Submitted to *Econometrica*

ONLINE APPENDIX: DISCRETE CHOICE WITH GENERALIZED SOCIAL INTERACTIONS

OSCAR VOLPE

Department of Economics, Harvard University

November 20, 2025

This online appendix accompanies the paper "Discrete Choice with Generalized Social Interactions". It contains supplemental proofs, model extensions, additional details about identification and estimation, and robustness analyses for the empirical applications. All notation is consistent with the main text of the paper.

APPENDIX A: EQUILIBRIUM ANALYSIS AND MODEL EXTENSIONS

A.1. Justification for Supplemental Claims

CLAIM: *If $J_{kj_1}J_{j_1j_2}\cdots J_{j_Mk}=0$ for all $k, j_1, \dots, j_M$, then $\rho(\mathbf{D}(m))=0$ for all $m\in[0,1]^K$.*

PROOF: Let $J_{kj_1}J_{j_1j_2}\cdots J_{j_Mk}=0$ for all types $k, j_1, \dots, j_M$. It follows that $(\mathbf{J}^M)_{kk}=0$ for all types k and integers $M\geq 1$. Since this property also applies to $\mathbf{D}(m)$, it must be that:

$$\sum_{k=1}^{K}[\lambda_k(m)]^M=\text{tr}([\mathbf{D}(m)]^M)=\sum_{k=1}^{K}([\mathbf{D}(m)]^M)_{kk}=0,$$

for every $M\geq 1$, where $\{\lambda_k(m)\}_k$ are the eigenvalues of $\mathbf{D}(m)$. This equation holds only if $\lambda_k(m)=0$ for all k. Given that $\rho(\mathbf{D}(m))=\max_k\{|\lambda_k(m)|\}$, it follows that $\rho(\mathbf{D}(m))=0$. *Q.E.D.*

CLAIM: *If $\sum_{\ell=1}^{K}|J_{k\ell}|<S_k$ for some $\{S_k\}_k$, the equilibrium is unique and locally stable.*

PROOF: By Horn and Johnson (1985), Thm. 5.6.9, it follows that $\rho(\mathbf{D}(m))\leq||\mathbf{D}(m)||_\infty$, where $||\cdot||_\infty$ represents the p-norm defined by the maximum absolute row sum of $\mathbf{D}(m)$:

$$||\mathbf{D}(m)||_\infty=\max_k\left\{\sum_{\ell=1}^{K}|[\mathbf{D}(m)]_{k\ell}|\right\}=\max_k\left\{f_{\varepsilon|k}\left(h_k+\sum_{\ell=1}^{K}J_{k\ell}m^\ell\right)\sum_{\ell=1}^{K}|J_{k\ell}|\right\}.$$

Oscar Volpe: ovolpe@fas.harvard.edu

So, it must be that $\rho(\mathbf{D}(m)) < 1$ if $\sum_{\ell=1}^{K} |J_{k\ell}| < S_k$, where $S_k = f_{\varepsilon|k}^{-1}\big(h_k + \sum_{\ell=1}^{K} J_{k\ell} m^{\ell}\big)$. Note that $\{m_k\}_k$ can be expressed in terms of the model parameters by solving system (5).

Q.E.D.

CLAIM: *If* $\mathbf{J}$ *is symmetric and if its eigenvalues all have non-positive real parts, then there is a unique equilibrium, which is stable if* $\rho(\mathbf{D}(m)) < 1$ *and unstable if* $\rho(\mathbf{D}(m)) > 1$.

PROOF: Let $\mathbf{J}$ be symmetric with eigenvalues that all have non-positive real parts. As $\boldsymbol{\beta}(m)$ is a diagonal matrix with only positive elements, $\mathbf{J}$ is congruent to $\boldsymbol{\beta}^{1/2}(m)\mathbf{J}\boldsymbol{\beta}^{1/2}(m)$, which is similar to $\boldsymbol{\beta}^{1/2}(m)[\boldsymbol{\beta}^{1/2}(m)\mathbf{J}\boldsymbol{\beta}^{1/2}(m)]\boldsymbol{\beta}^{-1/2}(m) = \boldsymbol{\beta}(m)\mathbf{J} = \mathbf{D}(m)$. By Sylvester's law of inertia, $\mathbf{D}(m)$ also has eigenvalues with non-positive real parts. Therefore, it follows that: $0 < \prod_{k=1}^{K}\big(1-\lambda_k(m)\big) = \det\big(I - \mathbf{D}(m)\big) = \det\big(\mathbf{D}_{\mathcal{H}}(m)\big)$ at any equilibrium m. By the index theorem, the model always has a unique equilibrium. By Lemma 1, this equilibrium will be locally stable whenever $\rho(\mathbf{D}(m)) < 1$, and it will be unstable whenever $\rho(\mathbf{D}(m)) > 1$.

Q.E.D.

CLAIM: *No Negative Feedback holds iff* $\mathbf{BJB}^{-1}$ *is nonnegative for a diagonal matrix* $\mathbf{B}$.

PROOF: "$\Rightarrow$" Suppose that $\mathbf{J}$ is similar to a non-negative matrix $\mathbf{A}$ by way of a diagonal change-of-basis matrix $\mathbf{B}$. That is, there exists some diagonal $\mathbf{B}$ where $\mathbf{A} = \mathbf{BJB}^{-1} \geq \mathbf{0}$. Since $\mathbf{B}$ is diagonal, the elements of $\mathbf{A}$ can be expressed as $A_{k\ell} = B_{kk}J_{k\ell}/B_{\ell\ell}$, for all k, ℓ. Therefore, for any choice of indices k and $\ell_1, \ldots, \ell_M$ in the set $\{1, \ldots, K\}$, it must be that:

$$
\begin{aligned}
0 &\leq A_{k\ell_1}A_{\ell_1\ell_2}\cdots A_{\ell_M k} \\
&= \Big(\frac{B_{kk}}{B_{\ell_1\ell_1}}J_{k\ell_1}\Big)\Big(\frac{B_{\ell_1\ell_1}}{B_{\ell_2\ell_2}}J_{\ell_1\ell_2}\Big)\cdots\Big(\frac{B_{\ell_M\ell_M}}{B_{kk}}J_{\ell_M k}\Big) \\
&= J_{k\ell_1}J_{\ell_1\ell_2}\cdots J_{\ell_M k}.
\end{aligned}
$$

"$\Leftarrow$" Suppose that NNF holds. I first restrict attention to the case where $\mathbf{J}$ is irreducible. For any k, define $\{\gamma_\ell^k\}_{\ell=1}^{K}$ so that $\gamma_\ell^k = 1$ if $J_{kj_1}J_{j_1j_2}\ldots J_{j_M\ell} \geq 0$ for all types $j_1, \ldots, j_M$, and $\gamma_\ell^k = -1$ otherwise. Next, fixing some reference type k_0, define the matrix $\mathbf{B}$ such that:

$$
\mathbf{B} = \operatorname{diag}\begin{bmatrix}\gamma_1^{k_0}\\ \vdots \\ \gamma_K^{k_0}\end{bmatrix}.
$$

