Peers as treatments: Understanding contextual peer effects *

Brian Krauth
Simon Fraser University

August 29, 2025

Abstract

This paper develops a general framework for interpreting linear regression estimates of contextual peer effects under random peer assignment. Rather than imposing the strong assumption that peers influence individual outcomes solely and directly through specific observed characteristics, the model considers social interaction with a given peer or group as a treatment with an unknown and variable treatment effect. In this setting, a wide variety of conventional peer effect regressions are informative and can be interpreted as measuring treatment effect heterogeneity along researcher-selected dimensions of interest. These regressions can also be used to predict the consequences of counterfactual peer group assignments. The relevance of the framework and results to empirical research is illustrated using an application to measuring classroom peer effects in Project STAR.

JEL codes: C21, C31
Keywords: Peer effects, social interactions, contextual effects

1 Introduction

Researchers investigating social influences on choices or outcomes often employ a behavioral model associated with[1] Manski (1993) in which an individual's outcome

*Email: bkrauth@sfu.ca. Revisions available at bvkrauth.github.io/publication/peertreat

[1] In Manski (1993), behavior responds to the conditional expectation of peer behavior and characteristics, but in most subsequent empirical work it is taken to respond to their observed values. Blume et al. (2011, p. 891-892) discuss this distinction and some of its implications.

depends on outcomes (endogenous effects) and characteristics (contextual effects) of others in some peer or reference group. Manski's formulation has inspired the development of methods for modeling endogenous effects and empirically distinguishing them from both contextual effects and endogenous peer selection. These methods have demanding data requirements, so applied researchers often estimate reduced form models that contain only contextual effects. Despite their central role in empirical research, the modeling of contextual effects has seen limited formal attention, resulting in a wide variety of ad hoc model specifications whose causal interpretation is unclear.

To simplify, a representative[2] empirical study considers the regression model:

$$y_i = \mathbf{x}_i \alpha_1 + \bar{\mathbf{x}}_i \alpha_2 + \epsilon_i \tag{1}$$

where y_i is an outcome, $\mathbf{x}_i$ is a vector of individual characteristics, and $\bar{\mathbf{x}}_i$ is the average characteristics of their peers. Peers are fellow members of some social group like classmates or co-workers, and all individual-level variables are predetermined characteristics like gender and family background rather than treatments that can be directly assigned by a policy maker. The coefficient α_2 is interpreted as the effect of $\bar{\mathbf{x}}_i$ — any intervention that changes $\bar{\mathbf{x}}_i$ by $\Delta\bar{\mathbf{x}}$ changes y_i by $\Delta\bar{\mathbf{x}}\alpha_2$ — and the case for this causal interpretation relies on some random mechanism that assigns peers. This model's simplicity is appealing but leaves important practical and conceptual issues to address.

The key practical issue is that researchers must choose which characteristics to include and whether to add interaction terms or nonlinearities, in a setting with limited data and high costs of model complexity. Should researchers aim to estimate complete models with all potentially-relevant characteristics, or parsimonious models that only include characteristics related to their research question?

The key conceptual issue is that the predetermined characteristics of peers cannot be directly and individually manipulated by a policy maker. Changes to one peer characteristic can only be implemented by moving specific individuals between groups, which may induce changes in many other characteristics. For example, reducing the number of boys in a classroom of fixed size requires replacing a specific boy with a specific girl. These two students may differ in motivation, disruptive behavior, academic ability, etc., as well as gender. As a result, the effect of a change to $\bar{\mathbf{x}}_i$ is not constant and depends on how individuals are selected to implement the change. Given this heterogeneity, what counterfactual policies can be evaluated using a research design based on random peer assignment?

[2]The benchmark "linear in means" specification is used here for illustration; the issues discussed here also apply to more flexible specifications that incorporate nonlinearities or interaction terms.

This paper develops an explicit potential outcomes framework to address these questions. The underlying model is nonparametric, and takes social interaction with a specific and potentially unique individual as the treatment of interest. The main results emphasize the use of simple linear regression models to measure heterogeneity of the associated treatment effect across identifiable sub-populations. For example, the coefficient α_2 in the linear in means model (1) can often be interpreted as the effect of replacing a randomly selected peer from one sub-population with a randomly selected peer from another sub-population. This interpretation holds under straightforward and testable conditions, and does not require researchers to estimate complete models. Richer regression models can be used to relax key assumptions, and to make additional counterfactual inferences on the effect of replacing an individual's entire peer group, or of reallocating peer groups across the entire population. Throughout the paper, emphasis is placed on using the simplest regression model that can answer a given research question. An illustrative application to estimating classroom peer effects in Project STAR (Bietenbeck, 2025) shows how the results provide guidance on specification and variable choice, and facilitate the interpretation of results from different specifications within a common conceptual framework and set of identifying assumptions.

1.1 Related literature

A traditional alternative to the potential outcomes framework is to interpret a model like equation (1) as the complete reduced form of the structural production function for outcomes, include as many potentially-relevant peer characteristics as the data allow, and hope that there are no relevant omitted variables. The limitations of this method for causal inference are well known, but are particularly acute in this setting. Every characteristic that affects one's own outcome also affects peer outcomes through the endogenous effect. If the endogenous effect is nonzero, the set of potentially-relevant peer characteristics in the reduced form is identical to the virtually unlimited set of potentially-relevant own characteristics. Unfortunately, limited sample sizes and exploitable variation in peer characteristics in the available data sets make it impractical to estimate a plausibly complete reduced form model.

Most papers that apply a potential outcomes framework to contextual effects address the measurement of treatment effects with interference/spillovers (Manski, 2013; Aronow and Samii, 2017). In that literature, the peer group is predetermined and the counterfactual of interest is an individual-level variable representing some treatment assignment. This contrasts sharply with the setting considered in this paper, where all individual-level variables are predetermined and the counterfactual of interest is an assignment of individuals to different peer groups.

Three recent papers also consider peer assignment in a potential outcomes framework. Li et al. (2019) and Basse et al. (2024) emphasize the complete model case (in the language of this literature, $\bar{x}_i$ is a correctly specified exposure mapping for the social network of individual i), though they have some results that relax this assumption. Graham et al. (2025) do not assume model completeness and treat observed peer characteristics as imperfect proxies for unobserved peer characteristics. The analysis in this paper is complementary to theirs, but emphasizes variable selection, specification choice, and what can be learned from simple parametric regression models. In addition, I consider a richer characteristics space and model observed characteristics as a function of unobserved heterogeneity rather than as orthogonal to it. These two modeling approaches are substantively equivalent (each can be mapped to the other by redefining variables), but the formulation here helps to separate practical specification questions from assumptions about causal mechanisms.

Finally, this paper is among several that use estimated peer effect models to predict the consequences of population-level peer group reallocations. Bhattacharya (2009) develops algorithms to find optimal assignments from model estimates. A field experiment by Carrell et al. (2013) shows a key limitation of this approach: peer effect estimates from one cohort of students were used to construct presumably optimal allocations for a later cohort, but the "optimal" allocation yielded poor results because it was off the support of the original data and thus highly sensitive to misspecification. Graham et al. (2025) also note that the large changes needed to reach an optimal group assignment are typically infeasible and emphasize tools for predicting the marginal effect of smaller and more feasible reallocations.

2 Classroom peer effects in Project STAR

The model and results are illustrated in a running example based on Bietenbeck (2025), who uses data from the Project STAR class size experiment to estimate classroom peer effects. This data set has been repeatedly used for this purpose because its experimental design randomly assigned students to classrooms. Bietenbeck (2025) estimates contextual effects for classmate motivation, while other studies use Project STAR to estimate endogenous effects (Boozer and Cacciola, 2001; Graham, 2008; Rose, 2017) and contextual effects for classmate gender (Whitmore, 2005; Graham et al., 2025), age (Cascio and Schanzenbach, 2007), economic disadvantage (Chetty et al., 2011), lagged achievement (Sojourner, 2013), and grade repeating (Bietenbeck, 2019).

Collectively, these studies exhibit several issues raised in the introduction:

- Multiple researchers have measured contextual effects for the same outcomes in the same data set, each using a different set of predetermined peer characteristics.
- These characteristics are correlated at the individual level, and are likely to be correlated with other outcome-relevant unobserved/omitted characteristics.
- There is evidence of endogenous effects in the same outcomes and data set, implying that the set of relevant peer characteristics in the reduced form is large.
- Random assignment produces too little variation in composition (Chetty et al., 2011) to estimate models that are rich enough to be plausibly complete.

The running example will illustrate how these earlier studies can be interpreted within the peers-as-treatments model, and the implications of different specification choices.

The analysis is based on Bietenbeck (2025) and its replication package (Bietenbeck, 2024). Following that paper's methodological choices, reading and math test scores are the main outcomes, the estimation sample is defined as students who have newly entered in grade 2 or 3, the peer group is defined as returning students in the entry grade classroom, and school-by-entry-grade fixed effects are included in all regressions to account for non-random assignment to schools. Although Project STAR classes vary in size, it will be expositionally convenient to assume that each class has exactly 16 students. Appendix A provides additional methodological details.

3 Model

This section develops the model, defining the general framework in Section 3.1, key maintained assumptions in Section 3.2, and additional conditions in Section 3.3.

3.1 Framework and notation

The model features a population of heterogeneous **individuals** arbitrarily indexed by $i \in \mathcal{I} \equiv \{1, 2, \ldots, I\}$. Each individual is characterized by an **unobserved type** $\tau_i \in \mathcal{T}$ and membership in some social **group** $g_i \in \mathcal{G} \equiv \{1, 2, \ldots, G\}$. The population as a whole is fully characterized by the random matrices $\mathbf{T} \in \mathcal{T}^I$ and $\mathbf{G} \in \mathcal{G}^I$, in the sense that all variables in the model are functions of $(\mathbf{T}, \mathbf{G})$.

The unobserved type has two essential features: it provides a complete description of everything about the individual that is potentially relevant, and it is predetermined[3]

[3] As discussed in Section 1.1, this feature distinguishes this paper from the literature on treatment effects with spillovers (Manski, 2013; Aronow and Samii, 2017).

in the sense that it cannot be modified by a policymaker. The type space $\mathcal{T}$ is defined abstractly, but one can interpret τ_i as representing a vector of characteristics.

Unlike unobserved types, group membership can be modified by a policymaker. Its observed value is determined by a **group assignment mechanism** that is represented by a discrete conditional PDF $f_{\mathbf{G}|\mathbf{T}} : \mathcal{G}^I \times \mathcal{T}^I \to [0, 1]$ such that:

$$f_{\mathbf{G}|\mathbf{T}}(\mathbf{G_0}, \mathbf{T_0}) \equiv \Pr(\mathbf{G} = \mathbf{G_0}|\mathbf{T} = \mathbf{T_0}) \tag{2}$$

All causal inferences in this paper relate to the consequences of counterfactual group assignments or mechanisms. Given a group assignment $\mathbf{G}$, group g has size:

$$n_g \equiv n(g, \mathbf{G}) \equiv \sum_{i=1}^{I} \mathbb{I}\left(g_i = g\right) \tag{3}$$

with support $\mathbb{S}_n \subset \mathcal{I} \cup \{0\}$, and individual i's **peer group** is:

$$\mathbf{p}_i \equiv \mathbf{p}(i, \mathbf{G}) \equiv \{j \neq i : g_j = g_i\} \tag{4}$$

and has size $|\mathbf{p}_i| = n_{g_i} - 1$.

Each individual experiences a scalar **outcome** of interest $y_i \in \mathbb{R}$ that depends on both their own type and that of other group members:

$$\mathbf{Y} \equiv \begin{bmatrix} y_1 \\ \vdots \\ y_I \end{bmatrix} \equiv \begin{bmatrix} y_1\left(\mathbf{T}, \mathbf{G}\right) \\ \vdots \\ y_I\left(\mathbf{T}, \mathbf{G}\right) \end{bmatrix} \equiv \mathbf{Y}(\mathbf{T}, \mathbf{G}) \tag{5}$$

The model does not include a direct causal effect of peer outcomes ("endogenous effects" in the language of Manski 1993) but can be interpreted as the reduced form of such a model. For ease of exposition, the outcome is a deterministic function of types and group assignments. Appendix B.2 shows that post-assignment shocks can be accommodated.

For each individual i, we observe the outcome y_i, the peer group g_i, and a vector of observed **characteristics of interest** $\mathbf{x}_i \equiv (x_{i1}, \ldots, x_{iK}) \in \mathbb{R}^K$ that depend on one's own type:

$$\mathbf{X} \equiv \begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_I \end{bmatrix} \equiv \begin{bmatrix} \mathbf{x}(\tau_1) \\ \vdots \\ \mathbf{x}(\tau_I) \end{bmatrix} \equiv \mathbf{X}(\mathbf{T}) \tag{6}$$

Like $\mathbf{T}$, $\mathbf{X}$ is predetermined. The observed characteristics in $\mathbf{X}$ do not appear directly in the outcome model (5) but $\mathbf{x}_i$ could be a proxy, summary, or subset of the full set of potentially relevant characteristics represented by τ_i. This is a key feature of this model: the observed characteristics in are not assumed to play any specific role in the "true" causal model, but rather have been chosen by the researcher based on data availability and their research question. The researcher's choice of characteristics remains critical because it determines which research questions can be addressed.

Given $(\mathbf{X}, \mathbf{G})$, **peer characteristics** for individual i are the multiset[4] $\{\mathbf{x}_j\}_{j \in \mathbf{p}_i}$ and **peer average characteristics** are the vector $\bar{\mathbf{x}}_i \equiv (\bar{x}_{i1}, \ldots, \bar{x}_{iK}) \in \mathbb{R}^K$ where:

$$\bar{\mathbf{x}}_i \equiv \bar{\mathbf{x}}\left(\{\mathbf{x}_j\}_{j \in \mathbf{p}_i}\right) \equiv \frac{1}{|\mathbf{p}_i|} \sum_{j \in \mathbf{p}_i} \mathbf{x}_j \tag{7}$$

For those who have no peers ($\mathbf{p}_i = \emptyset$ or $n_{g_i} = 1$), $\bar{\mathbf{x}}_i$ can be left undefined or set to zero.

Example 1 (Classmate effects on test scores)**.** *Section 2 describes several papers that estimate the effect of various classmate characteristics on test scores in Project STAR data. Abstracting from methodological details, these studies fit in the framework:*

$$\begin{aligned}
y_i &\equiv \textit{student } i\textit{'s test score} \\
g_i &\equiv \textit{classroom ID for student } i \\
\tau_i &\equiv \textit{all relevant characteristics of student } i \\
\mathbf{x}_i &\equiv \textit{researcher-selected characteristics of student } i \\
\{\mathbf{x}_j\}_{j \in \mathbf{p}_i} &\equiv \textit{researcher-selected characteristics of student } i\textit{'s classmates} \\
\bar{\mathbf{x}}_i &\equiv \textit{average characteristics of student } i\textit{'s classmates}
\end{aligned}$$

where the characteristics in $\mathbf{x}_i$ vary across studies but those represented by τ_i do not.

3.2 Maintained assumptions

The basic assumptions defined in this section will be maintained throughout the analysis.

Assumption 1 (Outcome model)**.** *The outcome for individual i is:*

$$y_i(\mathbf{T}, \mathbf{G}) = y\left(\tau_i, \{\tau_j\}_{g_i = g_j}\right) \tag{8}$$

where $y : \mathcal{T} \times \mathcal{M}_{\mathcal{T}} \to \mathbb{R}$ is an unknown function and $\mathcal{M}_{\mathcal{T}}$ is the set of multisets on $\mathcal{T}$

[4]A multiset is a set that can have repeated elements. Formally, it is a pair (S, m) where S is the underlying set and $m : S \to \mathbb{N}$ is a function giving the number of times each element of S appears. In the interest of readability, ordinary set notation, terminology, and operators are used whenever their generalization to multisets is straightforward.

Assumption 1 describes the complete outcome model at the individual level, which depends on the full set of own characteristics (represented by τ_i) and peer characteristics (represented by $\{\tau_j\}_{g_i=g_j}$). The complete model cannot be estimated directly since τ_i is unobserved, but it will be used to describe causal inferences that are feasible with observed data. The outcome model is nonparametric but imposes some restrictions: no cross-group spillovers, anonymous/exchangeable within-group spillovers, and no direct effects of group assignment itself. Direct effects of group assignment or a more general social network can in principle be accommodated in this model, but would require additional structure and assumptions that are beyond the scope of this paper.

Assumption 2 (Finite type space)**.** *The type space has finite size* $T \equiv |\mathcal{T}|$.

Assumption 2 limits mathematical complexity by allowing the use of elementary probability theory. It also allows types to be represented as finite scalars:

$$\mathcal{T} \equiv \{1, 2, \ldots, T\} \tag{9}$$

by assigning an integer to each unique value of the original type space. The ordering of types in (9) is arbitrary and need not convey ranking or similarity information, and the number of types can be large to represent a rich underlying characteristics space.

Assumption 3 (Independent types)**.** *Each individual's type is an independent draw from a common type distribution:*

$$\Pr(\mathbf{T} = \mathbf{T_0}) = \prod_{i=1}^{I} f_\tau(\tau_i(\mathbf{T_0})) \tag{10}$$

where $f_\tau : \mathcal{T} \rightarrow [0, 1]$ *is some unknown discrete PDF.*

Assumption 3 is innocuous: the indexing of individuals is arbitrary, so independence is supported by standard exchangeability arguments. This *unconditional* independence does not imply *conditional* independence of types given $(\mathbf{X}, \mathbf{G})$.

Assumption 4 (Rank condition)**.** $E(\mathbf{d}_i'\mathbf{d}_i)$ *is full rank where* $\mathbf{d}_i \equiv vec(1, \mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i'\bar{\mathbf{x}}_i)$.

Assumption 4 is the standard rank condition needed for identification of relevant regression (best linear predictor) coefficients from the joint distribution of $(y_i, \mathbf{x}_i, \bar{\mathbf{x}}_i)$.

Assumption 5 (Base type)**.** *The support of* $\mathbf{x}_i$ *includes* $\mathbf{e}_0$*, a* K*-vector of zeros.*

Assumption 5 is just a convenient normalization that simplifies some definitions.

Assumption 6 (Constant group size)**.** *Each peer group in* $\mathbf{G}$ *has exactly* n_0 *members:*

$$n_g = n_0 \qquad \forall g \in \mathcal{G} \tag{11}$$

where $n_0 \geq 2$ *is an integer and* $I = n_0 G$*.*

Assumption 6 is a standard assumption that simplifies exposition. Variable group size is typically handled in applied work by parametric assumptions, but can be accommodated nonparametrically in this setting by including group size as a conditioning/explanatory variable. See Appendix B.2 for details.

3.3 Additional conditions

The conditions defined in this section are *not* maintained as assumptions throughout the paper, but will play a key role in specific results.

Definition 1 (Simple random assignment)**.** *The group assignment mechanism* $f_{\mathbf{G}|\mathbf{T}}$ *satisfies* **simple random assignment (RA)** *if:*

$$\mathbf{G} \perp\!\!\!\perp \mathbf{T} \tag{\textbf{RA}}$$

i.e., peer group assignment does not depend on one's unobserved type or any other predetermined characteristics.

Definition 2 (Stratified random assignment)**.** *The group assignment mechanism* $f_{\mathbf{G}|\mathbf{T}}$ *satisfies* **stratified random assignment (SA)** *based on observed characteristics if:*

$$\mathbf{G} \perp\!\!\!\perp \mathbf{T}|\mathbf{X} \tag{\textbf{SA}}$$

i.e., peer group assignment may depend on one's observed characteristics but does not otherwise depend on one's unobserved type.

Causal inference on changes to peer group assignments will typically require some form of random group assignment. An important difference between these two forms of random assignment is that simple random assignment does not constrain the researcher's choice of characteristics to include in $\mathbf{x}_i$. In contrast, stratified random assignment requires $\mathbf{x}_i$ to include all characteristics used in stratification.

Definition 3 (Peer separability)**.** *Given Assumption 1, outcomes are* **peer-separable (PS)** *if the effect of replacing one peer with another does not depend on other peers:*

$$y\left(a, \left\{b'\right\} \cup \boldsymbol{\tau}\right) - y\left(a, \{b\} \cup \boldsymbol{\tau}\right) = y\left(a, \left\{b'\right\} \cup \boldsymbol{\tau}'\right) - y\left(a, \{b\} \cup \boldsymbol{\tau}'\right) \tag{\textbf{PS}}$$

for any $a, b, b' \in \mathcal{T}$ *and* $\boldsymbol{\tau}, \boldsymbol{\tau}' \in M_{\mathcal{T}}$ *such that* $|\boldsymbol{\tau}| = |\boldsymbol{\tau}'|$.

Definition 4 (Own separability)**.** *Given Assumption 1, outcomes are* **own-separable (OS)** *if the effect of changing peers does not depend on one's own type:*

$$y\left(a, \boldsymbol{\tau}'\right) - y\left(a, \boldsymbol{\tau}\right) = y\left(a', \boldsymbol{\tau}'\right) - y\left(a', \boldsymbol{\tau}\right) \quad (\mathbf{OS})$$

for any $a, a' \in \mathcal{T}$ *and* $\boldsymbol{\tau}, \boldsymbol{\tau}' \in M_{\mathcal{T}}$.

Outcomes that are neither own-separable nor peer-separable are **non-separable**. Both forms of separability are properties of the complete outcome function, and do not depend on the specific characteristics in $\mathbf{x}_i$. Separability simplifies the analysis but is not required for identification.

Definition 5 (Discrete characteristics)**.** *The researcher has* **discrete characteristics of interest (DC)** *if the support of* $\mathbf{x}_i$ *is:*

$$\mathbb{S}_{\mathbf{x}} \equiv \{\mathbf{e}_0, \mathbf{e}_1, \ldots, \mathbf{e}_K\} \quad (\mathbf{DC})$$

where $\mathbf{e}_k \in \{0,1\}^K$ *is the unit vector of length* $K \geq 1$ *containing one in column* k *and zero elsewhere; and its probability distribution*[5] *is fully described by:*

$$\mu_k \equiv \Pr(\mathbf{x}_i = \mathbf{e}_k) \quad \text{(for all } k \in 0, 1, \ldots, K)$$
$$\boldsymbol{\mu} \equiv E(\mathbf{x}_i) = \begin{bmatrix} \mu_1 & \cdots & \mu_K \end{bmatrix} \quad (12)$$

The main identification results in Sections 5.2 through 5.4 take discrete characteristics as given. Discrete characteristics facilitate the separation of causal inference and functional form issues because all functions of discrete $\mathbf{x}_i$ are linear and $(n_{g_i}, \bar{\mathbf{x}}_i)$ fully describes any $\{\mathbf{x}_j\}_{j \in \mathbf{p}_i}$ for discrete $\mathbf{x}_i$.

In many applications, the characteristics of interest are naturally discrete. Continuous characteristics can also be discretized when appropriate to the research question. Identification does not depend on whether the discretization is coarse (small K) or fine (large K), but there is a trade-off between model complexity (K) and precision in a finite sample. A natural data-driven approach to discretizing a continuous characteristic is to divide evenly by quantile (as in Example 2 below), with the number of levels chosen to minimize out of sample forecast error (estimated by cross-validation) or some information criterion. Tree-based approaches that vary both the number of levels and their placement may also be useful. Discretization implies information loss, so functional

[5] Note that $\mu_0 = 1 - \sum_{k=1}^{K} \mu_k$ is not included in the vector $\boldsymbol{\mu}$ but can be expressed as a function of it.

form assumptions may provide a better approximation to the researcher's ideal model. Section 5.5 develops a sieve-based framework for considering alternative approximations including discrete, polynomial, and spline.

Example 2 (Classmate gender and motivation)**.** *Whitmore (2005) and Graham et al. (2025) measure the effect of classmate gender ($K = 1$) which can be expressed as:*

$$\mathbf{x}_i \equiv [x_{i1}] \equiv \begin{cases} 1 & \textit{if student } i \textit{ is male} \\ 0 & \textit{if student } i \textit{ is female} \end{cases} \tag{13}$$

Bietenbeck (2025) measures the effect of classmate motivation, which is continuous but can be discretized. Table 7 in Bietenbeck (2025) reports results for regressions in which motivation is discretized by tercile ($K = 2$):

$$\mathbf{x}_i \equiv \begin{bmatrix} x_{i1} & x_{i2} \end{bmatrix} \equiv \begin{cases} \begin{bmatrix} 1 & 0 \end{bmatrix} & \textit{if low motivation (below 33rd percentile)} \\ \begin{bmatrix} 0 & 1 \end{bmatrix} & \textit{if high motivation (above 66th percentile)} \\ \begin{bmatrix} 0 & 0 \end{bmatrix} & \textit{otherwise} \end{cases} \tag{14}$$

Both of these choices satisfy (DC).

4 Defining social effects

Causal social effects can be defined in this setting by stating an explicit potential outcome function and the counterfactual outcomes of interest. As discussed in Section 1.1, characteristics are predetermined, so the applicable counterfactuals in this model relate to the peer group assignment, and not to the characteristics of any specific individual.

Definition 6 (Potential outcomes)**.** *Given Assumption 1, individual i's* **potential outcome function** *is:*

$$y_i(\mathbf{p}) \equiv y\left(\tau_i, \{\tau_j\}_{j\in\mathbf{p}}\right) \tag{15}$$

where $\mathbf{p}$ is any subset of $\mathcal{I} \setminus \{i\}$.

That is, the observed outcome for individual i is $y_i(\mathbf{p}_i)$, and the counterfactual outcome $y_i(\mathbf{p})$ is the outcome that would have occurred if i had been assigned peer group $\mathbf{p}$. The multiset $\{\tau_j\}_{j\in\mathbf{p}}$ serves as an exposure mapping (Aronow and Samii, 2017; Basse et al., 2024): a summary of the full social environment $(\mathbf{T}, \mathbf{G})$ that is sufficient to determine potential outcomes for individual i. The traditional model completeness assumption imposes the more restrictive exposure mapping $\{\mathbf{x}(\tau_j)\}_{j\in\mathbf{p}}$.

Counterfactual peer group assignments can be described as the effect of changing a single peer (**peer effects**), an entire peer group (**group effects**), or the peer group assignment mechanism itself (**reallocation effects**). These effects can be averaged over the population (**average effects**) or conditioned on the characteristics of the treated individual (**conditional effects**). The rest of this section defines each of these effect types in terms of the potential outcome function defined in (15).

4.1 Peer effects

Peer effects predict the effect of replacing a single peer with another peer who has different observed characteristics.

Definition 7 (Average peer effect)**.** *Given Assumptions 1, 5, and 6, the* **average peer effect** *of peers with characteristics* $\mathbf{x}^p \in \mathbb{S}_{\mathbf{x}}$ *is:*

$$APE(\mathbf{x}^p) \equiv E\left(y_i(\{j\} \cup \tilde{\mathbf{p}}) - y_i(\{j'\} \cup \tilde{\mathbf{p}})\middle|\, \mathbf{x}_j = \mathbf{x}^p, \mathbf{x}_{j'} = \mathbf{e}_0\right) \tag{16}$$

where $\tilde{\mathbf{p}}$ *is a purely random draw of* $n_0 - 2$ *peers from* $\mathcal{I} \setminus \{i, j, j'\}$*. When* $\mathbf{x}_i$ *is discrete (DC), the average peer effect of peers of observed type* ℓ *is:*

$$APE_\ell \equiv APE(\mathbf{e}_\ell) \tag{17}$$

where $APE(\cdot)$ *is as defined in equation (16).*

The average peer effect is similar to the "average spillover effect" in Graham et al. (2025). Although the definition looks complex, the concept is simple. Take a representative (random) individual (i) with a random peer group ($\{j'\} \cup \tilde{\mathbf{p}}$) that includes a peer ($j'$) of the base observed type ($\mathbf{x}_{j'} = \mathbf{e}_0$). Replace that peer with a random peer (j) of observed type $\mathbf{x}^p$ ($\mathbf{x}_j = \mathbf{x}^p$). The average peer effect $APE(\mathbf{x}^p)$ is the predicted change in this individual's outcome. The base type has $APE_0 = APE(\mathbf{e}_0) = 0$ by construction, and can be chosen for convenience without limiting the available comparisons.

Average peer effects are causal — the expected difference between two potential outcomes — but do not generally represent the causal impact of the peer characteristic itself. Instead, interaction with a specific individual is the treatment of interest, and average peer effects describe how the associated treatment effect varies across researcher-selected sub-populations. An analogy would be when a labour economist separately estimates the elasticity of labour supply for workers with and without children. The elasticity itself is causal, but its heterogeneity across the two sub-populations is not necesssarily the causal effect of having children on labour supply.

Rather than averaging across all treated individuals, researchers may also be interested in how peer effects vary with the treated individual's characteristics of interest.

Definition 8 (Conditional peer effect)**.** *Given Assumptions 1, 5, and 6, the* **conditional peer effect** *of peers with characteristics* $\mathbf{x}^p \in \mathbb{S}_{\mathbf{x}}$ *on treated individuals with characteristics* $\mathbf{x}^o \in \mathbb{S}_{\mathbf{x}}$ *is:*

$$CPE(\mathbf{x}^o, \mathbf{x}^p) \equiv E\left(y_i\left(\{j\} \cup \tilde{\mathbf{p}}\right) - y_i\left(\left\{j'\right\} \cup \tilde{\mathbf{p}}\right) \middle| \mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p, \mathbf{x}_{j'} = \mathbf{e}_0\right) \tag{18}$$

where $\tilde{\mathbf{p}}$ *is a purely random draw of* $n_0 - 2$ *peers from* $\mathcal{I} \setminus \{i, j, j'\}$*. When* $\mathbf{x}_i$ *is discrete (DC), the conditional peer effect of peers of observed type* ℓ *on treated individuals of observed type* k *is:*

$$CPE_{k\ell} \equiv CPE(\mathbf{e}_k, \mathbf{e}_\ell) \tag{19}$$

where $CPE(\cdot, \cdot)$ *is as defined in equation (18).*

Average and conditional peer effects are well-defined under the model's maintained assumptions without requiring additional conditions like random assignment, separability, or discrete characteristics. However, these additional conditions may play an important role in identification and interpretation.

4.2 Group effects

Peer **group effects** predict the effect of replacing the entire peer group. This effect differs from a simple aggregation of individual peer effects if outcomes are non-separable. For example, a low-motivation classmate may be more disruptive if there are other low-motivation students in the classroom, or classroom social dynamics may change abruptly when boys outnumber girls (Hoxby and Weingarth, 2005).

Group effects can be different for every value in the support of $\{\mathbf{x}_j\}_{j \in \mathbf{p}_i}$, denoted by $\mathbb{S}_{\{\mathbf{x}\}}$. This support can be large, but binning can be used for empirical tractability.

Definition 9 (Binned peer group variable)**.** *Given Assumption 6, let the* **binned peer group variable** $\mathbf{z}_i \in \{0, 1\}^B$ *be defined by the* **binning scheme** $\mathbf{z}(\cdot)$*:*

$$\mathbf{z}_i \equiv \mathbf{z}\left(\{\mathbf{x}_j\}_{j \in \mathbf{p}_i}\right) \equiv \sum_{b=1}^{B} \mathbf{e}_b \mathbb{I}\left(\{\mathbf{x}_j\}_{j \in \mathbf{p}_i} \in \mathbb{S}^b_{\{\mathbf{x}\}}\right) \tag{20}$$

where $\left(\mathbb{S}^0_{\{\mathbf{x}\}}, \mathbb{S}^1_{\{\mathbf{x}\}}, \ldots, \mathbb{S}^B_{\{\mathbf{x}\}}\right)$ *is a partition of* $\mathbb{S}_{\{\mathbf{x}\}}$ *into* $B + 1$ **bins***, and* $\mathbf{e}_b$ *is the unit vector of length* B *containing one in column* b *and zero elsewhere. Bin* b *is a* **singleton** *if* $\left|\mathbb{S}^b_{\{\mathbf{x}\}}\right| = 1$ *and* **pooled** *if* $\left|\mathbb{S}^b_{\{\mathbf{x}\}}\right| > 1$*.*

The binning scheme $\mathbf{z}(\cdot)$ is chosen by the researcher to reflect their research question, given the available variation in the data. As with discretization, the identification results in Section 5 apply to any binning scheme but there is a trade-off between model complexity and precision in a finite sample. The data-driven approaches to discretization discussed in Section 3.2 can also be applied to binning.

Example 3 (A binned variable for classmate gender)**.** *A researcher interested in the effect of having an unusually high proportion of male or female classmates may bin peer groups into mixed, female-dominated, and male-dominated ($B = 2$):*

$$\mathbf{z}_i = \mathbf{z}\left(\{\mathbf{x}_j\}_{j \in \mathbf{p}_i}\right) = \begin{cases} [\ 0 \quad 0\ ] & \text{if } 0.43 \leq \bar{\mathbf{x}}_i \leq 0.57 \\ [\ 1 \quad 0\ ] & \text{if } \bar{\mathbf{x}}_i < 0.43 \\ [\ 0 \quad 1\ ] & \text{if } \bar{\mathbf{x}}_i > 0.57 \end{cases} \tag{21}$$

where $\bar{\mathbf{x}}_i$ is the male share among classmates. The specific cutoffs are roughly the 25th and 75th percentiles of $\bar{\mathbf{x}}_i$ in the Project STAR data, and are chosen to reflect the research question of the effect of a "male-dominated" or "female-dominated" peer group.

Group effects can be defined for every value in $\mathbb{S}_{\{\mathbf{x}\}}$ or for each bin.

Definition 10 (Average group effect)**.** *Given Assumptions 1 and 6, the* **average group effect** *of a peer group with characteristics $\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}$ is:*

$$AGE(\mathbf{x}^p) \equiv E\left(y_i(\tilde{\mathbf{p}}) - y_i(\tilde{\mathbf{q}}) \,\middle|\, \{\mathbf{x}_j\}_{j \in \tilde{\mathbf{p}}} = \mathbf{x}^p, \mathbf{x}_j = \mathbf{e}_0 \text{ for all } j \in \tilde{\mathbf{q}}\right) \tag{22}$$

where $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{q}}$ are purely random draws of $n_0 - 1$ peers from $\mathcal{I} \setminus \{i\}$.

Given a binning scheme $\mathbf{z}(\cdot)$, the average group effect of a bin b peer group is:

$$AGE_b \equiv E\left(y_i(\tilde{\mathbf{p}}) - y_i(\tilde{\mathbf{q}}) \,\middle|\, \mathbf{z}\left(\{\mathbf{x}_j\}_{j \in \tilde{\mathbf{p}}}\right) = \mathbf{e}_b, \mathbf{z}\left(\{\mathbf{x}_j\}_{j \in \tilde{\mathbf{q}}}\right) = \mathbf{e}_0\right) \tag{23}$$

where $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{q}}$ are purely random draws[6] *of $n_0 - 1$ peers from $\mathcal{I} \setminus \{i\}$.*

With binning, the average group effect can be interpreted as the predicted change in outcome from replacing the average (random) peer group from one bin with the average peer group from another bin, where the bins are defined by the peer characteristics of interest. Group effects are defined relative to a base bin whose group effect is $AGE_0 = 0$

[6] Note that AGE_b and $CGE_{k,b}$ are defined in terms of a purely random draw of peers, and thus imposes a particular conditional distribution for $\Pr\left(\{\mathbf{x}_j\}_{j \in \mathbf{p}_i} \,\middle|\, \mathbf{z}_i\right)$. Proposition 7 in Section 5.4 shows that AGE_b and $CGE_{k,b}$ are only informative about peer group reallocations that preserve this conditional distribution. See Section 5.4 for additional details.

by construction, and the choice of base is without loss of generality: the average effect of replacing a bin b peer group with a bin b' peer group is $AGE_{b'} - AGE_b$.

