ONLINE APPENDIX

Quantifying Family, School, and Location Effects in the Presence of Complementarities and Sorting

Mohit Agrawal
Yale University

Joseph G. Altonji
Yale University and NBER

Richard K. Mansfield
University of Colorado-Boulder and NBER

Online Appendices

B1 Proof of Proposition 2:

In deviation from mean form, the model is

$$\begin{aligned} DY_i &= \boldsymbol{DX}_i\boldsymbol{\beta} + DM_i\boldsymbol{X}_g\boldsymbol{\rho}_1 + DM_i\boldsymbol{Z}_{2g}\boldsymbol{\rho}_2 \\ &\quad + DM_i\boldsymbol{X}_g^U\boldsymbol{\rho}_1^U + DM_i\boldsymbol{Z}_{2g}^U\boldsymbol{\rho}_2^U + Dx_i^U + D\eta_{gi} + D\xi_{gi}\ . \end{aligned}$$

Using (12), (13), (14), and (15), we can rewrite the above equation as

$$\begin{aligned} DY_i &= \boldsymbol{DX}_i[\boldsymbol{\beta} + \boldsymbol{\Pi}_{\boldsymbol{DX}_i\boldsymbol{DX}_i^U}\boldsymbol{\beta}^U + \Pi_{D\eta_g D\boldsymbol{X}_i}] + DM_i\boldsymbol{X}_g[\boldsymbol{\rho}_1 + \boldsymbol{\Pi}_{\boldsymbol{X}_g^U\boldsymbol{X}_g}\boldsymbol{\rho}_1^U + \boldsymbol{\Pi}_{\boldsymbol{Z}_{2g}^U\boldsymbol{X}_g}\boldsymbol{\rho}_2^U] \qquad (31) \\ &\quad + DM_i\boldsymbol{Z}_{2g}[\boldsymbol{\rho}_2 + \boldsymbol{\Pi}_{\boldsymbol{Z}_{2g}^U\boldsymbol{Z}_{2g}}\boldsymbol{\rho}_2^U] \\ &\quad + DM_i\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\rho}_2^U + D\tilde{x}_i^U + D\tilde{\eta}_{gi} + D\xi_{gi}\ . \end{aligned}$$

From basic regression theory, the probability limit of the OLS estimator of the coefficients on the regressors $[\boldsymbol{DX}_i, DM_i\boldsymbol{X}_g, DM_i\boldsymbol{Z}_{2g}]$ in (31) equals the actual coefficients if the regressors are all uncorrelated with the composite error term $DM_i\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\rho}_2^U + D\tilde{x}_i^U + D\tilde{\eta}_{gi} + D\xi_{gi}$. We now show that this is the case, considering the error components one at a time. $D\xi_{gi}$ is uncorrelated with all variables in the model by definition of a shock. $D\tilde{\eta}_{gi}$ is also uncorrelated with $DM_i\boldsymbol{X}_g$ and $DM_i\boldsymbol{Z}_{2g}$ by definition of $\boldsymbol{\rho}_1, \boldsymbol{\rho}_1^U, \boldsymbol{\rho}_2$ and $\boldsymbol{\rho}_2^U$ (see Section 3).

Next we consider $D\tilde{x}_i^U$. $D\tilde{x}_i^U$ is uncorrelated with $\boldsymbol{DX}_i$ by definition of $D\tilde{x}_i^U$. $Cov(D\tilde{x}_i^U, DM_i\boldsymbol{Z}_{2g}) = Cov(D\tilde{x}_i^U DM_i, \boldsymbol{Z}_{2g}) = 0$ by A6. Similarly, $Cov(D\tilde{x}_i^U, DM_i\boldsymbol{X}_g) = Cov(D\tilde{x}_i^U DM_i, \boldsymbol{X}_g) = 0$ by A6.