Notice that $\mathbf{B}$ is involutory, i.e. $\mathbf{B}^{-1} = \mathbf{B}$. Therefore, for all (g, ℓ), NNF would ensure that:

$$
[\mathbf{BJB}^{-1}]_{k,\ell} = [\mathbf{BJB}]_{k,\ell} = \gamma_k^{k_0}\gamma_\ell^{k_0}J_{k\ell} = \gamma_{k_0}^{k}\gamma_\ell^{k_0}J_{k\ell} = \gamma_\ell^{k}J_{k\ell} = |J_{k\ell}|
$$

It follows that $\mathbf{BJB}^{-1}$ equals the absolute value of $\mathbf{J}$. Therefore, $\mathbf{BJB}^{-1}$ is non-negative. Finally, if $\mathbf{J}$ is not irreducible, then this same reasoning applies to all irreducible blocks of $\mathbf{J}$. So, it must always be that $\mathbf{BJB}^{-1}$ is non-negative for some diagonal matrix $\mathbf{B} \in \mathbb{R}^{K\times K}$.

Q.E.D.

CLAIM: *No Negative Feedback holds if* $\operatorname{sgn}(J_{k\ell}) = \operatorname{sgn}(J_{km}J_{m\ell})$ *for all* $k, \ell, m \in \{1, \dots, K\}$.

PROOF: To begin, suppose that $\operatorname{sgn}(J_{k\ell}) = \operatorname{sgn}(J_{km}J_{m\ell})$ for any types k, ℓ, m. I prove by induction that $\operatorname{sgn}(J_{j_0j_1}J_{j_1j_2}\cdots J_{j_Mj_0}) \geq 0$ for any arbitrary types $j_0, j_1, \dots, j_M$. First, note that $\operatorname{sgn}(J_{j_0j_2}) = \operatorname{sgn}(J_{j_0j_1}J_{j_1j_2})$. Next, let $\operatorname{sgn}(J_{j_0j_M}) = \operatorname{sgn}(J_{j_0j_1}J_{j_1j_2}\cdots J_{j_{M-1}j_M})$ for some positive integer $M \geq 1$. By construction, the following equation must be satisfied:

$$\operatorname{sgn}(J_{j_0j_{M+1}}) = \operatorname{sgn}(J_{j_0j_M}J_{j_Mj_{M+1}}) = \operatorname{sgn}(J_{j_0j_1}J_{j_1j_2}\cdots J_{j_{M-1}j_M}J_{j_Mj_{M+1}}).$$

By induction, $\operatorname{sgn}(J_{j_0\ell}) = \operatorname{sgn}(J_{j_0j_1}J_{j_1j_2}\cdots J_{j_M\ell})$ for any agent type ℓ. By setting $\ell = j_0$, I obtain that $\operatorname{sgn}(J_{j_0j_1}J_{j_1j_2}\cdots J_{j_Mj_0}) = \operatorname{sgn}(J_{j_0j_0})$ where $\operatorname{sgn}(J_{j_0j_0}) = \operatorname{sgn}(J_{j_0j_0}J_{j_0j_0}) \geq 0$.

Q.E.D.

A.2. Extensions to Alternative Network-Based Models

The properties laid out in Section 2 are not specific to binary choice models. They also have implications for a much broader class of models where agents interact in a network. To see how, consider a game with K players. Each player k chooses a_k from a compact action space $A_k \in \mathbb{R}$. Given a profile of actions $a \in A_1 \times A_2 \times \cdots \times A_K$, each player's best response is $a_k^* = q_k\big(\sum_{\ell=1}^K J_{k\ell}a_\ell\big)$, where $q_k(\cdot)$ is a non-decreasing function mapping from $\mathbb{R}$ to A_k.

This game encompasses a variety of economic models. For example, the action a_k could represent a person's investment into a public good, where everyone benefits from how much their neighbors contribute.[1] Alternatively, a_k could represent the output produced by a firm that competes in an oligopoly, where every firm's action affects the market price. Both these models are well-studied, and they both involve negative interaction effects between agents.

Another interpretation of the game is that each player is a community of individuals. In the binary choice model, a player is a social group whose members make one of two choices subject to social influences. Agents act noncooperatively, and a_k is the average choice in group k. The same logic applies when residents of a state or local institution k make a collective choice a_k. For example, one could model policy spillovers across U.S. states, where

[1] Bramoullé et al. (2014) study a public goods game that is nested within this framework. In their paper, players choose from an interval $[0,1]$ and the best responses are $a_k = \max\{0, 1 - \delta\sum_{\ell\neq k} g_{k\ell}a_\ell\}$, where $\delta \in [0,1]$ and $g_{k\ell} \in \{0,1\}$ indicates whether players k and ℓ are linked. Note that this game involves pure strategic substitutes.

voters support more liberal or conservative positions in response to laws adopted elsewhere.

I focus on pure strategy Nash equilibria, defined as the action profiles a from which no player k wishes to deviate. It is well known that an equilibrium exists if there are continuous best responses and compact, convex action spaces. Moreover, if the best responses are differentiable at each equilibrium a, then the spectral radius of the Jacobian matrix is all that is needed to determine uniqueness. In particular, let a^* be the equilibrium at which $\rho(\mathbf{D}(a))$ is greatest. If $\rho(\mathbf{D}(a^*)) < 1$, then a^* is the unique, locally stable equilibrium. Moreover, if Assumption A is satisfied, then there are multiple, locally stable equilibria if $\rho(\mathbf{D}(a^*)) > 1$.

The set of equilibria is generally not well-ordered. In particular, it need not form a lattice structure, which yields a single equilibrium that is broadly favorable to all players. As seen from Property 4, negative interactions can lead different players to favor different equilibria, making it impossible to jointly maximize the welfare of all players at the same equilibrium.

One convenient feature of this game is that every player's best response either (weakly) increases or decreases monotonically in another player's action. Thus, if best responses are differentiable, the entries of $\mathbf{D}$ retain the same sign for all values of a. This assumption may be weakened. For example, suppose that a player's current action determines whether they will conform to or deviate from others. Then, the characterization of equilibrium properties requires checking that Assumption A holds locally in certain regions of the support of a.

A.3. Endogenous Network Formation with Generalized Interactions

In the paper, it is assumed that agents respond to the expected average action for each type. I now consider an alternative model, where utility depends on the expected composition of agents who choose an action; that is, I assume agents care about $\mathrm{E}(k|\omega_i)$ instead of $\mathrm{E}(\omega_i|k)$.

This alternative model specification can be used to analyze how social identity influences network formation. For example, suppose agents must decide whether to enter a new environment, such as a school, and they care about the types of people they expect to encounter. This situation invariably leads to negative interaction effects, because preferring one group to be in the majority is equivalent to preferring another group to be in the minority. These preferences can be analyzed by using the same techniques outlined in Section 2 of the paper.