Definition 11 (Conditional group effects). *Given Assumptions 1 and 6, the* **conditional group effect** *of a peer group with characteristics* $\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}$ *on treated individuals with characteristics* $\mathbf{x}^o \in \mathbb{S}_{\mathbf{x}}$ *is:*

$$CGE(\mathbf{x}^o, \mathbf{x}^p) \equiv E\left(y_i(\tilde{\mathbf{p}}) - y_i(\tilde{\mathbf{q}}) \middle| \begin{array}{l} \mathbf{x}_i = \mathbf{x}^o, \{\mathbf{x}_j\}_{j \in \tilde{\mathbf{p}}} = \mathbf{x}^p, \\ \mathbf{x}_j = \mathbf{e}_0 \text{ for all } j \in \tilde{\mathbf{q}} \end{array} \right) \tag{24}$$

where $\tilde{\mathbf{p}}$ *and* $\tilde{\mathbf{q}}$ *are purely random draws of* $n_0 - 1$ *peers from* $\mathcal{I} \setminus \{i\}$.

Given discrete characteristics (DC) and a binning scheme $\mathbf{z}(\cdot)$, *the conditional group effect of a bin* b *peer group on treated individuals of observed type* k *is:*

$$CGE_{kb} \equiv E\left(y_i(\tilde{\mathbf{p}}) - y_i(\tilde{\mathbf{q}}) \middle| \mathbf{x}_i = \mathbf{e}_k, \mathbf{z}\left(\{\mathbf{x}_j\}_{j \in \tilde{\mathbf{p}}}\right) = \mathbf{e}_b, \mathbf{z}\left(\{\mathbf{x}_j\}_{j \in \tilde{\mathbf{q}}}\right) = \mathbf{e}_0 \right) \tag{25}$$

where $\tilde{\mathbf{p}}$ *and* $\tilde{\mathbf{q}}$ *are purely random draws of* $n_0 - 1$ *peers from* $\mathcal{I} \setminus \{i\}$.

As with average and conditional peer effects, average and conditional group effects are well-defined under the maintained assumptions of the model, though their identification and interpretation may depend on additional conditions.

An important special case of binning is when the peer group variable is **saturated**, i.e., each value in $\mathbb{S}_{\{\mathbf{x}\}}$ has its own bin.

Definition 12 (Saturated peer group variable). *Given Assumption 6 and discrete characteristics (DC), the* **saturated peer group variable** $\mathbf{z}_i^S$ *is:*

$$\mathbf{z}_i^S \equiv \mathbf{z}^S\left(\{\mathbf{x}_j\}_{j \in \mathbf{p}_i}\right) \equiv \sum_{s=1}^{S} \mathbf{e}_s \mathbb{I}\left(m(\bar{\mathbf{x}}_i) = s\right) \tag{26}$$

where $m : \mathbf{S}_{\bar{\mathbf{x}}} \to \{0, 1, \ldots, S\}$ *is a strict ordering on* $\mathbf{S}_{\bar{\mathbf{x}}}$ *such that* $m(\mathbf{e}_0) = 0$, *and* $\mathbf{e}_s$ *is the unit vector of length* $S = |\mathbf{S}_{\bar{\mathbf{x}}}| - 1$ *containing one in column* s *and zero elsewhere.*

Saturated peer group variables allow precise identification results with minimal assumptions, but the support of $\bar{\mathbf{x}}_i$ is typically too large for $\mathbf{z}_i^S$ to be a practical explanatory variable. Empirical researchers may prefer a more parsimonious regression model based on substantially fewer bins, or an approximation to the unrestricted model using functional form restrictions (see Section 5.5 and Appendix B.3).

Example 4 (A saturated group variable for classmate gender). *With classroom size* $n_0 = 16$, *the male share* $\bar{\mathbf{x}}_i$ *of student* i*'s classroom has support* $\mathbf{S}_{\bar{\mathbf{x}}} = \left\{0, \frac{1}{15}, \frac{2}{15}, \ldots, 1\right\}$

which has $|\mathbf{S}_{\bar{\mathbf{x}}}| = 16$ elements. It fully describes the gender composition of student i's classmates and can be represented by the saturated variable:

$$\mathbf{z}_i^S = \mathbf{z}^S\left(\{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) = \begin{cases} [\ 0 \quad 0 \quad \cdots \quad 0\ ] & \text{if } \bar{\mathbf{x}}_i = 0.0 \\ [\ 1 \quad 0 \quad \cdots \quad 0\ ] & \text{if } \bar{\mathbf{x}}_i = \frac{1}{15} \\ & \vdots \\ [\ 0 \quad 0 \quad \cdots \quad 1\ ] & \text{if } \bar{\mathbf{x}}_i = 1.0 \end{cases}$$

This saturated variable is also a binned variable with $B + 1 = 16$ bins.

4.3 Reallocation effects

Reallocation effects predict the effect of changing the entire social network, accounting for the constraint that changes to one peer group imply offsetting changes to other peer groups. For example, increasing the number of boys in one classroom requires a corresponding reduction in the number of boys in other classrooms.

As with peer and group effects, the first step is to define the relevant counterfactual scenario, which in this case is a (possibly stochastic) rule for assigning individuals to groups. That counterfactual rule will be called a reallocation mechanism, and the resulting assignment will be called a reallocation.

Definition 13 (Reallocation)**.** *A* **reallocation mechanism** *is a function $\mathbf{G}_R : \mathcal{T}^I \times \mathcal{G}^I \times [0,1]^{\mathbb{N}} \to \mathcal{G}^I_{n_0}$. A* **reallocation** *is a random vector:*

$$\tilde{\mathbf{G}}_R \equiv \mathbf{G}_R(\mathbf{T}, \mathbf{G}, \boldsymbol{\rho}) \equiv \begin{bmatrix} g_{1R}(\mathbf{T}, \mathbf{G}, \boldsymbol{\rho}) \\ \vdots \\ g_{IR}(\mathbf{T}, \mathbf{G}, \boldsymbol{\rho}) \end{bmatrix} \tag{27}$$

where $\mathbf{G}_R$ is a reallocation mechanism and $\boldsymbol{\rho} = (\rho_1, \rho_2, \ldots) \perp\!\!\!\perp \mathbf{T}, \mathbf{G}$ is a sequence of independent $U(0,1)$ random variables.

Reallocation mechanisms are deterministic functions, but can use a random number generator or similar device (represented by $\boldsymbol{\rho}$). They can be based on observed characteristics $\mathbf{X} = \mathbf{X}(\mathbf{T})$, outcomes $\mathbf{Y} = \mathbf{Y}(\mathbf{T}, \mathbf{G})$, or other relevant information.

Example 5 (Reallocations by gender)**.** *Suppose for convenience there is a large population of half boys and half girls allocated to classrooms of size $n_0 = 16$. Examples of reallocation mechanisms in this setting include:*

- *Evenly-divided: randomly select 8 boys and 8 girls for each classroom.*
- *6/10 divided: randomly select 6 boys and 10 girls for half of the classrooms, and randomly select 10 boys and 6 girls for the rest.*
- *Random: randomly select 16 students for each classroom.*
- *60/40 random: randomly place students in the "60/40" sub-population with probability 0.6 for boys and 0.4 for girls, and place the rest in the "40/60" sub-population. Then randomly assign students to classrooms within each sub-population.*
- *Single-gender: randomly assign students to all-boy and all-girl classrooms.*

Appendix A.1 defines $\mathbf{G}_R(\cdot)$ *functions to implement each mechanism.*

Each reallocation mechanism implies a probability distribution over the counterfactual group assignment $\tilde{\mathbf{G}}_R$ and counterfactual outcomes $\mathbf{Y}(\mathbf{T}, \tilde{\mathbf{G}}_R)$, so the average outcome can be compared across any two mechanisms.

Definition 14 (Reallocation effects)**.** *Given Assumptions 1, 5, and 6, the* **average reallocation effect** *of the reallocation mechanism* $\mathbf{G}_R$ *is:*

$$ARE(\mathbf{G}_R) \equiv E\left(y_i(\mathbf{p}(i, \tilde{\mathbf{G}}_R)) - y_i(\tilde{\mathbf{p}})\right) \tag{28}$$

and its **conditional reallocation effect** *on treated individuals of observed type* k *is:*

$$CRE_k(\mathbf{G}_R) \equiv E\left(y_i(\mathbf{p}(i, \tilde{\mathbf{G}}_R)) - y_i(\tilde{\mathbf{p}})\middle| \mathbf{x}_i = \mathbf{e}_k\right) \tag{29}$$

where $\tilde{\mathbf{G}}_R = \mathbf{G}_R(\mathbf{T}, \mathbf{G}, \boldsymbol{\rho})$ *and* $\tilde{\mathbf{p}}$ *is a purely random draw of* $n_0 - 1$ *peers from* f_τ*.*

Reallocation effects are defined relative to a benchmark mechanism of simple random assignment, but any two reallocation mechanisms $\mathbf{G}_0$ and $\mathbf{G}_1$ can be compared by calculating $ARE(\mathbf{G}_1) - ARE(\mathbf{G}_0)$ or $CRE_k(\mathbf{G}_1) - CRE_k(\mathbf{G}_0)$.

5 Main results

This section demonstrates the relevant properties of the model. Section 5.1 establishes some preliminary results, and the next three sections provide identification results for peer effects (5.2), group effects (5.3), and reallocation effects (5.4) under the assumption of discrete characteristics. Section 5.5 addresses continuous characteristics.

5.1 Preliminaries

Proposition 1 below shows that simple causal effects are weighted averages of more complex effects, with weights derived from the probability distribution of $\mathbf{x}_i$.

Proposition 1 (Aggregation)**.** *Given Assumptions 1–6 and discrete characteristics (DC):*

1. *Average effects are a weighted average of conditional effects:*

$$APE_\ell = \sum_{k=0}^{K} \mu_k CPE_{k\ell} \tag{30}$$

$$AGE_b = \sum_{k=0}^{K} \mu_k CGE_{kb} \tag{31}$$

$$ARE(\mathbf{G}_R) = \sum_{k=0}^{K} \mu_k CRE_k(\mathbf{G}_R) \tag{32}$$

where $\mu_k = E(x_{ik}) = \Pr(\mathbf{x}_i = \mathbf{e}_k)$ *as defined earlier.*

2. *Binned group effects are a weighted average of saturated group effects:*

$$AGE_b = \sum_{k=0}^{K} \sum_{s=1}^{S} \mu_k \left(w_{sb}^G(\boldsymbol{\mu}) - w_{s0}^G(\boldsymbol{\mu}) \right) CGE_{ks}^S \tag{33}$$

$$CGE_{kb} = \sum_{s=1}^{S} \left(w_{sb}^G(\boldsymbol{\mu}) - w_{s0}^G(\boldsymbol{\mu}) \right) CGE_{ks}^S \tag{34}$$

where CGE_{ks}^S *is the conditional group effect for bin* s *of the saturated variable* $\mathbf{z}_i^S$*,*

$$w_{sb}^G(\boldsymbol{\mu}) \equiv \frac{\sum_{\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}} M\left(\bar{\mathbf{x}}\left(\mathbf{x}^p\right), n_0 - 1, \boldsymbol{\mu}\right) \mathbb{I}\left(\mathbf{z}^S\left(\mathbf{x}^p\right) = \mathbf{e}_s\right) \mathbb{I}\left(\mathbf{z}\left(\mathbf{x}^p\right) = \mathbf{e}_b\right)}{\sum_{\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}} M(\bar{\mathbf{x}}\left(\mathbf{x}^p\right), n_0 - 1, \boldsymbol{\mu}) \mathbb{I}\left(\mathbf{z}\left(\mathbf{x}^p\right) = \mathbf{e}_b\right)} \tag{35}$$

is a weighting function, and:

$$M(\bar{\mathbf{x}}, n, \boldsymbol{\mu}) \equiv \frac{n!}{\prod_{k=0}^{K} (n\bar{x}_{\cdot k})!} \prod_{k=0}^{K} \mu_k^{n\bar{x}_{\cdot k}} \tag{36}$$

is the probability of drawing the value $n\bar{\mathbf{x}}$ *from a multinomial distribution with* n *trials and categorical probability vector* $\boldsymbol{\mu}$*.*

3. *Peer effects are a weighted average of saturated group effects:*

$$APE_\ell = \sum_{k=0}^{K} \sum_{s=1}^{S} \mu_k \left(w^P_{s\ell}(\boldsymbol{\mu}) - w^P_{s0}(\boldsymbol{\mu})\right) CGE^S_{ks} \tag{37}$$

$$CPE_{k\ell} = \sum_{s=1}^{S} \left(w^P_{s\ell}(\boldsymbol{\mu}) - w^P_{s0}(\boldsymbol{\mu})\right) CGE^S_{ks} \tag{38}$$

where:

$$w^P_{s\ell}(\boldsymbol{\mu}) \equiv \sum_{\mathbf{x}^p \in \mathcal{M}_{\mathbf{x}} : |\mathbf{x}^p| = n_0 - 2} M\left(\bar{\mathbf{x}}\left(\mathbf{x}^p\right), n_0 - 2, \boldsymbol{\mu}\right) \mathbb{I}\left(\mathbf{z}^S\left(\mathbf{x}^p \cup \{\mathbf{e}_\ell\}\right) = \mathbf{e}_s\right) \tag{39}$$

is a weighting function and $\mathcal{M}_{\mathbf{x}}$ *is the set of all multisets on* $\mathbb{S}_{\mathbf{x}}$*.*

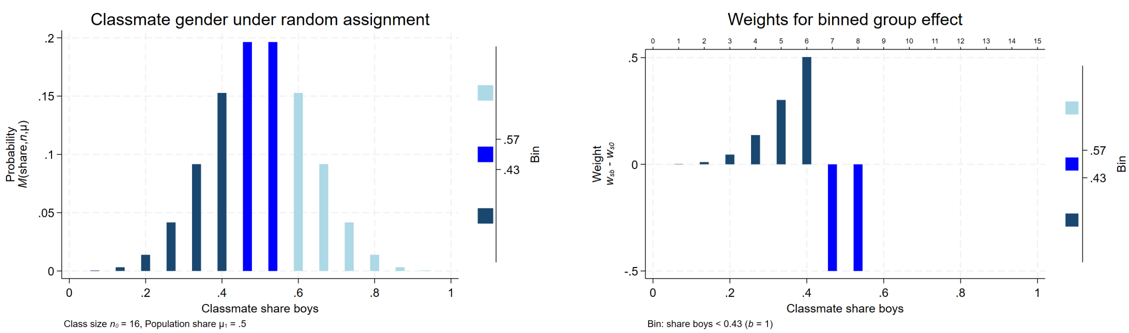

Figure 1: Weights for binned group effects, bin $b = (< 43\%$ boys). Weights are proportional to random assignment probabilities, positive within the bin, and negative in the base bin.

Example 6 (Weights for binned group effects)**.** *Figure 1 shows the weights relating group effects for the binned variable defined in Example 3 to those for the saturated variable defined in Example 4. The first graph shows the probability distribution of classmate share boys* $\bar{\mathbf{x}}_i$ *under random assignment, i.e., the function* $M(\cdot, n, \boldsymbol{\mu})$*. The second graph shows the weights* $w^G_{sb}(\boldsymbol{\mu}) - w^G_{s0}(\boldsymbol{\mu})$ *for bin* $b = 1$ *(< 43% boys). As the figure shows, each value in the bin receives a positive weight proportional to its probability under random assignment, and each value in the base bin receives a negative weight.*

Proposition 2 below restates a standard result that allows individual characteristics to be added to or dropped from certain regressions.

Proposition 2 (Implications of simple random assignment). *Given Assumption 3, simple random assignment (RA) implies:*

$$L(y_i|\mathbf{c}_i, \bar{\mathbf{x}}_i) = L(y_i|\mathbf{c}_i) + L(y_i|\bar{\mathbf{x}}_i) + constant \tag{40}$$

$$L(y_i|\mathbf{c}_i, \mathbf{z}_i) = L(y_i|\mathbf{c}_i) + L(y_i|\mathbf{z}_i) + constant \tag{41}$$

for any vector $\mathbf{c}_i = \mathbf{c}(\tau_i)$ of own characteristics.

Lemma 1 below shows that stratified random assignment produces the same conditional expectation function (if not the same best linear predictor) as simple random assignment. It will be used to prove identification under stratified random assignment.

Lemma 1 (Implications of stratified random assignment). *Given Assumptions 1–6 and discrete characteristics (DC), stratified random assignment (SA) implies:*

$$E(y_i|\mathbf{x}_i = \mathbf{x}^o, \bar{\mathbf{x}}_i = \mathbf{x}^p) = E\left(y_i(\tilde{\mathbf{p}}) \middle| \mathbf{x}_i = \mathbf{x}^o, \bar{\mathbf{x}}\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right) = \mathbf{x}^p\right) \tag{42}$$

where $\tilde{\mathbf{p}}$ is a purely random draw of $(n_0 - 1)$ peers from $\mathcal{I} \setminus \{i\}$.

Proposition 3 below shows that any peer-separable outcome function can be written as a sum of latent variables, and peer effects as their conditional expectations.

Proposition 3 (Implications of separability). *Given Assumptions 1–6, let $PE : \mathcal{T}^2 \times \{2, 3, \ldots\} \to \mathbb{R}$ be defined:*

$$PE(a, b, n) \equiv \frac{y\left(a, \left\{b^{[n-1]}\right\}\right)}{n-1} \tag{43}$$

where $b^{[n-1]}$ is $n-1$ copies of b, and let $PE_{ij} \equiv PE(\tau_i, \tau_j, n_0)$. Then:

1. *Peer separability (PS) implies that, for any $\mathbf{p}$ of size $|\mathbf{p}| = n_0 - 1$:*

$$y_i(\mathbf{p}) = \sum_{j\in\mathbf{p}} PE_{ij} \tag{44}$$

and conditional and average peer effects can be expressed as:

$$CPE(\mathbf{x}^o, \mathbf{x}^p) = E(PE_{ij}|\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p) - E(PE_{ij}|\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{e}_0) \tag{45}$$

$$APE(\mathbf{x}^p) = E(PE_{ij}|\mathbf{x}_j = \mathbf{x}^p) - E(PE_{ij}|\mathbf{x}_j = \mathbf{e}_0) \tag{46}$$

for all $\mathbf{x}^o, \mathbf{x}^p \in \mathbb{S}_{\mathbf{x}}$.

2. *Peer separability (PS) and own separability (OS) imply that, for any* $\mathbf{p}$ *of size* $|\mathbf{p}| = n_0 - 1$*:*

$$y_i(\mathbf{p}) = \sum_{j \in \mathbf{p}} (OE_i + PE_j) \tag{47}$$

where $OE_i \equiv PE(\tau_i, 1, n_0) + c$*,* $PE_j \equiv PE(1, \tau_j, n_0) - PE(1, 1, n_0) - c$*, and* c *is an arbitrary constant. Conditional and average peer effects can be expressed as:*

$$CPE(\mathbf{x}^o, \mathbf{x}^p) = APE(\mathbf{x}^p) = E(PE_j|\mathbf{x}_j = \mathbf{x}^p) - E(PE_j|\mathbf{x}_j = \mathbf{e}_0) \tag{48}$$

for all $\mathbf{x}^o, \mathbf{x}^p \in \mathbb{S}_{\mathbf{x}}$*.*

The pair-specific latent variable $PE_{i,j}$ is defined without assuming separability, and describes individual i's outcome if all peers were identical to peer j. Peer separability has the additional implication that replacing peer j with new peer j' will change individual i's outcome by $PE_{i,j'} - PE_{i,j}$. Own separability allows this pairwise effect to be further decomposed into an own effect OE_i and peer effect PE_j: replacing peer j with new peer j' will change i's outcome by $PE_{j'} - PE_j$, and the difference $OE_i - OE_j$ is the difference in outcomes between individuals i and j if they had the same peers. This decomposition is defined up to an additive constant that cancels out in such comparisons.

5.2 Identifying peer effects

Proposition 4 below shows identification of average and conditional peer effects under the asssumptions of peer separability and (stratified) random assignment. The identification results are constructive and suggest simple linear regression estimators whose implementation is described in more detail in Appendix B.1.

Proposition 4 (Identification of peer effects)**.** *Given Assumptions 1–6 and discrete characteristics (DC):*

1. *Simple random assignment (RA) and peer separability (PS) imply that peer effects are identified from the joint distribution of* $(y_i, \mathbf{x}_i, \bar{\mathbf{x}}_i)$*:*

$$APE_\ell = \frac{\alpha_{1\ell}}{n_0 - 1} \tag{49}$$

$$CPE_{k\ell} = \frac{\beta_{2\ell} + \beta_{3k\ell}}{n_0 - 1} \tag{50}$$

where $(\alpha_{1\ell}, \beta_{2\ell}, \beta_{3k\ell})$ *are coefficients from the best linear predictors:*

$$L(y_i|\bar{\mathbf{x}}_i) \equiv \alpha_0 + \bar{\mathbf{x}}_i \boldsymbol{\alpha}_1 \tag{51}$$

$$L(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i'\bar{\mathbf{x}}_i) \equiv \beta_0 + \mathbf{x}_i\boldsymbol{\beta}_1 + \bar{\mathbf{x}}_i\boldsymbol{\beta}_2 + \mathbf{x}_i\boldsymbol{\beta}_3\bar{\mathbf{x}}_i' \tag{52}$$

i.e., $\alpha_{1\ell}$ *is element* ℓ *of* $\boldsymbol{\alpha}_1$*,* $\beta_{2\ell}$ *is element* ℓ *of* $\boldsymbol{\beta}_2$*,* $\beta_{3k\ell}$ *is the element in row* k *and column* ℓ *of* $\boldsymbol{\beta}_3$ *for all* $k > 0$*, and* $\beta_{30\ell} \equiv 0$ *for all* ℓ*.*

2. *Stratified random assignment (SA) and peer separability (PS) imply that peer effects are identified from the joint distribution of* $(y_i, \mathbf{x}_i, \bar{\mathbf{x}}_i)$*:*

$$APE_\ell = \sum_{k=0}^{K} \mu_k \frac{\beta_{2\ell} + \beta_{3k\ell}}{n_0 - 1} \tag{53}$$

$$CPE_{k\ell} = \frac{\beta_{2\ell} + \beta_{3k\ell}}{n_0 - 1} \tag{54}$$

where $(\beta_{2\ell}, \beta_{3k\ell})$ *are defined as in equation (52).*

Returning to the issues raised in the introduction, result (49) shows what can be learned from the simple linear in means model: under the two key assumptions of peer separability and simple random assignment, this model can be used to estimate average peer effects for any discrete characteristic(s). Proposition 2 also implies that own characteristics can be included as in the canonical linear in means model (1), excluded as in equation (51), or replaced with any vector of individual-level characteristics. These results do not require the assumption of own separability, or that the regression model is complete in the sense of including all outcome-relevant characteristics.

Two key assumptions each play a distinct role in this result. Simple random assignment allows average effects to be recovered without including an interaction between own and peer characteristics: any interaction effect averages out. Peer separability adds linearity in average characteristics: outcomes are sums of pairwise latent variables and their conditional expectations are linear in the discrete peer characteristics.

Adding own characteristics and an interaction term (52) allows the researcher to assume stratified rather than simple random assignment (53) or to measure conditional rather than average peer effects (50, 54). Stratified random assignment produces the same conditional expectation function as simple random assignment (by Lemma 1), but not the same best linear predictor since $\mathbf{x}_i$ and $\bar{\mathbf{x}}_i$ are no longer independent. As a result, average peer effects can only be recovered from regression model (52) which is rich enough to recover the full CEF.

Example 7 (Gender peer effects). *Table 1 reports estimates of gender peer effects. Given Assumptions 1–6, simple random assignment (RA) into classrooms of size* $n_0 = 16$*, and peer separability (PS), results (49) and (50) in Proposition 4 can be applied to the regression coefficients in (1) and (2) to predict that replacing a randomly-selected girl with a random boy:*

- *reduces the average student's reading score by*
  - $|APE_1| = \frac{0.342}{16-1} = 0.023$ *standard deviations (by column 1).*
  - $|APE_1| = \frac{(0.453\times 0.268)+0.547\times(0.268+0.133)}{16-1} = 0.023$ *SD (by column 2).*
- *reduces the average girl's reading score by* $|CPE_{01}| = \frac{0.268}{16-1} = 0.018$ *SD.*
- *reduces the average boy's reading score by* $|CPE_{11}| = \frac{0.268+0.133}{16-1} = 0.027$ *SD.*

Columns (5) and (6) have the same derivation and interpretation for math scores. The other columns in Table 1 are discussed in Examples 10 and 11 of Section 5.3.

	Reading score				Math score			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share male peers	−0.342	−0.268		-0.224	−0.532**	−0.279		−0.303
	(0.286)	(0.372)		(0.725)	(0.226)	(0.338)		(0.624)
Male x Share male peers		−0.133		-0.142		−0.455		-0.453
		(0.390)		(0.390)		(0.412)		(0.411)
Share male peers < 0.43			-0.024	-0.066			0.090	0.012
			(0.066)	(0.109)			(0.065)	(0.100)
Share male peers > 0.57			-0.119*	-0.083			-0.045	0.021
			(0.072)	(0.110)			(0.065)	(0.100)
Sample size (# students)	2,185	2,185	2,185	2,185	2,196	2,196	2,196	2,196
# clusters	147	147	147	147	148	148	148	148
Average effect of a male peer:								
All students (APE_1)	-0.023	-0.023			-0.035**	-0.035**		
Girls (CPE_{01})		-0.018				-0.019		
Boys (CPE_{11})		-0.027				-0.049***		
p-value for test of:								
own separability		0.733				0.271		
peer separability				0.421				0.954

Table 1: Gender peer and group effects in Project STAR. Additional control variables include own gender and a school/grade fixed effect. Cluster-robust standard errors in parentheses, * = 0.1, ** = 0.05, *** = 0.01. See Appendix A for additional details.

The results in Proposition 4 do not depend on the researcher's choice of which (discrete) characteristics to include in the model. An alternative choice of characteristics is equally valid but addresses a different set of counterfactual questions.

Example 8 (Motivation peer effects). *Table 2 reports estimates of motivation peer effects. Column (1) reports regression coefficients for the discrete motivation variable defined in equation (14) of Example 2, and are comparable to results in Bietenbeck (2025) Table 7. Given the assumptions from Example 7, the results in column (1) imply:*

- *replacing a random medium-motivation classmate with a random low-motivation student reduces the average student's reading score by* $|APE_1| = \frac{0.426}{16-1} = 0.028$ *SD.*
- *replacing a random medium-motivation classmate with a random high-motivation student increases the average student's reading score by* $|APE_2| = \frac{0.220}{16-1} = 0.014$ *SD.*
- *replacing a random high-motivation classmate with a random low-motivation student reduces the average student's reading score by* $|APE_1 - APE_2| = \frac{0.426+0.220}{16-1} = 0.043$ *SD.*

Column (5) has the same derivation and interpretation for math scores. The other columns in Table 2 are discussed in Example 9 below.

	Reading score				Math score			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share low-motivation (LM) peers	-0.426^{**}	-0.540^{***}	-0.523^{***}	-0.396^{*}	-0.281	-0.320^{*}	-0.285	-0.292
(< 33rd percentile)	(0.167)	(0.136)	(0.137)	(0.219)	(0.189)	(0.185)	(0.187)	(0.269)
Share high-motivation (HM) peers	0.220				0.076			
(> 66th percentile)	(0.213)				(0.283)			
Share male peers			-0.242	-0.144			-0.477^{**}	-0.482^{*}
			(0.276)	(0.301)			(0.229)	(0.271)
Share male LM peers				-0.243				0.014
				(0.345)				(0.341)
Sample size (# students)	2,185	2,185	2,185	2,185	2,196	2,196	2,196	2,196
# clusters	147	147	147	147	148	148	148	148
Avg effect of an LM peer:								
replacing MM peers	-0.028^{**}				-0.019			
replacing HM peers	-0.043^{***}				-0.024			
replacing MM/HM peers		-0.036^{***}	-0.036^{***}	-0.036^{***}		-0.021^{*}	-0.022^{*}	-0.022^{*}
Avg effect of a male LM peer:								
replacing male MM/HM peers			-0.035^{***}	-0.043^{***}			-0.019	-0.019
replacing female MM/HM peers			-0.051^{***}	-0.052^{***}			-0.051^{***}	-0.051^{***}
replacing female LM peers			-0.016	-0.026			-0.032^{**}	-0.031
Avg effect of a female LM peer:								
replacing male MM/HM peers			-0.019	-0.017			0.013	0.013
replacing female MM/HM peers			-0.035^{***}	-0.026^{*}			-0.019	-0.019
replacing male LM peers			0.016	0.026			0.032^{**}	0.031

Table 2: Motivation peer effects in Project STAR. Additional control variables include own gender and a school/grade fixed effect. Cluster-robust standard errors in parentheses, $^{*} = 0.1, ^{**} = 0.05, ^{***} = 0.01$. See Appendix A for additional details.

The framework can accommodate multiple characteristics without requiring that those characteristics constitute a complete causal model.

Example 9 (Classmate gender and motivation)**.** *Table 2 also reports estimated peer effects for the combination of gender and motivation. Column (2) is similar to column (1), but with motivation coded as a single binary variable for expositional convenience. Column (3) adds peer gender, while column (4) adds the interaction of peer gender and motivation. Regressions (1), (2), and (4) directly fit into Proposition 4, while (3) can be interpreted as a restricted version of (4). Columns (5) through (8) repeat the analysis with math scores. Gender and motivation are not likely to be the only relevant peer characteristics, so including both characteristics is not enough to make the model complete. Instead, these richer models open up an additional set of comparisons.*

The results for reading predict a smaller effect of replacing a random girl with a random boy if both students have similar motivation levels, and a larger effect of replacing a random medium/high-motivation student with a random low-motivation student if both students are boys.

5.3 Identifying group effects

Proposition 5 below shows that group affects can be identified without requiring the strong assumption of peer separability. As with Proposition 4, the identification results are constructive and can in some cases imply very simple regression models.

Proposition 5 (Identification of group effects)**.** *Given Assumptions 1–6 and discrete characteristics (DC):*

1. *Simple random assignment (RA) implies that binned group effects are identified from the joint distribution of* $(y_i, \mathbf{x}_i, \mathbf{z}_i)$*:*

$$AGE_b = \gamma_{1b} \tag{55}$$

$$CGE_{kb} = \delta_{2b} + \delta_{3kb} \tag{56}$$

where $(\gamma_{1b}, \delta_{2b}, \delta_{3kb})$ *are coefficients from the best linear predictors:*

$$L(y_i|\mathbf{z}_i) \equiv \gamma_0 + \mathbf{z}_i\boldsymbol{\gamma}_1 \tag{57}$$

$$L(y_i|\mathbf{x}_i, \mathbf{z}_i, \mathbf{x}_i'\mathbf{z}_i) \equiv \delta_0 + \mathbf{x}_i\boldsymbol{\delta}_1 + \mathbf{z}_i\boldsymbol{\delta}_2 + \mathbf{x}_i\boldsymbol{\delta}_3\mathbf{z}_i' \tag{58}$$

i.e., γ_{1b} *is element* b *of* $\boldsymbol{\gamma}_1$*,* δ_{2b} *is element* b *of* $\boldsymbol{\delta}_2$*,* δ_{3kb} *is the element in row* k *and column* b *of* $\boldsymbol{\delta}_3$ *for all* $k > 0$*, and* $\delta_{30b} \equiv 0$ *for all* b*.*

2. *Stratified random assignment (SA) implies that saturated group effects are identified from the joint distribution of* $(y_i, \mathbf{x}_i, \bar{\mathbf{x}}_i)$*:*

$$CGE^S_{ks} = \lambda_{2s} + \lambda_{3ks} \tag{59}$$

where $(\lambda_{2s}, \lambda_{3ks})$ *are coefficients from the best linear predictor:*

$$L(y_i|\mathbf{x}_i, \mathbf{z}^S_i, \mathbf{x}'_i\mathbf{z}^S_i) \equiv \lambda_0 + \mathbf{x}_i\boldsymbol{\lambda}_1 + \mathbf{z}^S_i\boldsymbol{\lambda}_2 + \mathbf{x}_i\boldsymbol{\lambda}_3\mathbf{z}^{S\prime}_i \tag{60}$$

i.e., λ_{2s} *is element* s *of* $\boldsymbol{\lambda}_2$*,* λ_{3ks} *is the element in row* k *and column* s *of* $\boldsymbol{\lambda}_3$ *for all* $k > 0$*, and* $\lambda_{30s} \equiv 0$ *for all* s*. Peer effects and binned group effects are also identified:*

$$AGE_b = \sum_{k=0}^{K}\sum_{s=1}^{S}\mu_k\left(w^G_{sb}(\boldsymbol{\mu}) - w^G_{s0}(\boldsymbol{\mu})\right)(\lambda_{2s} + \lambda_{3ks}) \tag{61}$$

$$CGE_{kb} = \sum_{s=1}^{S}\left(w^G_{sb}(\boldsymbol{\mu}) - w^G_{s0}(\boldsymbol{\mu})\right)(\lambda_{2s} + \lambda_{3ks}) \tag{62}$$

$$APE_\ell = \sum_{k=0}^{K}\sum_{s=1}^{S}\mu_k\left(w^P_{s\ell}(\boldsymbol{\mu}) - w^P_{s0}(\boldsymbol{\mu})\right)(\lambda_{2s} + \lambda_{3ks}) \tag{63}$$

$$CPE_{k\ell} = \sum_{s=1}^{S}\left(w^P_{s\ell}(\boldsymbol{\mu}) - w^P_{s0}(\boldsymbol{\mu})\right)(\lambda_{2s} + \lambda_{3ks}) \tag{64}$$

where $w^G_{sb}(\cdot)$ *and* $w^P_{s\ell}(\cdot)$ *are defined in Proposition 1.*

As in Proposition 4, simple random assignment allows the researcher to recover informative causal effects from a simple regression model. For example, result (55) in Proposition 5 implies that average group effects can be recovered from regression model (57) with a coarse binning scheme and no individual characteristics or interaction terms. Such a model can be estimated with reasonable precision given limited data.

Example 10 (Group effects for classmate gender)**.** *Given Assumptions 1–6 and simple random assignment (RA) into classrooms of size* $n_0 = 16$*, the results in column (3) of Table 1 imply that:*

- *moving from the average mixed classroom (43% to 57% boys) to the average female-dominated classroom (< 43% boys) reduces the average student's reading score by* $|AGE_1| = 0.024$ *SD.*
- *moving from the average mixed classroom to the average male-dominated classroom (> 57% boys) reduces the average student's reading score by* $|AGE_2| = 0.119$ *SD.*

The results in column (7) have the a similar interpretation for math scores.

Proposition 5 also shows general identification of both peer and group effects under stratified random assignment. However, the saturated regression model (60) required to recover these effects is likely to be impractical in most applications. A researcher may choose to impose functional form assumptions as a feasible approximation to the saturated model. Appendix B.3 discusses this alternative in more detail.

Since separability assumptions imply functional form restrictions but are not required for identification, they are testable. Proposition 6 below describes two simple implications that can be tested using linear regression coefficients.

Proposition 6 (Testable implications of separability)**.** *Given Assumptions 1–6, discrete characteristics (DC), and stratified random assignment (SA):*

1. *Peer separability (PS) implies:*

$$L(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i'\bar{\mathbf{x}}_i, \mathbf{z}_i) = L(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i'\bar{\mathbf{x}}_i) \tag{65}$$

2. *Peer separability (PS) and own separability (OS) imply:*

$$L(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i'\bar{\mathbf{x}}_i) = L(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i) \tag{66}$$

The implications in Proposition 6 hold for any $(\mathbf{x}_i, \mathbf{z}_i)$ since separability is a property of the outcome function $y(\cdot, \cdot)$ and not of the researcher's chosen explanatory variables or binning scheme. These researcher choices do affect the power of the test.

Example 11 (Separability tests)**.** *Columns (4) and (8) in Table 1 implement the peer separability test described in equation (65). The coefficients on the two binned peer group variables are jointly insignificant, so we cannot reject peer separability. The statistically insignificant interaction terms in columns (2) and (6) imply we cannot reject own separability using the test described in equation (66). Both tests have low power due to the sample size, so they provide only weak evidence in favor of separability.*

5.4 Identifying reallocation effects

Proposition 7 below shows that reallocation effects can be derived from peer and/or group effects.