This leaves $DM_i\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\rho}_2^U$. $Cov(\boldsymbol{DX}_i, DM_i\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\rho}_2^U) = \boldsymbol{E}(\boldsymbol{DX}_i DM_i\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\rho}_2^U)$ because $\boldsymbol{E}(\boldsymbol{DX}_i) = \boldsymbol{0}$. A7 and the fact that DM_i is a function of $\boldsymbol{DX}_i$ imply that $\boldsymbol{DX}_i DM_i$ is independent of $\tilde{\boldsymbol{Z}}_{2g}^U$ Thus $Cov(\boldsymbol{DX}_i, DM_i\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\rho}_2^U) = \boldsymbol{E}(\boldsymbol{DX}_i DM_i)\boldsymbol{E}(\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\rho}_2^U)$. The last term is zero because the mean of the residual $\tilde{\boldsymbol{Z}}_{2g}^U = 0$. Similar arguments using A7 establish that the covariance between $DM_i\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\rho}_2^U$ and $[DM_i\boldsymbol{X}_g, DM_i\boldsymbol{Z}_{2g}]$ are 0. This completes the proof.

B2 Proof of Proposition 3

Note first that because $\boldsymbol{X}_g$, $\boldsymbol{Z}_{2g}$, $M_g\boldsymbol{X}_g$, and $M_g\boldsymbol{Z}_{2g}$ do not vary within groups, $\boldsymbol{G}_1$, $\boldsymbol{G}_2$, $\boldsymbol{G}_3$, and $\boldsymbol{G}_4$ are identified exclusively from between-group variation. Thus, the OLS coefficients $\boldsymbol{G}_1$, $\boldsymbol{G}_2$, $\boldsymbol{G}_3$, and $\boldsymbol{G}_4$ are numerically identical to the coefficients of the projection of the adjusted group g mean of Y_{gi}, $Y_g - [\boldsymbol{X}_g\boldsymbol{B} + M_g\boldsymbol{X}_g\boldsymbol{r}_1 + M_g\boldsymbol{Z}_{2g}\boldsymbol{r}_2]$, onto $\boldsymbol{X}_g$, $\boldsymbol{Z}_{2g}$, $M_g\boldsymbol{X}_g$, and $M_g\boldsymbol{Z}_{2g}$.

Using (9), we obtain

$$
\begin{aligned}
Y_g - [\boldsymbol{X}_g\boldsymbol{B} + M_g\boldsymbol{X}_g\boldsymbol{r}_1 + M_g\boldsymbol{Z}_{2g}\boldsymbol{r}_2] &= \boldsymbol{X}_g[\boldsymbol{\beta} - \boldsymbol{B} + \boldsymbol{\Gamma}_1] + \boldsymbol{Z}_{2g}\boldsymbol{\Gamma}_2 + M_g\boldsymbol{X}_g[\boldsymbol{\rho}_1 - \boldsymbol{r}_1] + M_g\boldsymbol{Z}_{2g}[\boldsymbol{\rho}_2 - \boldsymbol{r}_2] \\
&\quad + M_g\boldsymbol{X}_g^U\boldsymbol{\rho}_1^U + M_g\boldsymbol{Z}_{2g}^U\boldsymbol{\rho}_2^U + x_g^U + z_g^U + \xi_g \qquad (32)
\end{aligned}
$$

Recall that under assumptions A1-A5, $\boldsymbol{X}_g^U = [\boldsymbol{\Pi}_{\boldsymbol{X}^U\boldsymbol{X}} + \boldsymbol{Var}(\boldsymbol{X}_i)^{-1}\boldsymbol{R}'\boldsymbol{Var}(\tilde{\boldsymbol{X}}_i^U)] \equiv \boldsymbol{X}_g\boldsymbol{\Pi}_{\boldsymbol{X}_g^U\boldsymbol{X}_g}$, so $x_g^U = \boldsymbol{X}_g\boldsymbol{\Pi}_{\boldsymbol{X}_g^U\boldsymbol{X}_g}\boldsymbol{\beta}^U = \boldsymbol{X}_g\boldsymbol{\Pi}_{x_g^U\boldsymbol{X}_g}$ where $\boldsymbol{\Pi}_{x_g^U\boldsymbol{X}_g} \equiv \boldsymbol{\Pi}_{\boldsymbol{X}_g^U\boldsymbol{X}_g}\boldsymbol{\beta}^U$. Recall also that $z_g^U \equiv \boldsymbol{X}_g^U\boldsymbol{\Gamma}_1^U + \boldsymbol{Z}_{2g}^U\boldsymbol{\Gamma}_2^U$ Using these facts and also using (15) to substitute for $\boldsymbol{Z}_{2g}^U$ in the term $M_g\boldsymbol{Z}_{2g}^U\boldsymbol{\rho}_2^U$, one may rewrite (32) as

