to be equal (i.e. $\phi^E = \phi^N = \phi^{NE} = \phi^{EN}$) in Proposition 3 in the Supplementary Appendix S. Third, observe that the intuition of our strategy is similar in spirit to Lee (2007) and subsequent empirical papers (see e.g., Boucher et al., 2014). However, while the model in Lee (2007) is designed for studying observational data, our model is designed for analyzing experimental data. The fact that we propose a model for policy evaluation of randomized experiments changes the modeling and estimation strategy in Lee (2007) along three main lines. First, in Lee (2007) identification requires variation in group size. We do not need it. Even if the groups have the same size, we can still identify our model parameters provided that the share of treated varies across groups. Second, in Lee (2007) the underlying identifying assumption is that the size of the group is not correlated with the error term (conditional on observed characteristics and unobserved group fixed effects). In our case, the randomization procedure guarantees that the share of treated peers is random, and thus by construction is orthogonal to the error term. Thus, it is not essential to include a full set of contextual effects (observed peer characteristics) and control for correlated effects (group fixed effects) since we do not need to deal with troubling unobserved factors. Third, Lee (2007) deals with one endogenous variable. Our model is able to deal with two (or more) of them because it is designed to estimate heterogeneous externalities. This is not a trivial extension because it implies different conditions for identification.

3.3 Monte Carlo Experiments

In this section, we use simulated data to investigate the performance of the proposed estimator in finite samples. We conduct a Monte Carlo simulation based on the following model

$$y_{ic}^{E} = \phi^{E}\bar{y}_{-ic}^{E} + \phi^{EN}\bar{y}_{ic}^{N} + \delta t_{ic} + \epsilon_{ic}^{E}, \tag{17}$$

$$y_{ic}^{N} = \phi^{N}\bar{y}_{-ic}^{N} + \phi^{NE}\bar{y}_{ic}^{E} + \epsilon_{ic}^{N}, \tag{18}$$

where $\epsilon_c^E, \epsilon_c^N \sim N(0, \sigma)$ with $\sigma = 1$. We generate $\bar{c}$ groups with size $m_c = k$; we then split each group c into two sub-populations of eligibles (E) and ineligibles (N). The share of eligibles is allocated at random across groups and according to a uniform distribution. Specifically, we draw e_c from an interval $[e_{min}, e_{max}]$, where e_{min} and e_{max} are two parameters such that $e_{min} > 0$ and $e_{max} < m_c$. The group treatment status t_c is generated by a Bernoulli distribution with probability $P(t_c = 1) = p = 0.7$. Following the design of PROGRESA, every eligible is treated if the village is drawn, i.e. $t_{ic} = t_c$. We estimate model (17)-(18) 3,000 times for each experiment. In equation (17) the set of empirical IVs is (Q_E, Q_{EN}), while in equation (18) this set is (Q_N, Q_{NE}). We report the mean point estimate, the relative standard error, and the root mean square error for the EIV, TIV, and OLS estimators.

Table 1 shows how the estimators are affected by the variation in the share of the eligible population. We set $m_c = 50$ and $\bar{c} = 60$, having 3,000 observations in each experiment, and gradually decrease the variability of the share of the eligible population, narrowing the interval $[e_{min}, e_{max}]$ from $[1, 49]$ (which covers about 96 percent of the possible shares in a population of 50 individuals) to $[24, 27]$ (which covers about 6 percent). In one experiment we also use the range of the share of eligibles in the PROGRESA program, that is $[6, 49]$ (84 percent coverage). Table 1 shows that the performance of the estimators improves as the variability increases (from the fourth to the first panel). Indeed, the root mean square error of ϕ^E increases more than tenfold for both the TIV and EIV when the share interval shrinks, while it increases by more than twofold for the OLS. It is worth noticing that the bias generated by limited variation in the share of eligible agents is more severe for the parameters of the eligible outcome equation. This is because the treatment is not excluded from the eligible outcome equation. Thus, all of the instruments rely only on the variation in the share of eligible agents to identify the effects of externalities.

In the Supplementary Appendix S, we study the performance of the estimators for different sample sizes and different values of the parameters. In *Panel (a)* of Table S.1, we study the performance of the estimators as the sample size increases from 500 to 3,000, holding constant the numerosity of each group at $k = 50$. The performance of the EIV is close to the TIV, and both the standard errors and mean squared errors significantly decrease as the number of observations increases. The IV point estimates are very close to the real parameter values even in small samples. The OLS is biased regardless of the sample size. This evidence shows that the chosen EIV (with the IV vectors in (12)-(15)) performs well, even if the number of terms used for the approximation is low. We also ran simulations in which we varied the combination and number of instruments used in the approximation of the theoretical IVs. The performance remains remarkably stable, thus showing no specific efficiency loss when using a limited set of IVs. In *Panel (b)* of Table S.1, we vary the values of the parameters. The IV point estimates always remain close to the real parameter values, while the OLS estimators are always biased. This evidence shows that the performance of the estimators does not depend on the specific set of parameter values.

4 Treatment Effects with Heterogeneous Externalities

Let us now highlight the importance of our framework for the analysis of treatment response with spillovers. In our model, the Stable Unit Treatment Value Assumption (SUTVA) (Rubin, 1986) does not hold because (i) spillovers are at work, and (ii) spillovers are heterogeneous. To the best of our knowledge, there are no studies that consider violations of the SUTVA because of (ii). In this section, we map our model to Rubin's potential outcomes model. The detailed derivation of the estimands defined in this section is contained in the Supplementary Appendix S. The aim is to show how our structural model enables us to not only estimate treatment effects, but also to obtain a policy relevant decomposition of these effects.

Average Treatment Effect Let $Y_{i,c}(1)$ and $Y_{i,c}(0)$ be the potential outcomes of an eligible unit i in group c when it receives the treatment and when it does not respectively. Given that the SUTVA does not hold, the potential outcomes of unit i depend on the treatment status of other units. Let $T^{i,c,1}$ and $T^{i,c,0}$ be two $e_c \times 1$ treatment vectors such that $T_i^{i,c,1} = 1$ and $T_i^{i,c,0} = 0$ and