Consider a model with two social groups: a and b.[2] Let λ_a denote the share of people in group a, and let $s_a(\omega_i)$ denote the probability of being in group a given that one chooses ω_i. The payoff from a choice ω_i depends on the expected share of agents who make that choice.

[2] I focus on the two-group case for simplicity, but this framework may be generalized to cases with many groups.

$$U_i(\omega_i|k) = v_k(\omega_i) + J_k s_a(\omega_i) + \epsilon_i(\omega_i), \quad \text{for } k \in \{a, b\}.$$

Under this framework, the parameter J_k indicates how much people in group k benefit from associating with members of group a. Both $v_k(\cdot)$ and $\epsilon_i(\cdot)$ are specified exactly as before.

In equilibrium, the expected composition of agents making a choice must be consistent with individually optimal decisionmaking. By Bayes' rule, any equilibrium should satisfy:

$$s_a(\omega_i) = \frac{\lambda_a \, \mathrm{P}(\omega_i|a)}{\lambda_a \, \mathrm{P}(\omega_i|a) + (1-\lambda_a)\, \mathrm{P}(\omega_i|b)}, \quad \text{for } \omega_i \in \{0, 1\},$$

where $\mathrm{P}(\omega_i = 1|k) = F_{\varepsilon|k}\big(h_k + J_k(s_a(1) - s_a(0))\big)$. By Brouwer's FPT, a solution exists.

Uniqueness and dynamic stability of equilibria depend on the Jacobian matrix $\mathbf{D} \in \mathbb{R}^{2\times 2}$ of the system (1). Letting $\beta_k = f_{\varepsilon|k}\big(h_k + J_k(s_a(1) - s_a(0))\big)$, I define this matrix so that:

$$\mathbf{D} = \begin{bmatrix} D_{11} & -D_{11} \\ -D_{22} & D_{22} \end{bmatrix}, \quad \text{where: } \begin{cases} D_{11} &= s_a(0)[1-s_a(0)]\left(\frac{J_a\beta_a}{\mathrm{P}(0|a)} - \frac{J_b\beta_b}{\mathrm{P}(0|b)}\right) \\ D_{22} &= s_a(1)[1-s_a(1)]\left(\frac{J_a\beta_a}{\mathrm{P}(1|a)} - \frac{J_b\beta_b}{\mathrm{P}(1|b)}\right) \end{cases}$$

Suppose that all agents prefer to associate with their own group, i.e., $J_a \geq 0$ and $J_b \leq 0$. In this case, the matrix $\mathbf{D}$ satisfies Assumption A.1. Hence, there is almost always a locally stable equilibrium—even if the spillover effects are strong enough to generate multiplicity.

If agents prefer not to associate with their own social group, i.e., if $J_a < 0$ or $J_b > 0$, then the model may not have any locally stable equilibria. Global instability occurs when agents' behavior is self-undermining, such that a group always seeks to deviate from its own action. As an example, consider Allende (2019), who studies how peer effects drive student sorting in Peru. In her model, all parents prefer to that their children attend schools that are made up of wealthy, high-achieving peers. If the peer effects are sufficiently weak, then a unique, stable equilibrium would exist. However, if low-performing students overwhelmingly seek to associate with their high-achieving peers, then a locally stable equilibrium may not exist.

In this model, the amount of diversity in a network is tied to stability of an equilibrium. To see why, note that the rows of the Jacobian matrix $\mathbf{D}$ are scaled by $s_a(\omega_i)[1 - s_a(\omega_i)]$. This term grows larger as $s_a(\omega_i)$ approaches $1/2$, and it tends to zero as $s_a(\omega_i)$ approaches 0 or 1. So, if the agents who choose an action are more diverse, then the Jacobian is more expansive, and the realized equilibrium is more likely to be unstable. Conversely, if one group makes up the majority of people choosing an action, then the equilibrium is more likely to be locally stable. Of course, this property only holds for static models, where group membership is fixed. In the school choice example, some traits like student achievement are likely to change over time, while other traits, e.g., racial identity, may be fixed. One area for future work is to explore how the dynamics of identity affect the formation of networks.

APPENDIX B: DETAILS ON IDENTIFICATION AND ESTIMATION

B.1. Modifying Identification Results to Allow for Covariates

Suppose that the outcome equation is modified to allow for exogenous covariates $X_i \in \mathbb{R}^r$. Let $\omega_i = \mathbb{1}\{X_i'c + h_k + \alpha_n + \sum_{\ell=1}^K J_{k\ell} m_n^\ell + \varepsilon_i \geq 0\}$ where $\mathrm{P}(\varepsilon_i \leq z | X_i, k, \alpha_n) = F_{\varepsilon|k}(z)$ and $\mathrm{P}(X_i \leq x | k, \alpha_n) = F_{X|k}(x)$. Then $m_n^k = \int \mathrm{E}(\omega_i | X_i, k, \alpha_n, \{m_n^\ell\}_{\ell=1}^K) dF_{X|k}$, where:

$$\mathrm{E}(\omega_i | X_i, k, \alpha_n, \{m_n^\ell\}_{\ell=1}^K) = F_{\varepsilon|k}\left(h_k + \alpha_n + X_i'c + \sum_{\ell=1}^K J_{k\ell} m_n^\ell\right), \quad \text{for } k = 1, \ldots, K.$$

In the presence of covariates, Lemma 2 must be adapted. I do so in the following way.

LEMMA 2—Sufficiency Claim: *For any agents i and j in groups k_1 and k_2, respectively, and network n:* $\mathrm{E}(\omega_i | X_i, k_1, \alpha_n, \{m_n^\ell\}_\ell) = \mathrm{E}(\omega_i | X_i, X_j, k_1, \{m_n^\ell\}_\ell, \mathrm{E}(\omega_j | X_j, k_2, \alpha_n, \{m_n^\ell\}_\ell))$.

PROOF: As $F_{\varepsilon|k_2}$ is strictly increasing, its inverse exists. By this property, I can write:

$$\alpha_n = F_{\varepsilon|k_2}^{-1}\big(\mathrm{E}(\omega_j | X_j, k_2, \alpha_n, \{m_n^\ell\}_\ell)\big) - h_{k_2} - X_j'c - \sum_{\ell=1}^K J_{k_2\ell} m_n^\ell.$$

By plugging this expression for α_n into the definition of $\mathrm{E}(\omega_i | X_i, k_1, \alpha_n, \{m_n^\ell\}_\ell)$, I obtain:

$$\begin{aligned}\mathrm{E}(\omega_i | X_i, k_1, \alpha_n, \{m_n^\ell\}_\ell) = F_{\varepsilon|k_1}\Bigg(& h_{k_1} - h_{k_2} + (X_i - X_j)'c \\ & + \sum_{\ell=1}^K (J_{k_1\ell} - J_{k_2\ell}) m_n^\ell + F_{\varepsilon|k_2}^{-1}\big(\mathrm{E}(\omega_j | X_j, k_2, \alpha_n, \{m_n^\ell\}_\ell)\big)\Bigg).\end{aligned}$$

Q.E.D.