$$
\begin{aligned}
Y_g - [\boldsymbol{X}_g\boldsymbol{B} + M_g\boldsymbol{X}_g\boldsymbol{r}_1 + M_g\boldsymbol{Z}_{2g}\boldsymbol{r}_2] &= \boldsymbol{X}_g[\boldsymbol{\beta} - \boldsymbol{B} + \boldsymbol{\Gamma}_1 + \boldsymbol{\Pi}_{x_g^U\boldsymbol{X}_g} + \boldsymbol{\Pi}_{X_g^UX_g}\boldsymbol{\Gamma}_1^U + \boldsymbol{\Pi}_{Z_{2g}^UX_g}\boldsymbol{\Gamma}_2^U] + \boldsymbol{Z}_{2g}[\boldsymbol{\Gamma}_2 + \boldsymbol{\Pi}_{Z_{2g}^UZ_{2g}}\boldsymbol{\Gamma}_2^U] \\
&\quad + M_g\boldsymbol{X}_g[\boldsymbol{\rho}_1 - \boldsymbol{r}_1 + \boldsymbol{\Pi}_{X_g^UX_g}\boldsymbol{\rho}_1^U + \boldsymbol{\Pi}_{\boldsymbol{Z}_{2g}^U\boldsymbol{X}_g}\boldsymbol{\rho}_2^U] \\
&\quad + M_g\boldsymbol{Z}_{2g}[\boldsymbol{\rho}_2 - \boldsymbol{r}_2 + \boldsymbol{\Pi}_{\boldsymbol{Z}_{2g}^U\boldsymbol{Z}_{2g}}\boldsymbol{\rho}_2^U] + M_g\widetilde{\boldsymbol{Z}_{2g}^U}\boldsymbol{\rho}_2^U + \tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\Gamma}_2^U + \xi_g. \qquad (33)
\end{aligned}
$$

The post high school shocks ξ_g are uncorrelated with all variables in the model by definition of a shock. Consider next the projection of the error component $M_g\widetilde{\boldsymbol{Z}_{2g}^U}\boldsymbol{\rho}_2^U$ onto $\boldsymbol{X}_g$, $\boldsymbol{Z}_{2g}$, $M_g\boldsymbol{X}_g$, and $M_g\boldsymbol{Z}_{2g}$. Recall that A7 states that $\widetilde{\boldsymbol{Z}_{2g}^U}$ is independent of $\boldsymbol{X}_g$ and $\boldsymbol{Z}_{2g}$ and not simply uncorrelated with them. Also recall that M_g is a linear function of $\boldsymbol{X}_g$. Consequently, $\boldsymbol{X}_g$, $\boldsymbol{Z}_{2g}$, $\boldsymbol{M}_g\boldsymbol{X}_g$, and $\boldsymbol{M}_g\boldsymbol{Z}_{2g}$ are all uncorrelated with $M_g\widetilde{\boldsymbol{Z}_{2g}^U}\boldsymbol{\rho}_2^U$. To elaborate slightly, $E(M_g\widetilde{\boldsymbol{Z}_{2g}^U}\boldsymbol{\rho}_2^U|\boldsymbol{X}_g,\boldsymbol{Z}_{2g},M_g\boldsymbol{X}_g,M_g\boldsymbol{Z}_{2g}) = 0$ by A7 , the fact that M_g is function of $\boldsymbol{X}_g$ and is thus also independent of $\widetilde{\boldsymbol{Z}_{2g}^U}$, and the fact that the expectation of the product of two independent random variables is the product of the expectations.