Table 1: Monte Carlo simulations - variance of group size

	(1) TIV	(2) EIV	(3) OLS
$e_c \in [1, 49]$ on $m_c = 50$ (96% span)			
$\phi^E = 0.8$	0.800(0.014)[0.014]	0.803(0.014)[0.014]	0.832(0.011)[0.034]
$\phi^{EN} = 0.9$	0.898(0.047)[0.047]	0.910(0.045)[0.046]	1.016(0.029)[0.120]
$\delta = 1.7$	1.704(0.134)[0.134]	1.667(0.130)[0.134]	1.353(0.087)[0.358]
$\phi^N = 0.8$	0.799(0.025)[0.025]	0.804(0.024)[0.024]	0.848(0.020)[0.052]
$\phi^{NE} = 0.9$	0.901(0.015)[0.015]	0.899(0.015)[0.015]	0.884(0.014)[0.022]
$e_c \in [6, 49]$ on $m_c = 50$ (84% PROGRESA design)			
$\phi^E = 0.8$	0.799(0.021)[0.021]	0.806(0.020)[0.021]	0.832(0.018)[0.036]
$\phi^{EN} = 0.9$	0.900(0.011)[0.011]	0.898(0.011)[0.012]	0.892(0.011)[0.014]
$\delta = 1.7$	1.706(0.153)[0.153]	1.604(0.136)[0.167]	1.275(0.088)[0.434]
$\phi^N = 0.8$	0.799(0.015)[0.015]	0.808(0.013)[0.016]	0.836(0.010)[0.037]
$\phi^{NE} = 0.9$	0.898(0.052)[0.052]	0.931(0.046)[0.055]	1.040(0.028)[0.143]
$e_c \in [13, 37]$ on $m_c = 50$ (48% span)			
$\phi^E = 0.8$	0.797(0.023)[0.023]	0.808(0.018)[0.020]	0.822(0.014)[0.026]
$\phi^{EN} = 0.9$	0.874(0.151)[0.153]	0.976(0.101)[0.126]	1.113(0.034)[0.216]
$\delta = 1.7$	1.773(0.399)[0.406]	1.499(0.270)[0.336]	1.127(0.091)[0.580]
$\phi^N = 0.8$	0.800(0.029)[0.029]	0.804(0.028)[0.028]	0.834(0.026)[0.042]
$\phi^{NE} = 0.9$	0.900(0.018)[0.018]	0.898(0.017)[0.018]	0.886(0.016)[0.022]
$e_c \in [24, 27]$ on $m_c = 50$ (6% span)			
$\phi^E = 0.8$	0.814(0.204)[0.205]	0.814(0.204)[0.205]	0.813(0.086)[0.087]
$\phi^{EN} = 0.9$	1.168(0.582)[0.641]	1.168(0.582)[0.641]	1.143(0.092)[0.260]
$\delta = 1.7$	1.000(1.513)[1.668]	1.000(1.513)[1.668]	1.049(0.119)[0.662]
$\phi^N = 0.8$	0.772(0.186)[0.189]	0.812(0.175)[0.176]	1.083(0.099)[0.300]
$\phi^{NE} = 0.9$	0.918(0.125)[0.127]	0.892(0.118)[0.118]	0.714(0.067)[0.198]

Notes. Point estimate (standard error) [root mean squared error]. $\sigma = 1$, number of replications = 3000, number of observations = 3000, group size = 50, probability of being in a treated group = 0.7. In PROGRESA the share of elibgibles varies between 13 and 97 percent, which translates into 6 and 49 eligibles when the population is 50 people.

$T_j^{i,c,0} = T_j^{i,c,1} = 1, 0, \ \forall j \neq i$. The treatment effect of unit i is thus given by

$$E[Y_{i,c}(1) - Y_{i,c}(0)] = E[Y_{i,c}(1)] - E[Y_{i,c}(0)] = E[Y_{i,c}(T^{i,c,1})] - E[Y_{i,c}(T^{i,c,0})]. \tag{19}$$

Given (2)-(3) and from the derivations in the proof of Proposition 1, it follows that

$$E[Y_{i,c}(T^{i,c,1})] - E[Y_{i,c}(T^{i,c,0})] \approxeq \tag{20}$$
$$\delta \left[1 + \left(\frac{(e_c - 1)(\phi^E)^2}{(m_c - 1)^2} + \frac{n_c \phi^{EN} \phi^{NE}}{(m_c - 1)^2}\right) \sum_{v=0}^{\infty} \sum_{s=0}^{\infty} \sum_{r=0}^{\infty} \frac{[(e_c - 1)\phi^E]^v (e_c n_c \phi^{EN} \phi^{NE})^s [(n_c - 1)\phi^N]^r}{(m_c - 1)^{2s+v+r}}\right].$$

Let $E_C[.]$ denote the expectation operator across groups and units. The average treatment effect is

$$ATE = E_C[Y_{i,c}(1)] - E_C[Y_{i,c}(0)] = \delta(1 + E_C[g_{e_c,m_c,n_c}]), \tag{21}$$

where g_{e_c,m_c,n_c} is a function of $(e_c, m_c, n_c, \phi^E, \phi^N, \phi^{EN}, \phi^{NE})$ that is approximated in the second addend of (20). The ATE can be decomposed into two parts

$$ATE = \underbrace{\delta}_{\text{DTE}} + \underbrace{\delta E_C[g_{e_c,m_c,n_c}]}_{\text{FLTE}}. \tag{22}$$

The first part is the *Direct Treatment Effect* (hereafter DTE), while the second part is the effect of the treatment due to the interactions among agents, i.e. the effect of i's treatment that impacts i through other agents. We denote the latter effect as the *Feedback Loop Treatment Effect* (hereafter $FLTE$). This decomposition highlights that the program impact for unit i can be large if she/he is in a group with a high level of social interaction (i.e. if ϕ^E and $\frac{e_c}{m_c}$ are high), even if the direct treatment effect (δ) is low. Observe that, while $\phi^E, \phi^N, \phi^{EN}, \phi^{NE}$ are constant parameters, e_c, m_c, n_c vary across groups. Even if $\delta_i = \delta$ (i.e. the direct treatment effect is homogeneous), the ATE can be heterogeneous because of the different group composition.

The $FTLE$ can be further decomposed into three parts: (i) $w_{e_c,m_c,\phi^E} = \frac{(e_c-1)(\phi^E)^2}{(m_c-1)^2} \sum_{r=0}^{\infty} (\phi^E \frac{e_c-1}{m_c-1})^r$, the pure loop effect within the treated (WTE); (ii) $b_{n_c,e_c,m_c,\phi^{NE},\phi^{EN}} = \frac{n_c \phi^{EN} \phi^{NE}}{(m_c-1)^2} \sum_{r=0}^{\infty} (\phi^{EN} \phi^{NE} \frac{e_c n_c}{(m_c-1)^2})^r$, the loop effect between the treated and the untreated (BTE); and (iii) $j_{e_c,n_c,m_c,\phi^E,\phi^N,\phi^{NE},\phi^{EN}} = g_{e_c,m_c,n_c} - (w_{e_c,m_c,\phi^E} + b_{n_c,e_c,m_c,\phi^{NE},\phi^{EN}})$, the residual effect ($RTE$). The first component is generated by loops within same-type agents and reflects both the strength of interactions among them and their relative share in the population. The second component is proportional to the intensity of between-group interactions weighted by the relative shares of the groups. The third term is a residual component. The ATE can thus be decomposed as

$$ATE = \underbrace{\delta}_{\text{DTE}} + \underbrace{\delta E_C[w_{e_c,m_c,\phi^E}]}_{\text{WTE}} + \underbrace{\delta E_C[b_{n_c,e_c,m_c,\phi^{NE},\phi^{EN}}]}_{\text{BTE}} + \underbrace{\delta E_C[j_{e_c,n_c,m_c,\phi^E,\phi^N,\phi^{NE},\phi^{EN}}]}_{\text{RTE}}. \tag{23}$$

Indirect Treatment Effect Let $Y_{i,c}(1)$ and $Y_{i,c}(0)$ be the potential outcomes of an ineligible unit i in group c when the treatment is administered to all eligibles in her group and administered

to no eligibles if she is not in the treated group. According to (2)-(3) and the derivations in the proof of Proposition 1, the treatment effect of unit i is given by

$$\begin{aligned} & E[Y_{i,c}(1) - Y_{i,c}(0)] = E[Y_{i,c}(1)] - E[Y_{i,c}(0)] \\ \approx \; & \delta\left(1 + \frac{(n_c - 1)\phi^N}{(m_c - 1)}\right) \sum_{v=0}^{\infty}\sum_{s=0}^{\infty}\sum_{q=0}^{\infty} \frac{e_c(e_c n_c \phi^{EN}\phi^{NE})^s[(e_c - 1)\phi^E]^q[(n_c - 1)\phi^N]^v}{(m_c - 1)^{2s+q+v+1}}. \end{aligned} \tag{24}$$

The indirect treatment effect is

$$ITE = E_C[Y_{i,c}(1)] - E_C[Y_{i,c}(0)] = \delta E_C[h_{e_c,m_c,n_c}], \tag{25}$$

where h_{e_c,m_c,n_c} is a function of $(e_c, m_c, n_c, \phi^E, \phi^N, \phi^{EN}, \phi^{NE})$ that is approximated in (24). This implies that, when interactions are at work, the ITE can be heterogeneous because of the different group compositions. For example, under the assumption that the number of treated units is constant, when ineligible agents have greater within-group social interactions than eligible agents ($\phi^N > \phi^E$), and ineligible agents have greater within-group than between-group interactions ($\phi^N > \phi^{NE}$,$\phi^N > \phi^{EN}$) the ITE is increasing in n_c/m_c.