Proposition 7 (Identification of reallocation effects)**.** *Let $\mathbf{G}_R$ be a reallocation mechanism such that $\tilde{\mathbf{G}}_R = \mathbf{G}_R(\mathbf{T}, \mathbf{G}, \boldsymbol{\rho})$ satisfies Assumption 6 and stratified random assignment (SA). Then given Assumptions 1–6 and discrete characteristics (DC):*

1. If $(\mathbf{S}^1_{\bar{\mathbf{x}}}, \ldots, \mathbf{S}^B_{\bar{\mathbf{x}}})$ *are singletons, and* $\Pr(\bar{\mathbf{x}}_i(\mathbf{X}, \tilde{\mathbf{G}}_R) \in \mathbf{S}^0_{\bar{\mathbf{x}}}) = 0$*, then:*

$$ARE(\mathbf{G}_R) = \sum_{k=0}^{K} \sum_{b=1}^{B} \mu_k \Delta z_{kb}(\mathbf{G}_R) CGE_{kb} \tag{67}$$

$$CRE_k(\mathbf{G}_R) = \sum_{b=1}^{B} \Delta z_{kb}(\mathbf{G}_R) CGE_{kb} \tag{68}$$

where:

$$\begin{aligned}\Delta z_{kb}(\mathbf{G}_R) \equiv\ & \Pr\left(\bar{\mathbf{x}}_i(\mathbf{X}, \mathbf{G}_R(\mathbf{T}, \mathbf{G}, \boldsymbol{\rho})) \in \mathbf{S}^b_{\bar{\mathbf{x}}} | \mathbf{x}_i = \mathbf{e}_k\right) \\ & - \Pr\left(\bar{\mathbf{x}}_i(\tilde{\mathbf{p}}) \in \mathbf{S}^b_{\bar{\mathbf{x}}}\right)\end{aligned} \tag{69}$$

2. Peer separability (PS) implies:

$$ARE(\mathbf{G}_R) = \sum_{k=0}^{K} \sum_{\ell=1}^{K} \mu_k \Delta \bar{x}_{k\ell}(\mathbf{G}_R) CPE_{k\ell}(n_0 - 1) \tag{70}$$

$$CRE_k(\mathbf{G}_R) = \sum_{\ell=1}^{K} \Delta \bar{x}_{k\ell}(\mathbf{G}_R) CPE_{k\ell}(n_0 - 1) \tag{71}$$

where:

$$\Delta \bar{x}_{k\ell}(\mathbf{G}_R) \equiv E\left(\bar{x}_{i\ell}(\mathbf{X}, \mathbf{G}_R(\mathbf{T}, \mathbf{G}, \boldsymbol{\rho})) | \mathbf{x}_i = \mathbf{e}_k\right) - \mu_\ell \tag{72}$$

3. Peer separability (PS) and own separability (OS) imply:

$$ARE(\mathbf{G}_R) = 0 \tag{73}$$

$$CRE_k(\mathbf{G}_R) = \sum_{\ell=1}^{K} \Delta \bar{x}_{k\ell}(\mathbf{G}_R) APE_\ell(n_0 - 1) \tag{74}$$

Identification of reallocation effects thus requires stratified random assignment for both the observed group assignment $\mathbf{G}$ (so that Proposition 4 or 5 applies and peer/group effects are identified) and the counterfactual reallocation $\tilde{\mathbf{G}}_R$ (so that Proposition 7 applies and reallocation effects can be derived from peer/group effects). Constructing reallocation mechanisms that satisfy (SA) is straightforward: mechanisms based entirely on $\mathbf{X}$ will work, as will mechanisms that use the observed $\mathbf{G}$ if it satisfies (SA). The requirement that the counterfactual reallocation satisfies Assumption 6 is a support condition: without further restrictions, outcomes cannot be predicted for reallocations whose group size does not appear in the data.

The details of Proposition 7 imply two additional practical constraints. First, average peer/group effects are insufficient to evaluate any reallocation mechanism; estimates of conditional effects are required. Second, in the absence of peer separability (PS), reallocation effects can only be measured for reallocation mechanisms that place all of their weight on singleton bins. With a moderate sample size, only a few values of $\bar{\mathbf{x}}_i$ will be observed often enough under random assignment to estimate their conditional group effects with reasonable precision.

Example 12 (Reallocation effects for classmate gender)**.** *Table 3 below shows estimated reallocation effects for each of the reallocation mechanisms described in Example 5. Details are available in Appendix A.1.*

The first panel assumes peer separability (PS) and applies equations (70)–(72) to columns (2) and (6) of Table 1. The results show monotonic reallocation effects: more integrated classrooms produce better average outcomes because male peers have a stronger negative effect on other boys than on girls.

The second panel applies equations (67)–(69) to binned group effect estimates that do not assume separability. The evenly-divided reallocation effect is estimated using two singleton bins $\left(\bar{\mathbf{x}}_i \approx \frac{7}{15}, \bar{\mathbf{x}}_i \approx \frac{8}{15}\right)$ *and one pooled bin (all other values), and produces similar predictions to the separable model. The 6/10 divided reallocation effect is estimated using four singleton bins and suggests an important difference from the first panel: the cost of a classroom with many boys exceeds the benefit of a classroom with few. This is broadly consistent with the group effect estimates in Table 1. No reallocation effects are reported for the two mechanisms that have nonzero probability for all 16 values of* $\bar{\mathbf{x}}_i$ *because some values are not observed in the data.*

5.5 Approximation and functional form

The identification results in Sections 5.2–5.4 use discretization and binning to abstract from functional form restrictions. When this approach is adequate for the research question, it provides simple regression models with clear interpretations that can be estimated precisely from limited data. In other settings, continuity may be essential to the research question. For example, a researcher may wish to estimate $APE(\mathbf{x}^p)$ for specific values of some continuous and/or high-dimensional characteristic $\mathbf{x}^p$, or to estimate marginal effects. For these research questions, the discretized model is only an approximation to the continuous model of interest. Discretizing implies information loss, so other approximation methods may use information more efficiently.

This section describes a sieve-based method that allows $\mathbf{x}_i$ to be any K-vector

			Reading score			Math score		
Reallocation	$\Delta\bar{x}_{01}$	$\Delta\bar{x}_{11}$	CRE_0	CRE_1	ARE	CRE_0	CRE_1	ARE
Separable model:								
Evenly-divided	0.033	−0.033	−0.009	0.013	0.002	−0.009	0.024***	0.008
6/10 divided	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Random	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60/40 random	−0.020	0.020	0.005	−0.008	−0.001	0.006	−0.015***	−0.005
Single-gender	−0.500	0.500	0.134	−0.200	−0.033	0.139	−0.367***	−0.114
Binned model:								
Evenly-divided	0.033	−0.033	−0.016	0.081	0.032	0.036	−0.023	0.006
6/10 divided	0.000	0.000	−0.007	−0.149**	−0.078*	−0.070	−0.018	−0.044
Random	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60/40 random	−0.020	0.020	— insufficient data —					
Single-gender	−0.500	0.500	— insufficient data —					

Table 3: Reallocation effects for classmate gender in Project STAR. Reallocation effects for simple random assignment are reported for comparison and are always zero. Cluster-robust p-values: $^{*} = 0.1, ^{**} = 0.05, ^{***} = 0.01$. See Appendix A.1 for additional calculation details.

while still exploiting the dimension-reducing implications of separability and random assignment for measuring average peer effects. Sieve methods can also be useful in measuring conditional and group effects; see Appendix B.3 for details.

By Proposition 3, peer separability allows the unrestricted $E\left(y_i \middle| \{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right)$ — an unknown function of $(n_0-1)K$ variables — to be expressed in terms of $E\left(PE_{ij}\left|\mathbf{x}_j\right.\right)$, an unknown function of K variables.

$$(PS, RA) \implies \qquad E\left(y_i \middle| \{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) = \sum_{j\in\mathbf{p}_i} E\left(PE_{ij}\left|\mathbf{x}_j\right.\right) \tag{75}$$

The unrestricted $E\left(PE_{ij}\left|\mathbf{x}_j\right.\right)$ can be approximated by a sieve (Hansen, 2014): a series of flexible linear models whose complexity is increasing in the sample size. Average peer effects can then be recovered by applying equation (46) in Proposition 3.

Definition 15 (Sieve approximation)**.** *Given Assumptions 1–6, a sieve approximation of order m to the function $a:\mathbb{R}^K\to\mathbb{R}$ is a known function $a_m:\mathbb{R}^K\to\mathbb{R}^m$ such that:*

$$a_m(\mathbf{x}^p)\bar{\boldsymbol{\pi}}_m \approx a(\mathbf{x}^p) \tag{76}$$

for the unknown parameter vector $\bar{\boldsymbol{\pi}}_m \equiv E\left(a_m(\mathbf{x}_i)'a_m(\mathbf{x}_i)\right)^{-1}E\left(a_m(\mathbf{x}_i)'a(\mathbf{x}_i)\right)$.

As shown in Proposition 8 below, the linearity of the approximation in (76) carries through the sum in (75), producing an approximation to the unrestricted CEF whose coefficients can be estimated by OLS and used to estimate average peer effects.

Proposition 8 (Sieve model for average peer effects)**.** *Given Assumptions 1–6 and peer separability (PS), let* a_m *be a sieve approximation of order* m *to* $a(\mathbf{x}^p) \equiv E\left(PE_{ij}\,|\mathbf{x}_j = \mathbf{x}^p\right)$ *and let:*

$$\boldsymbol{\pi} \equiv (\pi_1, \ldots, \pi_m) \equiv E(\bar{\mathbf{a}}_i'\bar{\mathbf{a}}_i)^{-1} E(\bar{\mathbf{a}}_i' y_i) \qquad \text{where } \bar{\mathbf{a}}_i \equiv \frac{1}{n_0 - 1} \sum_{j \in \mathbf{p}_i} a_m(\mathbf{x}_j) \tag{77}$$

Then simple random assignment (RA) implies:

$$E\left(y_i \,\middle|\, \{\mathbf{x}_j\}_{j \in \mathbf{p}_i}\right) = \sum_{j \in \mathbf{p}_i} a(\mathbf{x}_j) \approx \bar{\mathbf{a}}_i \boldsymbol{\pi} \tag{78}$$

$$APE(\mathbf{x}^p) = a(\mathbf{x}^p) - a(\mathbf{0}) \approx \left(\frac{a_m(\mathbf{x}^p) - a_m(\mathbf{0})}{n_0 - 1}\right) \boldsymbol{\pi} \tag{79}$$

where the approximation errors in (78) and (79) are proportional to the approximation error in (76).

Returning to the linear in means model (1) from the introduction, Proposition 8 allows that regression to be interpreted as measuring average peer effects for continuous $\mathbf{x}_i$ using the approximating function $a_2(\mathbf{x}) = (1, \mathbf{x})$. Other functional form assumptions such as discretization can also be interpreted as sieve approximations with a different choice of approximating function.

Example 13 (Three models of motivation peer effects)**.** *Table 4 reports coefficient estimates from three sieve approximations to* $a(\mathbf{x}) \equiv E\left(PE_{i,j}\,|\mathbf{x}_j = \mathbf{x}\right)$*:*

$$a_2(\mathbf{x}) = \begin{bmatrix} 1 & \mathbf{x} \end{bmatrix} \qquad \text{(linear model)}$$

$$a_3(\mathbf{x}) = \begin{bmatrix} 1 & \mathbb{I}\left(\mathbf{x} < p_{33}\right) & I(\mathbf{x} > p_{66}) \end{bmatrix} \qquad \text{(discretized model)}$$

$$a_4(\mathbf{x}) = \begin{bmatrix} 1 & \mathbf{x} & (\mathbf{x} - p_{33})I(\mathbf{x} > p_{33}) & (\mathbf{x} - p_{66}) * I(\mathbf{x} > p_{66}) \end{bmatrix} \qquad \text{(2-knot spline)}$$

where $\mathbf{x}$ *is the continuous measure of motivation used in Bietenbeck (2025) and* p_N *is its* N*th percentile. The linear model is comparable to Bietenbeck (2025) Table 5 and can also be interpreted as a spline with zero knots. The discretized model is similar to column (1) of Table 2, but represents a discrete CEF in Table 2 and an approximation to a continuous CEF here. 2-knot spline coefficients are reported as slopes* $(\pi_2, \pi_2 + \pi_3, \pi_2 + \pi_3 + \pi_4)$ *for comparison, and are estimated by applying equation (77).*

For comparison, Table 4 also reports approximate predictions for the discrete counterfactuals that are directly estimated in Table 2. The 2-knot spline implies similar *predictions to the direct estimates, while the linear model predicts smaller effects.*

Figure 2 shows the average peer effects function $APE(\cdot)$ for each model, derived by applying equation (79) to the regression results, and Table 4 reports the associated marginal effects $\frac{\partial APE(\mathbf{x}^p)}{\partial \mathbf{x}^p}$. The 2-knot spline results suggest that the marginal impact of a more motivated peer is high in the middle of the motivation distribution, and is low or even negative in the tails. Motivation scores are self-reported, so scores in the tails could be driven by non-classical measurement error, and only weakly related to actual motivation. This feature cannot be captured by the linear model.

	Reading score			Math score		
	(1)	(2)	(3)	(4)	(5)	(6)
Peer average motivation	0.082^{***}			0.039		
	(0.024)			(0.032)		
Share low-motivation (LM) peers		-0.433^{**}			-0.282	
(< 33rd percentile)		(0.171)			(0.189)	
Share high-motivation (HM) peers		0.221			0.077	
(> 66th percentile)		(0.218)			(0.283)	
Peer motivation, 2-knot linear spline:						
Low-motivation peer			0.015			-0.012
			(0.046)			(0.047)
Medium-motivation peer			1.031^{**}			0.769
			(0.448)			(0.515)
High-motivation peer			-0.249			-0.224
			(0.315)			(0.434)
Sample size (# students)	2,185	2,185	2,185	2,196	2,196	2,196
# clusters	147	147	147	148	148	148
Avg effect of LM peers:						
replacing MM peers	-0.018^{***}	-0.029^{**}	-0.030^{***}	-0.009	-0.019	-0.018^{*}
replacing HM peers	-0.032^{***}	-0.044^{***}	-0.039^{***}	-0.015	-0.024	-0.020
Marginal effect of more motivated peer:						
replacing LM peers	0.005^{***}		0.001	0.003		-0.001
replacing MM peers	0.005^{***}		0.069^{**}	0.003		0.051
replacing HM peers	0.005^{***}		-0.017	0.003		-0.015
Model selection:						
AIC statistic	5626.07	5623.97	5624.05	5663.50	5662.79	5663.83
Leave-one-out MSE	0.8933	0.8924	0.8928	0.8967	0.8963	0.8964
JMA weight	0.3760		0.6240	0.1336		0.8664

Table 4: Sieve model estimates of motivation peer effects in Project STAR. Additional control variables include a school/grade fixed effect. Cluster-robust standard errors in parentheses, $^{*} = 0.1, ^{**} = 0.05, ^{***} = 0.01$. See Appendix A.2 for additional details.

Although Proposition 8 facilitates the interpretation of ad hoc peer effects regressions, a more systematic approach will have better econometric properties. Sieve models are typically estimated by selecting the approximating function $a_m(\cdot)$ from a family of basis functions such as polynomials or splines that can reproduce an arbitrary smooth function as $m \to \infty$, and selecting the order m using a data-driven procedure that increases in the sample size. This procedure will consistently estimate the unrestricted model under conditions described in Hansen (2014). The basis function family is usually chosen for convenience and desired properties such as differentiability and tail behavior. The order can be selected acccording to an information criterion or by cross-validation. Model averaging is an alternative to model selection that is more efficient and less sensitive to small performance variations across dissimilar models. Hansen (2014) discusses the relative merits of different selection criteria and averaging procedures.

Example 14 (Model selection for classmate motivation)**.** *Table 4 reports two model selection criteria: the Akaike Information Criterion (AIC) statistic and the mean squared prediction error estimated by leave-one-out cross-validation. The 2-knot spline beats the linear model by both criteria for reading and by cross-validation for math, while the linear model is preferred by AIC for math. The discretized model performs best of all.*

Table 4 also reports optimal jackknife model averaging (JMA) weights for the spline models (Hansen and Racine, 2012; Hansen, 2014). 1-knot and 3-knot splines were also considered but received zero weight; see Appendix A.2 for details. The thick line in Figure 2 shows the JMA estimate of average peer effects, which preserves but smooths out the nonlinear pattern shown in the 2-knot spline.

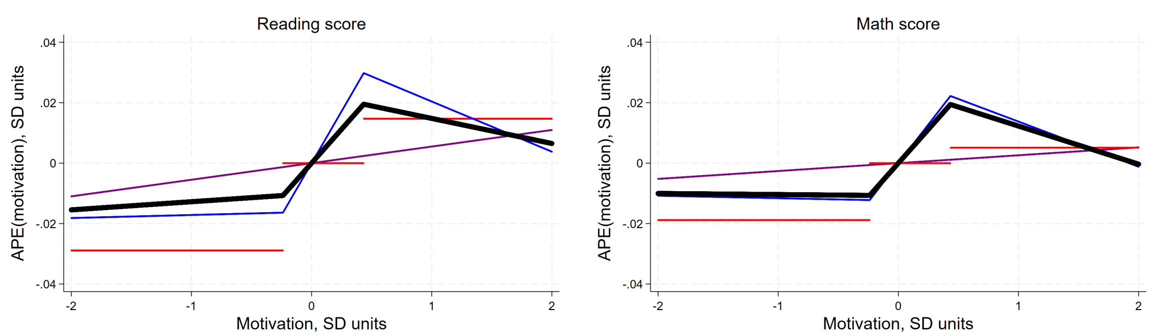

Figure 2: Average peer effects for peer motivation: linear model (purple), 3-category discretized model (red), 2-knot linear spline (blue), and jackknife model average of linear splines (thick black line). By construction, $APE(0) = 0$ in all models.

6 Conclusion

This paper establishes a simple framework for thinking about contextual effects when estimating a complete model is infeasible, clarifies common empirical procedures, and suggests simple extensions to those procedures. Returning to the issues raised in the introduction, the results here have several implications for empirical research.

The first implication is that simple specifications using categorical explanatory variables can have a clear and robust causal interpretation. A minimal specification that is linear in the peer group average of a a single binary characteristic (high/low income, black/white, male/female, etc.) measures the difference in average peer effects across the two categories under straightforward and testable assumptions. A researcher can choose the characteristic(s) based on their research question, and researchers with the same data but other research questions can choose other characteristics. In contrast, a regression with many related peer characteristics is difficult to interpret without imposing the implausible assumption that the regression model is complete.

The second implication is that specific extensions to this simple model allow the researcher to address additional causal questions or relax some assumptions. Adding an interaction term allows the researcher to estimate conditional peer effects and reallocation effects, or to relax the assumption of simple random assignment. Binning can be used to relax the assumption of peer separability and estimate group effects while still maintaining a tractable regression model. Discretization or sieve approximations can be used to handle continuous and/or high-dimensional characteristics while still exploiting the dimension-reducing implications of peer separability and random assignment.

A third implication is that the randomization mechanism is important in ways that are not often appreciated. For example, average peer effects describe the effect of replacing a randomly selected peer from one category with a randomly selected peer from another category. This effect does not in general correspond to the precise effect of replacing any peer from one category with any peer from the other category. As a result, not all potentially interesting reallocation effects can be identified.

Simple models and methods are central to empirical research, and are the focus of this paper. However, the framework developed here provides avenues for further research that develops or applies more novel econometric methods. For example, the ability to analyze treatment effect heterogeneity along any dimension of interest opens up the risk of unstructured regression fishing. Tools for systematically analyzing treatment effect heterogeneity (Wager and Athey, 2018) can be adapted to this setting, and may be useful in constructing robust data-driven predictors of peer and group effects. A second avenue for further research is to investigate more complex social networks than the group-based structure considered here. Preliminary work in progress suggests that many results in this paper extend to a general network structure. Other work in progress considers the commonly used multiple-cohort research design (Hoxby, 2000), in which students are assumed to be randomly assigned to entry/birth cohorts within non-randomly selected schools. This feature substantially constrains the set of counterfactual comparisons that can be made in the absence of strong restrictions on potential outcomes.

References

Aronow, Peter M. and Cyrus Samii, "Estimating average causal effects under general interference, with application to a social network experiment," *The Annals of Applied Statistics*, 2017, *11* (4), 1912–1947.

Basse, Guillaume, Peng Ding, Avi Feller, and Panos Toulis, "Randomization tests for peer effects in group formation experiments," *Econometrica*, 2024, *92* (2), 567–590.

Bhattacharya, Debopam, "Inferring optimal peer assignment from experimental data," *Journal of the American Statistical Association*, 2009, *104* (486), 486–500.

Bietenbeck, Jan, "The long-term impacts of low-achieving childhood peers: Evidence from Project STAR," *Journal of the European Economic Association*, 01 2019, *18* (1), 392–426.

— , "Replication package for: 'Do motivated classmates matter for educational success?'," doi:10.5281/zenodo.11911078 June 2024.

— , "Do motivated classmates matter for educational success?," *The Economic Journal*, 2025, *135* (665), 36–58.

Blume, Lawrence E., William A. Brock, Steven N. Durlauf, and Yannis M. Ioannides, "Identification of social interactions," in Jess Benhabib, Matthew O. Jackson, and Alberto Bisin, eds., *Handbook of Social Economics Volume 1B*, Elsevier, 2011.

Boozer, Michael A. and Stephen E. Cacciola, "Inside the black box of Project STAR: Estimation of peer effects using experimental data," Discussion Paper 832, Yale University Economic Growth Center 2001.

Carrell, Scott E., Bruce I. Sacerdote, and James E. West, "From natural variation to optimal policy? The importance of endogenous peer group formation," *Econometrica*, 2013, *81* (3), 855–882.

Cascio, Elizabeth and Diane Whitmore Schanzenbach, "First in the class? Age and The education production function," Working Paper 13663, NBER 2007.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan, "How does your Kindergarten classroom affect your earnings? Evidence from Project STAR," *The Quarterly Journal of Economics*, 11 2011, *126* (4), 1593–1660.

Graham, Bryan S., "Identifying social interactions through conditional variance restrictions," *Econometrica*, 2008, *76* (3), 643–660.

_ , Guido W. Imbens, and Geert Ridder, "Measuring the effects of segregation in the presence of social spillovers: A nonparametric approach," *Journal of Econometrics*, 2025. Forthcoming, DOI: 10.1016/j.jeconom.2025.106006.

Hansen, Bruce E., "Nonparametric sieve regression: Least squares, averaging least squares, and cross-validation," in Jeffrey S. Racine, Liangjun Su, and Aman Ullah, eds., *The Oxford Handbook of Applied Nonparametric and Semiparametric Econometrics and Statistics*, Oxford University Press, 2014, chapter 8, pp. 214–248.

_ and Jeffrey S. Racine, "Jackknife model averaging," *Journal of Econometrics*, 2012, *167*, 38–46.

Hoxby, Caroline M., "Peer effects in the classroom: Learning from gender and race variation," Working Paper 7867, NBER 2000.

_ and Gretchen Weingarth, "Taking race out of the equation: School reassignment and the structure of peer effects," Working Paper 2005.

Li, Xinran, Peng Ding, Qian Lin, Dawei Yang, and Jun S. Liu, "Randomization inference for peer effects," *Journal of the American Statistical Association*, 2019, *114* (528), 1651–1664.

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

_ , "Identification of treatment response with social interactions," *The Econometrics Journal*, 2013, *16*, S1–S23.

Rose, Christiern D., "Identification of peer effects through social networks using variance restrictions," *The Econometrics Journal*, 2017, *20* (3), S47–S60.

Sojourner, Aaron, "Identification of peer effects with missing peer data: Evidence from Project STAR," *The Economic Journal*, 06 2013, *123* (569), 574–605.

Wager, Stefan and Susan Athey, "Estimation and inference of heterogeneous treatment effects using random forests," *Journal of the American Statistical Association*, 2018, *113* (523), 1228–1242.

Whitmore, Diane, "Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment," *American Economic Review*, May 2005, *95* (2), 199–203.

Not-for-publication Appendix

A Application details

As described in Section 2 of the main text, the analysis of motivation peer effets is adapted from Bietenbeck (2025) using the replication package provided by its author (Bietenbeck, 2024). The analysis of gender peer effects uses the same sample, peer group definition, and other methodological choices to facilitate comparison.

Bietenbeck makes the following methodological choices, which I take as given:

- The treated population is students who enter in grades 2 or 3.
- The peer group is returning students in the treated student's entry-grade classroom.
- Nonrandom school selection is addressed by including school/grade fixed effects.
- Peer motivation is measured using the peer's motivation score in the previous grade, in standard deviation units. Treated students do not have previous-grade motivation scores since they were not present in the previous grade. As a result, the treated student's own motivation is not included in any regression models.

A few additional modifications are made to simplify the analysis and fit it into this paper's framework:

- Peer characteristics other than motivation are dropped from the model.
  - This allows for the coefficient on peer motivation to be interpreted as the average peer effect associated with peer motivation.
- Class size and characteristics of the treated student (other than gender in some regressions) are dropped from the model.
  - This simplifies the model. As discussed in the main text, their inclusion or exclusion does not affect identification since students are randomly assigned to classrooms.
- Classroom size is taken to be $n_0 = 16$ when calculating average and conditional peer effects. This is based on the median number of peers in the data.

Table 5 below shows that these modifications do not have a substantial effect on the results. Column (1) is a direct replication of column (1) in Bietenbeck's Table 5, while column (3) is a direct replication of column (1) in Bietenbeck's Table 7. Column (2) replicates column (1) in Table 4 of this paper, while column (4) replicates column (1) in Table 2 and column (5) replicates column (2) in Table 4. Columns (6)–(10) have a similar interpretation for math scores.

	Reading score					Math score				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Peer avg motivation	0.081***	0.082***				0.036	0.039			
	(0.023)	(0.024)				(0.032)	(0.032)			
Share with low motivation			−0.429***	−0.426**	−0.433***			-0.222	-0.281	-0.282
(< 33rd percentile)			(0.157)	(0.167)	(0.171)			(0.174)	(0.189)	(0.189)
Share with high motivation			0.136	0.220	0.221			0.074	0.076	0.077
(> 66th percentile)			(0.187)	(0.213)	(0.218)			(0.295)	(0.283)	(0.283)
Additional controls:										
School/grade fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Own gender	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	No
Other demographics	Yes	No	Yes	No	No	Yes	No	Yes	No	No
Class size	Yes	No	Yes	No	No	Yes	No	Yes	No	No
Peer achievement	No	No	Yes	No	No	No	No	Yes	No	No
Peer demographics	No	No	Yes	No	No	No	No	Yes	No	No
Sample size (# students)	2,185	2,185	2,185	2,185	2,185	2,196	2,196	2,196	2,196	2,196
# clusters	147	147	147	147	147	148	148	148	148	148

Table 5: Motivation peer effects in Project STAR, comparison of results with Bietenbeck (2025). Cluster-robust standard errors in parentheses, $^* = 0.1, ^{**} = 0.05, ^{***} = 0.01$.

A.1 Reallocation effect details

This section provides additional calculation details for the reallocation effect results in Sections 4.3 and 5.4

Example 15 below gives a specific $\mathbf{G}_R$ function that implements each reallocation mechanism described in Example 5. Each function uses the random vector $\boldsymbol{\rho}$ to randomly sort individuals and then sequentially fills classrooms from the sorted list(s) to produce the desired properties.

Example 15 (Reallocation mechanisms)**.** *The reallocation mechanisms described in Example 5 can be implemented as follows:*

- *The evenly-divided mechanism can implemented by:*

$$g_{iR}(\mathbf{X}, \boldsymbol{\rho}) \equiv \begin{cases} \text{ceil}\left(\frac{r_{i1}}{8}\right) & \text{if } r_{i1} \leq 8g^* \\ \text{ceil}\left(\frac{r_{i2}}{16}\right) & \text{if } r_{i1} > 8g^* \end{cases} \tag{80}$$

where $r_{i1} \equiv \sum_{j=1}^{I} \mathbb{I}(\mathbf{x}_j = \mathbf{x}_i)\,\mathbb{I}(\rho_j \leq \rho_i)$, $g^* \equiv \text{floor}\left(\frac{\min(\sum_{i=1}^{I} \mathbf{x}_i, \sum_{i=1}^{I} 1-\mathbf{x}_i)}{8}\right)$, *and* $r_{i2} = \sum_{j=1}^{I} \mathbb{I}(r_{j1} > 8g^*)\,\mathbb{I}(\rho_j \leq \rho_i)$. *That is, we sort the boys and girls separately based on* ρ_i, *fill in group 1 with the first 8 boys and the first 8 girls from their respective lists, etc. The realized* $\mathbf{X}$ *will not necessarily have an equal number of boys and girls, so there may be one or two* $(G - g^*)$ *"spillover" classrooms that are not evenly mixed.*

- *The 6/10 divided mechanism can implemented by:*

$$g_{iR}(\mathbf{X},\boldsymbol{\rho}) \equiv \begin{cases} \text{ceil}\left(\frac{r_i}{10-4\mathbf{x}_i}\right) & \text{if } r_i \leq (10-4\mathbf{x}_i)g^* \\ \min\left(g^* + \text{ceil}\left(\frac{r_i-(10-4\mathbf{x}_i)g^*}{6+4\mathbf{x}_i}\right), G\right) & \text{if } r_i > (10-4\mathbf{x}_i)g^* \end{cases} \tag{81}$$

where $r_i \equiv \sum_{j=1}^{I} \mathbb{I}(\mathbf{x}_j = \mathbf{x}_i)\,\mathbb{I}(\rho_j \leq \rho_i)$, *and* $g^* \equiv \text{floor}\left(\frac{10G-\sum_{i=1}^{I}\mathbf{x}_i}{4}\right)$. *Classroom* G *is a spillover classroom when* $\sum_{i=1}^{I}\mathbf{x}_i$ *is not a multiple of 16.*

- *The random mechanism can implemented by:*

$$g_{iR}(\mathbf{X},\boldsymbol{\rho}) \equiv \text{ceil}\left(\frac{r_i}{16}\right) \tag{82}$$

where $r_i \equiv \sum_{j=1}^{I} \mathbb{I}(\rho_j \leq \rho_i)$.

- *The 60/40 random mechanism can implemented by:*

$$g_{iR}(\mathbf{X},\boldsymbol{\rho}) \equiv \text{ceil}\left(\frac{r_{i1}}{16}\right) \tag{83}$$

where $r_{i1} = \sum_{j=1}^{I} r_{i2}(1-r_{j2}) + \mathbb{I}(r_{i2} = r_{j2})\,\mathbb{I}(\rho_j \leq \rho_i)$ *and* $r_{i2} = \mathbb{I}(\rho_{I+i} \leq (0.4+0.2\mathbf{x}_i))$. *Classroom* $g^* \equiv \text{ceil}\left(\sum_{i=1}^{I} r_{i2}/16\right)$ *is a spillover classroom when* $\sum_{i=1}^{I} r_{i2}$ *is not a multiple of 16.*

- *The single-gender mechanism can implemented by:*

$$g_{iR}(\mathbf{X},\boldsymbol{\rho}) \equiv \text{ceil}\left(\frac{r_i}{16}\right) \tag{84}$$

where $r_i \equiv \sum_{j=1}^{I} \mathbf{x}_i(1-\mathbf{x}_j) + \mathbb{I}(\mathbf{x}_j = \mathbf{x}_i)\,\mathbb{I}(\rho_j \leq \rho_i)$. *This mechanism will produce one mixed-gender spillover classroom when the number of boys is not a multiple of 16.*

Note that the spillover classroom(s) account for a negligible proportion of students as the population increases.

Example 16 below shows how the net changes (Δx and/or Δz) required to calculate the reallocation effects in Table 3 are determined.

Example 16 (Reallocation effect calculations, part 1)**.** *The net changes defined in equations (69) and (72) of Proposition 7 are calculated for each example as follows:*

- *In the evenly-divided mechanism, all girls have 8 male classmates out of 15, and* *all boys have 7 out of 15. Therefore:*

$$\Delta x_{01} = E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 0) - \mu_1 \approx 8/15 - 0.5 \approx 0.033 \qquad \text{(evenly-divided)}$$
$$\Delta x_{11} = E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 1) - \mu_1 \approx 7/15 - 0.5 \approx -0.033$$
$$\mathbf{z}_i \equiv \begin{bmatrix} \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 7) & \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 8) \end{bmatrix} \qquad \text{(evenly-divided bins)}$$
$$\Delta z_{01} = 0 - M(7/15, 15, 0.5) \approx -0.196 \qquad \text{(evenly-divided bin changes)}$$
$$\Delta z_{02} = 1 - M(8/15, 15, 0.5) \approx 0.804$$
$$\Delta z_{11} = 1 - M(7/15, 15, 0.5) \approx 0.804$$
$$\Delta z_{12} = 0 - M(8/15, 15, 0.5) \approx -0.196$$

where $M(\cdot)$ is the multinomial/binomial probability defined in equation (35).

- *In the 6/10 divided mechanism, 10/16 of girls have 6 male classmates out of 15, 6/16 of girls have 10 male classmates, 10/16 of boys have 9 male classmates, and 6/16 have 5 male classmates. Therefore:*

$$\Delta x_{01} = E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 0) - \mu_1 \approx \frac{10}{16} \times \frac{6}{15} + \frac{6}{16} \times \frac{10}{15} - 0.5 \approx 0.0 \qquad \text{(6/10 divided)}$$
$$\Delta x_{11} = E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 1) - \mu_1 \approx \frac{10}{16} \times \frac{9}{15} + \frac{6}{16} \times \frac{5}{15} - 0.5 \approx 0.0$$
$$\mathbf{z}_i \equiv \begin{bmatrix} \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 5) & \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 6) & \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 9) & \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 10) \end{bmatrix}$$
$$\text{(6/10 divided bins)}$$
$$\Delta z_{01} = 0 - M(5/15, 15, 0.5) \approx -0.092 \qquad \text{(6/10 divided bin changes)}$$
$$\Delta z_{02} = 10/16 - M(6/15, 15, 0.5) \approx 0.472$$
$$\Delta z_{03} = 0 - M(9/15, 15, 0.5) \approx -0.153$$
$$\Delta z_{04} = 6/16 - M(10/15, 15, 0.5) \approx 0.283$$
$$\Delta z_{11} = 06/16 - M(5/15, 15, 0.5) \approx 0.283$$
$$\Delta z_{12} = 0 - M(6/15, 15, 0.5) \approx -0.153$$
$$\Delta z_{13} = 10/16 - M(9/15, 15, 0.5) \approx 0.472$$
$$\Delta z_{14} = 0 - M(10/15, 15, 0.5) \approx -0.092$$

- *In the random mechanism, $\bar{\mathbf{x}}_i$ and $\mathbf{x}_i$ are independent, so $E(\bar{\mathbf{x}}_i|\mathbf{x}_i) = E(\bar{\mathbf{x}}_i) = \mu_1$.*

Therefore:

$$\Delta x_{01} = E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 0) - \mu_1 = 0.5 - 0.5 = 0 \qquad \text{(random)}$$
$$\Delta x_{11} = E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 1) - \mu_1 = 0.5 - 0.5 = 0$$
$$\mathbf{z}_i \equiv \begin{bmatrix} \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 0) & \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 1) & \cdots & \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 15) \end{bmatrix} \qquad \text{(random bins)}$$
$$\Delta z_{kb} = 0 \qquad \textit{for all } k, b \qquad \text{(random bin changes)}$$

- *In the 60/40 random mechanism,* $\bar{\mathbf{x}}_i$ *and* $\mathbf{x}_i$ *are independent conditional on classroom type. Therefore:*

$$E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 0) = E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 0, type = 60/40)\Pr(type = 60/40|\mathbf{x}_i = 0) + E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 0, type = 40/60)\Pr(type = 40/60|\mathbf{x}_i = 0) \approx (0.6 \times 0.4) + (0.4 \times 0.6) \approx 0.48 \tag{85}$$

$$E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 1) = E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 1, type = 60/40)\Pr(type = 60/40|\mathbf{x}_i = 1) + E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 1, type = 40/60)\Pr(type = 40/60|\mathbf{x}_i = 1) \approx (0.6 \times 0.6) + (0.4 \times 0.4) \approx 0.52 \tag{86}$$

which implies:

$$\Delta x_{01} = E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 0) - \mu_1 \approx 0.48 - 0.5 \approx -0.02 \qquad \text{(60/40 random)}$$
$$\Delta x_{11} = E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 1) - \mu_1 \approx 0.52 - 0.5 \approx 0.02$$
$$\mathbf{z}_i \equiv \begin{bmatrix} \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 0) & \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 1) & \cdots & \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 15) \end{bmatrix} \qquad \text{(60/40 random bins)}$$
$$\Delta z_{kb} > 0 \qquad \textit{for all } k, b \qquad \text{(60/40 random bin changes)}$$

- *In the single-gender mechanism,* $\bar{\mathbf{x}}_i \approx \mathbf{x}_i$ *for all* i*. Therefore:*

$$\Delta x_{01} = E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 0) - \mu_1 \approx 0 - 0.5 \approx -0.5 \qquad \text{(single-gender)}$$
$$\Delta x_{11} = E(\bar{\mathbf{x}}_i|\mathbf{x}_i = 1) - \mu_1 \approx 1 - 0.5 \approx 0.5$$
$$\mathbf{z}_i \equiv \begin{bmatrix} \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 0) & \mathbb{I}(n_0\bar{\mathbf{x}}_i \approx 15) \end{bmatrix} \qquad \text{(single-gender bins)}$$
$$\Delta z_{01} = 1 - M(0/15, 15, 0.5) \approx 1 \qquad \text{(single-gender bin changes)}$$
$$\Delta z_{02} = 0 - M(15/15, 15, 0.5) \approx 0$$
$$\Delta z_{11} = 0 - M(0/15, 15, 0.5) \approx 0$$
$$\Delta z_{12} = 1 - M(15/15, 15, 0.5) \approx 1$$

Note that these calculations are approximations that ignore the "spillover" classroom(s). The approximation error will be negligible as long as the classroom size is small relative to the population.