Next note that $\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\Gamma}_2^U$ is independent of $\boldsymbol{X}_g$, $\boldsymbol{Z}_{2g}$ by A7 and is independent of $M_g\boldsymbol{X}_g$ and $M_g\boldsymbol{Z}_{2g}$ by A7 and the fact that M_g is a linear function of $\boldsymbol{X}_g$. Consequently, all of the regressors are also uncorrelated with $\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\Gamma}_2^U$ .Collecting terms from equation (33) and using $\boldsymbol{\Pi}_{z_{2g}^UX_g} = \boldsymbol{\Pi}_{Z_{2g}^UX_g}\boldsymbol{\Gamma}_2^U$, we conclude that

$$
\begin{aligned}
\boldsymbol{G}_1 &= [(\boldsymbol{\beta} - \boldsymbol{B}) + \boldsymbol{\Pi}_{x_g^UX_g}] + [\boldsymbol{\Gamma}_1 + \boldsymbol{\Pi}_{X_g^UX_g}\boldsymbol{\Gamma}_1^U + \boldsymbol{\Pi}_{z_{2g}^UX_g}] \qquad (34) \\
\boldsymbol{G}_2 &= \boldsymbol{\Gamma}_2 + \boldsymbol{\Pi}_{z_g^UZ_{2g}} \qquad (35) \\
\boldsymbol{G}_3 &= [\boldsymbol{\rho}_1 - \boldsymbol{r}_1 + \boldsymbol{\Pi}_{X_g^UX_g}\boldsymbol{\rho}_1^U + \boldsymbol{\Pi}_{\boldsymbol{Z}_{2g}^U\boldsymbol{X}_g}\boldsymbol{\rho}_2^U] \qquad (36) \\
\boldsymbol{G}_4 &= [\boldsymbol{\rho}_2 - \boldsymbol{r}_2] + \boldsymbol{\Pi}_{Z_{2g}^UZ_{2g}}\boldsymbol{\rho}_2^U. \qquad (37)
\end{aligned}
$$

But recall the results of Proposition 2:

$$
\boldsymbol{r}_1 = \boldsymbol{\rho}_1 + \boldsymbol{\Pi}_{X_g^UX_g}\boldsymbol{\rho}_1^U + \boldsymbol{\Pi}_{Z_{2g}^UX_g}\boldsymbol{\rho}_2^U
$$

$$
\boldsymbol{r}_2 = \boldsymbol{\rho}_2 + \boldsymbol{\Pi}_{Z_{2g}^UZ_{2g}}\boldsymbol{\rho}_2^U
$$

implies that both $\boldsymbol{G}_3$ and $\boldsymbol{G}_4$ are zero.

Combining these insights we obtain:

$$\begin{aligned}
\boldsymbol{G}_1 &= [(\boldsymbol{\beta}-\boldsymbol{B})+\boldsymbol{\Pi}_{x_g^U X_g}]+[\boldsymbol{\Gamma}_1+\boldsymbol{\Pi}_{X_g^U X_g}\boldsymbol{\Gamma}_1^U+\boldsymbol{\Pi}_{z_{2g}^U X_g}] \\
\boldsymbol{G}_2 &= \boldsymbol{\Gamma}_2+\boldsymbol{\Pi}_{z_{2g}^U Z_{2g}} \\
\boldsymbol{G}_3 &= 0 \\
\boldsymbol{G}_4 &= 0
\end{aligned}$$

This completes the proof.

B3 Additional Details of the Estimation Procedure

B3.1 Notes on Step 1 and 2

In the first step of the estimation procedure, we impose the restrictions that the interactions operate through the same regressor indices as the main group effects as follows. First we choose initial values $\boldsymbol{B}^0$, $\boldsymbol{G}_1^{N,0}$, $\boldsymbol{G}_1^{S,0}$, $\boldsymbol{G}_2^{S,0}$, and $\boldsymbol{G}_2^{C,0}$. Then, letting k denote the iteration number, we implement an iterative estimation procedure in which (temporary) main effect parameters $\boldsymbol{B}^k$, $\boldsymbol{G}_1^{N,k}$, $\boldsymbol{G}_1^{S,k}$, $\boldsymbol{G}_2^{S,k}$, and $\boldsymbol{G}_2^{C,k}$ and the interaction coefficients $r_1^{N,k}$, $r_1^{S,k}$, $\boldsymbol{r}_2^{S,k}$ and $\boldsymbol{r}_2^{C,k}$ are estimated while holding fixed the regressor indices entering the interaction terms at their values from the previous iteration ($\boldsymbol{X}_i\boldsymbol{B}^{k-1}$, $\boldsymbol{X}_n\boldsymbol{G}_1^{N,k-1}$, $\boldsymbol{X}_s\boldsymbol{G}_1^{S,k-1}$, $\boldsymbol{Z}_{2s}^S\boldsymbol{G}_2^{S,k-1}$, and $\boldsymbol{Z}_{2c}^C\boldsymbol{G}_2^{C,k-1}$). The routine ends when successive iterations produce sufficiently similar parameter estimates.