As for the ATE, an analog further decomposition can be made for the ITE

$$ITE = \underbrace{\delta\phi^{NE}\frac{e_c}{m_c - 1}}_{\text{DSE}} + \underbrace{\delta E_C[h_{e_c,m_c,n_c}]}_{\text{ISE}} \tag{26}$$

where the DSE represents the direct spillover from treated to untreated and the ISE accounts for the indirect spillover effect generated by the DSE. The ISE can also be further decomposed into three parts: (i) the pure loop effect within the untreated population (WUE), (ii) the loop effect between untreated and treated populations (BUE), and (iii) the residual loop effect (RUE). The ITE can thus be decomposed as

$$ITE = \underbrace{\delta\phi^{NE}\frac{e_c}{m_c - 1}}_{\text{DSE}} + \underbrace{\delta E_C[r_{n_c,m_c,\phi^N}]}_{\text{WUE}} + \underbrace{\delta E_C[q_{n_c,e_c,m_c,\phi^{NE},\phi^{EN}}]}_{\text{BUE}} + \underbrace{\delta E_C[z_{e_c,n_c,m_c,\phi^E,\phi^N,\phi^{NE},\phi^{EN}}]}_{\text{RUE}}, \tag{27}$$

where $r_{n_c,m_c,\phi^N} = \delta\phi^{NE}\frac{e_c}{m_c-1}\sum_{r=1}^{\infty}(\phi_N\frac{n_c-1}{m_c-1})^r$, $q_{n_c,e_c,m_c,\phi^{NE},\phi^{EN}} = \delta\phi^{NE}\frac{e_c}{m_c-1}b_{n_c,e_c,m_c,\phi^{NE},\phi^{EN}}$, and $z_{e_c,n_c,m_c,\phi^E,\phi^N,\phi^{NE},\phi^{EN}} = h_{e_c,m_c,n_c} - (r_{n_c,m_c,\phi^N} + q_{n_c,e_c,m_c,\phi^{NE},\phi^{EN}})$.

5 Empirical Application

5.1 The PROGRESA program

PROGRESA was a partial-population experiment which was implemented in Mexico in 1998. The program offered conditional cash transfers to the rural poor in exchange for sending their children to school and for regular visits at health clinics and health talks. The education component of PROGRESA consisted of providing subsidies, ranging from 70 to 255 pesos per month (depending on the child's gender and grade level), to children attending school in grades 3 to 9. PROGRESA was targeted in two stages: (1) the selection of villages where PROGRESA operated; (2) the selection of eligible households within the selected villages. Stage (1) was implemented by identifying communities with high scores on a "village marginality index" based on the socioeconomic information collected in a pre-program survey in 1997. A total of 506 communities across seven states were selected. Among these communities, 320 were randomly assigned into a treatment group, with the remaining 186 communities serving as a control group. Stage (2) was implemented by selecting poor households on the basis of a "household poverty index" constructed from the baseline survey in 1997. A cutoff value of the poverty index defined poor families eligible for PROGRESA. While household eligibility was determined within all (treatment and comparison group) communities, only households classified as eligible and within the 320 villages assigned to treatment became program beneficiaries. Our focus is on the effects of PROGRESA on school enrollment decisions.

5.2 Empirical model

The reduced form effect of the PROGRESA status of the village on school enrollment of children whose families are not part of the program using a peer effect model has been estimated by Bobonis and Finan (2009) and Lalive and Cattaneo (2009). Lalive and Cattaneo (2009) further improves upon Bobonis and Finan (2009) in four important respects. First, they refine the peer group by considering as peers all children in the same village and grade. Second, they consider changes in school attendance rather than levels, thus controlling for time-invariant unobserved heterogeneity in children. Third, they refine the IV strategy by using the share of treated households in a village rather than a dummy indicating the PROGRESA status of a village. Fourth, they decompose the average treatment effect into a direct effect that arises due to financial incentives (DTE in

our equation (22)) and a residual effect due to the presence of social interactions ($FLTE$ in our equation (22)).

We replicate the analysis performed by Lalive and Cattaneo (2009) when including heterogeneous externalities. In the parlance of our model (2)-(3), Lalive and Cattaneo (2009) use the following specification:

$$\Delta y_{igv}^{N} = \phi \Delta \bar{y}_{-igv} + x_{igv}^{N} \beta + \epsilon_{igv}^{N}. \tag{28}$$

The dependent variable is the change in school enrollment between October 1997 and October 1998. School enrollment is a binary indicator taking value 1 if the child attends school at the date of the interview and 0 otherwise. The peer group (c) of child i is defined as all children in the same grade g and village v, excluding i. They identify ϕ using $\frac{e_{gv}}{m_{gv}} T_{gv}$ as an instrument for $\Delta \bar{y}_{-igv}$. This empirical framework allows for the identification of the total effect of PROGRESA on the schooling attendance of eligible and ineligible children under the assumption that social interactions are equally important among poor and nonpoor households. Our framework relaxes this assumption, and further allows social interactions to be different both within and between groups. Specifically, our empirical model (2)-(3) takes the form:

$$\Delta y_{igv}^{E} = \phi^{E} \Delta \bar{y}_{-igv}^{E} + \phi^{EN} \Delta \bar{y}_{igv}^{N} + \delta t_{v} + x_{igv}^{E} \beta^{E} + \epsilon_{igv}^{E}, \tag{29}$$

$$\Delta y_{igv}^{N} = \phi^{N} \Delta \bar{y}_{-igv}^{N} + \phi^{NE} \Delta \bar{y}_{ic}^{E} + x_{igv}^{N} \beta^{N} + \epsilon_{igv}^{N}. \tag{30}$$