Example 17 below shows the calculated reallocation effects.

Example 17 (Reallocation effect calculations, part 2)**.** *Assuming peer separability, the results in columns (2) and (6) of Table 1 imply that:*

$$CPE_{01}(n_0 - 1) = -0.268 \qquad \text{(reading)}$$
$$CPE_{11}(n_0 - 1) = -0.268 + (-0.133) = -0.401$$
$$CPE_{01}(n_0 - 1) = -0.279 \qquad \text{(math)}$$
$$CPE_{11}(n_0 - 1) = -0.279 + (-0.455) = -0.734$$

Proposition 7 implies implies that:

$$CRE_0 = \Delta x_{01} CPE_{01}(n_0 - 1) \qquad \text{(by (71))}$$
$$CRE_1 = \Delta x_{11} CPE_{11}(n_0 - 1) \qquad \text{(by (71))}$$
$$ARE = 0.5\, CRE_0 + 0.5\, CRE_1 \qquad \text{(by (70))}$$

The estimated reallocation effects in the first panel of Table 3 can then be calculated by substituting in the Δx values calculated in Example 16.

The estimated reallocation effects in the second panel of Table 3 can be calculated by estimating the binned model corresponding to each reallocation, looking up the Δz values calculated in Example 16, and applying equations (67) and (68) from Proposition 7. No estimates are available for the 60/40 random and single-gender reallocations because they place nonzero weight on bins that are not present in the data.

A.2 Sieve estimate details

For expositional convenience, the sieve results reported in Section 5.5 omit several calculation details that can be explained more fully here:

1. Table 4 reports linear spline results for zero-knot (linear) and 2-knot cases only. Table 6 below reports results for these cases along with the 1-knot and 3-knot cases. As the table shows, the 1-knot and 3-knot models are not selected by either criterion and receive zero JMA weights for both reading and math.
2. The average peer effects in Figure 2 are calculated by applying equation (79) in

Proposition 8:

$$APE(\mathbf{x}^p) \approx \frac{\mathbf{x}^p \pi_2}{n_0 - 1} \qquad \text{(linear model)}$$

$$APE(\mathbf{x}^p) \approx \begin{cases} \frac{\pi_2}{n_0-1} & \text{if } \mathbf{x}^p < p_{33} \\ 0 & \text{if } p_{33} \leq \mathbf{x}^p \leq p_{66} \\ \frac{\pi_3}{n_0-1} & \text{if } \mathbf{x}^p > p_{66} \end{cases} \qquad \text{(discretized model)}$$

$$APE(\mathbf{x}^p) \approx \begin{cases} \frac{\mathbf{x}^p \pi_2 + p_{33}\pi_3}{n_0-1} & \text{if } \mathbf{x}^p < p_{33} \\ \frac{\mathbf{x}^p(\pi_2+\pi_3)}{n_0-1} & \text{if } p_{33} \leq \mathbf{x}^p \leq p_{66} \\ \frac{\mathbf{x}^p(\pi_2+\pi_3+\pi_4) - p_{66}\pi_4}{n_0-1} & \text{if } \mathbf{x}^p > p_{66} \end{cases} \qquad \text{(2-knot spline)}$$

3. The marginal effects in Table 4 are calculated by taking derivatives:

$$\frac{\partial APE(\mathbf{x}^p)}{\partial \mathbf{x}^p} \approx \frac{\pi_2}{n_0 - 1} \qquad \text{(linear model)}$$

$$\frac{\partial APE(\mathbf{x}^p)}{\partial \mathbf{x}^p} \approx \begin{cases} \frac{\pi_2}{n_0-1} & \text{if } \mathbf{x}^p < p_{33} \\ \frac{\pi_2+\pi_3}{n_0-1} & \text{if } p_{33} < \mathbf{x}^p < p_{66} \\ \frac{\pi_2+\pi_3+\pi_4}{n_0-1} & \text{if } \mathbf{x}^p > p_{66} \end{cases} \qquad \text{(2-knot spline)}$$

The discretized model is flat or non-differentiable everywhere, and so does not have meaningful marginal effects.

4. Table 4 also reports the effect of replacing a randomly-selected medium-motivation peer with a random low-motivation peer:

$$APE_{medium \to low} \equiv E(APE(\mathbf{x}_i)|\mathbf{x}_i < p_{33}) - E(APE(\mathbf{x}_i)|p_{33} < \mathbf{x}_i < p_{66}) \tag{87}$$
$$\approx (E(a_m(\mathbf{x}_i)|\mathbf{x}_i < p_{33}) - E(a_m(\mathbf{x}_i)|p_{33} < \mathbf{x}_i < p_{66}))\, \boldsymbol{\pi} \tag{88}$$

as well as the effect of replacing a random high-motivation peer with a random low-motivation peer:

$$APE_{high \to low} \equiv E(APE(\mathbf{x}_i)|\mathbf{x}_i < p_{33}) - E(APE(\mathbf{x}_i)|\mathbf{x}_i > p_{66}) \tag{89}$$
$$\approx (E(a_m(\mathbf{x}_i)|\mathbf{x}_i < p_{33}) - E(a_m(\mathbf{x}_i)|\mathbf{x}_i > p_{66}))\, \boldsymbol{\pi} \tag{90}$$

where the conditional expectations are estimated by conditional averages.

	Reading score				Math score			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Number of knots	0	1	2	3	0	1	2	3
Peer motivation								
< 25th percentile	0.082***	0.079**	0.015	0.000	0.039	0.039	-0.012	-0.026
	(0.024)	(0.038)	(0.046)	(0.051)	(0.032)	(0.042)	(0.047)	(0.056)
25th-33rd percentile	↓	↓	↓	0.654	↓	↓	↓	0.591
				(0.623)				(0.648)
33rd-50th percentile			1.031**	↓			0.769	↓
			(0.448)				(0.515)	
50th-66th percentile		0.107	↓	0.875		0.040	↓	0.507
		(0.230)		(0.802)		(0.323)		(0.817)
66th-75th percentile		↓	-0.249	↓		↓	-0.224	↓
			(0.315)				(0.434)	
> 75th percentile			↓	-0.326			↓	-0.233
				(0.442)				(0.550)
Sample size (# students)	2,185	2,185	2,185	2,185	2,196	2,196	2,196	2,196
# clusters	147	147	147	147	148	148	148	148
Model selection:								
AIC statistic	5626.07	5628.06	5624.05	5625.58	5663.50	5665.50	5663.83	5665.82
Leave-one-out MSE	0.8933	0.8941	0.8928	0.8936	0.8967	0.8975	0.8964	0.8974
JMA weight	0.3760	0.0000	0.6240	0.0000	0.1336	0.0000	0.8664	0.0000

Table 6: Linear spline estimates of motivation peer effects in Project STAR, with 0 to 3 knots. Additional control variables include a school/grade fixed effect. Cluster-robust standard errors in parentheses, $^{*} = 0.1, ^{**} = 0.05, ^{***} = 0.01$.

B Additional results

This section reports additional results that are omitted from the main text in the interest of clarity and space.

B.1 Estimation and inference

The identification results in Section 5 are constructive and suggest simple plug-in estimators that are easily implemented in standard statistical packages. This section provides an informal discussion of estimation and inference in this setting.

B.1.1 Estimating peer and group effects

Suppose the researcher has a sample of N observations on $(y_i, \mathbf{x}_i, \bar{\mathbf{x}}_i)$ from a large population that satisfies the model assumptions. Sampling models vary in the applied literature, so rather than specifying the details of the sampling scheme we simply assume it satisfies all conditions required for:

$$\sqrt{N}\left(\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}\right) \xrightarrow{D} N(\mathbf{0}, \boldsymbol{\Sigma}) \tag{91}$$

where $\boldsymbol{\psi} \equiv (\boldsymbol{\mu}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}, \boldsymbol{\lambda})$ is a vector of previously-defined population means and best linear predictor coefficients, and $\hat{\boldsymbol{\psi}} \equiv (\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\delta}}, \hat{\boldsymbol{\lambda}})$ is a consistent and asymptotically normal estimator of $\boldsymbol{\psi}$. In most applications, the researcher will have a cluster sample of size $N = n_0 G$ constructed from data on all n_0 members of G randomly selected groups, $\hat{\boldsymbol{\mu}}$ will be the sample average of $\mathbf{x}_i$, and $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\delta}}, \hat{\boldsymbol{\lambda}})$ will be the OLS regression coefficients. In other applications, the researcher may observe data on $(y_i, \mathbf{x}_i)$ for a random sample of individuals, each of whom can be linked to some aggregate data source such as census tract characteristics to construct $\bar{\mathbf{x}}_i$.

If peers are randomly assigned, Propositions 4 and 5 show that peer and group effects correspond to best linear predictor coefficients or linear combinations of those coefficients, and can therefore be estimated by:

$$\widehat{APE}_\ell = \frac{\hat{\alpha}_{1\ell}}{n_0 - 1} \qquad \text{if (PS, RA)} \tag{92}$$

$$\widehat{CPE}_{k\ell} = \frac{\hat{\beta}_{2\ell} + \hat{\beta}_{3k\ell}}{n_0 - 1} \qquad \text{if (PS, RA)} \tag{93}$$

$$\widehat{AGE}_b = \hat{\gamma}_{1b} \qquad \text{if (RA)} \tag{94}$$

$$\widehat{CGE}_{kb} = \hat{\delta}_{2b} + \hat{\delta}_{3kb} \qquad \text{if (RA)} \tag{95}$$

With stratified random assignment, peer and group effects can be expressed as linear combinations of best linear predictor coefficients or as weighted averages of those coefficients. As a result, they can be estimated by:

$$\widehat{APE}_\ell = \sum_{k=0}^{K} \hat{\mu}_k \frac{\hat{\beta}_{2\ell} + \hat{\beta}_{3k\ell}}{n_0 - 1} \qquad \text{if (PS, SA)} \tag{96}$$

$$\widehat{CPE}_{k\ell} = \frac{\hat{\beta}_{2\ell} + \hat{\beta}_{3k\ell}}{n_0 - 1} \qquad \text{if (PS, SA)} \tag{97}$$

$$\widehat{AGE}_b = \sum_{k=0}^{K} \sum_{s=1}^{S} \hat{\mu}_k \left(w_{sb}^G(\hat{\boldsymbol{\mu}}) - w_{s0}^G(\hat{\boldsymbol{\mu}})\right) \left(\hat{\lambda}_{2s} + \hat{\lambda}_{3ks}\right) \qquad \text{if (SA)} \tag{98}$$

$$\widehat{CGE}_{kb} = \sum_{s=1}^{S} \left(w_{sb}^G(\hat{\boldsymbol{\mu}}) - w_{s0}^G(\hat{\boldsymbol{\mu}})\right) \left(\hat{\lambda}_{2s} + \hat{\lambda}_{3ks}\right) \qquad \text{if (SA)} \tag{99}$$

Five of these eight estimators are just linear combinations of OLS coefficients, so the researcher can apply standard cluster-robust asymptotic inference procedures to construct standard errors and confidence intervals, or to perform hypothesis tests.

Inference is slightly more complicated for the three estimators that include weights based on $\hat{\boldsymbol{\mu}}$, as their asymptotic variance depends on the *joint* distribution of $\hat{\boldsymbol{\mu}}$ and the regression coefficients. A straightforward general approach is to define $\hat{\boldsymbol{\psi}}$ as the just-identified GMM estimator[7] for the vector of moment conditions:

$$E\left(\begin{bmatrix} \mathbf{x}_i - \boldsymbol{\mu} \\ y_i - \alpha_0 - \bar{\mathbf{x}}_i \boldsymbol{\alpha}_1 \\ \mathbf{x}_i'(y_i - \alpha_0 - \bar{\mathbf{x}}_i \boldsymbol{\alpha}_1) \\ etc. \end{bmatrix}\right) = \mathbf{0} \tag{100}$$

and $\hat{\boldsymbol{\Sigma}}$ as the associated (cluster-robust) GMM variance matrix. Under the usual GMM regularity conditions:

$$\hat{\boldsymbol{\Sigma}} \overset{P}{\to} \boldsymbol{\Sigma} \tag{101}$$

The parameter (vector) of interest can then be defined as $\boldsymbol{\theta} = \boldsymbol{\theta}(\boldsymbol{\psi})$ for some differentiable function $\boldsymbol{\theta}(\cdot)$, and its estimator can be defined as $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}(\hat{\boldsymbol{\psi}})$. Then $\hat{\boldsymbol{\theta}}$ has the asymptotic distribution:

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \overset{D}{\to} N(\mathbf{0}, (\nabla\boldsymbol{\theta}(\boldsymbol{\psi}))\boldsymbol{\Sigma}(\nabla\boldsymbol{\theta}(\boldsymbol{\psi}))') \tag{102}$$

where $\nabla\boldsymbol{\theta}(\boldsymbol{\psi})$ is the Jacobian matrix of $\boldsymbol{\theta}(\boldsymbol{\psi})$, and the asymptotic variance can be

[7] Note that the GMM estimator here is identical to the OLS estimator; the purpose of applying GMM here is to estimate the full $\boldsymbol{\Sigma}$ matrix including the asymptotic covariance of $\hat{\boldsymbol{\mu}}$ with the regression coefficients using commonly-available tools.

estimated:

$$(\nabla\boldsymbol{\theta}(\hat{\boldsymbol{\psi}}))\hat{\boldsymbol{\Sigma}}(\nabla\boldsymbol{\theta}(\hat{\boldsymbol{\psi}}))' \xrightarrow{P} (\nabla\boldsymbol{\theta}(\boldsymbol{\psi}))\boldsymbol{\Sigma}(\nabla\boldsymbol{\theta}(\boldsymbol{\psi}))' \tag{103}$$

Similarly, a hypothesis of the form $\boldsymbol{\theta}(\boldsymbol{\psi}) = \mathbf{0}$ can be tested using the Wald statistic:

$$H_0 : \boldsymbol{\theta}(\boldsymbol{\psi}) = \mathbf{0} \qquad \Longrightarrow \qquad \boldsymbol{\theta}(\hat{\boldsymbol{\psi}})'\Big((\nabla\boldsymbol{\theta}(\hat{\boldsymbol{\psi}}))\hat{\boldsymbol{\Sigma}}\Big(\nabla\boldsymbol{\theta}(\hat{\boldsymbol{\psi}})\Big)'\Big)^{-1}\boldsymbol{\theta}(\hat{\boldsymbol{\psi}}) \xrightarrow{D} \chi^2(r) \tag{104}$$

where r is the number of restrictions imposed by the null. Each of these steps is standard, and can be implemented by commonly-available software (e.g., the `gmm`, `nlcom`, and `testnl` commands in Stata).

Estimators based on sieve or other flexible approximations to the CEF can be constructed in a similar matter, and their asymptotic properties follow from results in Hansen (2014).

B.1.2 Estimating reallocation effects

Proposition 7 provides a starting point for estimating reallocation effects by a plug-in method:

$$\widehat{ARE}(\mathbf{G}_R) = \begin{cases} 0 & \text{if (PS, OS)} \\ \sum_{k=0}^{K} \hat{\mu}_k \widehat{CRE}_k(\mathbf{G}_R) & \text{if (PS, SA) or (singletons, SA)} \end{cases} \tag{105}$$

$$\widehat{CRE}_k(\mathbf{G}_R) = \begin{cases} \sum_{\ell=1}^{K} \Delta\bar{x}_{k\ell}(\mathbf{G}_R, \hat{\boldsymbol{\mu}})\widehat{CPE}_{k\ell}(n_0 - 1) & \text{if (PS, SA)} \\ \sum_{b=1}^{B} \Delta z_{kb}(\mathbf{G}_R, \hat{\boldsymbol{\mu}})\widehat{CGE}_{kb} & \text{if (singletons, SA)} \end{cases} \tag{106}$$

where:

$$\Delta\bar{x}_{k\ell}(\mathbf{G}_R, \hat{\boldsymbol{\mu}}) \equiv E\left(\bar{x}_{i\ell}\left(\mathbf{X}, \mathbf{G}_R(\mathbf{T}, \mathbf{G}, \boldsymbol{\rho})\right) \middle| \mathbf{x}_i = \mathbf{e}_k, \frac{\sum_{j\neq i}\mathbf{x}_j}{I-1} = \hat{\boldsymbol{\mu}}\right) - \hat{\mu}_\ell \tag{107}$$

$$\begin{aligned}\Delta z_{kb}(\mathbf{G}_R, \hat{\boldsymbol{\mu}}) \equiv\ & \Pr\left(\bar{\mathbf{x}}_i(\mathbf{X}, \mathbf{G}_R(\mathbf{T}, \mathbf{G}, \boldsymbol{\rho})) \in \mathbf{S}^b_{\bar{\mathbf{x}}} \middle| \mathbf{x}_i = \mathbf{e}_k, \frac{\sum_{j\neq i}\mathbf{x}_j}{I-1} = \hat{\boldsymbol{\mu}}\right) \\ & - \Pr\left(\bar{\mathbf{x}}_i(\tilde{\mathbf{p}}) \in \mathbf{S}^b_{\bar{\mathbf{x}}} \middle| \mathbf{x}_i = \mathbf{e}_k, \frac{\sum_{j\neq i}\mathbf{x}_j}{I-1} = \hat{\boldsymbol{\mu}}\right)\end{aligned} \tag{108}$$

Both $\Delta\bar{x}_{k\ell}(\mathbf{G}_R, \hat{\boldsymbol{\mu}})$ and $\Delta z_{kb}(\mathbf{G}_R, \hat{\boldsymbol{\mu}})$ can be calculated by enumeration, or approximated by simulation. The asymptotic properties of the estimators defined in (105) and (106) depend on the specific reallocation mechanism $\mathbf{G}_R$ chosen by the researcher, and how the resulting value of $\Delta\bar{x}_{k\ell}(\mathbf{G}_R, \hat{\boldsymbol{\mu}})$ and/or $\Delta z_{kb}(\mathbf{G}_R, \hat{\boldsymbol{\mu}})$ depends on $\hat{\boldsymbol{\mu}}$. For example, the delta method can be appplied if $\Delta\bar{x}_{kb}(\mathbf{G}_R, \hat{\boldsymbol{\mu}})$ is a differentiable function of $\hat{\boldsymbol{\mu}}$.

B.2 Variable group size and post-assignment shocks

To simplify exposition, the main results in this paper are established under the assumption that there are no post-assignment shocks (equation (8)) and that group size is constant (Assumption 6). This section relaxes those assumptions.

To accommodate post-assignment shocks, let $\boldsymbol{\epsilon} \equiv (\boldsymbol{\eta}, \boldsymbol{\nu}) \perp\!\!\!\perp \mathbf{T}, \mathbf{G}$, where $\boldsymbol{\eta} \in \mathbb{R}^I$ is a vector of IID individual-level (η_i affects individual i) shocks with finite support $\mathbb{S}_\eta$ and $\boldsymbol{\nu} \in \mathbb{R}^G$ is a vector of IID group-level (ν_g affects all individuals in group g) shocks with finite support $\mathbb{S}_\nu$. For convenience, assume that both $\mathbb{S}_\eta$ and $\mathbb{S}_\nu$ include zero. Replace equation (5) with:

$$\mathbf{Y} \equiv \begin{bmatrix} y_1 \\ \vdots \\ y_I \end{bmatrix} \equiv \begin{bmatrix} y_1(\mathbf{T}, \mathbf{G}, \boldsymbol{\epsilon}) \\ \vdots \\ y_I(\mathbf{T}, \mathbf{G}, \boldsymbol{\epsilon}) \end{bmatrix} \equiv \mathbf{Y}(\mathbf{T}, \mathbf{G}, \boldsymbol{\epsilon}) \qquad \text{where } \boldsymbol{\epsilon} \perp\!\!\!\perp \mathbf{T}, \mathbf{G} \tag{5$'$}$$

and replace equation (8) in Assumption 1 with:

$$y_i(\mathbf{T}, \mathbf{G}, \boldsymbol{\epsilon}) = y\left(\tau_i, \{\tau_j\}_{g_i = g_j}, \epsilon_i\right) \tag{8$'$}$$

where $\epsilon_i \equiv (\eta_i, \nu_{g_i})$. The model in the main text can be interpreted as a special case of this model in which $\boldsymbol{\epsilon}$ is constant ($\mathbb{S}_\eta = \mathbb{S}_\nu = \{0\}$) or does not affect the outcome.

With variable group size, the size of group g is a random variable $n_g = n(g, \mathbf{G})$. Let:

$$f_n(n) \equiv \Pr(n_{g_i} = n | n_{g_i} \geq 2) \tag{11$'$}$$

be the probability distribution of group size across individuals, excluding those individuals who have no peers, and let $\mathbb{S}_n \equiv \{n \in \mathbb{N} : f_n(n) > 0\}$ be the associated support. The model in the main text can be interpreted as a special case in which $f_n(n) = \mathbb{I}(n = n_0)$.

Given these modifications to the main model, several definitions can be generalized accordingly.

Definition 16 (Potential outcomes)**.** *Given Assumption 1, individual i's* **potential outcome function** *is defined as:*

$$y_i(\mathbf{p}) \equiv y\left(\tau_i, \{\tau_j\}_{j \in \mathbf{p}}, \epsilon_i\right) \tag{15$'$}$$

where $\mathbf{p}$ *is any subset of* $\mathcal{I} \setminus \{i\}$*.*

Definition 17 (Average peer effect)**.** *Given Assumptions 1 and 5, the* **average peer**

effect *in groups of size* $n \in \mathbb{S}_n$ *for peers with characteristics* $\mathbf{x}^p \in \mathbb{S}_{\mathbf{x}}$ *is:*

$$APE(\mathbf{x}^p, n) \equiv E\left(y_i(\{j\} \cup \tilde{\mathbf{p}}) - y_i(\{j'\} \cup \tilde{\mathbf{p}})\middle| \mathbf{x}_j = \mathbf{x}^p, \mathbf{x}_{j'} = \mathbf{e}_0\right) \tag{16a'}$$

where $\tilde{\mathbf{p}}$ *is a purely random draw of* $n-2$ *peers from* $\mathcal{I} \setminus \{i, j, j'\}$*, and the average peer effect of peers with characteristics* $\mathbf{x}^p$ *is:*

$$APE(\mathbf{x}^p) \equiv \sum_{n \in \mathbb{S}_n} APE(\mathbf{x}^p, n) f_n(n) \tag{16b'}$$

When $\mathbf{x}_i$ *is discrete (DC), the average peer effect from peers of observed type* ℓ *is:*

$$APE_\ell(n) \equiv APE(\mathbf{e}_\ell, n) \tag{17a'}$$

$$APE_\ell \equiv \sum_{n \in \mathbb{S}_n} APE_\ell(n) f_n(n) \tag{17b'}$$

where $APE(\cdot)$ *is as defined in equation (16a′).*

Definition 18 (Conditional peer effect)**.** *Given Assumptions 1 and 5, the* **conditional peer effect** *in groups of size* $n \in \mathbb{S}_n$ *from peers with characteristics* $\mathbf{x}^p \in \mathbb{S}_{\mathbf{x}}$ *on treated individuals with characteristics* $\mathbf{x}^o \in \mathbb{S}_{\mathbf{x}}$ *is:*

$$CPE(\mathbf{x}^o, \mathbf{x}^p, n) \equiv E\left(y_i\left(\{j\} \cup \tilde{\mathbf{p}}\right) - y_i\left(\{j'\} \cup \tilde{\mathbf{p}}\right)\middle| \mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p, \mathbf{x}_{j'} = \mathbf{e}_0\right) \tag{18a'}$$

where $\tilde{\mathbf{p}}$ *is a purely random draw of* $n-2$ *peers from* $\mathcal{I} \setminus \{i, j, j'\}$*, and the conditional peer effect of peers with characteristics* $\mathbf{x}^p$ *on treated individuals with characteristics* $\mathbf{x}^o$ *is:*

$$CPE(\mathbf{x}^o, \mathbf{x}^p) \equiv \sum_{n \in \mathbb{S}_n} CPE(\mathbf{x}^o, \mathbf{x}^p, n) f_n(n) \tag{18b'}$$

When $\mathbf{x}_i$ *is discrete (DC), the conditional peer effect is:*

$$CPE_{k\ell}(n) \equiv CPE(\mathbf{e}_k, \mathbf{e}_\ell, n) \tag{19a'}$$

$$CPE_{k\ell} \equiv \sum_{n \in \mathbb{S}_n} CPE_{k\ell}(n) f_n(n) \tag{19b'}$$

where $CPE(\cdot, \cdot)$ *is as defined in equation (18a′).*

The original definitions in the main text are a special case of these definitions:

- Equation (11′) applies with $f_n(n) = \mathbb{I}(n = n_0)$ and $\mathbb{S}_n = \{n_0\}$.
- Equations (16a') and (16b') reduce to equation (16).

- Equations ($17a'$) and ($17b'$) reduce to equation (17).
- Equations ($18a'$) and ($18b'$) reduce to equation (18).
- Equations ($19a'$) and ($19b'$) reduce to equation (19).

When group size exhibits nontrivial variation, a researcher can choose to report $APE_\ell(n)$ for selected values of n or to average across the group size distribution to get APE_ℓ.

Group effects can also be generalized, with the main complication being the absence of a natural base group.

Definition 19 (Group effects)**.** *Given Assumption 1, the* **average group effect** *of a peer group with characteristics* $\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}$ *is:*

$$AGE(\mathbf{x}^p) \equiv E\left(y_i(\tilde{\mathbf{p}}) - y_i(\tilde{\mathbf{q}}) \middle| \begin{array}{l} \{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}} = \mathbf{x}^p, \\ \mathbf{x}_j = \mathbf{e}_0 \textit{ for all } j \in \tilde{\mathbf{q}} \end{array} \right) \tag{22'}$$

the average group effect of a bin b peer group under the binning scheme $\mathbf{z}(\cdot)$ *is:*

$$AGE_b \equiv E\left(y_i(\tilde{\mathbf{p}}) - y_i(\tilde{\mathbf{q}}) \middle| \mathbf{z}\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right) = \mathbf{e}_b, \mathbf{z}\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right) = \mathbf{e}_0 \right) \tag{23'}$$

and the associated **conditional group effects** *are:*

$$CGE(\mathbf{x}^o, \mathbf{x}^p) \equiv E\left(y_i(\tilde{\mathbf{p}}) - y_i(\tilde{\mathbf{q}}) \middle| \begin{array}{l} \mathbf{x}_i = \mathbf{x}^o, \{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}} = \mathbf{x}^p \\ \mathbf{x}_j = \mathbf{e}_0 \textit{ for all } j \in \tilde{\mathbf{q}} \end{array} \right) \tag{24'}$$

$$CGE_{kb} \equiv E\left(y_i(\tilde{\mathbf{p}}) - y_i(\tilde{\mathbf{q}}) \middle| \begin{array}{l} \mathbf{x}_i = \mathbf{e}_k, \mathbf{z}\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right) = \mathbf{e}_b \\ \mathbf{z}\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right) = \mathbf{e}_0 \end{array} \right) \tag{25'}$$

where $\tilde{n}_p$ and $\tilde{n}_q$ are independent random draws from f_n and $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{q}}$ are independent random draws of $\tilde{n}_p - 1$ and $\tilde{n}_q - 1$ peers from $\mathcal{I} \setminus \{i\}$.

Definition 20 (Peer separability)**.** *Given Assumption 1, outcomes are* **peer-separable (PS′)** *if the effect of replacing one peer with another does not depend on one's other peers:*

$$y\left(\tau_i, \left\{\tau_j', \boldsymbol{\tau}\right\}, \epsilon\right) - y\left(\tau_i, \left\{\tau_j, \boldsymbol{\tau}\right\}, \epsilon\right) = y\left(\tau_i, \left\{\tau_j', \boldsymbol{\tau}'\right\}, \epsilon\right) - y\left(\tau_i, \left\{\tau_j, \boldsymbol{\tau}'\right\}, \epsilon\right) \tag{$\mathbf{PS'}$}$$

for any $a, b, b' \in \mathcal{T}$ and $\boldsymbol{\tau}, \boldsymbol{\tau}' \in M_{\mathcal{T}}$ such that $|\boldsymbol{\tau}| = |\boldsymbol{\tau}'|$, and for all $\epsilon \in \mathbb{S}_\epsilon$.

Definition 21 (Own separability)**.** *Given Assumption 1, outcomes are* **own-separable (OS′)** *if the effect of replacing one peer group with another does not depend on one's* *own type:*

$$y\left(\tau_i,\left\{\boldsymbol{\tau}'\right\},\epsilon\right)-y\left(\tau_i,\left\{\boldsymbol{\tau}\right\},\epsilon\right)=y\left(\tau_i',\left\{\boldsymbol{\tau}'\right\},\epsilon'\right)-y\left(\tau_i',\left\{\boldsymbol{\tau}\right\},\epsilon'\right) \qquad (\mathbf{OS}')$$

for all $a, a' \in \mathcal{T}$, *all* $\boldsymbol{\tau}, \boldsymbol{\tau}' \in M_{\mathcal{T}}$ *and all* $\epsilon, \epsilon' \in \mathbb{S}_\epsilon$.

Note that both sides of equation ($\mathbf{PS}'$) refer to the same treated individual, so both sides also have the same shocks. In contrast, each side of equation ($\mathbf{OS}'$) refers to a different treated individual, so each side has different shocks.

Given these modifications, Propositions 9–12 and Lemma 2 show that the results in the main text generalize in a straightforward manner.

Proposition 9 (Aggregation with variable group size and post-assignment shocks)**.** *Given Assumptions 1–5 and discrete characteristics (DC):*

1. *Average effects are a weighted average of conditional effects:*

$$APE_\ell(n)=\sum_{k=0}^{K}\mu_k CPE_{k\ell}(n) \qquad (109)$$

$$APE_\ell=\sum_{k=0}^{K}\mu_k CPE_{k\ell} \qquad (30')$$

$$AGE_b=\sum_{k=0}^{K}\mu_k CGE_{kb} \qquad (31')$$

where $\mu_k = E(x_{ik}) = \Pr(\mathbf{x}_i = \mathbf{e}_k)$ *as defined earlier.*

2. *Binned group effects are a weighted average of saturated group effects:*

$$AGE_b=\sum_{k=0}^{K}\sum_{s=1}^{S}\mu_k\left(w_{sb}^G(\boldsymbol{\mu})-w_{s0}^G(\boldsymbol{\mu})\right)CGE_{ks}^S \qquad (33')$$

$$CGE_{kb}=\sum_{s=1}^{S}\left(w_{sb}^G(\boldsymbol{\mu})-w_{s0}^G(\boldsymbol{\mu})\right)CGE_{ks}^S \qquad (34')$$

where CGE_{ks}^S *is the conditional group effect for bin* s *of the saturated variable* $\mathbf{z}_i^S$,

$$w_{sb}^G(\boldsymbol{\mu})=\frac{\sum_{\mathbf{x}^p\in\mathbb{S}_{\{\mathbf{x}\}}}M\left(\bar{\mathbf{x}}\left(\mathbf{x}^p\right),|\mathbf{x}^p|,\boldsymbol{\mu}\right)\mathbb{I}\left(\mathbf{z}^S\left(\bar{\mathbf{x}}\left(\mathbf{x}^p\right)\right)=\mathbf{e}_s\right)\mathbb{I}\left(\mathbf{z}\left(\mathbf{x}^p\right)=\mathbf{e}_b\right)f_n(|\mathbf{x}^p|+1)}{\sum_{\mathbf{x}^p\in\mathbb{S}_{\{\mathbf{x}\}}}M(\bar{\mathbf{x}}\left(\mathbf{x}^p\right),|\mathbf{x}^p|,\boldsymbol{\mu})\mathbb{I}\left(\mathbf{z}\left(\mathbf{x}^p\right)=\mathbf{e}_b\right)f_n(|\mathbf{x}^p|+1)} \qquad (35')$$

is a weighting function, and:

$$M(\bar{\mathbf{x}}, n, \boldsymbol{\mu}) \equiv \frac{n!}{\prod_{k=0}^{K}(n\bar{x}_{\cdot k})!} \prod_{k=0}^{K} \mu_k^{n\bar{x}_{\cdot k}} \tag{36'}$$

is the probability of drawing the value $n\bar{\mathbf{x}}$ from a multinomial distribution with n trials and categorical probability vector $\boldsymbol{\mu}$.

3. *Peer effects are a weighted average of saturated group effects:*

$$APE_\ell = \sum_{k=0}^{K} \sum_{s=1}^{S} \mu_k \left(w_{s\ell}^P(\boldsymbol{\mu}) - w_{s0}^P(\boldsymbol{\mu}) \right) CGE_{ks}^S \tag{37'}$$

$$CPE_{k\ell} = \sum_{s=1}^{S} \left(w_{s\ell}^P(\boldsymbol{\mu}) - w_{s0}^P(\boldsymbol{\mu}) \right) CGE_{ks}^S \tag{38'}$$

where:

$$w_{s\ell}^P(\boldsymbol{\mu}) \equiv \sum_{\mathbf{x}^p \in \mathcal{M}_{\mathbf{x}}} M\left(\bar{\mathbf{x}}\left(\mathbf{x}^p\right), |\mathbf{x}^p|, \boldsymbol{\mu}\right) \mathbb{I}\left(\mathbf{z}^S\left(\mathbf{x}^p \cup \{\mathbf{e}_\ell\}\right) = \mathbf{e}_s\right) f_n(|\mathbf{x}^p| + 2) \tag{39'}$$

is a weighting function and $\mathcal{M}_{\mathbf{x}}$ is the set of multisets on $\mathbb{S}_{\mathbf{x}}$.