In the second step, we first reparameterize (23) model so that the indices $\boldsymbol{X}_i\hat{\boldsymbol{B}}$, $\boldsymbol{X}_n\hat{\boldsymbol{G}}_1^N$, $\boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S$, $\boldsymbol{Z}_{2s}^S\hat{\boldsymbol{G}}_2^S$, and $\boldsymbol{Z}_{2c}^C\hat{\boldsymbol{G}}_2^C$ are expressed in standard deviation units. Specifically, we estimate

$$\begin{aligned}
Y_i &= \alpha_0+\alpha_1\frac{\boldsymbol{X}_i\hat{\boldsymbol{B}}}{sd(\boldsymbol{X}_i\hat{\boldsymbol{B}})}+\alpha_2\frac{\boldsymbol{X}_n\hat{\boldsymbol{G}}_1^N}{sd(\boldsymbol{X}_n\hat{\boldsymbol{G}}_1^N)}+\alpha_3\frac{\boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S}{sd(\boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S)}+\alpha_4\frac{\boldsymbol{Z}_{2s}^S\hat{\boldsymbol{G}}_2^S}{sd(\boldsymbol{Z}_{2s}^S\hat{\boldsymbol{G}}_2^S)}+\alpha_5\frac{\boldsymbol{Z}_{2c}^C\hat{\boldsymbol{G}}_2^C}{sd(\boldsymbol{Z}_{2c}^C\hat{\boldsymbol{G}}_2^C)}+ \\
&+r_1^N\frac{\boldsymbol{X}_i\hat{\boldsymbol{B}}}{sd(\boldsymbol{X}_i\hat{\boldsymbol{B}})}\frac{\boldsymbol{X}_n\hat{\boldsymbol{G}}_1^N}{sd(\boldsymbol{X}_n\hat{\boldsymbol{G}}_1^N)}+r_1^S\frac{\boldsymbol{X}_i\hat{\boldsymbol{B}}}{sd(\boldsymbol{X}_i\hat{\boldsymbol{B}})}\frac{\boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S}{sd(\boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S)}+[\hat{\boldsymbol{M}}_i\otimes\frac{\boldsymbol{Z}_{2s}^S\hat{\boldsymbol{G}}_2^S}{sd(\boldsymbol{Z}_{2s}^S\hat{\boldsymbol{G}}_2^S)}]\boldsymbol{r}_2^S+[\hat{\boldsymbol{M}}_i\otimes\frac{\boldsymbol{Z}_{2c}^C\hat{\boldsymbol{G}}_2^C}{sd(\boldsymbol{Z}_{2c}^C\hat{\boldsymbol{G}}_2^C)}]\boldsymbol{r}_2^C \\
&+v_c+(v_s-v_c)+(v_n-v_s)+(v_i-v_n), \qquad (38)
\end{aligned}$$