5.3 Data description

The data from the PROGRESA program consists of repeated observations (panel data) for 24,000 households from 506 villages (320 in the treatment group and 186 in the control group) across seven states over five rounds of surveys (baseline surveys in October 1997 and October 1998 and follow-ups in May 1999, June 1999, and November 1999). We thus have information on enrollment during three consecutive academic years (1997-1998, 1998-1999, and 1999-2000). Following the sample selection used by Lalive and Cattaneo (2009), we focus on information in the first two years and select children living with their mother who had completed grades 3 to 6 of primary school in October 1997, and for whom there is complete information on school attendance in 1997-1998

and 1998-1999. In order to allow for the presence of heterogeneous externalities, we then restrict our sample to children in peer groups (village-grade groups) with at least 2 poor households and 2 nonpoor households. We end up with a sample of 8,682 children, of whom 5,076 live in treated villages and 3,606 live in untreated villages. Table S.2 in the Supplementary Appendix S provides the descriptive statistics for the characteristics of the children in our sample. *Panel (a)* shows the raw evidence on the effects of the cash subsidy on school enrollment. We report school enrollment for children in eligible and ineligible households before and after the treatment, distinguishing between treated and untreated villages. In October 1997, school attendance of children in poor households is roughly the same in control and treated villages (roughly 78 percent vs 79 percent, respectively, with an insignificant difference). One year later, school attendance is 6 percentage points lower for eligible children in control villages (roughly 72 percent), whereas in treated village it remains roughly unchanged. This means that the program increased school enrollment by 6 percentage points. The data also indicate that the program had an effect on children in untreated households in treated villages. On average, about 77 percent of ineligible children in grades 3 to 6 attend school in control villages in October 1997. In treated villages, school attendance for ineligible children is about 80 percent, but the difference between the treated and untreated villages is not statistically significant. By October 1998, school attendance for ineligible children dropped by 5 points in control villages but only by roughly 3 points in treated villages. However, this difference is not statistically significant. This suggests that the spillover effect is at the most weak. *Panel (b)* in Table S.2 compares characteristics of eligible and ineligible children between treatment and control villages at baseline. Consistent with the random assignment of villages in the PROGRESA program, there are no statistically significant differences in the observed characteristics of children on all dimensions. This evidence resembles the descriptive analysis in Lalive and Cattaneo (2009), Table 2. We show the comparison in the Supplementary Appendix S. Table S.3 replicates the content of Table S.2 when using the sample in Lalive and Cattaneo (2009). The means of all variables are remarkably close. A formal comparison of our sample with the one used by Lalive and Cattaneo (2009) is in Table S.4, showing no relevant difference.

Table S.5 reports the effect of PROGRESA on eligible and ineligible children's change in school enrollment using a regression analysis that allows us to control for the observable characteristics of children (within-individual difference-in-difference analysis). The results confirm the descriptive evidence, showing that PROGRESA decreases the downward trend in school attendance for the

treated children by 5.5 percentage points. They also show an estimated increase (i.e. decrease of downward trend) of 2.1 percentage points for untreated children in treated villages. The estimate is now statistically significant, indicating the existence of spillover effects.

5.4 Estimation results

We begin our analysis by providing evidence in support of our identification strategy. We show in Section 3 that heterogeneous treatment externalities can be identified if there is variation in the share of eligible households. Figure S.1 in the Supplementary Appendix S depicts the distribution of peer groups by the share of eligible households. The graph reveals the presence of significant variation in group size. For each endogenous variable, the proposed IVs are combinations of four vectors, which are (nonlinear) functions of eligible group size. Figure S.2 shows that those vectors are not collinear. In fact, for both equations (29) and (30) one can see a relationship between vectors that is markedly nonlinear. This guarantees that the instrumental matrix has full rank in our empirical application. The share of eligible females in the same group is also used in the instrumental variable set for the eligible outcome equation to increase the power of the instruments. In PROGRESA, the grants awarded to females are higher than the ones awarded to males. Thus, attendance rates are plausibly correlated with the share of females in the group. Details on the construction of the empirical IVs can be found in the Supplementary Appendix S.

Table 2 displays our main results. It reports IV estimates which are obtained using a subset of the vectors (12)-(15) as instrumental variables (EIV). *Panel (a)* collects the parameter estimates of model (29), whereas *Panel (b)* collects the parameter estimates for model (30). We report the results for the entire set of control variables in Table S.6. In the Supplementary Appendix S, we show that the evidence on the existence of heterogeneous effects in the estimation of the externality persists if we use a different subset of instruments (Table S.7), when contextual treatment effects are added to the model specification (Table S.8), when other contextual effects are included (Table S.9), and when the contextual effects are allowed to be different for eligible and ineligible individuals (Table S.10).

The central result is that the IV estimates in *Panel (b)* reveal that the indirect effect on ineligible households is not entirely due to spillovers from eligible households. On the contrary, this spillover effect is small when compared to the externality produced within ineligible households. A 10 percentage point increase in school enrollment of eligible students is associated with a 4 percentage

point increase in ineligible students' school attendance, whereas this effect more than doubles (8.8 percentage points) for a 10 percentage point increase in school attendance of ineligible students. The difference between the two parameters is statistically different from zero at the 10 percent level. A standard Wald test is used. We performed weak instrument F-tests in models with multiple endogenous variables as described in Sanderson and Windmeijer (2016). The complete set of first stage results, together with the reduced form estimates, are reported in Tables S.11 and S.12 of the Supplementary Appendix S. The first stage F-statistics show that the instruments are informative, although not extremely strong for model (29). This is perhaps not surprising given that the treatment is not excluded from model (29). For this reason, we also report the Limited Information Maximum Likelihood (LIML) results in the second column of each panel. The results remain qualitatively unchanged, and the point estimates are similar. The LIML estimates thus lend credibility to the IV estimates by eliminating a suspected weak instrument bias (see Flores-Lagunes, 2007 or Angrist and Pischke, 2008, pages 209-216, for further details). While there may be several mechanisms underlying differences in peer effects across groups, our data does not allow us to pin down the exact channel.

Table 2: Externalities of PROGRESA for eligible and ineligible children

	Panel (a)		*Panel (b)*	
Dependent variable: change in attendance for	Eligible children		Ineligible children	
	IV	LIML	IV	LIML
	(1)	(2)	(3)	(4)
Change in peer group attendance for				
Ineligible children	0.8874**	1.0761**	0.8750***	0.8536***
	(0.4117)	(0.4404)	(0.3001)	(0.3067)
Eligible children	0.7343*	0.7290*	0.3966*	0.3781*
	(0.4077)	(0.4286)	(0.2376)	(0.2242)
Individual and family characteristics	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes
F-tests				
Peer group attendance of ineligible children	6.23		8.25	
Peer group attendance of eligible children	6.61		9.37	
Observations	5,387	5,387	3,295	3,295

Notes. Standard errors are clustered at the village level. * : $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$. The controls are listed in Table S.6. The Sanderson-Windmeijer (SW) multivariate first-stage F statistics are reported for both the endogenous terms. The instrumental variables used in *Panel (a)* are q_{Ef1}, q_{Ef3}, q_{EN3}, and q_{EN4}, in *Panel (b)* are q_{NE3}, q_{N1}, and q_{N4}. The subscript f denotes the IVs based on the share of eligible girls. The construction of the instrumental variables is detailed in Appendix S1.

5.5 Estimands decomposition

When heterogeneous interactions among units are allowed, the total effects of a program depend on the estimates of the endogenous social interactions and on the share of the eligible population. In *Panel (a)* of Table 3 we show the total effects of PROGRESA when externalities are allowed to be heterogeneous. The estimands are derived in Section 4. The ATE is equal to 0.0546, while the ITE is 0.0213 (see Table S.5). More than half of the ATE is due to externalities (that is to effects coming from other eligible and ineligible households directly or indirectly affected by the treatment). The main novelty of our empirical framework is that it allows us to investigate the nature of such externalities. It appears that, for eligible students, the externalities are mostly within eligible households (WTE is equal to 30 per cent and BTE is equal to 4 percent). For ineligible students, the externalities generated by the program within this group (WUE) and between eligibles and ineligibles (BUE) are roughly the same (11 percent), although the estimate of the social interaction parameter within ineligibles is much larger than the one capturing interactions between groups (Table 2, *Panel (b)*). This is due to the fact that the average share of eligible children in PROGRESA is very high (roughly 60 percent). In the majority of other conditional cash transfer programs, the share of the eligible population is much lower. For example, the target population in the World Bank programs in Latin America varies from less than 10 percent in Bolivia, Costa Rica, Paraguay, and Peru, to about 25 percent in Brazil, Colombia, Guatemala, and Mexico (Grosh et al., 2014). In the partial population program in Bangladesh that transfers livestock assets and skills to the poorest women described in Bandiera et al. (2017), the share of eligible women is about 6 percent. The decomposition of the ITE in *Panel (b)* of Table 3 shows that, if the average share of eligible children in PROGRESA were 20 percent, the treatment externalities generated within nonpoor households (WUE) would be more than double the spillover from poor to nonpoor households (DSE).