To prove Theorem 1 (semiparametric identification) with covariates, I require that there is some k such that supp$(X|k)$ is not contained in the proper linear subspace of $\mathbb{R}^r$. Note that this condition is weaker than Assumption B.3. Yet, it is sufficient to show that c, $h_{k_1} - h_{k_2}$ and $\{J_{k_1\ell} - J_{k_2\ell}\}_\ell$ are identified for all k_1 and k_2. Consider the following result and proof.

THEOREM 1: *Suppose Assumptions B.1 and B.2 hold, and assume m_n^k is observed for all networks n and agent types k. Also, suppose that there is some type k where supp$(X|k)$ is not contained in a proper linear subspace of $\mathbb{R}^r$. If the functions $\{F_{\varepsilon|k}\}_k$ are known, then:*

(i) Without further assumptions, $\{h_{k_1} - h_{k_2}\}_{k_1,k_2}$ and $\{J_{k_1\ell} - J_{k_2\ell}\}_{k_1,k_2,\ell}$ are identified.

(ii) If $\alpha_n = W_n'd$ for some observed vector W_n, then d, $\{h_k\}_k$, and $\{J_{k\ell}\}_{k,\ell}$ are identified.

PROOF: Fix some network n and define the term $\zeta_n^k = h_k + \alpha_n + \sum_{\ell=1}^K J_{k\ell} m_n^\ell$. I write:

$$\mathrm{E}(\omega_i | X_i, k, \alpha_n, \{m_n^\ell\}_{\ell=1}^K) = F_{\varepsilon|k}(\zeta_n^k + X_i'c)$$

To recover c, I must show that $F_{\varepsilon|k}(\zeta_n^k + X_i'c) = F_{\varepsilon|k}(\hat{\zeta}_n^k + X_i'\hat{c})$ implies $c = \hat{c}$. Since $F_{\varepsilon|k}$ is known, this property holds by Proposition 5 of Manski (1988). It follows that c is identified.

Having demonstrated that c can be recovered, I now focus on the other parameters. Consider any two social groups k_1 and k_2, and then define the function $\phi : \mathbb{R} \to [0,1]$ so that:

$$\phi(\nu) = \int F_{\varepsilon|k_1}\big((X_i - X_j)'c + \nu\big) dF_{X|k_1},$$

where X_j is chosen from supp$(X|k_2)$ and the integral is evaluated over the conditional support of X_i given k_1. By definition, $\phi(\cdot)$ is nonlinear and monotonically increasing in ν.

For any n, let $\nu_n = h_{k_1} - h_{k_2} + \sum_{\ell=1}^K (J_{k_1\ell} - J_{k_2\ell}) m_n^\ell + F_{\varepsilon|k_2}^{-1}(\mathrm{E}(\omega_j | X_j, k_2, \alpha_n, \{m_n^\ell\}_{\ell=1}^K))$. By Lemma 2, the expected average choice $m_n^{k_1}$ equals $\phi(\nu_n)$. Because $\phi(\cdot)$ is monotonic:

$$\begin{aligned} m_n^{k_1} &= \int F_{\varepsilon|k_1}\Bigg(h_{k_1} - h_{k_2} + (X_i - X_j)'c + \sum_{\ell=1}^K (J_{k_1\ell} - J_{k_2\ell}) m_n^\ell + F_{\varepsilon|k_2}^{-1}(\mathrm{E}(\omega_j | X_j, k_2, \alpha_n, \{m_n^\ell\}_{\ell=1}^K))\Bigg) dF_{X|k_1} \\ &= \int F_{\varepsilon|k_1}\Bigg(\overline{h_{k_1} - h_{k_2}} + (X_i - X_j)'c + \sum_{\ell=1}^K \overline{(J_{k_1\ell} - J_{k_2\ell})} m_n^\ell + F_{\varepsilon|k_2}^{-1}(\mathrm{E}(\omega_j | X_j, k_2, \alpha_n, \{m_n^\ell\}_{\ell=1}^K))\Bigg) dF_{X|k_1} \end{aligned}$$

is satisfied if and only if $\sum_{\ell=1}^K \big[(\overline{J_{k_1\ell} - J_{k_2\ell}}) - (J_{k_1\ell} - J_{k_2\ell})\big] m_n^\ell = (h_{k_1} - h_{k_2}) - (\overline{h_{k_1} - h_{k_2}})$ for all networks $n \in \{1, \ldots, N\}$. Since the expected average choices are nonlinear functions of one another, sufficient variation in the equilibrium outcomes across networks implies:

$$h_{k_1} - h_{k_2} = \overline{h_{k_1} - h_{k_2}} \quad \text{and} \quad J_{k_1\ell} - J_{k_2\ell} = \overline{J_{k_1\ell} - J_{k_2\ell}},$$

for all $\ell \in \{1, \ldots, K\}$. Also, as k_1 and k_2 are chosen arbitrarily, this holds for all type pairs. *Q.E.D.*

For Theorem 2, no modifications are needed, since this result already leverages individual-level covariates. For Theorem 3, all modifications follow directly from Theorems 1 and 2. Additionally, the IV estimands β_{k_1,k_2} and β_k can be readily adapted to allow for covariates.

B.2. Asymptotic Properties of the IV Estimator

First, consider the estimator $\hat{\beta}_k$, corresponding to the estimand β_k in Theorem 3*(ii)*. I write:

$$\hat{\beta}_k = \left(\frac{1}{N}\sum_{n=1}^{N} Z_n X_n'\right)^{-1} \left(\frac{1}{N}\sum_{n=1}^{N} Z_n Y_n^k\right).$$

Note that $\hat{\beta}_k \xrightarrow{p} \beta_k$ as the number of networks N grows large.[3] Additionally, it follows that:

$$\sqrt{N}(\hat{\beta}_k - \beta_k) = \left(\frac{1}{N}\sum_{n=1}^{N} Z_n X_n'\right)^{-1} \left(\frac{1}{\sqrt{N}}\sum_{n=1}^{N} Z_n [Y_n^k - X_n'\beta_k]\right) \xrightarrow{d} \mathcal{N}(0, Q\Omega Q'),$$

where $Q = \mathrm{E}(Z_n X_n')^{-1}$ and $\Omega = \mathrm{Var}(Z_n[Y_n^k - X_n'\beta_k])$. By the sample analogue principle, I can construct a consistent estimator for the asymptotic variance of $\hat{\beta}_k$ from $\hat{Q}\hat{\Omega}\hat{Q}'$, where:

$$\hat{Q} = \left(\frac{1}{N}\sum_{n=1}^{N} Z_n X_n'\right)^{-1} \quad \text{and} \quad \hat{\Omega} = \frac{1}{N}\sum_{n=1}^{N} Z_n [Y_n^k - X_n'\hat{\beta}_k]^2 Z_n'.$$