Lemma 2 (Implications of stratified random assignment with variable group size and post-assignment shocks)**.** *Given Assumptions 1–5 and discrete characteristics (DC), stratified random assignment (SA) implies:*

$$E(y_i | \mathbf{x}_i = \mathbf{x}^o, \bar{\mathbf{x}}_i = \mathbf{x}^p, n_{g_i} = n) = E\left(y_i(\tilde{\mathbf{p}}) \middle| \mathbf{x}_i = \mathbf{x}^o, \bar{\mathbf{x}}\left(\{\mathbf{x}_j\}_{j \in \tilde{\mathbf{p}}} \right) = \mathbf{x}^p \right) \tag{42'}$$

where $\tilde{\mathbf{p}}$ is a purely random draw of $(n-1)$ peers from $\mathcal{I} \setminus \{i\}$.

Proposition 10 (Implications of separability with variable group size and post-assignment shocks)**.** *Given Assumptions 1–5, let $PE : \mathcal{T}^2 \times \mathbb{R}^2 \times \{2, 3, \ldots\} \to \mathbb{R}$ be defined:*

$$PE(a, b, e, n) \equiv \frac{y\left(a, \left\{b^{[n-1]}\right\}, e\right)}{n-1} \tag{43'}$$

where $b^{[n-1]}$ is $n-1$ copies of b. If outcomes are peer-separable (PS), then for any $\mathbf{p}$ of size $|\mathbf{p}| \geq 1$:

$$y_i(\mathbf{p}) = \sum_{j \in \mathbf{p}} PE(\tau_i, \tau_j, \epsilon_i, |\mathbf{p}| + 1) \qquad \text{for all } \mathbf{p} \subset \mathcal{I} \setminus \{i\} \tag{44'}$$

Conditional and average peer effects can also be expressed in terms of these pairwise *latent variables:*

$$CPE(\mathbf{x}^o, \mathbf{x}^p, n) = E\left(PE(\tau_i, \tau_j, \epsilon_i, n)|\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p\right) \qquad \textit{for all } n \geq 2 \tag{45'}$$
$$- E\left(PE(\tau_i, \tau_j, \epsilon_i, n)|\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{e}_0\right)$$
$$APE(\mathbf{x}^p, n) = E\left(PE(\tau_i, \tau_j, \epsilon_i, n)|\mathbf{x}_j = \mathbf{x}^p\right) \qquad \textit{for all } n \geq 2 \tag{46'}$$
$$- E\left(PE(\tau_i, \tau_j, \epsilon_i, n)|\mathbf{x}_j = \mathbf{e}_0\right)$$

for all $\mathbf{x}^o, \mathbf{x}^p \in \mathbb{S}_{\mathbf{x}}$*.*

Proposition 11 (Identification of peer effects with variable group size and post-assignment shocks)**.** *Given Assumptions 1–5 and discrete characteristics (DC):*

1. *Simple random assignment (RA) and peer separability (PS) imply that peer effects are identified from the joint distribution of* $(y_i, \mathbf{x}_i, \bar{\mathbf{x}}_i, n_{g_i})$*:*

$$APE_\ell(n) = \frac{\alpha^n_{1\ell}}{n-1} \qquad \textit{for all } n \in \mathbb{S}_n \setminus \{0, 1\} \tag{49a'}$$
$$APE_\ell = \sum_{n \in \mathbb{S}_n} \frac{\alpha^n_{1\ell}}{n-1} f_n(n) \tag{49b'}$$
$$CPE_{k\ell}(n) = \frac{\beta^n_{2\ell} + \beta^n_{3k\ell}}{n-1} \qquad \textit{for all } n \in \mathbb{S}_n \tag{50a'}$$
$$CPE_{k\ell} = \sum_{n \in \mathbb{S}_n} \frac{\beta^n_{2\ell} + \beta^n_{3k\ell}}{n-1} f_n(n) \tag{50b'}$$

where $\boldsymbol{\alpha}^n = (\alpha^n_0, \boldsymbol{\alpha}^n_1)$ *and* $\boldsymbol{\beta}^n = (\beta^n_0, \boldsymbol{\beta}^n_1, \boldsymbol{\beta}^n_2, \boldsymbol{\beta}^n_3)$ *are the vectors of best linear predictor coefficients:*

$$L(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i; n_{g_i}) \equiv \sum_{n \in \mathbb{S}_n} (\alpha^n_0 + \bar{\mathbf{x}}_i \boldsymbol{\alpha}^n_1)\mathbb{I}\left(n_{g_i} = n\right) \tag{51'}$$
$$L(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i'\bar{\mathbf{x}}_i; n_{g_i}) \equiv \sum_{n \in \mathbb{S}_n} (\beta^n_0 + \mathbf{x}_i\boldsymbol{\beta}^n_1 + \bar{\mathbf{x}}_i\boldsymbol{\beta}^n_2 + \mathbf{x}_i\boldsymbol{\beta}^n_3\bar{\mathbf{x}}_i')\mathbb{I}\left(n_{g_i} = n\right) \tag{52'}$$

in the sub-population of individuals[8] *with* $n_{g_i} = n$*.*

2. *Stratified random assignment (SA) and peer separability (PS) imply that peer*

[8]That is: $\boldsymbol{\alpha}(n) \equiv E(\mathbf{d}_i'\mathbf{d}_i|n_{g_i} = n)^{-1} E(\mathbf{d}_i' y_i|n_{g_i} = n)$ where $\mathbf{d}_i \equiv (1, \mathbf{x}_i, \bar{\mathbf{x}}_i)$.

effects are identified from the joint distribution of $(y_i, \mathbf{x}_i, \bar{\mathbf{x}}_i, n_{g_i})$*:*

$$APE_\ell(n) = \sum_{k=0}^{K} \mu_k \frac{\beta_{2\ell}^n + \beta_{3k\ell}^n}{n-1} \tag{53a'}$$

$$APE_\ell = \sum_{n\in\mathbb{S}_n} \sum_{k=0}^{K} \mu_k \frac{\beta_{2\ell}^n + \beta_{3k\ell}^n}{n-1} f_n(n) \tag{53b'}$$

$$CPE_{k\ell}(n) = \frac{\beta_{2\ell}^n + \beta_{3k\ell}^n}{n-1} \tag{54a'}$$

$$CPE_{k\ell} = \sum_{n\in\mathbb{S}_n} \frac{\beta_{2\ell}^n + \beta_{3k\ell}^n}{n-1} f_n(n) \tag{54b'}$$

where $(\beta_{2\ell}^n, \beta_{3k\ell}^n)$ *are defined as in equation (52′).*

Proposition 12 (Identification of group effects with variable group size and post-assignment shocks)**.** *Given Assumptions 1–5 and discrete characteristics (DC):*

1. *Simple random assignment (RA) implies that binned group effects are identified from the joint distribution of* $(y_i, \mathbf{x}_i, \mathbf{z}_i)$*:*

$$AGE_b = \gamma_{1b} \tag{55'}$$

$$CGE_{kb} = \delta_{2b} + \delta_{3kb} \tag{56'}$$

where $(\gamma_{1b}, \delta_{2b}, \delta_{3kb})$ *are coefficients from the best linear predictors:*

$$L(y_i|\mathbf{x}_i, \mathbf{z}_i) \equiv \gamma_0 + \mathbf{z}_i\boldsymbol{\gamma}_1 \tag{57'}$$

$$L(y_i|\mathbf{x}_i, \mathbf{z}_i, \mathbf{x}_i'\mathbf{z}_i) \equiv \delta_0 + \mathbf{x}_i\boldsymbol{\delta}_1 + \mathbf{z}_i\boldsymbol{\delta}_2 + \mathbf{x}_i\boldsymbol{\delta}_3\mathbf{z}_i' \tag{58'}$$

i.e., γ_{1b} *is element* b *of* $\boldsymbol{\gamma}_1$*,* δ_{2b} *is element* b *of* $\boldsymbol{\delta}_2$*,* δ_{3kb} *is the element in row* k *and column* b *of* $\boldsymbol{\delta}_3$ *for all* $k > 0$*, and* $\delta_{30b} \equiv 0$ *for all* b*.*

2. *Stratified random assignment (SA) implies that saturated group effects are identified from the joint distribution of* $(y_i, \mathbf{x}_i, \mathbf{z}_i^S)$*:*

$$CGE_{ks}^S = \lambda_{2s} + \lambda_{3ks} \tag{59'}$$

where $(\lambda_{2s}, \lambda_{3ks})$ *are coefficients from the best linear predictor:*

$$L(y_i|\mathbf{x}_i, \mathbf{z}_i^S, \mathbf{x}_i'\mathbf{z}_i^S) \equiv \lambda_0 + \mathbf{x}_i\boldsymbol{\lambda}_1 + \mathbf{z}_i^S\boldsymbol{\lambda}_2 + \mathbf{x}_i\boldsymbol{\lambda}_3\mathbf{z}_i^{S\prime} \tag{60'}$$

i.e., λ_{2s} *is element* s *of* $\boldsymbol{\lambda}_2$*,* λ_{3ks} *is the element in row* k *and column* s *of* $\boldsymbol{\lambda}_3$ *for all* $k > 0$*, and* $\lambda_{30s} \equiv 0$ *for all* s*. Peer effects and binned group effects are also* *identified:*

$$AGE_b = \sum_{k=0}^{K} \sum_{s=1}^{S} \mu_k \left(w_{sb}^G(\boldsymbol{\mu}) - w_{s0}^G(\boldsymbol{\mu}) \right) (\lambda_{2s} + \lambda_{3ks}) \tag{61$'$}$$

$$CGE_{kb} = \sum_{s=1}^{S} \left(w_{sb}^G(\boldsymbol{\mu}) - w_{s0}^G(\boldsymbol{\mu}) \right) (\lambda_{2s} + \lambda_{3ks}) \tag{62$'$}$$

$$APE_\ell = \sum_{k=0}^{K} \sum_{s=1}^{S} \mu_k \left(w_{s\ell}^P(\boldsymbol{\mu}) - w_{s0}^P(\boldsymbol{\mu}) \right) (\lambda_{2s} + \lambda_{3ks}) \tag{63$'$}$$

$$CPE_{k\ell} = \sum_{s=1}^{S} \left(w_{s\ell}^P(\boldsymbol{\mu}) - w_{s0}^P(\boldsymbol{\mu}) \right) (\lambda_{2s} + \lambda_{3ks}) \tag{64$'$}$$

where $w_{sb}^G(\cdot)$ and $w_{s\ell}^P(\boldsymbol{\mu})$ are defined in Proposition 9.

In principle, the group size specific regression coefficients described in Propositions 11 and 12 are identified and can be estimated from data by linear regression. However, estimating such a model is likely to be impractical in most applications. An obvious alternative is to impose plausible assumptions on how $\boldsymbol{\alpha}^n$ varies with n.

Example 18 (Gender peer effects with variable group size)**.** *Table 7 below shows estimates of gender peer effects that account for variations in group size. Each column shows a different specification for $\boldsymbol{\alpha}^n$ and reports selected coefficients and average peer effect estimates for two selected group sizes $APE_1(16)$ and $APE_1(20)$ and averaging over the distribution of group sizes APE_1.*

- *Columns (1) and (6) show results for the regression model:*

$$L(y_i|\bar{\mathbf{x}}_i) = \alpha_0 + \bar{\mathbf{x}}_i \boldsymbol{\alpha}_1 \tag{110}$$

as well as average peer effects $APE_1(16) = \frac{\alpha_1}{16-1}$, $APE_1(20) = \frac{\alpha_1}{20-1}$, and $APE_1 = \alpha_1 E\left(\frac{1}{n_{g_i}-1}\right)$ under the assumption $(\alpha_0^n, \boldsymbol{\alpha}_1^n) = (\alpha_0, \boldsymbol{\alpha}_1)$. This model assumes that the effect of peer group composition does not vary with peer group size, and corresponds to the typical handling of group size variation in the literature.

- *Columns (2) and (7) show results for the regression model:*

$$L(y_i|n_{g_i}, (n_{g_i}-1)\bar{\mathbf{x}}_i) = \alpha_{00} + \alpha_0 n_{g_i} + (n_{g_i}-1)\bar{\mathbf{x}}_i \boldsymbol{\alpha}_1 \tag{111}$$

and average peer effects $APE_1(16) = APE_1(20) = APE_1 = \alpha_1$ under the assumption $(\alpha_0^n, \boldsymbol{\alpha}_1^n) = (\alpha_{00} + \alpha_0 n, \boldsymbol{\alpha}_1(n-1))$. This model assumes that the additive effect of an individual peer $(PE(\tau_i, \tau_j, n))$ does not vary with peer group size.

- *Columns (3) and (8) show results for the regression model:*

$$L(y_i|n_{g_i},(n_{g_i}-1)\bar{\mathbf{x}}_i,\bar{\mathbf{x}}_i)=\alpha_{00}+\alpha_0 n_{g_i}+(n_{g_i}-1)\bar{\mathbf{x}}_i\boldsymbol{\alpha}_1+\bar{\mathbf{x}}_i\boldsymbol{\alpha}_2 \tag{112}$$

  and average peer effects $APE_1(16)=\alpha_1+\frac{\alpha_2}{16-1}$, $APE_1(20)=\alpha_1+\frac{\alpha_2}{20-1}$ *and* $APE_1=\alpha_1+\alpha_2 E\left(\frac{1}{n_{g_i}-1}\right)$ *under the assumption* $(\alpha_0^n,\boldsymbol{\alpha}_1^n)=(\alpha_{00}+\alpha_0 n,\boldsymbol{\alpha}_1(n-1)+\boldsymbol{\alpha}_2)$. *This model nests models (110) and (111).*

- *Columns (4) and (9) show results for the mostly-unrestricted regression model:*

$$L(y_i|n_{g_i},\bar{\mathbf{x}}_i;n_{g_i})=\alpha_{00}+\alpha_0 n_{g_i}+\sum_{n\in\mathbb{S}_n}(\bar{\mathbf{x}}_i\boldsymbol{\alpha}_1^n)\,\mathbb{I}\,(n_{g_i}=n) \tag{113}$$

  and average peer effects $APE_1(16)=\frac{\alpha_1^{16}}{16-1}$, $APE_1(20)=\frac{\alpha_1^{20}}{20-1}$, *and* $APE=\sum_{n\in\mathbb{S}_n}\frac{\alpha_1^n}{n-1}f_n(n)$ *under the assumption* $(\alpha_0^n,\boldsymbol{\alpha}_1^n)=(\alpha_{00}+\alpha_0 n,\boldsymbol{\alpha}_1^n)$. *This model allows average peer effects to vary arbitrarily with group size, but restricts the other coefficients for tractability.*

- *Columns (5) and (10) show results for the unrestricted regression model:*

$$L(y_i|\bar{\mathbf{x}}_i;n_{g_i})=\sum_{n\in\mathbb{S}_n}(\alpha_0^n+\bar{\mathbf{x}}_i\boldsymbol{\alpha}_1^n)\,\mathbb{I}\,(n_{g_i}=n) \tag{114}$$

  and average peer effects $APE_1(16)=\frac{\alpha_1^{16}}{16-1}$, $APE_1(20)=\frac{\alpha_1^{20}}{20-1}$, *and* $APE=\sum_{n\in\mathbb{S}_n}\frac{\alpha_1^n}{n-1}f_n(n)$

As the results show, the first three models produce very similar average peer effects either for representative peer group sizes or when averaging across the group size distribution. While the unrestricted model produces highly variable estimates for specific group sizes, averaging across the size distribution produces results similar to those found in the other specifications.

B.3 More on approximation and functional form

Section 5.5 in the main text develops a sieve-based method for estimating average peer effects that exploits the dimension-reducing implications of random assignment and peer separability while allowing for general (continuous) characteristics. This appendix shows how similar methods can be applied to more complex cases.

	Reading score					Math score				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Share male peers	−0.342		-0.670			−0.532**		0.011		
	(0.286)		(1.012)			(0.226)		(0.994)		
Share male peers x (# peers)		−0.019	0.024				−0.037**	−0.038		
		(0.019)	(0.068)				(0.015)	(0.066)		
Share male peers x (15 peers)				-0.417	0.211				-0.956***	-0.163
				(0.346)	(0.784)				(0.342)	(0.765)
Share male peers x (19 peers)				-0.023	-1.996**				-0.528	-1.391
				(0.434)	(0.881)				(0.404)	(1.064)
Average peer effects:										
$APE(16)$	-0.023	−0.019	−0.021	-0.028	0.014	−0.035**	−0.037**	−0.037**	-0.064***	-0.011
	(0.019)	(0.019)	(0.019)	(0.023)	(0.052)	(0.015)	(0.015)	(0.015)	(0.023)	(0.051)
$APE(20)$	-0.018	−0.019	−0.012	-0.001	-0.105**	−0.028**	−0.037**	−0.037*	-0.028	-0.073
	(0.015)	(0.019)	(0.023)	(0.023)	(0.046)	(0.012)	(0.015)	(0.020)	(0.021)	(0.056)
APE	-0.024	−0.019	-0.024	-0.021	-0.037*	−0.038**	−0.037**	−0.037**	−0.032*	-0.039**
	(0.020)	(0.019)	(0.020)	(0.020)	(0.019)	(0.016)	(0.015)	(0.016)	(0.017)	(0.017)
Additional control variables:										
School/grade fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Own gender	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# peers	No	No	Linear	Linear	Dummy	No	No	Linear	Linear	Dummy
Sample size (# students)	2,185	2,185	2,185	2,185	2,185	2,196	2,196	2,196	2,196	2,196
# clusters	147	147	147	147	147	148	148	148	148	148

Table 7: Gender peer effects in Project STAR, by size of peer group. Cluster-robust standard errors in parentheses, $^{*} = 0.1, ^{**} = 0.05, ^{***} = 0.01$.

B.3.1 Sieve models for conditional peer effects

The dimension-reducing implications of separability and random assignment can also be exploited to estimate conditional peer effects. Proposition 13 below extends the sieve estimator to this case.

Proposition 13 (Sieve model for conditional peer effects)**.** *Given Assumptions 1–6 and peer separability (PS), suppose that:*

$$h(\mathbf{x}^o, \mathbf{x}^p) \equiv E\left(PE_{ij} \,|\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p\right) \approx h_m(\mathbf{x}^o, \mathbf{x}^p)\bar{\boldsymbol{\phi}}_m \tag{115}$$

for some known function $h_m : \mathbb{R}^{2K} \rightarrow \mathbb{R}^m$ *and unknown parameter vector* $\bar{\boldsymbol{\phi}}_m \equiv E\left(h_m(\mathbf{x}_i, \mathbf{x}_j)' h_m(\mathbf{x}_i, \mathbf{x}_j)\right)^{-1} E\left(h_m(\mathbf{x}_i, \mathbf{x}_j)' h(\mathbf{x}_i, \mathbf{x}_j)\right)^{-1}$*. Let:*

$$\boldsymbol{\phi} \equiv E(\bar{\mathbf{h}}_i' \bar{\mathbf{h}}_i)^{-1} E(\bar{\mathbf{h}}_i' y_i) \qquad \textit{where } \bar{\mathbf{h}}_i \equiv \frac{1}{n_0 - 1} \sum_{j \in \mathbf{p}_i} h_m(\mathbf{x}_i, \mathbf{x}_j) \tag{116}$$

Then simple random assignment (RA) implies:

$$E\left(y_i \,\middle|\, \mathbf{x}_i, \{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) = \sum_{j\in\mathbf{p}_i} h(\mathbf{x}_i,\mathbf{x}_j) \approx \bar{\mathbf{h}}_i \boldsymbol{\phi} \tag{117}$$

$$CPE(\mathbf{x}^o,\mathbf{x}^p) = h(\mathbf{x}^o,\mathbf{x}^p) - h(\mathbf{x}^o,\mathbf{0}) \approx \left(\frac{h_m(\mathbf{x}^o,\mathbf{x}^p) - h_m(\mathbf{x}^o,\mathbf{0})}{n_0 - 1}\right)\boldsymbol{\phi} \tag{118}$$

where the approximation errors in (117) and (118) are proportional to the approximation error in (115).

B.3.2 Sieve models for group effects

Sieve methods can also be used to estimate group effects, providing a practical alternative to saturated models. The dimension reduction implied by equation (75) is no longer available in the absence of peer separability, so the unrestricted conditional expectation function must be directly approximated. Proposition 14 below describes sieve estimators, but other standard nonparametric regression methods can be used.

Proposition 14 (Sieve model for group effects)**.** *Given Assumptions 1–6, let:*

$$a(\mathbf{x}^p) \equiv E\left(y_i \,\middle|\, \{\mathbf{x}_j\}_{j\in\mathbf{p}_i} = \mathbf{x}^p\right) \approx a_m(\mathbf{x}^p)\boldsymbol{\pi}_m \tag{119}$$

$$h(\mathbf{x}^o,\mathbf{x}^p) \equiv E\left(y_i \,\middle|\, \mathbf{x}_i = \mathbf{x}^o, \{\mathbf{x}_j\}_{j\in\mathbf{p}_i} = \mathbf{x}^p\right) \approx h_m(\mathbf{x}^o,\mathbf{x}^p)\boldsymbol{\phi}_m \tag{120}$$

for known functions $a_m : \mathbb{S}_{\{\mathbf{x}\}} \to \mathbb{R}^m$ and $h_m : (\mathbb{S}_{\mathbf{x}} \times \mathbb{S}_{\{\mathbf{x}\}}) \to \mathbb{R}^m$ and unknown vectors:

$$\boldsymbol{\pi}_m \equiv E\left(a_m\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right)' a_m\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right)\right)^{-1} E\left(a_m\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right)' a\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right)\right)$$

$$\boldsymbol{\phi}_m \equiv E\left(h_m\left(\mathbf{x}_i,\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right)' h_m\left(\mathbf{x}_i,\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right)\right)^{-1} E\left(h_m\left(\mathbf{x}_i,\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right)' h\left(\mathbf{x}_i,\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right)\right)$$

Then random assignment (RA) implies:

$$\begin{aligned} AGE(\mathbf{x}^p) &= a(\mathbf{x}^p) - a(\{\mathbf{e}_0,\ldots,\mathbf{e}_0\}) && (121)\\ &\approx (a_m(\mathbf{x}^p) - a_m(\{\mathbf{e}_0,\ldots,\mathbf{e}_0\}))\boldsymbol{\pi}_m \\ CGE(\mathbf{x}^o,\mathbf{x}^p) &= h(\mathbf{x}^o,\mathbf{x}^p) - h(\mathbf{x}^o,\{\mathbf{e}_0,\ldots,\mathbf{e}_0\}) && (122)\\ &\approx (h_m(\mathbf{x}^o,\mathbf{x}^p) - h_m(\mathbf{x}^o,\{\mathbf{e}_0,\ldots,\mathbf{e}_0\}))\boldsymbol{\phi}_m \end{aligned}$$

where the approximation errors in (121) and (122) are proportional to the approximation errors in (119) and (120) respectively.

As in Section 5.5, the sieve framework can be used to interpret ad-hoc functional form assumptions, or it can be used in combination with data-driven model selection and averaging to produce a flexible model that is nonparametric in the limit.

Example 19 (Average group effects for peer gender)**.** *Figure 3 below shows estimates of average group effects for share of male classmates. The specifications include:*

- *The linear model from columns (1) and (5) of Table 1:*

$$a_2\left(\{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) = \left[\begin{array}{cc} 1 & \bar{\mathbf{x}}_i \end{array}\right] \tag{123}$$

- *Linear splines with 1, 2, and 3 knots (the linear model is a zero-knot spline).*

$$a_{k+2}\left(\{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) = \left[\begin{array}{cccc} 1 & \bar{\mathbf{x}}_i & (\bar{\mathbf{x}}_i - knot_1)\mathbb{I}\left(\bar{\mathbf{x}}_i > knot_1\right) & \ldots \end{array}\right] \tag{124}$$

- *The 3-bin model from columns (3) and (7) of Table 1.*

$$a_3\left(\{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) = \left[\begin{array}{ccc} 1 & \mathbb{I}\left(\bar{\mathbf{x}}_i < 0.43\right) & \mathbb{I}\left(\bar{\mathbf{x}}_i > 0.57\right) \end{array}\right] \tag{125}$$

- *The jackknife model average (JMA) of the four linear spline models.*

The first panel in Table 8 shows model selection and weighting statistics for each specification. The 1-knot spline is the preferred model for the reading score by both AIC and cross-validation. In comparison to the linear model, the more flexible models imply a low marginal effect of boys when girls are in the majority, and a stronger negative effect when boys are in the majority. In contrast, the linear model is the preferred model for math, and makes similar predictions to those from richer models. The second panel in Table 8 is discussed in Example 20 below.

B.3.3 Sieve models for reallocation effects

In the absence of separability, conditional group effect estimates from a saturated model are required to recover many reallocation effects of interest. Unfortunately, estimating such richly parameterized models is rarely practical with limited data. Sieve approximations can provide a practical middle ground in this setting.

Example 20 (Sieve models of conditional group effects and reallocation effects)**.** *The second panel in Table 8 shows model selection and weighting statistics for conditional group effects models of the form:*

$$h_{2m}\left(\mathbf{x}_i, \{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) = \left[\begin{array}{cc} a_m\left(\{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) & \mathbf{x}_i a_m\left(\{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) \end{array}\right] \tag{126}$$

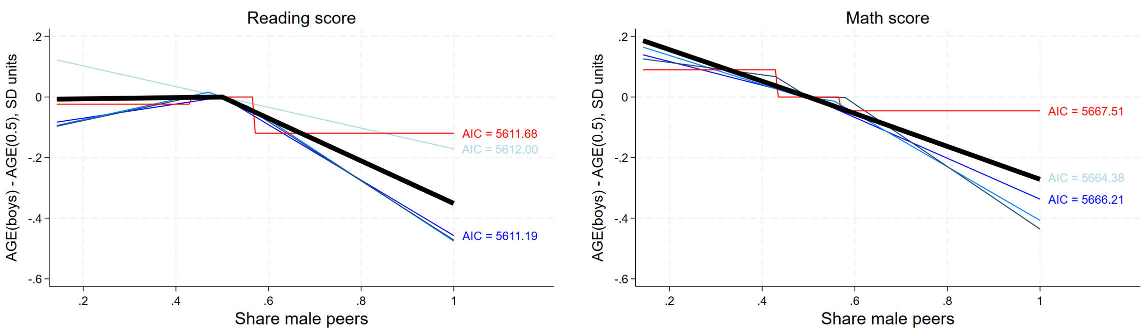

Figure 3: Average group effects for classmate gender relative to an evenly-mixed classroom, i.e., $AGE(\bar{\mathbf{x}}) - AGE(0.5)$. Specifications include 3-bin model (red) and linear spline with zero to three knots (blue). Jackknife model average of splines is depicted by thick black line.

	Reading score					Math score				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Model type	spline/ linear	spline	spline	spline	binned	spline/ linear	spline	spline	spline	binned
# knots/bins	0	1	2	3	3	0	1	2	3	3
AGE model selection:										
Model order (m)	2	3	4	5	3	2	3	4	5	3
AIC statistic	5612.00	5611.19	5612.98	5615.15	5611.68	5664.38	5666.21	5667.97	5669.67	5667.51
Leave-one-out MSE	0.8890	0.8885	0.8893	0.8903	0.8889	0.8976	0.8984	0.8993	0.9001	0.8987
JMA weight	0.3702	0.6298	0.0000	0.0000		0.9180	0.0820	0.0000	0.0000	
CGE model selection:										
Model order (m)	4	6	8	10	6	4	6	8	10	6
AIC statistic	5605.97	5605.97	5609.56	5613.78	5604.45	5658.83	5661.73	5665.87	5668.04	5661.00
Leave-one-out MSE	0.8864	0.8864	0.8881	0.8901	0.8859	0.8974	0.8986	0.9006	0.9014	0.8978
JMA weight	0.5945	0.4055	0.0000	0.0000		1.0000	0.0000	0.0000	0.0000	

Table 8: Model selection/weights for sieve estimates of group effects for classmate gender. See Examples 19 and 20 for model definitions and additional background.

for each $a_m(\cdot)$ model described in Example 19. Results are similar to those for Example 19. JMA weights put roughly equal weight on the linear and one-knot splines for reading, and all weight on the linear spline for math.

Table 9 shows reallocation effect estimates based on the JMA estimates. For reading, the JMA estimates imply a somewhat larger benefit from making classrooms more gender-balanced than is implied by the separable model estimates in Table 3. For math, the JMA estimates are identical to the separable model estimates since the data-driven JMA weights put all weight on the linear model.

			Reading score			Math score		
Reallocation	$\Delta\bar{x}_{01}$	$\Delta\bar{x}_{11}$	CRE_0	CRE_1	ARE	CRE_0	CRE_1	ARE
JMA of splines:								
Evenly-divided	0.033	−0.033	−0.002	0.038*	0.018	−0.009	0.024***	0.008
6/10 divided	0.000	0.000	−0.002	−0.007*	−0.004	0.000	0.000	0.000
Random	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60/40 random	−0.020	0.020	0.003	−0.018*	−0.007	0.006	−0.015***	−0.005
Single-gender	−0.500	0.500	0.100	−0.332*	−0.116	0.139	−0.367***	−0.114

Table 9: Reallocation effects for classmate gender in Project STAR using jackknife model average (JMA) of linear spline models. JMA weights are reported in the second panel of Table 8. Reallocation effects for simple random assignment are reported for comparison and are always zero. Cluster-robust p-values: * = 0.1, ** = 0.05, *** = 0.01.

C Proofs

This section provides proofs for all propositions in the main text and appendices.

Proof for Proposition 1

The conditions for Proposition 9 are met, so its results apply.

- (30) is a restatement of (30′) in Proposition 9.
- (31) is a restatement of (31′).

- (32) follows directly from the definitions:

$$\begin{aligned}
ARE(\mathbf{G}_R) &= E(y_i(\mathbf{p}(i,\mathbf{G}_R(\mathbf{X},\boldsymbol{\rho}))) - y_i(\tilde{\mathbf{p}})) && \text{(definition of } ARE\text{)} \\
&= \sum_{k=0}^{K} E\left(y_i(\mathbf{p}(i,\mathbf{G}_R(\mathbf{X},\boldsymbol{\rho}))) - y_i(\tilde{\mathbf{p}}) \middle| \mathbf{x}_i = \mathbf{e}_k\right) \Pr\left(\mathbf{x}_i = \mathbf{e}_k\right) \\
&= \sum_{k=0}^{K} \mu_k CRE_k(\mathbf{G}_R) && \text{(definition of } \mu \text{ and } CRE\text{)}
\end{aligned}$$

- (33) is a restatement of (33′).
- (34) is a restatement of (34′).
- (35′) simplifies to (35) since Assumption 6 implies that $|\mathbf{x}^p| = n_0 - 1$ and $f_n(|\mathbf{x}^p| + 1) = 1$ for all $\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}$. That is:

$$\begin{aligned}
w_{sb}^{G}(\boldsymbol{\mu}) &= \frac{\sum_{\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}} M\left(\bar{\mathbf{x}}(\mathbf{x}^p), |\mathbf{x}^p|, \boldsymbol{\mu}\right) f_n(|\mathbf{x}^p|+1)\mathbb{I}\left(\mathbf{z}^S(\mathbf{x}^p) = \mathbf{e}_s\right)\mathbb{I}\left(\mathbf{z}(\mathbf{x}^p) = \mathbf{e}_b\right)}{\sum_{\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}} M\left(\bar{\mathbf{x}}(\mathbf{x}^p), |\mathbf{x}^p|, \boldsymbol{\mu}\right) f_n(|\mathbf{x}^p|+1)\mathbb{I}\left(\mathbf{z}(\mathbf{x}^p) = \mathbf{e}_b\right)} && \text{(by (35′))} \\
&= \frac{\sum_{\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}} M\left(\bar{\mathbf{x}}(\mathbf{x}^p), n_0 - 1, \boldsymbol{\mu}\right) \mathbb{I}\left(\mathbf{z}^S(\mathbf{x}^p) = \mathbf{e}_s\right)\mathbb{I}\left(\mathbf{z}(\mathbf{x}^p) = \mathbf{e}_b\right)}{\sum_{\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}} M\left(\bar{\mathbf{x}}(\mathbf{x}^p), n_0 - 1, \boldsymbol{\mu}\right) \mathbb{I}\left(\mathbf{z}(\mathbf{x}^p) = \mathbf{e}_b\right)} && \text{(by substitution)}
\end{aligned}$$

  which is result (35).
- (36) is a restatement of (36′).
- (37) is a restatement of (37′).
- (38) is a restatement of (38′).
- (39′) simplifies to (39) since Assumption 6 implies that $|\mathbf{x}^p| = n_0 - 2$ and $f_n(|\mathbf{x}^p| + 2) = 1$ for all $\mathbf{x}^p \in \mathcal{M}_{\mathbf{x}}$.

Proof for Proposition 2

By construction, $\{\mathbf{x}_j\}_{j \in \mathbf{p}_i}$ is a function of $(\mathbf{T}_{-i}, \mathbf{G})$ where $\mathbf{T}_{-i}$ be the sub-matrix of $\mathbf{T}$ that excludes row i. Assumption 3 implies that τ_i is independent of $\mathbf{T}_{-i}$ and assumption (RA) implies that τ_i is independent of $\mathbf{G}$. Therefore $\mathbf{c}_i$ is independent of $\{\mathbf{x}_j\}_{j \in \mathbf{p}_i}$ as well as any function of $\{\mathbf{x}_j\}_{j \in \mathbf{p}_i}$ including $\bar{\mathbf{x}}_i$ or $\mathbf{z}_i$.

Results (40) and (41) are standard implications in regressions with uncorrelated explanatory variables. Let $L(y_i|\mathbf{c}_i, \bar{\mathbf{x}}_i) \equiv \zeta_0 + \mathbf{c}_i\boldsymbol{\zeta}_1 + \bar{\mathbf{x}}_i\boldsymbol{\zeta}_2$. Then:

$$\begin{aligned}
L(y_i|\bar{\mathbf{x}}_i) &= L(\zeta_0 + \mathbf{c}_i\boldsymbol{\zeta}_1 + \bar{\mathbf{x}}_i\boldsymbol{\zeta}_2|\bar{\mathbf{x}}_i) && \text{(law of iterated projections)} \\
&= \zeta_0 + L(\mathbf{c}_i|\bar{\mathbf{x}}_i)\boldsymbol{\zeta}_1 + \bar{\mathbf{x}}_i\boldsymbol{\zeta}_2 \\
&= \zeta_0 + E(\mathbf{c}_i)\boldsymbol{\zeta}_1 + \bar{\mathbf{x}}_i\boldsymbol{\zeta}_2 && (\text{RA} \implies L(\mathbf{c}_i|\bar{\mathbf{x}}_i) = E(\mathbf{c}_i)) \\
L(y_i|\mathbf{c}_i) &= L(\zeta_0 + \mathbf{c}_i\boldsymbol{\zeta}_1 + \bar{\mathbf{x}}_i\boldsymbol{\zeta}_2|\mathbf{c}_i) && \text{(law of iterated projections)} \\
&= \zeta_0 + \mathbf{c}_i\boldsymbol{\zeta}_1 + L(\bar{\mathbf{x}}_i|\mathbf{c}_i)\boldsymbol{\zeta}_2 \\
&= \zeta_0 + \mathbf{c}_i\boldsymbol{\zeta}_1 + E(\bar{\mathbf{x}}_i)\boldsymbol{\zeta}_2 && (\text{RA} \implies L(\bar{\mathbf{x}}_i|\mathbf{c}_i) = E(\bar{\mathbf{x}}_i)) \\
L(y_i|\mathbf{c}_i, \bar{\mathbf{x}}_i) &= \zeta_0 + \mathbf{c}_i\boldsymbol{\zeta}_1 + \bar{\mathbf{x}}_i\boldsymbol{\zeta}_2 && \text{(definition)} \\
&= \zeta_0 + (L(y_i|\mathbf{c}_i) - \zeta_0 - E(\bar{\mathbf{x}}_i)\boldsymbol{\zeta}_2) + (L(y_i|\bar{\mathbf{x}}_i) - \zeta_0 - E(\mathbf{c}_i)\boldsymbol{\zeta}_1) \\
& && \text{(substitution)} \\
&= L(y_i|\mathbf{c}_i) + L(y_i|\bar{\mathbf{x}}_i) + \underbrace{-(\zeta_0 + E(\mathbf{c}_i)\boldsymbol{\zeta}_1 + E(\bar{\mathbf{x}}_i)\boldsymbol{\zeta}_2)}_{(constant)}
\end{aligned}$$

which is result (40). Substituting $\mathbf{z}_i$ for $\bar{\mathbf{x}}_i$ in the argument above yields result (41).