Table 3: Estimands decomposition

	Panel (a): average share of eligible children = 60%								
ATE	0.0546				ITE	0.0213			
	DTE	0.0232	42%			DSE	0.0058	27%	
	FLTE	0.0314	58%			ISE	0.0155	73%	
		WTE	0.0165	30%			WUE	0.002	11%
		BTE	0.0025	4%			BUE	0.002	12%
	Panel (b): average share of eligible children = 20%								
ATE	0.0274				ITE	0.0090			
	DTE	0.0232	42%			DSE	0.0020	22%	
	FLTE	0.0042	8%			ISE	0.0070	78%	
		WTE	0.0017	3%			WUE	0.0042	46%
		BTE	0.0017	3%			BUE	0.0017	19%

Notes. The estimates of the social interactions parameters are in Table 2. ATE and ITE are reported in Table S.5. DTE is reported in Table S.6.

6 Discussion

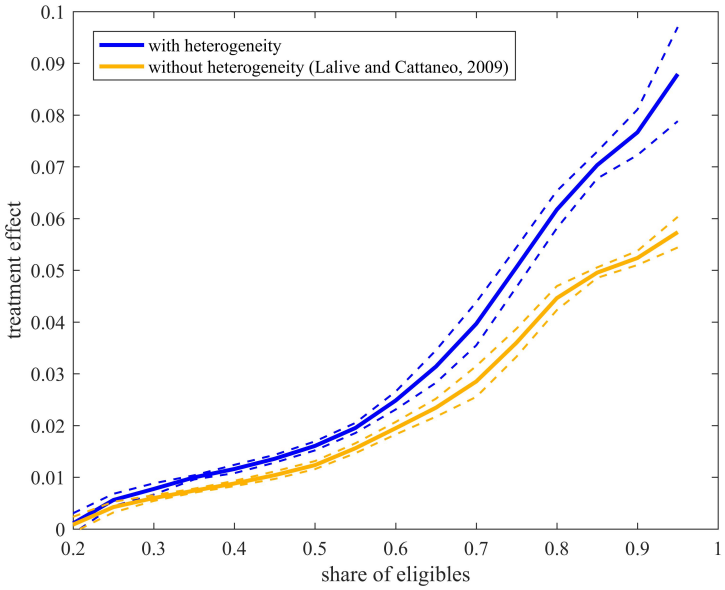
In this section, we discuss the implications of estimating heterogeneous externalities for the design of experiments and highlight the relevance of our methodology by identifying other contexts where our methods can be applied.

We begin by illustrating the importance of introducing heterogeneity in the externalities in terms of aggregate outcomes. In model (2)-(3) we have added heterogeneity to the linear-in-means model proposed by Lalive and Cattaneo (2009). To demonstrate the importance of this addition, we perform a numerical simulation and show how the estimated treatment response is biased when heterogeneity in the spillover effects is ignored.

We use the setting of the Monte Carlo simulation described in Section 5. Following the PROGRESA design, the share of eligibles is randomly chosen from a uniform $[0.13, 0.97]$, as in one experiment reported in Table 1. We then set the model parameters equal to our IV estimates in Table 2 (complete set of estimates in Table S.6), and generate outcomes. The blue line in Figure 1 represents the average outcome at the village level by the village's share of treated population, with 95 percent confidence intervals. Next, using the simulated data, we estimate the model without heterogeneity. In doing so, we implicitly set $\phi_N = \phi_{NE} = \phi_{EN} = \phi_E = \phi$. We obtain an estimate of ϕ equal to 0.557, which resembles the estimate presented by Lalive and Cattaneo (2009), $\gamma_N = 0.542$. Lalive and Cattaneo (2009) estimate this value using equation (30) only and aggregating peers' outcome at the village level, that is considering the treatment response of the

ineligibles in treated villages. We follow their approach here to estimate ϕ. They also present an estimate of ϕ from a pooled model, that is when including the eligibles. This estimate, however, is very close (equal to 0.505). The predicted average percentage increase in the outcome at the village level by the village's share of eligible population is plotted on the yellow line in Figure 1, with 95 percent confidence intervals. In Figure 1, the difference between the blue line and the yellow line represents the difference in the estimated treatment effects at the village level with and without heterogeneity of the peer effects. Perhaps unsurprisingly, it appears that the estimated treatment effects without heterogeneous externalities are biased downward, and the bias increases with the share of treated population in the village. It can be as large as one third of the real value. Intuitively, this is because the large within-treated externalities are attenuated when all of the spillovers are constrained to take a common value. In Figure 1 we plot the bias in the average treatment effect regardless of the eligibility status. The treatment effects conditional on eligibility, that is the ATE and ITE defined in Section 4, show a similar pattern. Being able to differentiate the magnitude of the spillovers between treated and untreated units is thus crucial for deciding how many treated units are needed to reach a certain goal. For example, Figure 1 reveals that if the policy maker aims to increase the aggregate outcome by 5 percent, then our model (blue line) would suggest assigning to treatment about 70 percent of the population in each village to treatment. The predictions of the model without heterogeneous externalities (yellow line) would instead suggest treating about 85 percent of the population in each village, a possible waste of resources.

Figure 1: Average outcome at the village level by village's share of eligible households



Notes. PROGRESA experiment design: $\epsilon_c^E, \epsilon_c^N \sim N(0, \sigma)$, $\sigma = 1$, $t_c \sim bernoulli(p)$, $p = 0.64$: the probability of beign treated for a village in PROGRESA, $c = 100$ villages with size $m_c = k = 60$. The shares of eligibles is randmoly chosen from a uniform $U[0.13, 0.97]$ to replicate the support of PROGRESA, every eligible is treated. The blue line is generated using the model with heterogeneous peer effects, when setting the coefficients equal to our IV estimates in Table 2 (and Table S.6). The yellow line depicts the estimated values with a model without heterogeneity in the peer effects.