where $sd(\boldsymbol{X}_i\hat{\boldsymbol{B}})$, $sd(\boldsymbol{X}_n\hat{\boldsymbol{G}}_1^N)$, $sd(\boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S)$, $sd(\boldsymbol{Z}_{2s}^S\hat{\boldsymbol{G}}_2^S)$, and $sd(\boldsymbol{Z}_{2c}^C\hat{\boldsymbol{G}}_2^C)$ are the student-weighted standard deviations of the regression indices evaluated using the slope coefficients from the first step. Note that we are abusing notation by continuing to use r_1^N, r_1^S, $\boldsymbol{r}_2^S$ and $\boldsymbol{r}_2^C$ as the interaction coefficients even though the regressors are in now standard deviation units. To understand the α parameters note that in the case of the wage model, aside from the effects of allowing for a multilevel random effects error structure, the second step estimate $\hat{\alpha}_1$ should equal $sd(\boldsymbol{X}_i\hat{\boldsymbol{B}})$, $\hat{\alpha}_2$ should equal $sd(\boldsymbol{X}_n\hat{\boldsymbol{G}}_1^N)$, $\hat{\alpha}_3$ should equal $sd(\boldsymbol{Z}_{2s}^S\hat{\boldsymbol{G}}_2^S)$, so on. In the case of the probit model, $\hat{\alpha}_1.\hat{\alpha}_2,\hat{\alpha}_3,\hat{\alpha}_4$ and $\hat{\alpha}_5$ should equal the first stage estimates $sd(\boldsymbol{X}_i\hat{\boldsymbol{B}})$, $sd(\boldsymbol{X}_n\hat{\boldsymbol{G}}_1^N)$, $sd(\boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S)$, $sd(\boldsymbol{Z}_{2s}^S\hat{\boldsymbol{G}}_2^S)$, and $sd(\boldsymbol{Z}_{2c}^C\hat{\boldsymbol{G}}_2^C)$ times the scale factor $[Var(v_c) + Var(v_s - v_c) + Var(v_n - v_s) + 1]^{0.5}$. The reason is that the first step probit normalizes $[v_c + (v_s - v_c) + (v_n - v_s) + (v_i - v_n)]^{0.5}$ to 1 while the second step normalizes $[var(v_i - v_n)]^{0.5}$ to 1. We freely estimate the α coefficients, which is what we mean when we say in the main text that in the second step we implicitly allow the elements of $\hat{\boldsymbol{B}}$ to update by a common factor of proportionately, and we do the same for $\hat{\boldsymbol{G}}_1^N$, $\hat{\boldsymbol{G}}_1^S$, $\hat{\boldsymbol{G}}_2^S$, $\hat{\boldsymbol{G}}_2^C$. In practice, first and second step estimates are very close to the implied values. We also report estimates of r_1^N, r_1^S, $\boldsymbol{r}_2^S$ and $\boldsymbol{r}_2^C$ from the second step.

B3.2 Bias Corrections for Error Component Variances

Step 3 of Section 6.2 describes how we implement the finite sample bias correction to remove sampling variance from our estimates of the variances and covariances of our observed regression indices. Here we discuss finite sample bias corrections for the error component variances. Consider the bias correction term $\frac{1}{N}\sum_i \boldsymbol{X}_{s(i)}\boldsymbol{Var}(\hat{\boldsymbol{G}}_1^S - \boldsymbol{G}_1^S)\boldsymbol{X}'_{s(i)}$ that is subtracted from $Var(\boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S)$ to estimate $Var(\boldsymbol{X}_s\boldsymbol{G}_1)$. Assuming that the outcome is measured without error, the expected sampling variance captured by this correction term reflects true inputs into Y_i that should have been allocated to the unobserved error components $v_i - v_n$, $v_n - v_s$, or $v_s - v_c$.

To determine the share of the bias correction to allocate to each error component, we ignore the heterogeneity in the number of sampled students per neighborhood, the number of sampled neighborhoods per school, and the number of sampled schools per commuting zone, and treat these as fixed scalar values $\frac{I}{N}$, $\frac{N}{S}$, and $\frac{S}{C}$, respectively (where I, N, S, and C are the number of sampled individuals, neighborhoods, schools, and commuting zones). We also treat the population number of students per neighborhood, neighborhoods per school, and schools per commuting zone as large, so that such sampling variance would disappear if we observed the full population of high school students in the United States. Then the variance in the sampling error among school averages Y_s within the same commuting zone (for schools each featuring $\frac{I}{S}$ sample members) is given by:

$$
\begin{aligned}
& Var(\frac{1}{I/S}\sum_{i\in s}[(v_i - v_{n(i)}) + (v_{n(i)} - v_{s(i)}) + v_s] \\
&= Var(\frac{1}{I/S}\sum_{i\in s}(v_i - v_n)) + Var(\frac{1}{N/S}\sum_{n'=1}^{\frac{N}{S}}(v_{n'} - v_s)) + Var(v_s) \\
&= \frac{1}{(I/S)^2}Var(\sum_{i\in s}(v_i - v_n)) + \frac{1}{(N/S)^2}Var(\sum_{n'=1}^{\frac{N}{S}}(v_{n'} - v_s)) + Var(v_s) \\
&= \frac{1}{(I/S)^2}(I/S)Var(v_i - v_n) + \frac{1}{(N/S)^2}(N/S)Var(v_n - v_s) + Var(v_s) \\
&= \frac{Var(v_i - v_n)}{I/S} + \frac{Var(v_n - v_s)}{N/S} + Var(v_s), \qquad (39)
\end{aligned}
$$

where we have assumed independence in the draws of $v_i - v_{n(i)}$, $v_n - v_s$, and v_s across individuals, neighborhoods and schools.