Next, consider the estimator $\hat{\beta}_{k_1,k_2}$, corresponding to the estimand β_{k_1,k_2} in Theorem 3*(i)*:

$$\hat{\beta}_{k_1,k_2} = \left(\frac{1}{N}\sum_{n=1}^{N} Z_n X_n'\right)^{-1} \left(\frac{1}{N}\sum_{n=1}^{N} Z_n [Y_n^{k_1} - Y_n^{k_2}]\right).$$

As in the previous case, $\hat{\beta}_{k_1,k_2} \xrightarrow{p} \beta_{k_1,k_2}$ as N grows large. Moreover, the estimator satisfies:

$$\sqrt{N}(\hat{\beta}_{k_1,k_2} - \beta_{k_1,k_2}) \xrightarrow{d} \mathcal{N}(0, Q\Omega Q'),$$

where $Q = \mathrm{E}(Z_n X_n')^{-1}$ and $\Omega = \mathrm{Var}(Z_n[Y_n^{k_1} - Y_n^{k_2} - X_n'\beta_{k_1,k_2}])$. As before, by the sample analogue principle, I can construct a consistent estimator $\hat{Q}\hat{\Omega}\hat{Q}'$ for $Q\Omega Q'$, where I define:

$$\hat{Q} = \left(\frac{1}{N}\sum_{n=1}^{N} Z_n X_n'\right)^{-1} \quad \text{and} \quad \hat{\Omega} = \frac{1}{N}\sum_{n=1}^{N} Z_n [Y_n^{k_1} - Y_n^{k_2} - X_n'\hat{\beta}_{k_1,k_2}]^2 Z_n'.$$

B.3. Specifications for Monte Carlo Simulations

For the first set of Monte Carlo simulations (Table 2, Specification A), choices are simulated from a data generating process (DGP) with $K = 1$ and $\alpha_n = W_n' d$. Outcomes are given by:

$$\omega_i = \mathbb{1}\left\{h + W_n' d + J m_n + \varepsilon_i \geq 0\right\}, \quad \text{for } n = 1, \ldots, N,$$

[3]Specifically, estimator consistency holds since $\frac{1}{N}\sum_{n=1}^{N} Z_n X_n' \xrightarrow{p} \mathrm{E}(Z_n X_n')$ and $\frac{1}{N}\sum_{n=1}^{N} Z_n Y_n^k \xrightarrow{p} \mathrm{E}(Z_n Y_n^k)$ by the Weak Law of Large Numbers. Therefore, it follows from the continuous mapping theorem that $\hat{\beta}_k \xrightarrow{p} \beta_k$.

where $m_n = F_{\varepsilon|k}(h + W_n' d + J m_n)$ and W_n is evenly distributed over the interval $[-2, 2]$. I set the parameter values: $J = 0.5$, $h = 1$, and $d = 1$. I also assume that $\varepsilon_i \overset{\text{i.i.d.}}{\sim} \text{Logistic}(0, 1)$. Note that the same DGP is used to make Figure 1, which plots the bias of the OLS estimand.

For the second set of Monte Carlo simulations (Table 2, Specification B), choices are simulated from a DGP with $K = 2$ and α_n treated as unobservable. Outcomes are given by:

$$\omega_i = \mathbb{1}\left\{h_k + \alpha_n + J_{k1} m_n^1 + J_{k2} m_n^2 + \varepsilon_i \geq 0\right\}, \quad \text{for } n = 1, \ldots, N,$$

where $m_n^k = F_{\varepsilon|k}(h_k + \alpha_n + J_{k1} m_n^1 + J_{k2} m_n^2)$ for $k \in \{1, 2\}$, and α_n is evenly distributed over the interval $[-2, 2]$. I set $h_1 = 1$ and $h_2 = 0$, and I define $\mathbf{J}$ such that: $J_{11} = J_{22} = 1.5$ and $J_{12} = J_{21} = -1$. As in the first DGP, I assume that $\varepsilon_i | k \overset{\text{i.i.d.}}{\sim} \text{Logistic}(0, 1)$ for $k = \{1, 2\}$.

To simulate agents' choices ω_i in equilibrium, I run a fixed point iteration on the equilibrium condition (5). This procedure relies on the fact that $\mathbf{J}$ satisfies Assumption A, which ensures there is at least one locally stable equilibrium (Property 2). Note that, although Assumption A is not required for the validity of the estimation procedure, it is useful in simulations because it makes the equilibrium computation feasible through fixed point iteration.

APPENDIX C: ROBUSTNESS ANALYSES FOR PROJECT STAR

C.1. Examining Alternative Outcome Variables

Table A.2 reports IV estimates for alternative outcomes, where students' math and reading scores are discretized to indicate whether they scored in the top 25% or top 75% statewide. These estimates are statistically indistinguishable from those shown in Table 3 of the paper.

TABLE A.1

PROJECT STAR: IV ESTIMATES FOR ALTERNATIVE OUTCOME VARIABLES

	Math Score				Reading Score			
	Top 25%		Top 75%		Top 25%		Top 75%	
$h_m - h_f$	-0.088	(0.56)	0.625	(2.69)	0.060	(0.15)	0.211	(0.47)
$J_{mm} - J_{fm}$	5.373**	(2.73)	2.021	(11.26)	5.157***	(0.98)	4.738***	(0.76)
$J_{ff} - J_{mf}$	-5.446**	(2.45)	-2.351	(9.51)	-5.293***	(1.20)	-4.988***	(0.79)
Observations	5,798		5,798		5,718		5,718	
School FE	Yes		Yes		Yes		Yes	
1st Stage F ($\bar{\omega}_n^m$)	8.74		15.53		24.53		15.50	
1st Stage F ($\bar{\omega}_n^f$)	13.72		15.24		17.32		6.69	

Notes. This table presents estimates of $\hat{\beta}_{f,m}$ for two alternative outcomes: scoring in the top 25% and 75%. All the specifications remain the same as in the full-sample analysis. *p<0.1; **p<0.05; ***p<0.01.

C.2. Estimates for a Model with Uniform Peer Effects

In Table A.3, I examine how the estimates change if I impose uniform peer effects, $J_{k\ell} = J$. Under this restriction, I estimate that girls have lower private effort costs than boys ($h_f > h_m$), although the difference becomes statistically insignificant after I include school fixed effects. Note that these results differ greatly from those in the generalized model. Therefore, allowing for nonuniform interaction leads to meaningfully different economic conclusions.

TABLE A.2

PROJECT STAR: ESTIMATES FOR A MODEL WITH UNIFORM PEER EFFECTS

	Top 50% Math Score		Top 50% Reading Score	
	(1)	(2)	(1)	(2)
$h_m - h_f$	-0.363*** (0.09)	-1.199 (1.73)	-0.313*** (0.10)	-1.95 (1.69)
$J_{mm} - J_{fm}$	—	—	—	—
$J_{ff} - J_{mf}$	—	—	—	—
Observations	5,798	5,760	5,718	5,681
School FE	No	Yes	No	Yes

Notes. This table reports estimates of $h_m - h_f$ for a model that imposes uniform peer effects, i.e., $J_{k\ell} = J$. All the specifications remain the same as in the full-sample analysis. *p<0.1; **p<0.05; ***p<0.01.