Table 4: Available data collected using Two-stage experimental Protocol

Policy Program	Treatment	Eligibles	Peers	Experimental Protocol
Zomba Cash Transfer Program				
YEAR: 2008-2009 COUNTRY: Malawi Baird et al. (2011, 2012)	Conditional and unconditional cash transfers	Never-married females, aged 13-22	Females in the same enumeration areas	Two-stage Randomization -First level: geographical area (enumeration areas) -Second level: individuals
BRAC's Targeting the Ultra-Poor (TUP) program				
YEAR: 2007-2014 COUNTRY: Bangladesh Bandiera et al. (2016)	Asset bundles: livestock and those relevant for tree nurseries and vegetable -Training and support package	Females in ultra-poor households	Other females in the same household	Partial Population -First level: geographical area (subdistrics) -Second level: households
One Laptop per Child program				
YEAR: 2011 COUNTRY: Peru Beuermannnet al.(2015)	XO laptops	Children between the first and sixth grade	Classmates	Two-stage Randomization -First level: school -Second level: individuals
Tax Deferred Account (TDA) retirement plan				
YEAR: 2000 COUNTRY:US Duflo and Saez (2003)	Invitation to Fair on TDA	Non-faculty employees at the university	Employees in the same department	Two-stage Randomization -First level: departments -Second level: individuals
The Bono de DesarrolloHumano Program				
YEAR: 2004-2006 COUNTRY: Ecuador Paxson and Schady (2010)	Unconditional cash transfer	Children aged 0-6 without older siblings in poor households	Children in the same parish	Partial Population -First level: geographical area (parishes) -Second level: households
The Dutch Postcode Lottery (PCL)				
YEAR: 2005 COUNTRY: Netherlands Kuhn et al. (2011)	Cash prize from the lottery	Owners of a lottery ticket	Postcode neighbors	Partial Population -First level: postcodes -Second level: individuals
The Voter Education Campaign				
YEAR: 2008 COUNTRY: Pakistan Ginï¿½ and Mansuri (2018)	Information to women on the balloting process and the importance of voting, through a door-to-door visit	Adult women	Women in the same geographical clusters Women registered in the same Polling station	Two-stage Randomization -First level: geographical clusters (consisted of segments of one or two contiguous streets) -Second level: households

Notes. Partial population experiment is a setting in which clusters are randomly assigned to treatment or control, and a subset of individuals are offered treatment within clusters assigned to treatment (Moffit, 2001). The individual treatment assignment within clusters is typically not random. In two-stage randomization protocols also the second stage is randomized (Hudgens and Halloran, 2008).

7 Concluding remarks

In this paper, we develop an empirical framework to identify and estimate the effects of treatment when the potential outcome of a unit depends on the eligibility status of other units. Although there is a clear conceptual innovation in the way the model parameters are identified, the proposed framework is intentionally built on a combination of existing tools that make the implementation of the methodology easy for the applied practitioner. The empirical relevance of our methodology is illustrated in the evaluation of the effects of the program PROGRESA on schooling enrollment. We show that the average indirect effect that is estimated in the existing literature hides information that is crucial for understanding the mechanisms underlying the policy impacts. We find that this aggregate statistic is only partially due to spillovers from eligible to ineligible households, as commonly assumed. Strong interactions within the ineligible population generate a large social multiplier that is able to generate the indirect treatment effect that is estimated in previous studies.

Appendix

In order to list our set of assumptions we need to introduce some matrix notation and the reduced form of the model in matrix form. Let $Y_c^E = \{y_{ic}\}_{i\in E}$, $Y_c^N = \{y_{ic}\}_{i\in N}$, $X_c^E = \{x_{-ic}\}_{i\in E}$, and $X_c^N = \{x_{ic}\}_{i\in N}$. Let $1_{a,b}$ be a matrix of ones of dimension $a\times b$. The adjacency group based matrices can be expressed as $\tilde{G}_c^E = \{g_{cij}^E\} = \frac{1}{e_c-1}(1_{e_c,e_c} - I_{e_c})$, $\tilde{G}_c^{EN} = \{g_{cij}^{EN}\} = \frac{1}{n_c}1_{e_c,n_c}$, $\tilde{G}_c^N = \{g_{cij}^N\} = \frac{1}{n_c-1}(1_{n_c,n_c} - I_{n_c})$, $\tilde{G}_c^{NE} = \{g_{cij}^{NE}\} = \frac{1}{e_c}1_{n_c,e_c}$. Let us define the share matrices as $S_c^E = \frac{e_c-1}{m_c-1}I_{e_c}$, $S_c^{EN} = \frac{n_c}{m_c-1}I_{n_c}$, $S_c^N = \frac{n_c-1}{m_c-1}I_{n_c}$, $S_c^{NE} = \frac{e_c}{m_c-1}I_{e_c}$.

Using these matrices, model (2)-(3) can be written in matrix form as

$$Y_c^E = \phi^E\tilde{G}_c^E S_c^E Y_c^E + \phi^{EN}\tilde{G}_c^{EN} S_c^{EN} Y_c^N + X_c^E\beta^E + \delta T_c + \tilde{G}_c^E S_c^E X_c^E\gamma^E + \tilde{G}_c^{EN} S_c^{EN} X_c^N\gamma^{EN} + \epsilon_c^E, \quad (31)$$

$$Y_c^N = \phi^N\tilde{G}_c^N S_c^N Y_c^N + \phi^{NE}\tilde{G}_c^{NE} S_c^{NE} Y_c^E + X_c^N\beta^N + \tilde{G}_c^N S_c^N X_c^N\gamma^N + \tilde{G}_c^{NE} S_c^{NE} X_c^E\gamma^{NE} + \epsilon_c^N. \quad (32)$$

Let us define the following vectors

$$A_c\theta^E = X_c^E\beta^E + G_c^E X_c^E\gamma^E + G_c^{EN} X_c^N\gamma^{EN} + \epsilon_c^E,$$

$$B_c\theta^N = +X_c^N\beta^N + G_c^N X_c^N\gamma^E + G_c^{NE} X_c^E\gamma^{NE} + \epsilon_c^N,$$

and set $G = \tilde{G}S$. Let us suppress the c index to ease the notation. The reduced form of model (31)-(32) is thus

$$Y^E = M^{E(-1)}(\phi^{EN}G^{EN}J^N\theta^N B + \theta^E A + \delta T), \quad (33)$$

$$Y^N = M^{N(-1)}(\phi^{NE}G^{NE}J^E(\theta^E A + \delta T) + \theta^N B), \quad (34)$$

where $M^E = (I_E - \phi^E G^E - \phi^{EN}\phi^{NE}C^E)$, $C^E = G^{EN}J^N G^{NE}$, $J^N = (I_N - \phi^N G^N)^{-1}$, $M^N = (I_N - \phi^N G^N - \phi^{EN}\phi^{NE}C^N)$, $C^N = G^{NE}J^E G^{EN}$, $J^E = (I_E - \phi^E G^E)^{-1}$, I_E and I_N are identity matrices of dimensions e and n, respectively.

In what follows we list the set of assumptions we need to prove Proposition 1.

Assumptions

1. Assignment mechanism: partial population design.

2. The M^E and M^N matrices are nonsingular.

Assumption 1 is needed in order to have that our instrument vectors, i.e. vectors of nonlinear functions of the share of eligibles, vary randomly across groups. Regarding Assumption 2, our model represents an equilibrium equation so M matrices are assumed to be invertible. In practice, this condition allows us to derive the reduced form of the model. Sufficient conditions for the nonsingularity of M^E and M^N are $|\phi^E| + |\phi^{NE}\phi^{EN}|||J^N||_\infty < 1$ and $|\phi^N| + |\phi^{NE}\phi^{EN}|||J^E||_\infty < 1$ where $||\cdot||_\infty$ is the row-sum matrix norm. To see this, let us consider a sufficient condition for nonsingularity (see, e.g. Liu, 2014) $|\phi^E|||G^E||_\infty + |\phi^{NE}\phi^{EN}|||G^{EN}J^NG^{NE}||_\infty < 1$. Using the fact that sociomatrices are row-normalized and applying the Schwarz matrix inequality, we obtain that

$$\begin{aligned} |\phi^E|||G^E||_\infty + |\phi^{NE}\phi^{EN}|||G^{EN}J^NG^{NE}||_\infty &\leq \\ |\phi^E| + |\phi^{NE}\phi^{EN}|||G^{EN}||_\infty||J^N||_\infty||G^{NE}||_\infty = |\phi^E| + |\phi^{NE}\phi^{EN}|||J^N||_\infty. \end{aligned}$$