Proof for Lemma 1

The conditions for Lemma 2 are met, so its results apply. Result (42) follows from result (42′) in Lemma 2, where $n = n_0$.

Proof for Proposition 3

1. The conditions for Proposition 10 are met, so its results apply.
   - (43) is a restatement of (43′).
   - (44) follows from (44′) where $PE_{ij} = PE(\tau_i, \tau_j, 0, n_0)$.
   - (45) follows from (45′) where $CPE(\mathbf{x}^o, \mathbf{x}^p) = CPE(\mathbf{x}^o, \mathbf{x}^p, n_0)$.
   - (46) follows from (46′) where $APE(\mathbf{x}^p) = CPE(\mathbf{x}^p, n_0)$.
2. For any i, j:

$$\begin{aligned}
OE_i + PE_j &= (PE(\tau_i, 1, 0, n_0) + c) + (PE(1, \tau_j, 0, n_0) - PE(1, 1, 0, n_0) - c) \\
&= (PE(\tau_i, 1, 0, n_0) + c) + (PE(\tau_i, \tau_j, 0, n_0) - PE(\tau_i, 1, 0, n_0) - c) \\
& \qquad \text{(by OS)} \\
&= PE(\tau_i, \tau_j, 0, n_0) \\
&= PE_{i,j} \qquad (127)
\end{aligned}$$

Then for any i:

$$\sum_{j\in\mathbf{p}} OE_i + PE_j = \sum_{j\in\mathbf{p}} PE_{i,j} tagby(127) \qquad (128)$$

$$= y_i(\mathbf{p}) \qquad \text{(by (44))}$$

which is the result in (47). To get the result in (48):

$$\begin{aligned}
APE(\mathbf{x}^p) &= E(PE_{ij}|\mathbf{x}_j = \mathbf{x}^p) - E(PE_{ij}|\mathbf{x}_j = \mathbf{e}_0) && \text{(by (46))}\\
&= E\left(OE_i + PE_j|\,\mathbf{x}_j = \mathbf{x}^p\right) - E\left(OE_i + PE_j|\,\mathbf{x}_j = \mathbf{e}_0\right) && \text{(by (127))}\\
&= E(PE_j|\mathbf{x}_j = \mathbf{x}^p) - E(PE_j|\mathbf{x}_j = \mathbf{e}_0) + E\left(OE_i|\mathbf{x}_j = \mathbf{x}^p\right) - E\left(OE_i|\mathbf{x}_j = \mathbf{e}_0\right)\\
&= E(PE_j|\mathbf{x}_j = \mathbf{x}^p) - E(PE_j|\mathbf{x}_j = \mathbf{e}_0) + E\left(OE_i\right) - E\left(OE_i\right)\\
& && \text{(by (10)} \implies OE_i \perp\!\!\!\perp \mathbf{x}_j)\\
&= E(PE_j|\mathbf{x}_j = \mathbf{x}^p) - E(PE_j|\mathbf{x}_j = \mathbf{e}_0)\\
CPE(\mathbf{x}^o, \mathbf{x}^p) &= E(PE_{ij}|\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p) - E(PE_{ij}|\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{e}_0) && \text{(by (45))}\\
&= E\left(OE_i + PE_j|\,\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p\right) - E\left(OE_i + PE_j|\,\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{e}_0\right)\\
& && \text{(by (127))}\\
&= E\left(PE_j|\,\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p\right) - E\left(PE_j|\,\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{e}_0\right) && (129)\\
&\quad + E\left(OE_i|\,\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p\right) - E\left(OE_i|\,\mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{e}_0\right)\\
&= E\left(PE_j|\,\mathbf{x}_j = \mathbf{x}^p\right) - E\left(PE_j|\,\mathbf{x}_j = \mathbf{e}_0\right) + E\left(OE_i|\,\mathbf{x}_i = \mathbf{x}^o\right) - E\left(OE_i|\,\mathbf{x}_i = \mathbf{x}^o\right)\\
& && \text{(by (10)} \implies (\mathbf{x}_i, OE_i) \perp\!\!\!\perp (\mathbf{x}_j, PE_j))\\
&= E(PE_j|\mathbf{x}_j = \mathbf{x}^p) - E(PE_j|\mathbf{x}_j = \mathbf{e}_0)\\
&= APE(\mathbf{x}^p)
\end{aligned}$$

which is the result in (48).

Proof for Proposition 4

The conditions for Proposition 11 are met, so its results apply.

- (49) and (51) follow from ($49b'$) and ($51'$) in Proposition 11 where $\mathbb{S}_n = \{n_0\}$ and $f_n(n_0) = 1$.
- (50) and (52) follow from ($50b'$) and ($52'$) in Proposition 11 where $\mathbb{S}_n = \{n_0\}$ and $f_n(n_0) = 1$.
- (53) and (54) follow from ($53b'$) and ($54b'$) in Proposition 11, where $\mathbb{S}_n = \{n_0\}$ and $f_n(n_0) = 1$, and ($52'$) simplifies to (52).

Proof for Proposition 5

The conditions for Proposition 12 are met, so its results apply.

- (55) and (57) are restatements of (55′) and (57′) in Proposition 12.
- (56) and (58) are restatements of (56′) and (58′) in Proposition 12.
- (61) and (62) are restatements of (61′) and (62′) in Proposition 12, where (35′) simplifies to (35) and (60) is a restatement of (60′).
- (63) and (64) are restatements of (63′) and (64′) in Proposition 12, where (39′) simplifies to (39) and (60) is a restatement of (60′).

Proof for Proposition 6

1. Let $\tilde{\mathbf{G}}$ be a purely random group assignment and let $\tilde{\mathbf{p}}_i = \mathbf{p}(i, \tilde{\mathbf{G}})$. Since $\mathbf{Y}(\cdot)$ satisfies (PS) and $\tilde{\mathbf{G}}$ satisfies (RA), Part 1 of Proposition 4 applies to the joint distribution of counterfactual outcomes $(\mathbf{Y}(\tilde{\mathbf{G}}), \mathbf{X}, \bar{\mathbf{X}}(\mathbf{X}, \tilde{\mathbf{G}}))$. Since $\mathbf{G}$ satisfies (SA), Lemma 1 applies to the joint distribution of actual outcomes $(\mathbf{Y}, \mathbf{X}, \bar{\mathbf{X}})$. Let the vector of best linear predictor coefficients $\boldsymbol{\zeta}$ be defined as in equation (157) of the proof for Proposition 11. Then:

$$
\begin{aligned}
E(y_i|\mathbf{x}_i = \mathbf{x}, \bar{\mathbf{x}}_i = \bar{\mathbf{x}}) &= E(y_i(\tilde{\mathbf{p}}_i)|\mathbf{x}_i = \mathbf{x}, \bar{\mathbf{x}}_i(\tilde{\mathbf{p}}_i) = \bar{\mathbf{x}}) && \text{(by (42) in Lemma 1)} \\
&= \zeta_0(n_0 - 1) + \mathbf{x}\boldsymbol{\zeta}_1(n_0 - 1) + \bar{\mathbf{x}}\boldsymbol{\zeta}_2(n_0 - 1) + \mathbf{x}\boldsymbol{\zeta}_3(n_0 - 1)\bar{\mathbf{x}}' \\
& && \text{(by (160) in the proof for Proposition 11)}
\end{aligned}
$$

Applying the law of iterated projections:

$$
\begin{aligned}
L(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i'\bar{\mathbf{x}}_i) &= L(E(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i)|\mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i'\bar{\mathbf{x}}_i) \quad \text{(law of iterated projections)} \\
&= L\left(\begin{array}{l} \zeta_0(n_0-1) + \mathbf{x}_i\boldsymbol{\zeta}_1(n_0-1) \\ \quad + \bar{\mathbf{x}}_i\boldsymbol{\zeta}_2(n_0-1) + \mathbf{x}_i\boldsymbol{\zeta}_3(n_0-1)\bar{\mathbf{x}}_i' \end{array} \middle| \mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i'\bar{\mathbf{x}}_i \right) \\
&\qquad \text{(result above)} \\
&= \zeta_0(n_0-1) + \mathbf{x}_i\boldsymbol{\zeta}_1(n_0-1) \qquad (130) \\
&\quad + \bar{\mathbf{x}}_i\boldsymbol{\zeta}_2(n_0-1) + \mathbf{x}_i\boldsymbol{\zeta}_3(n_0-1)\bar{\mathbf{x}}_i' \\
L(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i'\bar{\mathbf{x}}_i, \mathbf{z}_i) &= L(E(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i)|\mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i'\bar{\mathbf{x}}_i, \mathbf{z}_i) \\
&\qquad \text{(law of iterated projections)} \\
&= L\left(\begin{array}{l} \zeta_0(n_0-1) + \mathbf{x}_i\boldsymbol{\zeta}_1(n_0-1) \\ \quad + \bar{\mathbf{x}}_i\boldsymbol{\zeta}_2(n_0-1) + \mathbf{x}_i\boldsymbol{\zeta}_3(n_0-1)\bar{\mathbf{x}}_i' \end{array} \middle| \begin{array}{l} \mathbf{x}_i, \bar{\mathbf{x}}_i, \\ \mathbf{x}_i'\bar{\mathbf{x}}_i, \mathbf{z}_i \end{array} \right) \\
&\qquad \text{(result above)} \\
&= \zeta_0(n_0-1) + \mathbf{x}_i\boldsymbol{\zeta}_1(n_0-1) \qquad (131) \\
&\quad + \bar{\mathbf{x}}_i\boldsymbol{\zeta}_2(n_0-1) + \mathbf{x}_i\boldsymbol{\zeta}_3(n_0-1)\bar{\mathbf{x}}_i' \\
&= L(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i'\bar{\mathbf{x}}_i) \qquad \text{(by (130) and (131))}
\end{aligned}
$$

which is result (65).

2. The assumptions here (PS, OS, SA) imply that all results in Propositions 3 and 5 apply. Therefore:

$$
\begin{aligned}
APE_\ell &= CPE_{k\ell} \qquad \text{for all } k \qquad \text{(by (48) in Proposition 3)} \\
&= \frac{\beta_{2\ell} + \beta_{3k\ell}}{n_0 - 1} \qquad \text{(by (54) in Proposition 5)}
\end{aligned}
$$

which can only be true if $\beta_{3k\ell} = \beta_{30\ell} = 0$ for all k, ℓ.

Proof for Proposition 7

For convenience, let $y_i^R \equiv y_i(\mathbf{p}(i, \tilde{\mathbf{G}}_R))$, $\bar{\mathbf{x}}_i^R \equiv \bar{\mathbf{x}}(\mathbf{p}(i, \tilde{\mathbf{G}}_R))$, and $\mathbf{z}_i^R \equiv \mathbf{z}(\bar{\mathbf{x}}_i^R)$.

1. Since $\tilde{\mathbf{G}}_R$ satisfies (SA), Lemma 1 applies:

$$
E(y_i^R|\mathbf{x}_i = \mathbf{x}, \bar{\mathbf{x}}_i^R = \bar{\mathbf{x}}) = E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{x}, \bar{\mathbf{x}}_i(\tilde{\mathbf{p}}) = \bar{\mathbf{x}}) \qquad \text{(by (42) in Lemma 1)}
$$

Pick any $b > 0$. By assumption, $\mathbf{S}_{\bar{\mathbf{x}}}^b = \left\{\bar{\mathbf{x}}^b\right\}$ is a singleton, and the events $\bar{\mathbf{x}}_i^R = \bar{\mathbf{x}}^b$

and $\mathbf{z}_i^R = \mathbf{e}_b$ are identical. Therefore:

$$
\begin{aligned}
E(y_i^R|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}_i^R = \mathbf{e}_b) &= E(y_i^R|\mathbf{x}_i = \mathbf{e}_k, \bar{\mathbf{x}}_i^R = \bar{\mathbf{x}}^b) && \text{(identical events)} \\
&= E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \bar{\mathbf{x}}_i(\tilde{\mathbf{p}}) = \bar{\mathbf{x}}^b) && \text{(by Lemma 1)} \\
&= E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) && \text{(identical events)} \\
&= E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) + CGE_{kb} && (132)
\end{aligned}
$$

Summing over all values of $\mathbf{z}$:

$$
\begin{aligned}
E(y_i^R|\mathbf{x}_i = \mathbf{e}_k) &= \sum_{b=0}^{B} E(y_i^R|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}_i^R = \mathbf{e}_b) \Pr(\mathbf{z}_i^R = \mathbf{e}_b|\mathbf{x}_i = \mathbf{e}_k) \\
&= \sum_{b=1}^{B} E(y_i^R|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}_i^R = \mathbf{e}_b) \Pr(\mathbf{z}_i^R = \mathbf{e}_b|\mathbf{x}_i = \mathbf{e}_k) \\
&\qquad \text{(since } \Pr(\bar{\mathbf{x}}_i^R \in \mathbf{S}_{\bar{\mathbf{x}}}^0) = 0) \\
&= \sum_{b=1}^{B} \left(E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) + CGE_{kb}\right) \Pr(\mathbf{z}_i^R = \mathbf{e}_b|\mathbf{x}_i = \mathbf{e}_k) \\
&\qquad \text{(by (132))} \\
&= E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) \left(\underbrace{\sum_{b=1}^{B} \Pr(\mathbf{z}_i^R = \mathbf{e}_b|\mathbf{x}_i = \mathbf{e}_k)}_{1} \right) \\
&\quad + \sum_{b=1}^{B} CGE_{kb} \Pr(\mathbf{z}_i^R = \mathbf{e}_b|\mathbf{x}_i = \mathbf{e}_k) \\
&= E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) + \sum_{b=1}^{B} CGE_{kb} \Pr(\mathbf{z}_i^R = \mathbf{e}_b|\mathbf{x}_i = \mathbf{e}_k) \qquad (133)
\end{aligned}
$$

Similarly:

$$\begin{aligned}
E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k) &= \sum_{b=0}^{B} E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \Pr(\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b|\mathbf{x}_i = \mathbf{e}_k) \\
&= \sum_{b=0}^{B} E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \Pr(\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \\
&\qquad \text{(since } \tilde{\mathbf{p}} \perp\!\!\!\perp \mathbf{x}_i\text{)} \\
&= E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) \Pr(\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) \\
&\quad + \sum_{b=1}^{B} E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \Pr(\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \\
&= E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) \Pr(\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) \\
&\quad + \sum_{b=1}^{B} \left(E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) + CGE_{kb}\right) \Pr(\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \\
&\qquad \text{(by (132))} \\
&= E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) \left(\underbrace{\sum_{b=0}^{B} \Pr(\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b)}_{1} \right) \\
&\quad + \sum_{b=1}^{B} CGE_{kb} \Pr(\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \\
&= E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) + \sum_{b=1}^{B} CGE_{kb} \Pr(\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b)
\end{aligned} \tag{134}$$

Combining these results yields:

$$
\begin{aligned}
CRE_k(\mathbf{G}_R) &= E(y_i(\mathbf{p}(i,\tilde{\mathbf{G}}_R)) - y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k) && \text{(definition of } CRE\text{)} \\
&= E(y_i^R|\mathbf{x}_i = \mathbf{e}_k) - E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k) \\
&= \left(E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) + \sum_{b=1}^{B} CGE_{kb} \Pr(\mathbf{z}^R = \mathbf{e}_b|\mathbf{x}_i = \mathbf{e}_k) \right) \\
&\quad - \left(E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) + \sum_{b=1}^{B} CGE_{kb} \Pr(\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \right) \\
& && \text{(by (133) and (134))} \\
&= \sum_{b=1}^{B} CGE_{kb} \left(\Pr(\mathbf{z}_i^R = \mathbf{e}_b|\mathbf{x}_i = \mathbf{e}_k) - \Pr(\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \right) \\
&= \sum_{b=1}^{B} \Delta z_{kb}(\mathbf{G}_R) CGE_{kb}
\end{aligned}
$$

which is result (68). Result (67) follows by substituting (68) into result (32) of Proposition 1.

2. Given (PS), part 1 of Proposition 3 applies:

$$\begin{aligned}
E(y_i^R|\mathbf{X},\tilde{\mathbf{G}}_R) &= E(y_i(\mathbf{p}(i,\tilde{\mathbf{G}}_R))|\mathbf{X},\tilde{\mathbf{G}}_R) && \text{(definition)}\\
&= E\left(\sum_{j\in\mathbf{p}(i,\tilde{\mathbf{G}}_R)} PE_{ij}\middle|\mathbf{X},\tilde{\mathbf{G}}_R\right) && \text{(by Proposition 3)}\\
&= \sum_{j\neq i} E\left(PE_{ij}\middle|\mathbf{X},\tilde{\mathbf{G}}_R\right)\mathbb{I}\left(j\in\mathbf{p}(i,\tilde{\mathbf{G}}_R)\right)\\
&= \sum_{j\neq i} E\left(PE(\tau_i,\tau_j,n_0)\middle|\mathbf{X},\tilde{\mathbf{G}}_R\right)\mathbb{I}\left(j\in\mathbf{p}(i,\tilde{\mathbf{G}}_R)\right)\\
&= \sum_{j\neq i} E\left(PE(\tau_i,\tau_j,n_0)\,|\mathbf{X}\right)\mathbb{I}\left(j\in\mathbf{p}(i,\tilde{\mathbf{G}}_R)\right)\\
& && (\text{SA} \implies \tau_i,\tau_j \perp\!\!\!\perp \tilde{\mathbf{G}}_R|\mathbf{X})\\
&= \sum_{j\neq i} E\left(PE(\tau_i,\tau_j,n_0)\,|\mathbf{x}_i,\mathbf{x}_j\right)\mathbb{I}\left(j\in\mathbf{p}(i,\tilde{\mathbf{G}}_R)\right)\\
& && (\text{Assumption 3} \implies \tau_i,\mathbf{x}_i \perp\!\!\!\perp \tau_j,\mathbf{x}_j \text{ for } i\neq j)\\
&= \sum_{j\neq i}\left(\zeta_0+\mathbf{x}_i\boldsymbol{\zeta}_1+\mathbf{x}_j\boldsymbol{\zeta}_2+\mathbf{x}_i\boldsymbol{\zeta}_3\mathbf{x}_j'\right)\mathbb{I}\left(j\in\mathbf{p}(i,\tilde{\mathbf{G}}_R)\right)\\
& && (\text{where } \boldsymbol{\zeta} \text{ is defined as in (157)})\\
&= \zeta_0(n_0-1)+\mathbf{x}_i\boldsymbol{\zeta}_1(n_0-1)+\bar{\mathbf{x}}_i^R\boldsymbol{\zeta}_2(n_0-1)+\mathbf{x}_i\boldsymbol{\zeta}_3(n_0-1)\bar{\mathbf{x}}_i^{R\prime}
\end{aligned} \tag{135}$$

Averaging over values of $\bar{\mathbf{x}}$:

$$\begin{aligned}
E(y_i^R|\mathbf{x}_i=\mathbf{x}) &= E(E(y_i^R|\mathbf{X},\tilde{\mathbf{G}}_R)|\mathbf{x}_i=\mathbf{x}) && \text{(Law of iterated expectations)}\\
&= E\left(\begin{matrix}\zeta_0(n_0-1)+\mathbf{x}_i\boldsymbol{\zeta}_1(n_0-1)\\ +\bar{\mathbf{x}}_i^R\boldsymbol{\zeta}_2(n_0-1)+\mathbf{x}_i\boldsymbol{\zeta}_3(n_0-1)\bar{\mathbf{x}}_i^{R\prime}\end{matrix}\middle|\mathbf{x}_i=\mathbf{x}\right) && \text{(by (135))}\\
&= \zeta_0(n_0-1)+\mathbf{x}\boldsymbol{\zeta}_1(n_0-1) && (136)\\
&\quad + E(\bar{\mathbf{x}}_i^R|\mathbf{x}_i=\mathbf{x})\boldsymbol{\zeta}_2(n_0-1)+\mathbf{x}\boldsymbol{\zeta}_3(n_0-1)E(\bar{\mathbf{x}}_i^R|\mathbf{x}_i=\mathbf{x})'
\end{aligned}$$

This result also applies when $\tilde{\mathbf{G}}_R$ is purely random, so:

$$\begin{aligned}
E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{x}) &= \zeta_0(n_0-1) + \mathbf{x}\boldsymbol{\zeta}_1(n_0-1) + E(\bar{\mathbf{x}}_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{x})\boldsymbol{\zeta}_2(n_0-1) \\
&\quad + \mathbf{x}\boldsymbol{\zeta}_3(n_0-1)E(\bar{\mathbf{x}}_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{x})' \\
&\qquad\qquad \text{(by (136))} \\
&= \zeta_0(n_0-1) + \mathbf{x}\boldsymbol{\zeta}_1(n_0-1) + E(\bar{\mathbf{x}}_i(\tilde{\mathbf{p}}))\boldsymbol{\zeta}_2(n_0-1) \\
&\quad + \mathbf{x}\boldsymbol{\zeta}_3(n_0-1)E(\bar{\mathbf{x}}_i(\tilde{\mathbf{p}}))' \\
&\qquad\qquad (\text{RA} \implies \bar{\mathbf{x}}_i(\tilde{\mathbf{p}}) \perp\!\!\!\perp \mathbf{x}_i) \\
&= \zeta_0(n_0-1) + \mathbf{x}\boldsymbol{\zeta}_1(n_0-1) + \boldsymbol{\mu}\boldsymbol{\zeta}_2(n_0-1) + \mathbf{x}\boldsymbol{\zeta}_3(n_0-1)\boldsymbol{\mu}'
\end{aligned} \tag{137}$$

Combining these results:

$$\begin{aligned}
CRE_k(\mathbf{G}_R) &= E(y_i^R - y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{e}_k) \\
&= \begin{pmatrix} \zeta_0(n_0-1) + \mathbf{e}_k\boldsymbol{\zeta}_1(n_0-1) + E(\bar{\mathbf{x}}_i^R|\mathbf{x}_i = \mathbf{e}_k)\boldsymbol{\zeta}_2(n_0-1) \\ + \mathbf{e}_k\boldsymbol{\zeta}_3(n_0-1)E(\bar{\mathbf{x}}_i^R|\mathbf{x}_i = \mathbf{e}_k)' \end{pmatrix} \\
&\quad - \begin{pmatrix} \zeta_0(n_0-1) + \mathbf{e}_k\boldsymbol{\zeta}_1(n_0-1) + \boldsymbol{\mu}\boldsymbol{\zeta}_2(n_0-1) \\ + \mathbf{e}_k\boldsymbol{\zeta}_3(n_0-1)\boldsymbol{\mu}' \end{pmatrix} \\
&\qquad\qquad \text{(by (136) and (137))} \\
&= \left(E(\bar{\mathbf{x}}_i^R|\mathbf{x}_i = \mathbf{e}_k) - \boldsymbol{\mu}\right)\boldsymbol{\zeta}_2(n_0-1) + \mathbf{e}_k\boldsymbol{\zeta}_3(n_0-1)\left(E(\bar{\mathbf{x}}_i^R|\mathbf{x}_i = \mathbf{e}_k) - \boldsymbol{\mu}\right)' \\
&= (n_0-1)\sum_{\ell=1}^{K}\left(E(\bar{x}_{i\ell}^R|\mathbf{x}_i = \mathbf{e}_k) - \mu_\ell\right)(\zeta_{2\ell} + \zeta_{3k\ell}) \\
&= (n_0-1)\sum_{\ell=1}^{K}\Delta\bar{x}_{k\ell}(\mathbf{G}_R)CPE_{k\ell}
\end{aligned}$$

which is result (71). Result (70) follows by applying the law of total probability to (71).

3. Given (PS, OS), Part 2 of Proposition 3 applies. By equation (48) in Proposition 3, $CPE_{k\ell} = APE_\ell$ and so result (74) follows from (71) by substitution. Result (73) follows from the fact that individual and peer characteristics have the same

expected value in any reallocation:

$$\begin{aligned}
\sum_{k=0}^{K} \Delta\bar{x}_{k\ell}(\mathbf{G}_R)\mu_k &= \sum_{k=0}^{K} \left(E(\bar{x}_{i\ell}^{R}|\mathbf{x}_i = \mathbf{e}_k) - \mu_\ell\right)\mu_k && \text{(definition of } \Delta\bar{x}\text{)} \\
&= \sum_{k=0}^{K} E(\bar{x}_{i\ell}^{R}|\mathbf{x}_i = \mathbf{e}_k)\Pr(\mathbf{x}_i = \mathbf{e}_k) - E(x_{i\ell})\sum_{k=0}^{K}\Pr(\mathbf{x}_i = \mathbf{e}_k) \\
& && \text{(definition of } \boldsymbol{\mu}\text{)} \\
&= E(\bar{x}_{i\ell}^{R}) - E(x_{i\ell}) && \text{(law of total probability)} \\
&= 0 && (138)
\end{aligned}$$

Therefore:

$$\begin{aligned}
ARE(\mathbf{G}_R) &= (n_0 - 1)\sum_{k=0}^{K}\sum_{\ell=1}^{K} \mu_k \Delta\bar{x}_{k\ell}(\mathbf{G}_R) CPE_{k\ell} && \text{(by (70))} \\
&= (n_0 - 1)\sum_{k=0}^{K}\sum_{\ell=1}^{K} \mu_k \Delta\bar{x}_{k\ell}(\mathbf{G}_R) APE_{\ell} && \text{(by (48) in Proposition 3)} \\
&= (n_0 - 1)\sum_{\ell=1}^{K} APE_{\ell} \underbrace{\sum_{k=0}^{K} \Delta\bar{x}_{k\ell}(\mathbf{G}_R)\mu_k}_{=0 \text{ by } (138)} \\
&= 0
\end{aligned}$$

which is result (73).

Proof for Proposition 8

Proof. Let the approximation error in (76) be:

$$v_m(\mathbf{x}^p) \equiv a(\mathbf{x}^p) - a_m(\mathbf{x}^p)\bar{\boldsymbol{\pi}} \approx 0 \qquad (139)$$

Assumptions 1–6 and peer separability are given, so part (1) of Proposition 3 applies. Therefore:

$$y_i = y_i(\mathbf{p}_i) = \sum_{j \in \mathbf{p}_i} PE_{ij} \qquad \text{(by (44) in Proposition 3)}$$

Taking expectations:

$$
\begin{aligned}
E\left(y_i \middle| \{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) &= E\left(\sum_{j\in\mathbf{p}_i} PE_{ij} \middle| \{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) \\
&= \sum_{j\in\mathbf{p}_i} E\left(PE_{ij}|\mathbf{x}_j\right) && (\text{RA} \implies \tau_i, \tau_j \perp\!\!\!\perp \tau_{j'}) \\
&= \sum_{j\in\mathbf{p}_i} a(\mathbf{x}_j) && (\text{by (76)}) \\
&= \sum_{j\in\mathbf{p}_i} a_m(\mathbf{x}_j)\bar{\boldsymbol{\pi}} + \sum_{j\in\mathbf{p}_i} v_m(\mathbf{x}_j) && (\text{by (139)}) \\
&= (n_0 - 1)\bar{\boldsymbol{a}}_i\bar{\boldsymbol{\pi}} + \sum_{j\in\mathbf{p}_i} v_m(\mathbf{x}_j) && (140)
\end{aligned}
$$

By construction:

$$
\begin{aligned}
L(v_m(\mathbf{x}_i)|a_m(\mathbf{x}_i)) &= L(a(\mathbf{x}_i) - a_m(\mathbf{x}^p)\bar{\boldsymbol{\pi}}|a_m(\mathbf{x}_i)) \\
&= a_m(\mathbf{x}_i)\bar{\boldsymbol{\pi}} - a_m(\mathbf{x}_i)\bar{\boldsymbol{\pi}} \\
&= 0 && (141)
\end{aligned}
$$

Applying the law of iterated projections:

$$
\begin{aligned}
L\left(y_i|\bar{\boldsymbol{a}}_i\right) &= L\left(E\left(y_i \middle| \{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) \middle| \bar{\boldsymbol{a}}_i\right) \\
&= L\left((n_0 - 1)\bar{\boldsymbol{a}}_i\bar{\boldsymbol{\pi}} + \sum_{j\in\mathbf{p}_i} v_m(\mathbf{x}_j) \middle| \bar{\boldsymbol{a}}_i\right) && (\text{by (78)}) \\
&= (n_0 - 1)\bar{\boldsymbol{a}}_i\bar{\boldsymbol{\pi}} + \sum_{j\in\mathbf{p}_i} L\left(v_m(\mathbf{x}_j)\middle| \bar{\boldsymbol{a}}_i\right) \\
&= (n_0 - 1)\bar{\boldsymbol{a}}_i\bar{\boldsymbol{\pi}}
\end{aligned}
$$

which implies that:

$$
\boldsymbol{\pi} = (n_0 - 1)\bar{\boldsymbol{\pi}} \qquad (142)
$$

By part (1) of Proposition 3:

$$\begin{aligned}
APE(\mathbf{x}^p) &= E\left(PE_{ij}|\mathbf{x}_j = \mathbf{x}^p\right) - E\left(PE_{ij}|\mathbf{x}_j = \mathbf{e}_0\right) && \text{(by (46))} \\
&= a(\mathbf{x}^p) - a(\mathbf{e}_0) && \text{(by definition of } a(\cdot)) \\
&= a_m(\mathbf{x}^p)\bar{\boldsymbol{\pi}} - a_m(\mathbf{e}_0)\bar{\boldsymbol{\pi}} + v_m(\mathbf{x}^p) - v_m(\mathbf{e}_0) && \text{(by (139))} \\
&\approx \left(\frac{a_m(\mathbf{x}^p) - a_m(\mathbf{e}_0)}{n_0 - 1}\right)\boldsymbol{\pi} && \text{(by (142))}
\end{aligned}$$

which is result (79) in the proposition, with approximation error the same order of magnitude as the approximation error in (76) in the sense that:

$$\sup_{\mathbf{x}^p \in \mathbb{S}_{\mathbf{x}}} |v_m(\mathbf{x}^p)| < B \implies \sup_{\mathbf{x}^p \in \mathbb{S}_{\mathbf{x}}} |v_m(\mathbf{x}^p) - v_m(\mathbf{e}_0)| < 2B \tag{143}$$

for any finite constant $B > 0$. □

Proof for Proposition 9

1. Let $\tilde{\mathbf{p}}$ be a random draw of $n-2$ peers from $\mathcal{I} \setminus \{i\}$. Then:

$$\begin{aligned}
APE_\ell(n) &= E\left(y_i(\{j\} \cup \tilde{\mathbf{p}}) - y_i(\{j'\} \cup \tilde{\mathbf{p}}) \middle| \mathbf{x}_j = \mathbf{e}_\ell, \mathbf{x}_{j'} = \mathbf{e}_0\right) && \text{(definition)} \\
&= \sum_{k=0}^{K} E\left(y_i(\{j\} \cup \tilde{\mathbf{p}}) - y_i(\{j'\} \cup \tilde{\mathbf{p}}) \middle| \begin{array}{l}\mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{x}_j = \mathbf{e}_\ell, \mathbf{x}_{j'} = \mathbf{e}_0\end{array}\right) \Pr(\mathbf{x}_i = \mathbf{e}_k | \mathbf{x}_j = \mathbf{e}_\ell, \mathbf{x}_{j'} = \mathbf{e}_0) \\
&= \sum_{k=0}^{K} E\left(y_i(\{j\} \cup \tilde{\mathbf{p}}) - y_i(\{j'\} \cup \tilde{\mathbf{p}}) \middle| \begin{array}{l}\mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{x}_j = \mathbf{e}_\ell, \mathbf{x}_{j'} = \mathbf{e}_0\end{array}\right) \Pr(\mathbf{x}_i = \mathbf{e}_k) \\
& && \text{(Assumption 3} \implies \mathbf{x}_i \perp\!\!\!\perp \mathbf{x}_j, \mathbf{x}_{j'}) \\
&= \sum_{k=0}^{K} \mu_k CPE_{k\ell}(n)
\end{aligned}$$

which is result (109). Then:

$$
\begin{aligned}
APE_\ell &= \sum_{n\in\mathbb{S}_n} APE_\ell(n) f_n(n) && \text{(definition)}\\
&= \sum_{n\in\mathbb{S}_n}\left(\sum_{k=1}^{K} CPE_{k\ell}(n)\right) f_n(n) && \text{(by (109))}\\
&= \sum_{k=1}^{K}\sum_{n\in\mathbb{S}_n} CPE_{k\ell}(n) f_n(n)\\
&= \sum_{k=1}^{K} CPE_{k\ell} && \text{(definition of } CPE_{k,\ell})
\end{aligned}
$$

which is result (30′).

Similarly, let $\tilde{n}$ be a purely random draw from f_n, let $\tilde{\mathbf{p}}$ be a purely random draw of $\tilde{n}-1$ peers from $\mathcal{I}\setminus\{i\}$ and let $\mathbf{z}(\mathbf{p}) \equiv \mathbf{z}\left(\{\mathbf{x}_j\}_{j\in\mathbf{p}}\right)$ Then:

$$
\begin{aligned}
AGE_b &= E\left(y_i(\tilde{\mathbf{p}})\,|\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b\right) - E\left(y_i(\tilde{\mathbf{p}})\,|\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0\right) && \text{(definition)}\\
&= \sum_{k=0}^{K} E\left(y_i(\tilde{\mathbf{p}})\,|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b\right)\Pr\left(\mathbf{x}_i = \mathbf{e}_k\,|\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b\right)\\
&\quad - \sum_{k=0}^{K} E\left(y_i(\tilde{\mathbf{p}})\,|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0\right)\Pr\left(\mathbf{x}_i = \mathbf{e}_k\,|\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0\right)\\
&= \sum_{k=0}^{K}\begin{pmatrix} E\left(y_i(\tilde{\mathbf{p}})\,|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b\right)\\ - E\left(y_i(\tilde{\mathbf{p}})\,|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0\right)\end{pmatrix}\Pr\left(\mathbf{x}_i = \mathbf{e}_k\right)\\
& && \left(\text{Assumption 3} \implies \mathbf{x}_i \perp\!\!\!\perp \{\mathbf{x}_j\}_{j\neq i}, \tilde{\mathbf{p}}\right)\\
&= \sum_{k=0}^{K}\mu_k CGE_{kb}
\end{aligned}
$$

which is result (31′).

2. Let $\tilde{n}$ be a random draw from f_n and let $\tilde{\mathbf{p}}$ be a random draw of $\tilde{n}-1$ peers from $\mathcal{I}\setminus\{i\}$. By discrete characteristics (DC), $\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}$ is the result of $\tilde{n}-1$ independent draws from a categorical distribution with probability vector $\boldsymbol{\mu}$. So its probability distribution can be derived from the multinomial distribution:

$$
\begin{aligned}
\Pr\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}} = \mathbf{x}^p \,\middle|\, \tilde{n} = |\mathbf{x}^p| + 1\right) &= M(\bar{\mathbf{x}}(\mathbf{x}^p), |\mathbf{x}^p|, \boldsymbol{\mu})\\
\Pr\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}} = \mathbf{x}^p\right) &= M(\bar{\mathbf{x}}(\mathbf{x}^p), |\mathbf{x}^p|, \boldsymbol{\mu}) f_n(|\mathbf{x}^p| + 1) \qquad (144)
\end{aligned}
$$

where the function $M(\cdot)$ is defined in (36′).