C.3. Verifying Random Assignment to Classrooms

Under the experimental design, students within each school were randomly assigned to one of three class types. However, some schools had multiple classes of each type, which raises the possibility of nonrandom sorting within a given type. Although Graham (2008) argues that such sorting is unlikely, I conduct a robustness check where I estimate the model using only the schools that had one classroom of each type, thus ruling out any possibility of nonrandom assignment. These estimates have a weak first stage, particularly for math scores. However, I still find that they are statistically indistinguishable from the full-sample results.

C.4. Misspecification Tests

I conduct hypothesis tests to determine whether any of the type-specific parameters depend on observable class characteristics. Specifically, I test the null hypotheses that $J_{ff} - J_{mf}$, $J_{mm} - J_{fm}$, and $h_m - h_f$ (respectively) are different across any of the following criterion:

1. high poverty classrooms ($\geq 50\%$ FRPL) and low poverty classrooms ($< 50\%$ FRPL)
2. high minority classrooms ($< 75\%$ white) and low minority classrooms ($\geq 75\%$ white)
3. more experienced teachers (> 10 years) and less experienced teachers (≤ 10 years)
4. more educated teachers (graduate deg.) and less educated teachers (no graduate deg.)
5. rural classrooms (in rural district) and urban classrooms (in urban or suburban district)

TABLE A.3

PROJECT STAR: IV ESTIMATES FOR SCHOOLS WITH FEWER THAN 4 CLASSROOMS

	Top 50% Math Score		Top 50% Reading Score	
	(1)	(2)	(1)	(2)
$h_m - h_f$	0.819 (3.90)	-1.211 (9.40)	-0.089 (0.13)	-0.109 (0.49)
$J_{mm} - J_{fm}$	7.851 (16.93)	5.235 (4.18)	4.696*** (0.53)	4.172*** (0.76)
$J_{ff} - J_{mf}$	-9.676 (25.19)	-2.781 (13.27)	-4.539*** (0.76)	-3.933*** (1.32)
Observations	1,603	1,603	1,587	1,587
School FE	No	Yes	No	Yes
1st Stage F ($\bar{\omega}_n^m$)	6.18	6.41	18.99	19.63
1st Stage F ($\bar{\omega}_n^f$)	3.24	3.36	5.94	5.91

Notes. This table presents estimates of $\hat{\beta}_{f,m}$ for a restricted sample of schools with fewer than 4 classrooms. All the specifications remain the same as in the full-sample analysis. *p<0.1; **p<0.05; ***p<0.01.

The p-values from these hypothesis tests are shown in Tables A.2 and A.3. For both outcome variables, I find no evidence to reject the hypothesis that any type-specific parameters differ across networks. These results support the assumptions underlying the identification.

TABLE A.4

MISSPECIFICATION TESTS (*Outcome:* TOP 50% ON MATH EXAM)

	Large Share Poverty	Large Share Minority	Teacher Has >10yrs Experience	Teacher Has Higher Degree	Rural District
$h_m - h_f$	0.710	0.590	0.790	0.880	0.910
$J_{mm} - J_{fm}$	0.614	0.644	0.746	0.965	0.929
$J_{ff} - J_{mf}$	0.626	0.603	0.746	0.993	0.921

Notes. This table reports p-values corresponding to the one-dimensional hypothesis tests for whether $h_m - h_f$, $J_{ff} - J_{mf}$, and $J_{mm} - J_{fm}$ differ with respect to observed class characteristics.

C.5. Sensitivity of Estimates to the Partitioning Rule

When estimating the model, I define internal instruments by randomly splitting each classroom into two subsets, a and b, and using the outcomes in b as instruments for those in a. Although the resulting estimates will vary somewhat with the chosen partition, the estimator should remain consistent across all such partitions. To assess the sensitivity of the estimates to how the internal instruments are defined, I re-estimate the model $M = 1000$ times, each time drawing a new random partition, and I plot histograms of the resulting estimates.

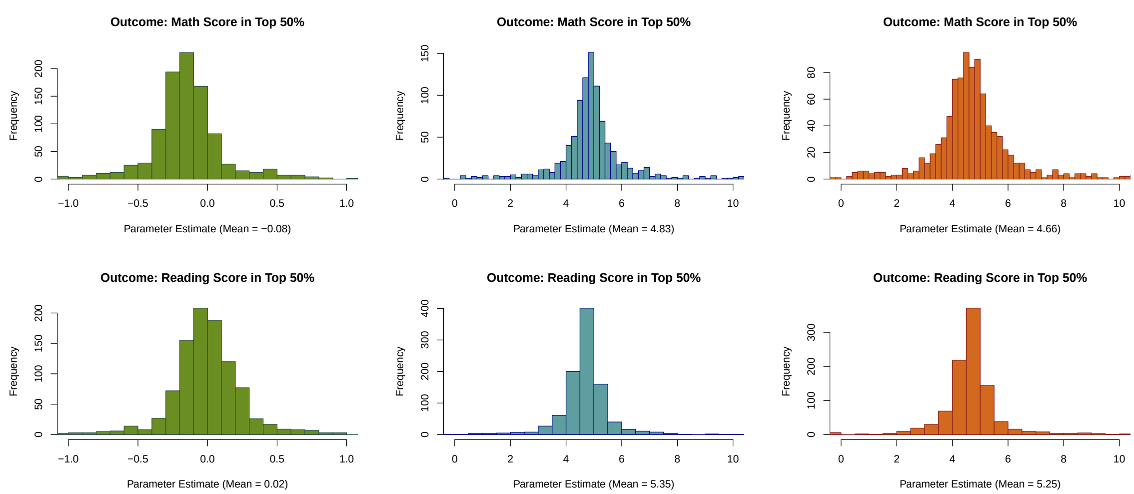
TABLE A.5

MISSPECIFICATION TESTS (*Outcome:* TOP 50% ON READING EXAM)

	Large Share Poverty	Large Share Minority	Teacher Has >10yrs Experience	Teacher Has Higher Degree	Rural District
$h_m - h_f$	0.996	0.611	0.760	0.983	0.834
$J_{mm} - J_{fm}$	0.442	0.756	0.798	0.987	0.523
$J_{ff} - J_{mf}$	0.622	0.711	0.790	0.990	0.659

Notes. This table reports p-values corresponding to the one-dimensional hypothesis tests for whether $h_m - h_f$, $J_{ff} - J_{mf}$, and $J_{mm} - J_{fm}$ differ with respect to observed class characteristics.

I report histograms of the coefficient estimates for each outcome variable in Figure A.1. Observe that the point estimates appear approximately normally distributed, and the amount of dispersion is not large enough to invalidate any of the qualitative findings in the Table 3. Moreover, the main IV estimates reported in the paper do not seem to be outliers, which indicates that many alternative partitions of classrooms would yield similar results. Thus, I interpret Figure A.1 as evidence that the main results in the empirical application are robust.