Proof of Proposition 1. From the reduced form of the model (33) and (34), we have

$$G^{NE}Y^E = G^{NE}(M^{E(-1)}(\phi^{EN}G^{EN}J^N\theta^NB + \theta^EA + \delta T)),$$

$$G^NY^N = G^N(M^{N(-1)}(\phi^{NE}G^{NE}J^E(\theta^EA + \delta T) + \theta^NB)).$$

if we use a series expansion we can write

$$G^{NE}Y^E = G^{NE}\left(\sum_{j=0}^{\infty}(\phi^EG^E + \phi^{EN}\phi^{NE}C^E)^j(\phi^{EN}G^{EN}\sum_{j=0}^{\infty}(\phi^NG^N)^j\theta^NB + \theta^EA + \delta T)\right), \quad (35)$$

$$G^NY^N = G^N\left(\sum_{j=0}^{\infty}(\phi^NG^N + \phi^{NE}\phi^{EN}C^N)^j(\phi^{NE}G^{NE}\sum_{j=0}^{\infty}(\phi^EG^E)^j(\theta^EA + \delta T) + \theta^NB)\right). \quad (36)$$

Using the binomial theorem, we can express equations (35)-(36) as

$$\begin{aligned} G^{NE}Y^{E} &= G^{NE}\left(\sum_{j=0}^{\infty}\sum_{p=0}^{j}\left(\binom{j}{p}(\phi^{E}G^{E})^{j-p}(\phi^{EN}\phi^{NE}C^{E})^{p}\right)\right. \\ &\times \left.\left(\phi^{EN}G^{EN}\sum_{j=0}^{\infty}(\phi^{N}G^{N})^{j}\theta^{N}B+\theta^{E}A+\delta T\right)\right), \end{aligned} \tag{37}$$

$$\begin{aligned} G^{N}Y^{N} &= G^{N}\left(\sum_{j=0}^{\infty}\sum_{p=0}^{j}\left(\binom{j}{p}(\phi^{N}G^{N})^{j-p}(\phi^{NE}\phi^{EN}C^{N})^{p}\right)\right. \\ &\times \left.\left(\phi^{NE}G^{NE}\sum_{j=0}^{\infty}(\phi^{E}G^{E})^{j}(\theta^{E}A+\delta T)+\theta^{N}B\right)\right). \end{aligned} \tag{38}$$

$E(G^{NE}Y^{E}|T)$ and $E(G^{N}Y^{N}|T)$ are valid instruments for $G^{NE}Y^{E}$ and $G^{N}Y^{N}$ since they are correlated with the endogenous terms but not with the error terms. Given (37)-(38), these two vectors can be represented as products of Gs and Ss times the treatment vector

$$E(G^{N}Y^{N}|T) = R_{N}^{\infty}T\mu^{*} = \sum_{r=1}^{\infty}\sum_{s=0}^{\infty}\sum_{q=0}^{\infty}\left\{(G^{N})^{r}[G^{NE}(G^{E})^{q}G^{EN}]^{s}G^{NE}\eta_{rsq}\right\}T,$$

$$E(G^{NE}Y^{E}|T) = R_{NE}^{\infty}T\iota^{*} = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{q=0}^{\infty}\left\{G^{NE}[(G^{E})^{q}G^{EN}(G^{N})^{r}G^{NE}]^{s}\iota_{rsq}\right\}T,$$

where R_{N}^{∞} and R_{NE}^{∞} are two sets of matrices containing all the combinations of products of powers of the adjacency matrices, and μ^{*} and ι^{*} are vectors containing the relative parameters, η_{rsq} ι_{rsq}, that in turn are products of δ and the endogenous effects (for each specific combination of r, s and q). It is easy to show that in the linear-in-means case these products of powers of the adjacency matrices are functions of products of the number of eligible agents, ineligible agents, and population size. We prove this by induction. Let us consider $E(G^{N}Y^{N}|T)$. (The proof remains valid for the other endogenous terms.) Given our definition of G, we can write

$$E(G^{N}Y^{N}|T) = \sum_{r=1}^{\infty}\sum_{s=0}^{\infty}\sum_{q=0}^{\infty}\left\{(\tilde{G}^{N}S^{N})^{r}[\tilde{G}^{NE}S^{NE}(\tilde{G}^{E}S^{E})^{q}\tilde{G}^{EN}S^{EN}]^{s}\tilde{G}^{NE}S^{NE}\eta_{rsq}\right\}T.$$

To ease the notation, let us define $\eta_{rsq} = \eta$. Without loss of generality let us also assume that $c = 1$. If $p = 0$, $q = 0$ and $T = 1_{e,1}$, every eligible is treated and we have

$$\begin{aligned}
\tilde{G}^N S^N \tilde{G}^{NE} S^{NE} T\eta &= \frac{1}{n-1}(1_{n,n} - I_n) I_n \frac{n-1}{m-1} 1_{n,e} \frac{1}{e} I_e \frac{e}{m-1} T\eta \\
&= \frac{1}{n-1}(1_{n,n} - I_n) I_n \frac{n-1}{m-1} \frac{1}{m-1} 1_{n,e} T\eta \\
&= \frac{1}{n-1}(1_{n,n} - I_n) I_n \frac{n-1}{m-1} \frac{e}{m-1} 1_{n,1}\eta \\
&= \frac{e}{(m-1)^2}(1_{n,1} n - 1_{n,1})\eta = \frac{(n-1)e}{(m-1)^2} 1_{n,1}\eta,
\end{aligned}$$

$$\begin{aligned}
(\tilde{G}^N S^N)^2 \tilde{G}^{NE} S^{NE} T\eta &= \frac{1}{n-1}(1_{n,n} - I_n) I_n \frac{n-1}{m-1} \frac{(n-1)e}{(m-1)^2} 1_{n,1}\eta \\
&= \frac{(n-1)^2 e}{(m-1)^3} 1_{n,1}\eta,
\end{aligned}$$

$$\begin{aligned}
(\tilde{G}^N S^N)^k \tilde{G}^{NE} S^{NE} T\eta &= \frac{1}{n-1}(1_{n,n} - I_n) I_n \frac{n-1}{m-1} \frac{(n-1)^{k-1} e}{(m-1)^k} 1_{n,1}\eta \\
&= \frac{(n-1)^k e}{(m-1)^{k+1}} 1_{n,1}\eta.
\end{aligned}$$

If we allow q to be different from zero we then have

$$(\tilde{G}^N S^N)^k \tilde{G}^{NE} S^{NE} (\tilde{G}^E S^E)^l T\eta = \frac{(n-1)^k e (e-1)^l}{(m-1)^{k+l+1}} 1_{n,1}\eta.$$

Finally, if we allow s to be greater than one we have

$$(\tilde{G}^N S^N)^k [\tilde{G}^{NE} S^{NE} (\tilde{G}^E S^E)^l \tilde{G}^{EN} S^{EN}]^v \tilde{G}^{NE} S^{NE} T\eta = \frac{(n-1)^k e^{v+1} (e-1)^{lv} n^v}{(m-1)^{[v(l+2)+k+1]}} 1_{n,1}\eta.$$