Let $\mathbf{z}(\mathbf{p}) \equiv \mathbf{z}\left(\left\{\mathbf{x}_j\right\}_{j\in\mathbf{p}}\right)$, let $\mathbf{z}^S(\mathbf{p}) \equiv \mathbf{z}^S\left(\bar{\mathbf{x}}\left(\left\{\mathbf{x}_j\right\}_{j\in\mathbf{p}}\right)\right)$, and let:

$$
\begin{aligned}
w_{sb}^G(\boldsymbol{\mu}) &\equiv \Pr(\mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s | \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \\
&= \frac{\Pr(\mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s \cap \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b)}{\Pr(\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b)} \\
&= \frac{\sum_{\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}} \Pr\left(\left\{\mathbf{x}_j\right\}_{j\in\tilde{\mathbf{p}}} = \mathbf{x}^p\right) \mathbb{I}\left(\mathbf{z}^S(\mathbf{x}^p) = \mathbf{e}_s\right) \mathbb{I}\left(\mathbf{z}(\mathbf{x}^p) = \mathbf{e}_b\right)}{\sum_{\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}} \Pr\left(\left\{\mathbf{x}_j\right\}_{j\in\tilde{\mathbf{p}}} = \mathbf{x}^p\right) \mathbb{I}\left(\mathbf{z}(\mathbf{x}^p) = \mathbf{e}_b\right)} \\
&= \frac{\sum_{\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}} M\left(\bar{\mathbf{x}}(\mathbf{x}^p), |\mathbf{x}^p|, \boldsymbol{\mu}\right) f_n(|\mathbf{x}^p|+1) \mathbb{I}\left(\mathbf{z}^S(\mathbf{x}^p) = \mathbf{e}_s\right) \mathbb{I}\left(\mathbf{z}(\mathbf{x}^p) = \mathbf{e}_b\right)}{\sum_{\mathbf{x}^p \in \mathbb{S}_{\{\mathbf{x}\}}} M\left(\bar{\mathbf{x}}(\mathbf{x}^p), |\mathbf{x}^p|, \boldsymbol{\mu}\right) f_n(|\mathbf{x}^p|+1) \mathbb{I}\left(\mathbf{z}(\mathbf{x}^p) = \mathbf{e}_b\right)} \\
&\hspace{10cm} \text{(by (144))}
\end{aligned}
$$

which is equation (35′).

For any (k, b):

$$
\begin{aligned}
E\left(y_i(\tilde{\mathbf{p}}) \middle| \begin{array}{r} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b \end{array}\right) &= \sum_{s=0}^{S} E\left(y_i(\tilde{\mathbf{p}}) \middle| \begin{array}{r} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s \end{array}\right) \Pr\left(\mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s \middle| \begin{array}{r} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b \end{array}\right) \\
&= E\left(y_i(\tilde{\mathbf{p}}) \middle| \begin{array}{r} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_0 \end{array}\right) \Pr(\mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_0 | \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \\
&\quad + \sum_{s=1}^{S} E\left(y_i(\tilde{\mathbf{p}}) \middle| \begin{array}{r} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s \end{array}\right) \Pr(\mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s | \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \\
&= E\left(y_i(\tilde{\mathbf{p}}) \middle| \begin{array}{r} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_0 \end{array}\right) \left(1 - \sum_{s=1}^{S} \Pr(\mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s | \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b)\right) \\
&\quad + \sum_{s=1}^{S} E\left(y_i(\tilde{\mathbf{p}}) \middle| \begin{array}{r} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s \end{array}\right) \Pr(\mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s | \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \\
&= E\left(y_i(\tilde{\mathbf{p}}) \middle| \begin{array}{r} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_0 \end{array}\right) + \sum_{s=1}^{S} CGE_{ks}^S \Pr(\mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s | \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b)
\end{aligned} \tag{145}
$$

Substituting result (145) into the definition of CGE_{kb} produces:

$$
\begin{aligned}
CGE_{kb} &= E\left(y_i(\tilde{\mathbf{p}}) \middle| \begin{matrix} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b \end{matrix}\right) - E\left(y_i(\tilde{\mathbf{p}}) \middle| \begin{matrix} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0 \end{matrix}\right) \\
&\hspace{20em} \text{(definition of } CGE_{kb}\text{)} \\
&= \left(E\left(y_i(\tilde{\mathbf{p}}) \middle| \begin{matrix} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_0 \end{matrix}\right) + \sum_{s=1}^{S} CGE_{ks}^S \Pr\left(\mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s \,|\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b\right)\right) \\
&\quad - \left(E\left(y_i(\tilde{\mathbf{p}}) \middle| \begin{matrix} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_0 \end{matrix}\right) + \sum_{s=1}^{S} CGE_{ks}^S \Pr(\mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s|\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0)\right) \\
&\hspace{20em} \text{(by (145))} \\
&= \sum_{s=1}^{S} CGE_{ks}^S \begin{pmatrix} \Pr(\mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s|\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) \\ - \Pr(\mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s|\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_0) \end{pmatrix} \\
&= \sum_{s=1}^{S} CGE_{ks}^S \left(w_{sb}^G(\boldsymbol{\mu}) - w_{s0}^G(\boldsymbol{\mu})\right) \qquad \text{(definition of } w_{sb}^G(\cdot) \text{ above)}
\end{aligned}
$$

which is result (34′). Result (33′) follows by substitution of (34′) into (31′).

3. Let $\tilde{n}$ be a random draw from f_n, let $\tilde{\mathbf{p}}$ be a random draw of $\tilde{n}-1$ peers from $\mathcal{I} \setminus \{i, j'\}$, and let $\tilde{\mathbf{q}}$ be a random draw of $\tilde{n}-2$ peers. By discrete characteristics (DC), $\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{q}}}$ is the result of $\tilde{n}-2$ independent draws from a categorical distribution with probability vector $\boldsymbol{\mu}$. So its probability distribution can be derived from the multinomial distribution. Since $\tilde{\mathbf{q}}$ is independent of $\mathbf{x}_i$ and $\mathbf{x}_{j'}$:

$$
\begin{aligned}
\Pr\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{q}}} = \mathbf{x}^p \middle| \begin{matrix} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{x}_{j'} = \mathbf{e}_\ell \end{matrix}\right) &= \Pr\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{q}}} = \mathbf{x}^p\right) \qquad (146) \\
&= \sum_{n\in\mathbb{S}_n} \Pr\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{q}}} = \mathbf{x}^p \middle| \tilde{n} = n\right) \Pr(\tilde{n} = n) \\
&\hspace{10em} \text{(law of total probability)} \\
&= \Pr\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{q}}} = \mathbf{x}^p \middle| \tilde{n} = |\mathbf{x}^p| + 2\right) \Pr(\tilde{n} = |\mathbf{x}^p| + 2) \\
&= M\left(\bar{\mathbf{x}}\left(\mathbf{x}^p\right), |\mathbf{x}^p|, \boldsymbol{\mu}\right) f_n(|\mathbf{x}^p| + 2) \qquad (147)
\end{aligned}
$$

Rearranging the definition of CGE:

$$
E\left(y_i\left(\tilde{\mathbf{p}}\right) \middle| \begin{matrix} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s \end{matrix}\right) = E\left(y_i\left(\tilde{\mathbf{p}}\right) \middle| \begin{matrix} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_0 \end{matrix}\right) + CGE_{ks}^S \qquad (148)
$$

For any $\mathbf{x}^p \in \mathcal{M}_{\mathbf{x}}$, let:

$$\begin{aligned}
A(\mathbf{x}^p) &\equiv E\left(y_i\left(\{j'\} \cup \tilde{\mathbf{q}}\right) \middle| \begin{array}{l} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{x}_{j'} = \mathbf{e}_\ell, \\ \{\mathbf{x}_j\}_{j \in \tilde{\mathbf{q}}} = \mathbf{x}^p \end{array} \right) \\
&= E\left(y_i\left(\tilde{\mathbf{p}}\right) \middle| \begin{array}{l} \mathbf{x}_i = \mathbf{e}_k, \\ \{\mathbf{x}_j\}_{j \in \tilde{\mathbf{p}}} = \mathbf{x}^p \cup \{\mathbf{e}_\ell\} \end{array} \right) && \text{(equivalent events)} \\
&= E\left(y_i\left(\tilde{\mathbf{p}}\right) \middle| \begin{array}{l} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S\left(\{\mathbf{x}_j\}_{j \in \tilde{\mathbf{p}}}\right) = \mathbf{z}^S\left(\mathbf{x}^p \cup \{\mathbf{e}_\ell\}\right) \end{array} \right) && \text{(equivalent events)} \\
&= \sum_{s=1}^{S} E\left(y_i\left(\tilde{\mathbf{p}}\right) \middle| \begin{array}{l} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_s \end{array} \right) \mathbb{I}\left(\mathbf{z}^S(\mathbf{x}^p \cup \{\mathbf{e}_\ell\}) = \mathbf{e}_s\right) \\
&= \sum_{s=1}^{S} \left(E\left(y_i\left(\tilde{\mathbf{p}}\right) \middle| \begin{array}{l} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_0 \end{array} \right) + CGE_{ks} \right) \mathbb{I}\left(\mathbf{z}^S(\mathbf{x}^p \cup \{\mathbf{e}_\ell\}) = \mathbf{e}_s\right) \\
& && \text{(by (148))} \\
&= E\left(y_i\left(\tilde{\mathbf{p}}\right) \middle| \begin{array}{l} \mathbf{x}_i = \mathbf{e}_k, \\ \mathbf{z}^S(\tilde{\mathbf{p}}) = \mathbf{e}_0 \end{array} \right) + \sum_{s=1}^{S} CGE_{ks} \mathbb{I}\left(\mathbf{z}^S(\mathbf{x}^p \cup \{\mathbf{e}_\ell\}) = \mathbf{e}_s\right) && (149)
\end{aligned}$$

For any ℓ, let:

$$
\begin{aligned}
B(\mathbf{e}_\ell) &\equiv E\left(y_i\left(\{j'\}\cup\tilde{\mathbf{q}}\right)\middle|\begin{array}{l}\mathbf{x}_i=\mathbf{e}_k,\\ \mathbf{x}_{j'}=\mathbf{e}_\ell\end{array}\right)\\
&=\sum_{\mathbf{x}^p\in\mathcal{M}_{\mathbf{x}}} E\left(y_i\left(\{j'\}\cup\tilde{\mathbf{q}}\right)\middle|\begin{array}{l}\mathbf{x}_i=\mathbf{e}_k,\\ \mathbf{x}_{j'}=\mathbf{e}_\ell,\\ \{\mathbf{x}_j\}_{j\in\tilde{\mathbf{q}}}=\mathbf{x}^p\end{array}\right)\Pr\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{q}}}=\mathbf{x}^p\middle|\begin{array}{l}\mathbf{x}_i=\mathbf{e}_k,\\ \mathbf{x}_j=\mathbf{e}_\ell\end{array}\right)\\
&\hspace{22em}\text{(law of total probability)}\\
&=\sum_{\mathbf{x}^p\in\mathcal{M}_{\mathbf{x}}} E\left(y_i\left(\{j'\}\cup\tilde{\mathbf{q}}\right)\middle|\begin{array}{l}\mathbf{x}_i=\mathbf{e}_k,\\ \mathbf{x}_{j'}=\mathbf{e}_\ell,\\ \{\mathbf{x}_j\}_{j\in\tilde{\mathbf{q}}}=\mathbf{x}^p\end{array}\right)\Pr\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{q}}}=\mathbf{x}^p\right)\ \text{(by (146))}\\
&=\sum_{\mathbf{x}^p\in\mathcal{M}_{\mathbf{x}}}\left[\begin{array}{l}E\left(y_i\left(\tilde{\mathbf{p}}\right)\middle|\begin{array}{l}\mathbf{x}_i=\mathbf{e}_k,\\ \mathbf{z}^S(\tilde{\mathbf{p}})=\mathbf{e}_0\end{array}\right)\\ +\sum_{s=1}^{S}CGE_{ks}\mathbb{I}\left(\mathbf{z}^S(\mathbf{x}^p\cup\{\mathbf{e}_\ell\})=\mathbf{e}_s\right)\end{array}\right]\Pr\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{q}}}=\mathbf{x}^p\right)\\
&\hspace{22em}\text{(by (149))}\\
&=E\left(y_i\left(\tilde{\mathbf{p}}\right)\middle|\begin{array}{l}\mathbf{x}_i=\mathbf{e}_k,\\ \mathbf{z}^S(\tilde{\mathbf{p}})=\mathbf{e}_0\end{array}\right)\sum_{\mathbf{x}^p\in\mathcal{M}_{\mathbf{x}}}\Pr\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{q}}}=\mathbf{x}^p\right)\\
&\quad+\sum_{s=1}^{S}CGE_{ks}\sum_{\mathbf{x}^p\in\mathcal{M}_{\mathbf{x}}}\mathbb{I}\left(\mathbf{z}^S(\mathbf{x}^p\cup\{\mathbf{e}_\ell\})=\mathbf{e}_s\right)\Pr\left(\{\mathbf{x}_j\}_{j\in\tilde{\mathbf{q}}}=\mathbf{x}^p\right)\\
&=E\left(y_i\left(\tilde{\mathbf{p}}\right)\middle|\begin{array}{l}\mathbf{x}_i=\mathbf{e}_k,\\ \mathbf{z}^S(\tilde{\mathbf{p}})=\mathbf{e}_0\end{array}\right)\\
&\quad+\sum_{s=1}^{S}CGE_{ks}\sum_{\mathbf{x}^p\in\mathcal{M}_{\mathbf{x}}}\mathbb{I}\left(\mathbf{z}^S(\mathbf{x}^p\cup\mathbf{e}_\ell)=\mathbf{e}_s\right)M(\bar{\mathbf{x}}\left(\mathbf{x}^p\right),|\mathbf{x}^p|,\boldsymbol{\mu})f_n(|\mathbf{x}^p|+2)\\
&\hspace{22em}\text{(by (147))}\\
&=E\left(y_i\left(\tilde{\mathbf{p}}\right)\middle|\begin{array}{l}\mathbf{x}_i=\mathbf{e}_k,\\ \mathbf{z}^S(\tilde{\mathbf{p}})=\mathbf{e}_0\end{array}\right)+\sum_{s=1}^{S}CGE_{ks}w^P_{s\ell}(\boldsymbol{\mu}) \qquad (150)
\end{aligned}
$$

Then by the definition of $CPE_{k\ell}$:

$$\begin{aligned}
CPE_{k\ell} &= E\left(y_i\left(\{j\}\cup\tilde{\mathbf{q}}\right) - y_i\left(\{j'\}\cup\tilde{\mathbf{q}}\right)\big|\, \mathbf{x}_i=\mathbf{e}_k, \mathbf{x}_j=\mathbf{e}_\ell, \mathbf{x}_{j'}=\mathbf{e}_0\right) \quad \text{(definition)} \\
&= \begin{array}{l} E\left(y_i\left(\{j\}\cup\tilde{\mathbf{q}}\right)\big|\, \mathbf{x}_i=\mathbf{e}_k, \mathbf{x}_j=\mathbf{e}_\ell\right) \\ -\, E\left(y_i\left(\{j\}\cup\tilde{\mathbf{q}}\right)\big|\, \mathbf{x}_i=\mathbf{e}_k, \mathbf{x}_j=\mathbf{e}_0\right) \end{array} \\
&= E\left(y_i\left(\tilde{\mathbf{p}}\right)\middle|\begin{array}{l}\mathbf{x}_i=\mathbf{e}_k,\\ \mathbf{z}^S(\tilde{\mathbf{p}})=\mathbf{e}_0\end{array}\right) + \sum_{s=1}^{S} CGE_{ks} w_{s\ell}^P(\boldsymbol{\mu}) \quad \text{(by (150))} \\
&\quad - E\left(y_i\left(\tilde{\mathbf{p}}\right)\middle|\begin{array}{l}\mathbf{x}_i=\mathbf{e}_k,\\ \mathbf{z}^S(\tilde{\mathbf{p}})=\mathbf{e}_0\end{array}\right) - \sum_{s=1}^{S} CGE_{ks} w_{s0}^P(\boldsymbol{\mu}) \\
&= \sum_{s=1}^{S} CGE_{ks}\left(w_{s\ell}^P(\boldsymbol{\mu}) - w_{s0}^P(\boldsymbol{\mu})\right)
\end{aligned}$$

which is result (38′). Result (37′) follows by substitution of (38′) into (30′).

Proof for Proposition 10

Let $a, b \in \mathcal{T}$ be unobserved types, let $e \in \mathbb{R}^2$ be unobserved post-assignment shocks, let $n \geq 2$ be a positive integer, and let $a^{[j]}$ represent j copies of a. Using unobserved type 1 as an arbitrary reference type, peer separability implies:

$$\begin{aligned}
y\left(a, \left\{b^{[n-1]}\right\}, e\right) &= y\left(a, \left\{b^{[n-1]}\right\}, e\right) + \sum_{j=1}^{n-1} y\left(a, \left\{b^{[j-1]}, 1^{[n-j]}\right\}, e\right) - y\left(a, \left\{1, b^{[j-1]}, 1^{[n-j-1]}\right\}, e\right) \\
&= y\left(a, \left\{1^{[n-1]}\right\}, e\right) + \sum_{j=1}^{n-1} y\left(a, \left\{b^{[j]}, 1^{[n-j-1]}\right\}, e\right) - y\left(a, \left\{1, b^{[j-1]}, 1^{[n-j-1]}\right\}, e\right) \\
&= y\left(a, \left\{1^{[n-1]}\right\}, e\right) + \sum_{j=1}^{n-1} y\left(a, \left\{b, 1^{[n-2]}\right\}, e\right) - y\left(a, \left\{1^{[n-1]}\right\}, e\right) \\
&\qquad \text{(by PS)} \\
&= y\left(a, \left\{1^{[n-1]}\right\}, e\right) + (n-1)\left(y\left(a, \left\{b, 1^{[n-2]}\right\}, e\right) - y\left(a, \left\{1^{[n-1]}\right\}, e\right)\right)
\end{aligned}$$

Rearranging this result:

$$\begin{aligned}
y\left(a, \left\{b, 1^{[n-2]}\right\}, e\right) - y\left(a, \left\{1^{[n-1]}\right\}, e\right) &= \frac{y\left(a, \left\{b^{[n-1]}\right\}, e\right) - y\left(a, \left\{1^{[n-1]}\right\}, e\right)}{n-1} \\
&= PE(a, b, e, n) - PE(a, 1, e, n) \qquad (151)
\end{aligned}$$

Let $\mathbf{B} \in \mathcal{T}^{n-1}$ be a vector of unobserved types, let B_j be element j of $\mathbf{B}$ and let $B_{j:k}$ be a multiset containing the elements j through k of $\mathbf{B}$ ($B_{j:k} = \emptyset$ for $j < k$). Then:

$$\begin{aligned}
y\left(a, \{B_{1:n-1}\}, e\right) &= y\left(a, \{B_{1:n-1}\}, e\right) + \sum_{j=1}^{n-1} y\left(a, \left\{B_{1:j-1}, 1^{[n-j]}\right\}, e\right) - y\left(a, \left\{B_{1:j-1}, 1^{[n-j]}\right\}, e\right) \\
&= y\left(a, \left\{1^{[n-1]}\right\}, e\right) + \sum_{j=1}^{n-1} y\left(a, \left\{B_{1:j}, 1^{[n-j-1]}\right\}, e\right) - y\left(a, \left\{B_{1:j-1}, 1^{[n-j]}\right\}, e\right) \\
&= y\left(a, \left\{1^{[n-1]}\right\}, e\right) + \sum_{j=1}^{n-1} y\left(a, \left\{B_j, 1^{[n-2]}\right\}, e\right) - y\left(a, \left\{1^{[n-1]}\right\}, e\right) \\
&\qquad \text{(by PS)} \\
&= y\left(a, \left\{1^{[n-1]}\right\}, e\right) + \sum_{j=1}^{n-1} PE(a, B_j, e, n) - PE(a, 1, e, n) \\
&\qquad \text{(by (151))} \\
&= y\left(a, \left\{1^{[n-1]}\right\}, e\right) - \sum_{j=1}^{n-1} PE(a, 1, e, n) + \sum_{j=1}^{n-1} PE(a, B_j, e, n) \\
&= \sum_{j=1}^{n-1} PE(a, B_j, e, n) \qquad (152)
\end{aligned}$$

Let $\mathbf{p}$ be a peer group of any size. Then result (152) implies:

$$\begin{aligned}
y_i(\mathbf{p}) &= y\left(\tau_i, \{\tau_j\}_{j \in \mathbf{p}}, \epsilon_i\right) \\
&= \sum_{j \in \mathbf{p}} PE(\tau_i, \tau_j, \epsilon_i, |\mathbf{p}| + 1)
\end{aligned}$$

which is result (44′).

Next, note that $(\tau_i, \tau_j) \perp\!\!\!\perp \tau_{j'}$ by equation (10) and $\epsilon_i \perp\!\!\!\perp \tau_{j'}$ so:

$$(PE(\tau_i, \tau_j, \epsilon_i, n), \mathbf{x}_i, \mathbf{x}_j) \perp\!\!\!\perp \mathbf{x}_{j'} \qquad (153)$$

Let $\tilde{\mathbf{q}}$ be a random draw of $n-2$ peers from $\mathcal{I} \setminus \{i, j\}$, and let:

$$
\begin{aligned}
CPE(\mathbf{x}^o, \mathbf{x}^p, n) &\equiv E\left(y_i(\{j\} \cup \tilde{\mathbf{q}}) - y_i(\{j'\} \cup \tilde{\mathbf{q}}) \middle| \mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p, \mathbf{x}_{j'} = \mathbf{e}_0\right) \\
&= E\left(\begin{array}{c} \left(PE(\tau_i, \tau_j, \epsilon_i, n) + \sum_{j'' \in \tilde{\mathbf{q}}} PE(\tau_i, \tau_{j''}, \epsilon_i, n)\right) \\ -\left(PE(\tau_i, \tau_{j'}, \epsilon_i, n) + \sum_{j'' \in \tilde{\mathbf{q}}} PE(\tau_i, \tau_{j''}, \epsilon_i, n)\right) \end{array} \middle| \begin{array}{c} \mathbf{x}_i = \mathbf{x}^o, \\ \mathbf{x}_j = \mathbf{x}^p, \\ \mathbf{x}_{j'} = \mathbf{e}_0 \end{array} \right) \\
&\qquad \text{(by (44'))} \\
&= E\left(PE(\tau_i, \tau_j, \epsilon_i, n) | \mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p, \mathbf{x}_{j'} = \mathbf{e}_0\right) \\
&\quad - E\left(PE(\tau_i, \tau_{j'}, \epsilon_i, n) \middle| \mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p, \mathbf{x}_{j'} = \mathbf{e}_0\right) \\
&= E\left(PE(\tau_i, \tau_j, \epsilon_i, n) | \mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p, \right) \qquad \text{(by (153))} \\
&\quad - E\left(PE(\tau_i, \tau_{j'}, \epsilon_i, n) \middle| \mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_{j'} = \mathbf{e}_0\right) \\
&= E\left(PE(\tau_i, \tau_j, \epsilon_i, n) | \mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p\right) \\
&\quad - E\left(PE(\tau_i, \tau_j, \epsilon_i, n) | \mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{e}_0\right)
\end{aligned}
$$

which is result (45′), and:

$$
\begin{aligned}
APE(\mathbf{x}^p, n) &\equiv E\left(y_i(\{j\} \cup \tilde{\mathbf{q}}) - y_i(\{j'\} \cup \tilde{\mathbf{q}}) \middle| \mathbf{x}_j = \mathbf{x}^p, \mathbf{x}_{j'} = \mathbf{e}_0\right) \\
&= E\left(\begin{array}{c} \left(PE(\tau_i, \tau_j, \epsilon_i, n) + \sum_{j'' \in \tilde{\mathbf{q}}} PE(\tau_i, \tau_{j''}, \epsilon_i, n)\right) \\ -\left(PE(\tau_i, \tau_{j'}, \epsilon_i, n) + \sum_{j'' \in \tilde{\mathbf{q}}} PE(\tau_i, \tau_{j''}, \epsilon_i, n)\right) \end{array} \middle| \begin{array}{c} \mathbf{x}_j = \mathbf{x}^p, \\ \mathbf{x}_{j'} = \mathbf{e}_0 \end{array} \right) \\
&\qquad \text{(by (44'))} \\
&= E\left(PE(\tau_i, \tau_j, \epsilon_i, n) | \mathbf{x}_j = \mathbf{x}^p, \mathbf{x}_{j'} = \mathbf{e}_0\right) \\
&\quad - E\left(PE(\tau_i, \tau_{j'}, \epsilon_i, n) \middle| \mathbf{x}_j = \mathbf{x}^p, \mathbf{x}_{j'} = \mathbf{e}_0\right) \\
&= E\left(PE(\tau_i, \tau_j, \epsilon_i, n) | \mathbf{x}_j = \mathbf{x}^p\right) \qquad \text{(by (153))} \\
&\quad - E\left(PE(\tau_i, \tau_{j'}, \epsilon_i, n) \middle| \mathbf{x}_{j'} = \mathbf{e}_0\right) \\
&= E\left(PE(\tau_i, \tau_j, \epsilon_i, n) | \mathbf{x}_j = \mathbf{x}^p\right) - E\left(PE(\tau_i, \tau_j, \epsilon_i, n) | \mathbf{x}_j = \mathbf{e}_0\right)
\end{aligned}
$$

which is result (46′).

Proof for Lemma 2

Let $\tilde{\mathbf{G}} \in \mathcal{G}^I$ be a random group assignment that satisfies (SA). Choose any $\mathbf{X_0} \in \mathbb{R}^{I\times K}$ and $\mathbf{G_0} \in \mathcal{G}^I$ such that $\mathbf{g}_i^0 \equiv (i, \mathbf{p}(i, \mathbf{G_0}))$ is a vector of length n. For any matrix $\mathbf{M}$ and vector $\mathbf{v}$, let $\mathbf{M}[\mathbf{v}]$ be the submatrix constructed from rows $\mathbf{v}$ in matrix $\mathbf{M}$. Then:

$$
\begin{aligned}
E\left(y_i(\mathbf{p}(i,\tilde{\mathbf{G}})) \middle| \begin{matrix}\mathbf{X}=\mathbf{X_0}, \\ \tilde{\mathbf{G}}=\mathbf{G_0}\end{matrix}\right) &= E\left(y\left(\tau_i, \{\tau_j\}_{j\in\mathbf{p}(i,\tilde{\mathbf{G}})}, \epsilon_i\right) \middle| \mathbf{X}=\mathbf{X_0}, \tilde{\mathbf{G}}=\mathbf{G_0}\right) \\
&= E\left(y\left(\tau_i, \{\tau_j\}_{j\in\mathbf{p}(i,\mathbf{G_0})}, \epsilon_i\right) \middle| \mathbf{X}=\mathbf{X_0}, \tilde{\mathbf{G}}=\mathbf{G_0}\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(conditioning rule)} \\
&= E\left(y\left(\tau_i, \{\tau_j\}_{j\in\mathbf{p}(i,\mathbf{G_0})}, \epsilon_i\right) \middle| \mathbf{X}=\mathbf{X_0}\right) \qquad\qquad \text{(by SA)} \\
&= \sum_{\mathbf{T_0}\in\mathcal{T}^n} \sum_{\epsilon_A\in\mathbb{S}_\epsilon} y\left(\mathbf{T_0}[1], \mathbf{T_0}[2:n], \epsilon_A\right) \\
&\qquad \times \Pr(\mathbf{T}[\mathbf{g}_i^0] = \mathbf{T_0}, \epsilon_i = \epsilon_A | \mathbf{X}=\mathbf{X_0}) \\
&= \sum_{\mathbf{T_0}\in\mathcal{T}^n} \sum_{\epsilon_A\in\mathbb{S}_\epsilon} y\left(\mathbf{T_0}[1], \mathbf{T_0}[2:n], \epsilon_A\right) \Pr(\epsilon_i = \epsilon_A) \\
&\qquad \times \left(\prod_{j=1}^n \Pr\left(\mathbf{T}[\mathbf{g}_i^0[j]] = \mathbf{T_0}[j] \,\middle|\, \mathbf{X}[\mathbf{g}_i^0[j]] = \mathbf{X_0}[\mathbf{g}_i^0[j]]\right)\right) \\
&\qquad\qquad \text{(since } \epsilon_i \perp\!\!\!\perp (\tau_i, \mathbf{x}_i) \perp\!\!\!\perp (\tau_j, \mathbf{x}_j) \text{ for all } i \neq j) \\
&= \sum_{\mathbf{T_0}\in\mathcal{T}^n} \sum_{\epsilon_A\in\mathbb{S}_\epsilon} y\left(\mathbf{T_0}[1], \mathbf{T_0}[2:n], \epsilon_A\right) \Pr(\epsilon_i = \epsilon_A) \\
&\qquad \times \prod_{j=1}^n \frac{f_\tau(\mathbf{T_0}[j]) \mathbb{I}\left(\mathbf{x}(\mathbf{T_0}[j]) = \mathbf{X_0}[\mathbf{g}_i^0[j]]\right)}{\sum_{\tau\in\mathcal{T}} f_\tau(\tau) \mathbb{I}\left(\mathbf{x}(\tau) = \mathbf{X_0}[\mathbf{g}_i^0[j]]\right)} \\
&\equiv \zeta(\mathbf{x}_i(\mathbf{X_0}), \bar{\mathbf{x}}_i(\mathbf{X_0}, \mathbf{G_0}), n) \qquad\qquad (154)
\end{aligned}
$$

Note that the last step in equation (154) makes use of the fact that $(\bar{\mathbf{x}}_i, n_{g_i})$ fully describes the frequency distribution of peer characteristics $\{\mathbf{x}_j\}_{j\in\mathbf{p}_i}$, and that the $\zeta(\cdot)$ function depends on the type distribution $f_\tau(\cdot)$ but not on the probability distribution of $\tilde{\mathbf{G}}$ other than through the (SA) condition.

Since a purely random group assignment and the true group assignment $\mathbf{G}$ both satisfy (SA), their conditional expectation functions are both given by (154).

$$
\begin{aligned}
E(y_i|\mathbf{x}_i = \mathbf{x}^o, \bar{\mathbf{x}}_i = \mathbf{x}^p, n_{g_i} = n) &= E(E(y_i|\mathbf{X}, \mathbf{G})|\mathbf{x}_i = \mathbf{x}^o, \bar{\mathbf{x}}_i = \mathbf{x}^p, n_{g_i} = n) \\
&\qquad \text{(law of iterated expectations)} \\
&= E(E(y_i(\mathbf{p}(i, \mathbf{G}))|\mathbf{X}, \mathbf{G})|\mathbf{x}_i = \mathbf{x}^o, \bar{\mathbf{x}}_i = \mathbf{x}^p, n_{g_i} = n) \\
&= E(\zeta(\mathbf{x}_i(\mathbf{X}), \bar{\mathbf{x}}_i(\mathbf{X}, \mathbf{G}), n_{g_i})|\mathbf{x}_i = \mathbf{x}^o, \bar{\mathbf{x}}_i = \mathbf{x}^p, n_{g_i} = n) \\
&\qquad \text{(by (154))} \\
&= \zeta(\mathbf{x}^o, \mathbf{x}^p, n) \qquad (155) \\
E(y_i(\tilde{\mathbf{p}})|\mathbf{x}_i = \mathbf{x}^o, \bar{\mathbf{x}}_i(\tilde{\mathbf{p}}) = \mathbf{x}^p) &= E(E(y_i(\tilde{\mathbf{p}})|\mathbf{X}, \tilde{\mathbf{G}})|\mathbf{x}_i = \mathbf{x}^o, \bar{\mathbf{x}}_i(\tilde{\mathbf{p}}) = \mathbf{x}^p) \\
&\qquad \text{(law of iterated expectations)} \\
&= E(\zeta(\mathbf{x}_i(\mathbf{X}), \bar{\mathbf{x}}_i(\mathbf{X}, \tilde{\mathbf{p}}), n)|\mathbf{x}_i = \mathbf{x}^o, \bar{\mathbf{x}}_i(\tilde{\mathbf{p}}) = \mathbf{x}^p) \\
&\qquad \text{(by (154))} \\
&= \zeta(\mathbf{x}^o, \mathbf{x}^p, n) \qquad (156)
\end{aligned}
$$

Therefore they are equal, which is result (42′).