(a) Parameter: $h_m - h_f$ (b) Parameter: $J_{mm} - J_{fm}$ (c) Parameter: $J_{ff} - J_{mf}$

FIGURE A.1.—Project STAR: Histograms of IV Estimates across $M = 1000$ Random Partitions

APPENDIX D: ROBUSTNESS ANALYSES FOR PRIMARY SCHOOL DEWORMING

D.1. Differences in Treatment Uptake within Schools

Figure A.2 shows histograms of the within-school gaps in treatment uptake rates by student type. For any two types k_1 and k_2, the size and direction of these gaps can vary a lot across schools. This suggests that factors at the school-level shape disparities in treatment uptake.

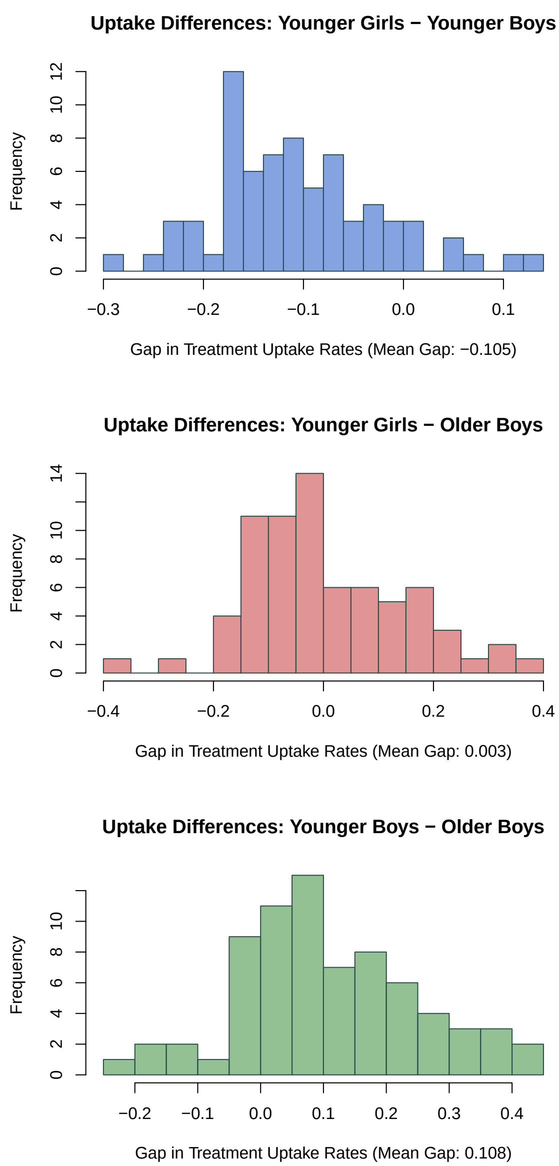


FIGURE A.2.—Distribution of Within-School Gaps in Treatment Uptake Rates by Student Type

D.2. IV Estimates for First Time Eligible Students

To better isolate contemporaneous interactions from pre-existing factors that can influence treatment decisions, I show a version of the IV estimates for a sample restricted to first time eligible students. The resulting estimates are largely consistent with the full-sample results.

TABLE A.6

DEWORMING IN SCHOOLS: IV ESTIMATES FOR FIRST TIME ELIGIBLE STUDENTS

	(1)		(2)		(3)		(4)	
Girls < 13 and Boys < 13								
$h_1 - h_2$	0.338	(1.31)	0.227	(1.41)	0.779	(2.19)	1.184	(2.16)
$J_{11} - J_{21}$	10.488***	(3.29)	10.777***	(3.38)	10.647***	(2.22)	10.973***	(1.81)
$J_{12} - J_{22}$	-10.527***	(2.57)	-11.105***	(2.63)	-13.414***	(2.77)	-14.942***	(2.62)
$J_{13} - J_{23}$	-0.402	(1.21)	-0.001	(1.23)	1.369	(1.46)	1.452	(1.56)
Girls < 13 and Boys $\geq$ 13								
$h_1 - h_3$	0.494	(1.00)	0.062	(0.95)	1.079	(1.09)	2.193**	(1.10)
$J_{11} - J_{31}$	5.215	(3.83)	6.338	(3.86)	5.645	(4.19)	5.226	(4.75)
$J_{12} - J_{32}$	-1.251	(3.51)	-3.497	(4.66)	-5.055	(6.07)	-5.898	(7.05)
$J_{13} - J_{33}$	-4.206***	(0.72)	-2.647*	(1.37)	-1.248	(1.57)	-1.086	(1.74)
Boys < 13 and Boys $\geq$ 13								
$h_2 - h_3$	0.156	(1.83)	-0.165	(1.67)	0.300	(1.71)	1.009	(2.20)
$J_{21} - J_{31}$	-5.273	(5.98)	-4.439	(5.97)	-5.003	(5.07)	-5.747	(6.08)
$J_{22} - J_{32}$	9.276*	(5.14)	7.608	(5.76)	8.359	(6.28)	9.043	(7.72)
$J_{23} - J_{33}$	-3.804***	(1.22)	-2.647	(1.70)	-2.617**	(1.31)	-2.538*	(1.31)
Year Fixed Effects	No		Yes		Yes		Yes	
Student Controls	No		No		Yes		Yes	
School Controls	No		No		No		Yes	
1st Stage F ($\bar{\omega}_n^1$)	9.32		6.97		12.64		10.34	
1st Stage F ($\bar{\omega}_n^2$)	10.02		8.05		12.08		9.74	
1st Stage F ($\bar{\omega}_n^3$)	32.17		13.42		11.08		10.55	

Notes. This table presents estimates of $\hat{\beta}_{k_1,k_2}$ for a restricted sample of students who are first time eligible. All specifications remain consistent with the full-sample analysis. *p<0.1; **p<0.05; ***p<0.01.

D.3. Misspecification Tests

I conduct hypothesis tests to determine whether any of the type-specific parameters depend on observable school characteristics. The school characteristics I consider are listed below:

1. treatment group (1998 group vs. 1999 group)
2. local infection rate (above vs. below median)
3. population density (above vs. below median)

The p-values from these hypothesis tests are reported in Tables A.7 and A.8, and I find no evidence to reject the hypothesis that any type-specific parameters differ across networks.