Thus, the individual, neighborhood, and school shares of the variance in the sampling error among school averages Y_s is given by:

$$Share_S^I = \frac{\frac{Var(v_i - v_n)}{I/S}}{\frac{Var(v_i - v_n)}{I/S} + \frac{Var(v_n - v_s)}{N/S} + Var(v_s)} \quad (40)$$

$$Share_S^N = \frac{\frac{Var(v_n - v_s)}{N/S}}{\frac{Var(v_i - v_n)}{I/S} + \frac{Var(v_n - v_s)}{N/S} + Var(v_s)} \quad (41)$$

$$Share_S^S = \frac{\frac{Var(v_s)}{N/S}}{\frac{Var(v_i - v_n)}{I/S} + \frac{Var(v_n - v_s)}{N/S} + Var(v_s)} \quad (42)$$

We assume that the sampling variance component of the estimated variance of each school-level regression index (or the estimated covariance among each pair of school-level regression indices) contains individual, neighborhood, and school subcomponents in the same proportions as the overall variance in sampling error among school averages Y_s. Thus, we allocate the estimated sampling variance $\frac{1}{N}\sum_i \boldsymbol{X}_{s(i)}\boldsymbol{Var}(\hat{\boldsymbol{G}}_1^S - \boldsymbol{G}_1^S)\boldsymbol{X}'_{s(i)}$ associated with $Var(\boldsymbol{X}_s\hat{\boldsymbol{G}}_1)$, for example, to the individual-level, neighborhood-level, and school-level error variances $Var(v_i - v_n)$, $Var(v_n - v_s)$ and $Var(v_s - v_c)$ according to the shares given in (40) - (42). We use analogous formulae to derive the individual and neighborhood shares used to allocate neighborhood-level sampling variance terms and to derive the individual, neighborhood, school, and commuting zone shares used to allocate commuting zone-level sampling variance terms.

B3.3 Details of Estimation of the Effect of Shifts in School and in Commuting Zone Quality

B3.3.1 The School Treatment and the Commuting Zone Treatment Estimators

The estimator of the expected outcome for a randomly chosen student who is assigned a school at the q-th percentile of quality is:

$$\begin{aligned} E[\hat{Y}^q] = \frac{1}{P}\sum_p \frac{1}{I}\sum_i \Phi(&\boldsymbol{X}_i\hat{\boldsymbol{B}} + \boldsymbol{X}_n\hat{\boldsymbol{G}}_1^N + \boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S + T^q + (\boldsymbol{Z}_{2c}^C\boldsymbol{G}_2^C)_p + (v_c)_p \\ &+ (\boldsymbol{X}_i\hat{\boldsymbol{B}})(\boldsymbol{X}_n^N\hat{\boldsymbol{G}}_1^N)\hat{r}_1^N + (\boldsymbol{X}_i\hat{\boldsymbol{B}})(\boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S)\hat{r}_1^S + \boldsymbol{M}_i \otimes (\boldsymbol{Z}_{2s}^S\boldsymbol{G}_2^S)_p\hat{\boldsymbol{r}}_2^S \\ &+ \boldsymbol{M}_i \otimes (\boldsymbol{Z}_{2c}^C\boldsymbol{G}_2^C)_p\hat{\boldsymbol{r}}_2^C)/(1 + Var(v_n - v_s)) \end{aligned} \quad (43)$$

where $(\boldsymbol{Z}_{2s}^S\boldsymbol{G}_2^S)_p$ represents the p-th draw from the conditional distribution $f(\boldsymbol{Z}_{2s}^S\boldsymbol{G}_2^S|T = T^q)$ and $(\boldsymbol{Z}_{2c}^C\boldsymbol{G}_2^C)_p$ and $(v_c)_p$ represent the p-th draws from the unconditional joint distribution $f(\boldsymbol{Z}_{2c}^C\boldsymbol{G}_2^C, v_c)$.