Proof for Proposition 11

1. By (PS), Part 1 of Proposition 10 applies. Let $\boldsymbol{\zeta}^n \equiv (\zeta_0^n, \boldsymbol{\zeta}_1^n, \boldsymbol{\zeta}_2^n, \boldsymbol{\zeta}_3^n)$ satisfy:

$$
E(PE(\tau_i, \tau_j, \epsilon_i, n)|\mathbf{x}_i = \mathbf{e}_k, \mathbf{x}_j = \mathbf{e}_\ell) = \zeta_0^n + \mathbf{e}_k\boldsymbol{\zeta}_1^n + \mathbf{e}_\ell\boldsymbol{\zeta}_2^n + \mathbf{e}_k\boldsymbol{\zeta}_3^n\mathbf{e}_\ell' \qquad (157)
$$

The linear functional form in (157) is without loss of generality since $\mathbf{x}$ is categorical. The parameter of interest $CPE_{k\ell}(n)$ can be expressed as a function of $\boldsymbol{\zeta}^n$:

$$
\begin{aligned}
CPE_{k\ell}(n) &= E(PE(\tau_i, \tau_j, \epsilon_i, n)|\mathbf{x}_i = \mathbf{e}_k, \mathbf{x}_j = \mathbf{e}_\ell) \\
&\quad - E(PE(\tau_i, \tau_j, \epsilon_i, n)|\mathbf{x}_i = \mathbf{e}_k, \mathbf{x}_j = \mathbf{e}_0) \\
&\qquad \text{(by (45′) in Proposition 10)} \\
&= \left(\zeta_0^n + \mathbf{e}_k\boldsymbol{\zeta}_1^n + \mathbf{e}_\ell\boldsymbol{\zeta}_2^n + \mathbf{e}_k\boldsymbol{\zeta}_3^n\mathbf{e}_\ell'\right) \qquad \text{(by (157))} \\
&\quad - \left(\zeta_0^n + \mathbf{e}_k\boldsymbol{\zeta}_1^n + \mathbf{e}_0\boldsymbol{\zeta}_2^n + \mathbf{e}_k\boldsymbol{\zeta}_3^n\mathbf{e}_0'\right) \\
&= \left(\zeta_0^n + \mathbf{e}_k\boldsymbol{\zeta}_1^n + \mathbf{e}_\ell\boldsymbol{\zeta}_2^n + \mathbf{e}_k\boldsymbol{\zeta}_3^n\mathbf{e}_\ell'\right) \qquad \text{(since } \mathbf{e}_0 = \mathbf{0}) \\
&\quad - (\zeta_0^n + \mathbf{e}_k\boldsymbol{\zeta}_1^n) \\
&= \mathbf{e}_\ell\boldsymbol{\zeta}_2^n + \mathbf{e}_k\boldsymbol{\zeta}_3^n\mathbf{e}_\ell' \\
&= \zeta_{2\ell}^n + \zeta_{3k\ell}^n \qquad (158)
\end{aligned}
$$

The next step is to show the relationship between the coefficients in $\boldsymbol{\zeta}$ and the coefficients in $\boldsymbol{\beta}$:

$$
\begin{aligned}
E(y_i|\mathbf{X},\mathbf{G}) &= E(y_i(\mathbf{p}_i)|\mathbf{X},\mathbf{G}) \\
&= E\left(\sum_{j\in\mathbf{p}_i} PE(\tau_i,\tau_j,\epsilon_i,n_{g_i})\middle|\mathbf{X},\mathbf{G}\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{(PS} \implies \text{(44) in Proposition 3)} \\
&= E\left(\sum_{j=1}^{I}\sum_{n=1}^{I} PE(\tau_i,\tau_j,\epsilon_i,n)\mathbb{I}\left(j\in\mathbf{p}_i\right)\mathbb{I}\left(n_{g_i}=n\right)\middle|\mathbf{X},\mathbf{G}\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{(where } \mathbb{I}\left(\cdot\right) \text{ is the indicator function)} \\
&= \sum_{n=1}^{I}\mathbb{I}\left(n_{g_i}=n\right)\sum_{j=1}^{I} E(PE(\tau_i,\tau_j,\epsilon_i,n)|\mathbf{X},\mathbf{G})\mathbb{I}\left(j\in\mathbf{p}_i\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{(since } \mathbf{p}_i \text{ is a function of } \mathbf{G}\text{)} \\
&= \sum_{n=1}^{I}\mathbb{I}\left(n_{g_i}=n\right)\sum_{j\in\mathbf{p}_i} E(PE(\tau_i,\tau_j,\epsilon_i,n)|\mathbf{X},\mathbf{G}) \\
&= \sum_{n=1}^{I}\mathbb{I}\left(n_{g_i}=n\right)\sum_{j\in\mathbf{p}_i} E(PE(\tau_i,\tau_j,\epsilon_i,n)|\mathbf{X}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{(since RA, (5}'\text{)} \implies (\mathbf{T},\boldsymbol{\epsilon},\mathbf{X}) \perp\!\!\!\perp (\mathbf{G},\mathbf{p}_i)\text{)} \\
&= \sum_{n=1}^{I}\mathbb{I}\left(n_{g_i}=n\right)\sum_{j\in\mathbf{p}_i} E(PE(\tau_i,\tau_j,\epsilon_i,n)|\mathbf{x}_i,\mathbf{x}_j) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{(since (10), (5}'\text{)} \implies (\tau_i,\tau_j,\epsilon_i) \perp\!\!\!\perp \left\{\tau_{j'}\right\}_{j'\notin\{i,j\}}\text{)} \\
&= \sum_{n=1}^{I}\mathbb{I}\left(n_{g_i}=n\right)\sum_{j\in\mathbf{p}_i} (\zeta_0^n+\mathbf{x}_i\boldsymbol{\zeta}_1^n+\mathbf{x}_j\boldsymbol{\zeta}_2^n+\mathbf{x}_i\boldsymbol{\zeta}_3^n\mathbf{x}_j') \qquad \text{(by (157))} \\
&= \zeta_0^{n_{g_i}}(n_{g_i}-1)+\mathbf{x}_i\boldsymbol{\zeta}_1^{n_{g_i}}(n_{g_i}-1)+\bar{\mathbf{x}}_i\boldsymbol{\zeta}_2^{n_{g_i}}(n_{g_i}-1)+\mathbf{x}_i\boldsymbol{\zeta}_3^{n_{g_i}}(n_{g_i}-1)\bar{\mathbf{x}}_i'
\end{aligned} \tag{159}
$$

Applying the law of iterated expectations to this result:

$$E(y_i|\mathbf{x}_i = \mathbf{x}, \bar{\mathbf{x}}_i = \bar{\mathbf{x}}, n_{g_i} = n) = E(E(y_i|\mathbf{X}, \mathbf{G})|\mathbf{x}_i = \mathbf{x}, \bar{\mathbf{x}}_i = \bar{\mathbf{x}}, n_{g_i} = n)$$

(law of iterated expectations)

$$= E\left(\begin{matrix} \zeta_0^{n_{g_i}}(n_{g_i}-1) + \mathbf{x}_i\boldsymbol{\zeta}_1^{n_{g_i}}(n_{g_i}-1) \\ + \bar{\mathbf{x}}_i\boldsymbol{\zeta}_2^{n_{g_i}}(n_{g_i}-1) + \mathbf{x}_i\boldsymbol{\zeta}_3^{n_{g_i}}(n_{g_i}-1)\bar{\mathbf{x}}_i' \end{matrix}\middle| \begin{matrix} \mathbf{x}_i = \mathbf{x}, \\ \bar{\mathbf{x}}_i = \bar{\mathbf{x}} \\ n_{g_i} = n \end{matrix}\right)$$

(by (159))

$$= \zeta_0^n(n-1) + \mathbf{x}\boldsymbol{\zeta}_1^n(n-1) + \bar{\mathbf{x}}\boldsymbol{\zeta}_2^n(n-1) + \mathbf{x}\boldsymbol{\zeta}_3^n(n-1)\bar{\mathbf{x}}' \tag{160}$$

Applying the law of iterated projections to this result:

$$L(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i\bar{\mathbf{x}}_i'; n) = L(E(y_i|\mathbf{x}_i, \bar{\mathbf{x}}_i, n_{g_i})|\mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i\bar{\mathbf{x}}_i'; n)$$

(law of iterated projections)

$$= L\left(\begin{matrix} \zeta_0^{n_{g_i}}(n_{g_i}-1) + \mathbf{x}_i\boldsymbol{\zeta}_1^{n_{g_i}}(n_{g_i}-1) \\ + \bar{\mathbf{x}}_i\boldsymbol{\zeta}_2^{n_{g_i}}(n_{g_i}-1) + \mathbf{x}_i\boldsymbol{\zeta}_3^{n_{g_i}}(n_{g_i}-1)\bar{\mathbf{x}}_i' \end{matrix}\middle| \mathbf{x}_i, \bar{\mathbf{x}}_i, \mathbf{x}_i\bar{\mathbf{x}}_i'; n\right)$$

(by (160))

$$= \underbrace{\zeta_0^n(n-1)}_{\beta_0^n} + \mathbf{x}_i\underbrace{\boldsymbol{\zeta}_1^n(n-1)}_{\boldsymbol{\beta}_1^n} + \bar{\mathbf{x}}_i\underbrace{\boldsymbol{\zeta}_2^n(n-1)}_{\boldsymbol{\beta}_2^n} + \mathbf{x}_i\underbrace{\boldsymbol{\zeta}_3^n(n-1)}_{\boldsymbol{\beta}_3^n}\bar{\mathbf{x}}_i' \tag{161}$$

So $\boldsymbol{\beta}_2^n = \boldsymbol{\zeta}_2^n(n-1)$, $\boldsymbol{\beta}_3^n = \boldsymbol{\zeta}_3^n(n-1)$ and:

$$\begin{aligned} CPE_{k,\ell}(n) &= \zeta_{2\ell}^n + \zeta_{3k\ell}^n && \text{(by (158))} \\ &= \frac{\beta_{2\ell}^n + \beta_{3k\ell}^n}{n-1} && \text{(by (161))} \end{aligned}$$

which is result $(50a')$. Result $(50b')$ follows by substituting $(50a')$ into the definition of $CPE_{k\ell}$ in equation $(19b')$.

To get result ($49a'$), first note that:

$$
\begin{aligned}
E(PE(\tau_i,\tau_j,\epsilon_i,n)|\mathbf{x}_j=\mathbf{x}) &= E(E(PE(\tau_i,\tau_j,\epsilon_i,n)|\mathbf{x}_i,\mathbf{x}_j)|\mathbf{x}_j=\mathbf{x}) \\
&\qquad\qquad \text{(law of iterated expectations)} \\
&= E(\zeta_0^n+\mathbf{x}_i\boldsymbol{\zeta}_1^n+\mathbf{x}_j\boldsymbol{\zeta}_2^n+\mathbf{x}_i\boldsymbol{\zeta}_3^n\mathbf{x}_j'|\mathbf{x}_j=\mathbf{x}) \quad \text{(by (157))} \\
&= \zeta_0^n+E(\mathbf{x}_i|\mathbf{x}_j=\mathbf{x})\boldsymbol{\zeta}_1^n+\mathbf{x}\boldsymbol{\zeta}_2^n+E(\mathbf{x}_i|\mathbf{x}_j=\mathbf{x})\boldsymbol{\zeta}_3^n\mathbf{x}' \\
&\qquad\qquad \text{(conditioning rule)} \\
&= \zeta_0^n+E(\mathbf{x}_i)\boldsymbol{\zeta}_1^n+\mathbf{x}\boldsymbol{\zeta}_2^n+E(\mathbf{x}_i)\boldsymbol{\zeta}_3^n\mathbf{x}' \\
&\qquad\qquad \text{(since (10)} \implies \mathbf{x}_i \perp\!\!\!\perp \mathbf{x}_j) \\
&= \left(\zeta_0^n+E(\mathbf{x}_i)\right)\boldsymbol{\zeta}_1^n+\mathbf{x}\left(\boldsymbol{\zeta}_2^n+\boldsymbol{\zeta}_3^{n\prime}E(\mathbf{x}_i')\right) \qquad (162)
\end{aligned}
$$

Equation ($46'$) from Proposition 10 implies:

$$
\begin{aligned}
APE_\ell(n) &= E(PE(\tau_i,\tau_j,\epsilon_i,n)|\mathbf{x}_j=\mathbf{e}_\ell)-E(PE(\tau_i,\tau_j,\epsilon_i,n)|\mathbf{x}_j=\mathbf{e}_0) \\
&\qquad\qquad \text{(PS} \implies (46') \text{ in Proposition 10)} \\
&= \left((\zeta_0^n+E(\mathbf{x}_i)\boldsymbol{\zeta}_1^n)+\mathbf{e}_\ell\left(\boldsymbol{\zeta}_2^n+\boldsymbol{\zeta}_3^{n\prime}E(\mathbf{x}_i')\right)\right) \qquad \text{(by (162))} \\
&\quad -\left((\zeta_0^n+E(\mathbf{x}_i)\boldsymbol{\zeta}_1^n)+\mathbf{e}_0\left(\boldsymbol{\zeta}_2^n+\boldsymbol{\zeta}_3^{n\prime}E(\mathbf{x}_i')\right)\right) \\
&= \left((\zeta_0^n+E(\mathbf{x}_i)\boldsymbol{\zeta}_1^n)+\mathbf{e}_\ell\left(\boldsymbol{\zeta}_2^n+\boldsymbol{\zeta}_3^{n\prime}E(\mathbf{x}_i')\right)\right) \qquad \text{(since } \mathbf{e}_0=0) \\
&\quad -((\zeta_0^n+E(\mathbf{x}_i)\boldsymbol{\zeta}_1^n)) \\
&= \mathbf{e}_\ell\left(\boldsymbol{\zeta}_2^n+\boldsymbol{\zeta}_3^{n\prime}E(\mathbf{x}_i')\right) \qquad (163)
\end{aligned}
$$

Having expressed $L(y_i|\bar{\mathbf{x}}_i;n)$ in terms of the the coefficients in $\boldsymbol{\alpha}^n$, it can also be expressed in terms of the coefficients in $\boldsymbol{\zeta}^n$:

$$\begin{aligned}
L(y_i|\bar{\mathbf{x}}_i;n) &= L(L(y_i|\mathbf{x}_i,\bar{\mathbf{x}}_i,\mathbf{x}_i'\bar{\mathbf{x}}_i;n)|\bar{\mathbf{x}}_i;n) && \text{(law of iterated projections)} \\
&= L\left(\begin{matrix}\zeta_0^n(n-1)+\mathbf{x}_i\boldsymbol{\zeta}_1^n(n-1)\\ +\bar{\mathbf{x}}_i\boldsymbol{\zeta}_2^n(n-1)+\mathbf{x}_i\boldsymbol{\zeta}_3^n(n-1)\bar{\mathbf{x}}_i'\end{matrix}\middle|\bar{\mathbf{x}}_i;n\right) && \text{(by (161))} \\
&= \zeta_0^n(n-1)+L(\mathbf{x}_i|\bar{\mathbf{x}}_i)\boldsymbol{\zeta}_1^n(n-1) \\
&\quad +\bar{\mathbf{x}}_i\boldsymbol{\zeta}_2^n(n-1)+L(\mathbf{x}_i\boldsymbol{\zeta}_3^n(n-1)\bar{\mathbf{x}}_i'|\bar{\mathbf{x}}_i) \\
& && \text{(property of linear projection)} \\
&= \zeta_0^n(n-1)+E(\mathbf{x}_i)\boldsymbol{\zeta}_1^n(n-1) && (\text{RA} \implies \mathbf{x}_i \perp\!\!\!\perp \bar{\mathbf{x}}_i) \\
&\quad +\bar{\mathbf{x}}_i\boldsymbol{\zeta}_2^n(n-1)+E(\mathbf{x}_i)\boldsymbol{\zeta}_3^n(n-1)\bar{\mathbf{x}}_i' \\
&= \underbrace{\zeta_0^n(n-1)+E(\mathbf{x}_i)\boldsymbol{\zeta}_1^n(n-1)}_{\alpha_0^n}+\bar{\mathbf{x}}_i\underbrace{\left(\boldsymbol{\zeta}_2^n(n-1)+\boldsymbol{\zeta}_3^{n\prime}E(\mathbf{x}_i')(n-1)\right)}_{\boldsymbol{\alpha}_1^n}
\end{aligned} \tag{164}$$

So $\boldsymbol{\alpha}_1^n = (\boldsymbol{\zeta}_2^n(n-1)+\boldsymbol{\zeta}_3^{n\prime}E(\mathbf{x}_i')(n-1))$ and:

$$\begin{aligned}
APE_\ell(n) &= \mathbf{e}_\ell\left(\boldsymbol{\zeta}_2^n+\boldsymbol{\zeta}_3^{n\prime}E(\mathbf{x}_i')\right) && \text{(by (163))} \\
&= \mathbf{e}_\ell\frac{\boldsymbol{\alpha}_1^n}{n-1} && \text{(by (164))} \\
&= \frac{\alpha_{1\ell}^n}{n-1}
\end{aligned}$$

which is the result in $(49a')$. Result $(49b')$ can be derived by substituting $(49a')$ into the definition of APE_ℓ in equation $(17b')$

2. Let $\tilde{\mathbf{G}}$ be a purely random group assignment whose group size distribution is f_n and let $\tilde{\mathbf{p}}_i = \mathbf{p}(i,\tilde{\mathbf{G}})$. Since outcomes are peer-separable (PS) and $\tilde{\mathbf{G}}$ satisfies (RA), Part 1 of Proposition 11 applies to the counterfactual outcomes. Therefore, the counterfactual CEF is linear as shown in equation (161):

$$\begin{aligned}
E\left(y_i(\tilde{\mathbf{p}}_i)\middle|\begin{matrix}\mathbf{x}_i=\mathbf{x}^o,\\ \bar{\mathbf{x}}(\tilde{\mathbf{p}}_i)=\mathbf{x}^p,\\ |\tilde{\mathbf{p}}_i|=n-1\end{matrix}\right) &= L\left(y_i(\tilde{\mathbf{p}}_i)\middle|\begin{matrix}\mathbf{x}_i=\mathbf{x}^o,\\ \bar{\mathbf{x}}(\tilde{\mathbf{p}}_i)=\mathbf{x}^p, \quad ;n\\ \mathbf{x}_i'\bar{\mathbf{x}}(\tilde{\mathbf{p}}_i)=\mathbf{x}'\bar{\mathbf{x}}\end{matrix}\right) && \text{(by (161))} \\
&= \tilde{\beta}_0^n+\mathbf{x}^o\tilde{\boldsymbol{\beta}}_1^n+\mathbf{x}^p\tilde{\boldsymbol{\beta}}_2^n+\mathbf{x}^o\tilde{\boldsymbol{\beta}}_3^n\mathbf{x}^{p\prime}
\end{aligned} \tag{165}$$

where $\bar{\mathbf{x}}(\mathbf{p}) \equiv \bar{\mathbf{x}}\left(\{\mathbf{x}_j\}_{j\in\mathbf{p}}\right)$ and $\tilde{\boldsymbol{\beta}}^n \equiv (\tilde{\beta}_0^n,\tilde{\boldsymbol{\beta}}_1^n,\tilde{\boldsymbol{\beta}}_2^n,\tilde{\boldsymbol{\beta}}_3^n)$ is the best linear predictor coefficients from the counterfactual regression. Proposition 11 also implies that:

$$CPE_{k\ell}(n) = \frac{\tilde{\beta}^n_{2\ell} + \tilde{\beta}^n_{3k\ell}}{n-1} \qquad \text{(by (50}a'\text{) in Proposition 11)}$$

Since $\mathbf{G}$ satisfies (SA), Lemma 2 applies. Therefore, the actual CEF is the same as the counterfactual CEF, and the same applies to the best linear predictor. Therefore, $\boldsymbol{\beta}^n = \tilde{\boldsymbol{\beta}}^n$ and:

$$CPE_{k\ell}(n) = \frac{\beta^n_{2\ell} + \beta^n_{3k\ell}}{n-1} \qquad \text{(since } \beta = \tilde{\beta}\text{)}$$

which is result $(54a')$. Result $(53a')$ follows from substitution of result $(54a')$ into result (109) of Proposition 9. Results $(53b')$ and $(54b')$ follow by applying $(53a')$ and $(54a')$ to the definitions of APE_ℓ and $CPE_{k\ell}$ respectively.

Proof for Proposition 12

1. By (RA), the actual peer group $\mathbf{p}_i$ is a purely random draw from the same distribution as $\tilde{\mathbf{p}}$, so the joint distribution of $\left(y_i(\tilde{\mathbf{p}}), \mathbf{x}_i, \{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right)$ is identical to the joint distribution of $\left(y_i(\mathbf{p}_i), \mathbf{x}_i, \{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right)$. Letting $\mathbf{z}(\mathbf{p}) \equiv \mathbf{z}\left(\{\mathbf{x}_j\}_{j\in\mathbf{p}}\right)$, this implies:

$$\begin{aligned}
CGE_{kb} = {} & E\left(y_i(\tilde{\mathbf{p}})\,|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b, \mathbf{z}(\tilde{\mathbf{q}}) = \mathbf{e}_0\right) && \text{(by (25}'\text{))} \\
& -E\left(y_i(\tilde{\mathbf{q}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b, \mathbf{z}(\tilde{\mathbf{q}}) = \mathbf{e}_0\right) \\
= {} & E\left(y_i(\tilde{\mathbf{p}})\,|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b\right) && \text{(since } \tilde{\mathbf{p}} \perp\!\!\!\perp \tilde{\mathbf{q}}\text{)} \\
& -E\left(y_i(\tilde{\mathbf{q}})|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\tilde{\mathbf{q}}) = \mathbf{e}_0\right) \\
= {} & E\left(y_i(\mathbf{p}_i)|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\mathbf{p}_i) = \mathbf{e}_b\right) && \text{(RA} \implies \text{same joint distribution)} \\
& -E\left(y_i(\mathbf{p}_i)|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}(\mathbf{p}_i) = \mathbf{e}_0\right) \\
= {} & E(y_i|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}_i = \mathbf{e}_b) - E(y_i|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}_i = \mathbf{e}_0) && (166)
\end{aligned}$$

Since $\mathbf{x}_i$ and $\mathbf{z}_i$ are categorical, $E(y_i|\mathbf{x}_i, \mathbf{z}_i)$ is trivially linear in $(\mathbf{x}_i, \mathbf{z}_i, \mathbf{x}_i'\mathbf{z}_i)$. Therefore:

$$\begin{aligned}
E(y_i|\mathbf{x}_i, \mathbf{z}_i) &= L(y_i|\mathbf{x}_i, \mathbf{z}_i, \mathbf{x}_i'\mathbf{z}_i) \\
&= \delta_0 + \mathbf{x}_i\boldsymbol{\delta}_1 + \mathbf{z}_i\boldsymbol{\delta}_2 + \mathbf{x}_i\boldsymbol{\delta}_3\mathbf{z}_i' && \text{(by (58}'\text{))}
\end{aligned}$$

Combining these two results produces:

$$\begin{aligned} CGE_{kb} &= E(y_i|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}_i = \mathbf{e}_b) - E(y_i|\mathbf{x}_i = \mathbf{e}_k, \mathbf{z}_i = \mathbf{e}_0) && \text{(by (166))} \\ &= \left(\delta_0 + \mathbf{e}_k\boldsymbol{\delta}_1 + \mathbf{e}_b\boldsymbol{\delta}_2 + \mathbf{e}_k\boldsymbol{\delta}_3\mathbf{e}_b'\right) - \left(\delta_0 + \mathbf{e}_k\boldsymbol{\delta}_1 + \mathbf{e}_0\boldsymbol{\delta}_2 + \mathbf{e}_k\boldsymbol{\delta}_3\mathbf{e}_0'\right) \\ & && \text{(result above)} \\ &= \left(\delta_0 + \mathbf{e}_k\boldsymbol{\delta}_1 + \mathbf{e}_b\boldsymbol{\delta}_2 + \mathbf{e}_k\boldsymbol{\delta}_3\mathbf{e}_b'\right) - (\delta_0 + \mathbf{e}_k\boldsymbol{\delta}_1) && \text{(since } \mathbf{e}_0 = \mathbf{0}) \\ &= \mathbf{e}_b\boldsymbol{\delta}_2 + \mathbf{e}_k\boldsymbol{\delta}_3\mathbf{e}_b' \\ &= \delta_{2b} + \delta_{3kb} \end{aligned}$$

which is result (56′). Result (55′) can be established by similar reasoning:

$$\begin{aligned} AGE_b = &\quad E(y_i(\tilde{\mathbf{p}})|\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b, \mathbf{z}(\tilde{\mathbf{q}}) = \mathbf{e}_0) && \text{(by (23′))} \\ &\quad - E(y_i(\tilde{\mathbf{q}})|\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b, \mathbf{z}(\tilde{\mathbf{q}}) = \mathbf{e}_0) \\ &= E(y_i(\tilde{\mathbf{p}})|\mathbf{z}(\tilde{\mathbf{p}}) = \mathbf{e}_b) - E(y_i(\tilde{\mathbf{q}})|\mathbf{z}(\tilde{\mathbf{q}}) = \mathbf{e}_0) && \text{(since } \tilde{\mathbf{p}} \perp\!\!\!\perp \tilde{\mathbf{q}}) \\ &= \quad E(y_i(\mathbf{p}_i)|\mathbf{z}(\mathbf{p}_i) = \mathbf{e}_b) && \text{(RA} \implies \text{same joint distribution)} \\ &\quad - E(y_i(\mathbf{p}_i)|\mathbf{z}(\mathbf{p}_i) = \mathbf{e}_0) \\ &= E(y_i|\mathbf{z}_i = \mathbf{e}_b) - E(y_i|\mathbf{z}_i = \mathbf{e}_0) && (167) \end{aligned}$$

Since $\mathbf{z}_i$ is categorical, $E(y_i|\mathbf{z}_i)$ is trivially linear in $\mathbf{z}_i$. Therefore:

$$E(y_i|\mathbf{z}_i) = L(y_i|\mathbf{z}_i) = \gamma_0 + \mathbf{z}_i\boldsymbol{\gamma}_1 \tag{168}$$

Combining these two results:

$$\begin{aligned} AGE_b &= E(y_i|\mathbf{z}_i = \mathbf{e}_b) - E(y_i|\mathbf{z}_i = \mathbf{e}_0) && \text{(by (167))} \\ &= (\gamma_0 + \mathbf{e}_b\boldsymbol{\gamma}_1) - (\gamma_0 + \mathbf{e}_0\boldsymbol{\gamma}_1) && \text{(by (168))} \\ &= \mathbf{e}_b\boldsymbol{\gamma}_1 && \text{(since } \mathbf{e}_0 = \mathbf{0}) \\ &= \gamma_{1b} \end{aligned}$$

which is result (55′).

2. Let $\tilde{\mathbf{G}}$ be a purely random group assignment whose group size distribution is f_n and let $\tilde{\mathbf{p}}_i = \mathbf{p}(i, \tilde{\mathbf{G}})$. Let $\tilde{\boldsymbol{\lambda}} \equiv (\tilde{\lambda}_0, \tilde{\boldsymbol{\lambda}}_1, \tilde{\boldsymbol{\lambda}}_2, \tilde{\boldsymbol{\lambda}}_3)$ be the best linear predictor coefficients from the counterfactual regression model:

$$L(y_i(\tilde{\mathbf{p}}_i)|\mathbf{x}_i, \mathbf{z}^S(\tilde{\mathbf{p}}_i), \mathbf{x}_i'\mathbf{z}^S(\tilde{\mathbf{p}}_i)) = \tilde{\lambda}_0 + \mathbf{x}_i\tilde{\boldsymbol{\lambda}}_1 + \mathbf{z}^S(\tilde{\mathbf{p}}_i)\tilde{\boldsymbol{\lambda}}_2 + \mathbf{x}_i\tilde{\boldsymbol{\lambda}}_3\mathbf{z}^S(\tilde{\mathbf{p}}_i)' \tag{169}$$

where $\mathbf{z}^S(\mathbf{p}) \equiv \mathbf{z}^S\left(\bar{\mathbf{x}}\left(\{\mathbf{x}_j\}_{j\in\mathbf{p}}\right)\right)$. Since $\tilde{\mathbf{G}}$ satisfies (RA), Part 1 of this proposition applies to the counterfactual outcomes:

$$CGE_{ks}^S = \tilde{\lambda}_{2s} + \tilde{\lambda}_{3ks} \qquad \text{(by (166))}$$

The counterfactual CEF is linear since $\mathbf{z}^S(\cdot)$ is saturated:

$$\begin{aligned} E(y_i(\tilde{\mathbf{p}}_i)|\mathbf{x}_i = \mathbf{x}, \bar{\mathbf{x}}(\tilde{\mathbf{p}}_i) = \bar{\mathbf{x}}) &= L(y_i(\tilde{\mathbf{p}}_i)|\mathbf{x}_i, \mathbf{z}^S(\tilde{\mathbf{p}}_i), \mathbf{x}_i'\mathbf{z}^S(\tilde{\mathbf{p}}_i)) && (170) \\ &= \tilde{\lambda}_0 + \mathbf{x}\tilde{\boldsymbol{\lambda}}_1 + \mathbf{z}^S(\bar{\mathbf{x}})\tilde{\boldsymbol{\lambda}}_2 + \mathbf{x}\tilde{\boldsymbol{\lambda}}_3\mathbf{z}^S(\bar{\mathbf{x}})' && (171) \end{aligned}$$

Since $\mathbf{G}$ satisfies (SA), Lemma 1 applies, which implies that:

$$\begin{aligned} E(y_i|\mathbf{x}_i = \mathbf{x}, \bar{\mathbf{x}}_i = \bar{\mathbf{x}}) &= E(y_i(\tilde{\mathbf{p}}_i)|\mathbf{x}_i = \mathbf{x}, \bar{\mathbf{x}}(\tilde{\mathbf{p}}_i) = \bar{\mathbf{x}}) && \text{(by (42) in Lemma 1)} \\ &= \tilde{\lambda}_0 + \mathbf{x}\tilde{\boldsymbol{\lambda}}_1 + \mathbf{z}^S(\bar{\mathbf{x}})\tilde{\boldsymbol{\lambda}}_2 + \mathbf{x}\tilde{\boldsymbol{\lambda}}_3\mathbf{z}^S(\bar{\mathbf{x}})' && (172) \end{aligned}$$

and $\boldsymbol{\lambda} = \tilde{\boldsymbol{\lambda}}$. Therefore:

$$CGE_{ks}^S = \lambda_{2s} + \lambda_{3ks} \qquad \text{(since } \boldsymbol{\lambda} = \tilde{\boldsymbol{\lambda}})$$

which is result (59′). Results (61′), (62′), (63′), and (64′) then follow by substitution of result (59′) into (33′), (34′), (37′), and (38′) in Proposition 9.

Proof for Proposition 13

Let the approximation error in (115) be:

$$v_m(\mathbf{x}^o, \mathbf{x}^p) \equiv h(\mathbf{x}^o, \mathbf{x}^p) - h_m(\mathbf{x}^o, \mathbf{x}^p)\bar{\boldsymbol{\phi}} \qquad (173)$$

Assumptions 1–6 and peer separability are given, so part (1) of Proposition 3 applies. Therefore:

$$y_i = y_i(\mathbf{p}_i) = \sum_{j\in\mathbf{p}_i} PE_{ij} \qquad \text{(by (44) in Proposition 3)}$$

Taking expectations:

$$
\begin{aligned}
E\left(y_i \middle| \mathbf{x}_i, \{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) &= E\left(\sum_{j\in\mathbf{p}_i} PE_{ij} \middle| \mathbf{x}_i, \{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) \\
&= \sum_{j\in\mathbf{p}_i} E\left(PE_{ij}|\mathbf{x}_i, \mathbf{x}_j\right) && (\text{RA} \implies \tau_i, \tau_j \perp\!\!\!\perp \tau_{j'}) \\
&= \sum_{j\in\mathbf{p}_i} h(\mathbf{x}_i, \mathbf{x}_j) && (\text{by (115)}) \\
&= \sum_{j\in\mathbf{p}_i} h_m(\mathbf{x}_i, \mathbf{x}_j)\bar{\boldsymbol{\phi}} + \sum_{j\in\mathbf{p}_i} v_m(\mathbf{x}_i, \mathbf{x}_j) && (\text{by (173)}) \\
&= (n_0 - 1)\bar{\mathbf{h}}_i\bar{\boldsymbol{\phi}} + \sum_{j\in\mathbf{p}_i} v_m(\mathbf{x}_i, \mathbf{x}_j) && (174)
\end{aligned}
$$

By construction:

$$
\begin{aligned}
L(v_m(\mathbf{x}_i, \mathbf{x}_j)|h_m(\mathbf{x}_i, \mathbf{x}_j)) &= L(h(\mathbf{x}_i, \mathbf{x}_j) - h_m(\mathbf{x}_i, \mathbf{x}_j)\bar{\boldsymbol{\phi}}|h_m(\mathbf{x}_i, \mathbf{x}_j)) \\
&= h_m(\mathbf{x}_i, \mathbf{x}_j)\bar{\boldsymbol{\phi}} - h_m(\mathbf{x}_i, \mathbf{x}_j)\bar{\boldsymbol{\phi}} \\
&= 0 && (175)
\end{aligned}
$$

Applying the law of iterated projections:

$$
\begin{aligned}
L\left(y_i|\bar{\mathbf{h}}_i\right) &= L\left(E\left(y_i \middle| \mathbf{x}_i, \{\mathbf{x}_j\}_{j\in\mathbf{p}_i}\right) \middle| \bar{\mathbf{h}}_i\right) \\
&= L\left((n_0 - 1)\bar{\mathbf{h}}_i\bar{\boldsymbol{\phi}} + \sum_{j\in\mathbf{p}_i} v_m(\mathbf{x}_i, \mathbf{x}_j) \middle| \bar{\mathbf{h}}_i\right) && (\text{by (117)}) \\
&= (n_0 - 1)\bar{\mathbf{h}}_i\bar{\boldsymbol{\phi}}
\end{aligned}
$$

which implies that:

$$
\boldsymbol{\phi} = (n_0 - 1)\bar{\boldsymbol{\phi}} \tag{176}
$$

By result (1) in Proposition 3:

$$
\begin{aligned}
CPE(\mathbf{x}^o, \mathbf{x}^p) &= E\left(PE_{ij} | \mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{x}^p\right) && \text{(by (45) in Proposition 3)} \\
&\quad - E\left(PE_{ij} | \mathbf{x}_i = \mathbf{x}^o, \mathbf{x}_j = \mathbf{e}_0\right) \\
&= h(\mathbf{x}^o, \mathbf{x}^p) - h(\mathbf{x}^o, \mathbf{e}_0) && \text{(by (115))} \\
&\approx h_m(\mathbf{x}^o, \mathbf{x}^p)\bar{\boldsymbol{\phi}} - h_m(\mathbf{x}^o, \mathbf{e}_0)\bar{\boldsymbol{\phi}} + v_m(\mathbf{x}^o, \mathbf{x}^p) - v_m(\mathbf{x}^o, \mathbf{e}_0) && \text{(by (173))} \\
&\approx \left(\frac{h_m(\mathbf{x}^o, \mathbf{x}^p) - h_m(\mathbf{x}^o, \mathbf{e}_0)}{n_0 - 1}\right)\boldsymbol{\phi} && \text{(by (176))}
\end{aligned}
$$

which is result (118) in the proposition, with approximation error the same order of magnitude as the approximation error in (115)

Proof for Proposition 14

Let the approximation error in (119) and (120) be:

$$v(\mathbf{x}^p) \equiv a(\mathbf{x}^p) - a_m(\mathbf{x}^p)\boldsymbol{\pi} \tag{177}$$

$$v(\mathbf{x}^o, \mathbf{x}^p) \equiv h(\mathbf{x}^o, \mathbf{x}^p) - h_m(\mathbf{x}^o, \mathbf{x}^p)\boldsymbol{\phi} \tag{178}$$

Let $\tilde{\mathbf{p}}$ be a purely random draw of $n_0 - 1$ peers from $\mathcal{I} \setminus \{i\}$. By (RA), $\mathbf{p}_i$ is a random draw from the same distribution, so $\left(y_i, \mathbf{x}_i, \{\mathbf{x}_j\}_{j \in \mathbf{p}_i}\right)$ has the same joint distribution as $\left(y_i(\tilde{\mathbf{p}}), \mathbf{x}_i, \{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}}\right)$. The result follows by substitution:

$$
\begin{aligned}
AGE(\mathbf{x}^p) &= E\left(y_i(\tilde{\mathbf{p}}) \middle| \{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}} = \mathbf{x}^p\right) && \text{(definition)}\\
&\quad - E\left(y_i(\tilde{\mathbf{p}}) \middle| \{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}} = \{\mathbf{e}_0, \ldots, \mathbf{e}_0\}\right)\\
&= E\left(y_i \middle| \{\mathbf{x}_j\}_{j\in\mathbf{p}_i} = \mathbf{x}^p\right) && \text{(same joint distribution)}\\
&\quad - E\left(y_i \middle| \{\mathbf{x}_j\}_{j\in\mathbf{p}_i} = \{\mathbf{e}_0, \ldots, \mathbf{e}_0\}\right)\\
&= a\left(\mathbf{x}^p\right) - a\left(\{\mathbf{e}_0, \ldots, \mathbf{e}_0\}\right) && \text{(by (119))}\\
&= a_m\left(\mathbf{x}^p\right)\boldsymbol{\pi} - a_m\left(\{\mathbf{e}_0, \ldots, \mathbf{e}_0\}\right)\boldsymbol{\pi} + v(\mathbf{x}^p) - v\left(\{\mathbf{e}_0, \ldots, \mathbf{e}_0\}\right) && \text{(by (177))}\\
&\approx \left(a_m\left(\mathbf{x}^p\right) - a_m\left(\{\mathbf{e}_0, \ldots, \mathbf{e}_0\}\right)\right)\boldsymbol{\pi} && \text{(by (119))}\\
CGE(\mathbf{x}^o, \mathbf{x}^p) &= E\left(y_i(\tilde{\mathbf{p}}) \middle| \mathbf{x}_i = \mathbf{x}^o, \{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}} = \mathbf{x}^p\right) && \text{(definition)}\\
&\quad - E\left(y_i(\tilde{\mathbf{p}}) \middle| \mathbf{x}_i = \mathbf{x}^o, \{\mathbf{x}_j\}_{j\in\tilde{\mathbf{p}}} = \{\mathbf{e}_0, \ldots, \mathbf{e}_0\}\right)\\
&= E\left(y_i \middle| \mathbf{x}_i = \mathbf{x}^o, \{\mathbf{x}_j\}_{j\in\mathbf{p}_i} = \mathbf{x}^p\right) && \text{(same joint distribution)}\\
&\quad - E\left(y_i \middle| \mathbf{x}_i = \mathbf{x}^o, \{\mathbf{x}_j\}_{j\in\mathbf{p}_i} = \{\mathbf{e}_0, \ldots, \mathbf{e}_0\}\right)\\
&= h\left(\mathbf{x}^o, \mathbf{x}^p\right) - h\left(\mathbf{x}^o, \{\mathbf{e}_0, \ldots, \mathbf{e}_0\}\right) && \text{(by (120))}\\
&= h_m\left(\mathbf{x}^o, \mathbf{x}^p\right)\phi - h_m\left(\mathbf{x}^o, \{\mathbf{e}_0, \ldots, \mathbf{e}_0\}\right)\boldsymbol{\phi} + v(\mathbf{x}^o, \mathbf{x}^p) - v\left(\mathbf{x}^o, \{\mathbf{e}_0, \ldots, \mathbf{e}_0\}\right)\\
& && \text{(by (178))}\\
&\approx \left(h_m\left(\mathbf{x}^o, \mathbf{x}^p\right) - h_m\left(\mathbf{x}^o, \{\mathbf{e}_0, \ldots, \mathbf{e}_0\}\right)\right)\boldsymbol{\phi} && \text{(by (120))}
\end{aligned}
$$

which are results (121) and (122).