It follows that the expected value of $G^N Y^N$ conditional on treatment can be approximated by

$$E(G^N Y^N | T) \approxeq Q_N^\infty \mu = \sum_{v=1}^{\infty} \sum_{s=0}^{\infty} \sum_{q=0}^{\infty} \frac{(n_c - 1)^v e_c (e_c - 1)^q (e_c n_c)^s}{(m_c - 1)^{2s+1+v+q}} 1_{n,1} \mu_{vrsq}, \tag{39}$$

where we use a more flexible combination of sums of nonlinear functions of the shares. Applying

the same procedure, we can compute the same approximations for all the endogenous terms

$$E(G^{NE}Y^{E}|T) \approxeq Q_{NE}^{\infty}\iota = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{q=0}^{\infty}\frac{e_c(e_c-1)^q(e_cn_c)^s(n_c-1)^r}{(m_c-1)^{2s+1+r+q}}1_{n,1}\iota_{rsq}, \tag{40}$$

$$E(G^{E}Y^{E}|T) \approxeq Q_{E}^{\infty}\psi = \sum_{\substack{v=1\\ r>0}}^{\infty}\sum_{\substack{r=0\\ if s>0}}^{\infty}\sum_{s=0}^{\infty}\frac{(e_c-1)^v(e_cn_c)^s(n_c-1)^r}{(m_c-1)^{2s+r+v}}1_{e,1}\psi_{vrsq}, \tag{41}$$

$$E(G^{EN}Y^{N}|T) \approxeq Q_{EN}^{\infty}\zeta = \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\sum_{q=0}^{\infty}\frac{n_ce_c(e_c-1)^q(e_cn_c)^s(n_c-1)^r}{(m_c-1)^{2s+2+r+q}}1_{e,1}\zeta_{rsq}. \tag{42}$$

Observe that these terms are approximations of the optimal instruments for the endogenous variables. It is straightforward to see that if e_c, n_c, and m_c vary across c (groups), $E(G^NY^N|T)$, $E(G^{NE}Y^E)$, $E(G^EY^E)$, and $E(G^{EN}Y^N)$ are linearly independent to T and thus $E(Z_E)$ and $E(Z_N)$, where $Z_E = [G^EY^E, G^{EN}Y^N, A, T]$ and $Z_N = [G^NY^N, G^{NE}Y^E, B]$ have full column rank. ■

SUPPLEMENTARY MATERIAL

Further theoretical results: In Appendix S1, we provide identification conditions for alternative specifications of model (2)-(3) and the detailed derivation of the instrumental variables and the estimands defined respectively in Sections 3.2 and 4.

Additional tables and figures: In Appendix S2, we collect additional results on the application of our methodology described in Section 5.2.

Matlab and Stata codes: This appendix contains the Matlab and Stata codes to implement our methodology, perform the simulations described in Section 3.3, and additional simulation results.

ACKNOWLEDGMENTS

For useful comments and discussions we thank Badi Baltagi, Giacomo De Giorgi, Herman K. van Dijk, Chihwa Kao, Alfonso Flores- Lagunes, Kajal Lahiri, Yoonseok Lee, Xiaodong Liu, Bin Peng, Fa Wang, Ian Wright, and participants at the 2017 Conference on Causal Inference, Program Evaluation and External Validity in Luxembourg, the 16th ZEW Summer Workshop for Young Economists: Advances in Microeconometrics and Programme Evaluation, the 6th Italian Congress

of Econometrics and Empirical Economics, the 13th CSEF-IGIER Symposium on Economics and Institutions, and seminar participants at the Stockholm University - SOFI and the University of Bologna.

References

Angelucci, M. and De Giorgi, G. (2009). Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption?, *American Economic Review* **99**(1): 486–508.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

Baird, S., Bohren, J. A., McIntosh, C. and Özler, B. (2018). Optimal design of experiments in the presence of interference, *Review of Economics and Statistics* (Forthcoming).

Bandiera, O., Burgess, R., Das, N., Gulesci, S., Rasul, I. and Sulaiman, M. (2017). Labor markets and poverty in village economies, *The Quarterly Journal of Economics* **132**(2): 811–870.

Barrera-Osorio, F., Bertrand, M., Linden, L. L. and Perez-Calle, F. (2011). Improving the design of conditional transfer programs: Evidence from a randomized education experiment in Colombia, *American Economic Journal: Applied Economics* **3**(2): 167–195.

Bobonis, G. J. and Finan, F. (2009). Neighborhood peer effects in secondary school enrollment decisions, *The Review of Economics and Statistics* **91**(4): 695–716.

Boucher, V., Bramoullé, Y., Djebbari, H. and Fortin, B. (2014). Do peers affect student achievement? evidence from canada using group size variation, *Journal of applied econometrics* **29**(1): 91–109.

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

Calvó-Armengol, A., Patacchini, E. and Zenou, Y. (2009). Peer effects and social networks in education, *The Review of Economic Studies* **76**(4): 1239–1267.

Djebbari, H. and Smith, J. (2008). Heterogeneous impacts in progresa, *Journal of Econometrics* **145**(1): 64–80.

Flores-Lagunes, A. (2007). Finite sample evidence of iv estimators under weak instruments, *Journal of Applied Econometrics* **22**(3): 677–694.

Forastiere, L., Airoldi, E. M. and Mealli, F. (2016). Identification and estimation of treatment and interference effects in observational studies on networks, *arXiv preprint arXiv:1609.06245* .

Grosh, M., Bussolo, M. and Freije, S. (2014). *Understanding the poverty impact of the global financial crisis in Latin America and the Caribbean*, World Bank Publications.

Hudgens, M. and Halloran, E. (2008). Toward causal inference with interference, *Journal of American Statistical Association* **103**: 832–842.

Imbens, G. and Woolridge, G. (2009). Recent developments in the econometrics of program evaluation., *Journal of Economic Literature* **47(1)**: 5–86.

Kelejian, H. and Prucha, I. R. (1998). 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.

Kelejian, H. and Prucha, I. R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model, *International economic review* **40**(2): 509–533.

Lalive, R. and Cattaneo, M. A. (2009). Social interactions and schooling decisions, *The Review of Economics and Statistics* **91**(3): 457–477.

Lee, L.-f. (2003). Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances, *Econometric Reviews* **22**(4): 307–335.

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

Lee, S. and Shaikh, A. M. (2014). Multiple testing and heterogeneous treatment effects: Re-evaluating the effect of progresa on school enrollment, *Journal of Applied Econometrics* **29**(4): 612–626.

Liu, L. and Hudgens, M. G. (2014). Large sample randomization inference of causal effects in the presence of interference, *Journal of the american statistical association* **109**(505): 288–301.

Liu, X. (2014). Identification and efficient estimation of simultaneous equations network models, *Journal of Business & Economic Statistics* **32**(4): 516–536.

Moffitt, R. A. et al. (2001). Policy interventions, low-level equilibria, and social interactions, *Social Dynamics* **4**: 45–82.

Rigdon, J. and Hudgens, M. G. (2015). Randomization inference for treatment effects on a binary outcome, *Statistics in medicine* **34**(6): 924–935.

Rubin, D. (1986). Which ifs have causal answers? discussion of hollands statistics and causal inference", *Journal of the American Statistical Association* **81**: 961–962.

Sanderson, E. and Windmeijer, F. (2016). A weak instrument F-test in linear iv models with multiple endogenous variables, *Journal of Econometrics* **190**(2): 212–221.

Sobel, M. E. (2006). What do randomized studies of housing mobility demonstrate? causal inference in the face of interference, *Journal of the American Statistical Association* **101**(476): 1398–1407.

Tchetgen, E. J. T. and VanderWeele, T. J. (2012). On causal inference in the presence of interference, *Statistical Methods in Medical Research* **21**(1): 55–75.