TABLE A.7

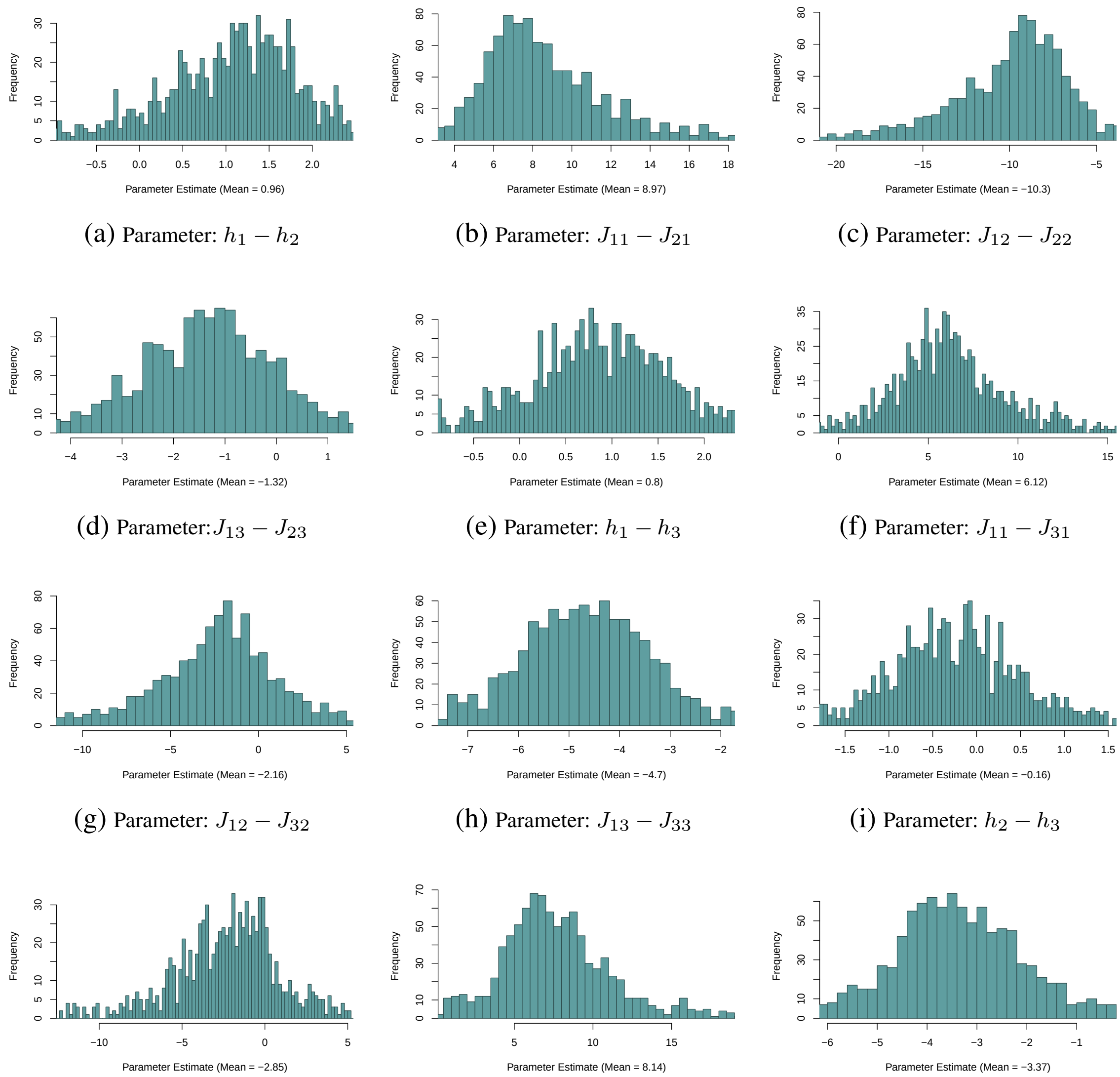
MISSPECIFICATION TESTS: PRIMARY SCHOOL DEWORMING

	Treatment Group	High Local Infection Rate	High Population Density (< 3km)	High Population Density (3-6km)
Girls < 13 and Boys < 13				
$h_1 - h_2$	0.954	0.914	0.053	0.748
$J_{11} - J_{21}$	0.304	0.722	0.583	0.231
$J_{12} - J_{22}$	0.120	0.772	0.381	0.122
$J_{13} - J_{23}$	0.983	0.735	0.852	0.650
Girls < 13 and Boys $\geq$ 13				
$h_1 - h_3$	0.637	0.831	0.886	0.727
$J_{11} - J_{31}$	0.158	0.806	0.277	0.281
$J_{12} - J_{32}$	0.049	0.799	0.395	0.053
$J_{13} - J_{33}$	0.348	0.754	0.746	0.715
Boys < 13 and Boys $\geq$ 13				
$h_2 - h_3$	0.343	0.751	0.367	0.999
$J_{21} - J_{31}$	0.580	0.934	0.682	0.886
$J_{22} - J_{32}$	0.866	0.866	0.993	0.709
$J_{23} - J_{33}$	0.278	0.835	0.795	0.433

Notes. This table reports p-values for one-dimensional hypothesis tests for whether $\{h_{k_1} - h_{k_2}\}$ and $\{J_{k_1\ell} - J_{k_2\ell}\}_{k_1,k_2,\ell}$ differ with respect to observed school characteristics, such as treatment group, above median local infection rate, and above median population density (both <3km and 3-6km from a school). In all cases, $h_k - h_\ell$ denotes $\frac{1}{T}\sum_{t=1}^{T}(h_{k,t} - h_{\ell,t})$. The regression specification includes year fixed effects but does not add student controls.

D.4. Sensitivity of Estimates to the Partitioning Rule

As in the first empirical application, I evaluate the sensitivity of the IV estimates to how the internal instruments are defined. To do so, I re-estimate the model $M = 1000$ times, each time drawing a new random partition, and then I plot histograms of the resulting estimates. Note that the baseline IV estimates reported in Table 4 do not appear to be outliers, and the amount of dispersion is not large enough to invalidate any of the qualitative findings in the paper. I interpret this finding as evidence that the baseline estimates in the paper are robust.



(a) Parameter: $h_1 - h_2$ (b) Parameter: $J_{11} - J_{21}$ (c) Parameter: $J_{12} - J_{22}$

(d) Parameter: $J_{13} - J_{23}$ (e) Parameter: $h_1 - h_3$ (f) Parameter: $J_{11} - J_{31}$

(g) Parameter: $J_{12} - J_{32}$ (h) Parameter: $J_{13} - J_{33}$ (i) Parameter: $h_2 - h_3$

(j) Parameter: $J_{21} - J_{31}$ (k) Parameter: $J_{22} - J_{32}$ (l) Parameter: $J_{23} - J_{33}$

FIGURE A.3.—Primary School Deworming: Histograms of IV Estimates across $M = 1000$ Random Partitions. *Note:* the estimation procedure follows the same specification used in Table 4, Column 4.

REFERENCES

ALLENDE, CLAUDIA (2019): "Competition Under Social Interactions and the Design of Education Policies," Working Paper. [5]

BRAMOULLÉ, YANN, RACHEL KRANTON, AND MARTIN D'AMOURS (2014): "Strategic Interaction and Networks," *American Economic Review*, 104 (3), 898–930. [3]

HORN, ROGER A. AND CHARLES R. JOHNSON (1985): *Matrix Analysis*, Cambridge University Press. [1]