Our estimator of the expected outcome for a randomly chosen student who is assigned a commuting zone at the q-th percentile of quality is:

$$E[\hat{Y}^q] = \frac{1}{P}\sum_p \frac{1}{I}\sum_i \Phi(\boldsymbol{X}_i\hat{\boldsymbol{B}} + \boldsymbol{X}_{1n}\hat{\boldsymbol{G}}_1^N + \boldsymbol{X}_{1s}\hat{\boldsymbol{G}}_1^S + (\boldsymbol{Z}_{2s}^S\boldsymbol{G}_2^S)_p + (v_s - v_c)_p + T^q$$
$$+ (\boldsymbol{X}_i\hat{\boldsymbol{B}})(\boldsymbol{X}_n\hat{\boldsymbol{G}}_1^N)\hat{r}_1^N + (\boldsymbol{X}_i\hat{\boldsymbol{B}})(\boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S)\hat{r}_1^S + \boldsymbol{M}_i \otimes (\boldsymbol{Z}_{2s}^S\boldsymbol{G}_2^S)_p\hat{\boldsymbol{r}}_2^S$$
$$+ \boldsymbol{M}_i \otimes (\boldsymbol{Z}_{2c}^C\boldsymbol{G}_2^C)_p\hat{\boldsymbol{r}}_2^C)/(1 + Var(v_n - v_s)) \tag{44}$$

where $(\boldsymbol{Z}_{2s}^S\boldsymbol{G}_2^S)_p$ and $(v_s - v_c)_p$ are the p-th draws from the unconditional joint distribution of $(\boldsymbol{Z}_{2s}^S\boldsymbol{G}_2^S)$ and $(v_s - v_c)$ and $(\boldsymbol{Z}_{2c}^C\boldsymbol{G}_2^C)_p$ and $(v_c)_p$ are the p-th draws from the conditional joint distribution $f(\boldsymbol{Z}_{2c}^C\boldsymbol{G}_2^C, v_c | T \equiv \boldsymbol{Z}_{2c}^C\boldsymbol{G}_2^C + v_c = T^q)$.

B3.3.2 Estimating Impacts of Shifts in School and Commuting Zone Quality for Particular Subpopulations

Here we provide more details about estimation of treatment effects for particular subpopulations. The most straightforward approach is simply to restrict the sample used for the counterfactual treatments to members of a particular subpopulation. We use the empirical distribution of individual and neighborhood inputs $\boldsymbol{X}_i\hat{\boldsymbol{B}} + \boldsymbol{X}_n\hat{\boldsymbol{G}}_1^N + \boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S$, so restricting the sample naturally imposes the chosen sample's joint distribution of observed individual and neighborhood inputs. Furthermore, recall that the unobserved components $v_i - v_n$ and $v_n - v_s$ are defined to be uncorrelated with all of the observable characteristics used to define the subpopulation. Thus, the formulas (30), (43) and (44) are still valid, with i and I now indexing the particular individual and number of individuals among the chosen subpopulation. All elements of $\boldsymbol{M}_i$ take on the values for i, so that the results for Hispanic students, for example, reflect not only the interaction terms involving the minority (non-Hispanic black or Hispanic) indicator but also differences across groups in the distribution of the other elements of $\boldsymbol{M}_i$, such as low income status, weighted by the corresponding elements of the interaction coefficients $\hat{\boldsymbol{r}}_2^S$ and $\hat{\boldsymbol{r}}_2^C$.

We compute treatment effects by ventile of the $\boldsymbol{X}_i\boldsymbol{B}$ distribution as follows. We fix $\boldsymbol{X}_i\boldsymbol{B}$ at each ventile dividing point $[.05, \ldots, .95]$ in its empirical distribution in the sample, and compute the change in expected outcome for each of our three counterfactual quality shifts ("School and CZ", "School only", and "CZ only", described above) for randomly chosen individuals at the chosen ventile of $\boldsymbol{X}_i\boldsymbol{B}$. We integrate over the joint distribution of $v_i - v_n$, $v_n - v_s$, $\boldsymbol{X}_n\boldsymbol{G}_1^N$ and $\boldsymbol{X}_s\boldsymbol{G}_1^S$. This means that we are not holding fixed the kind of neighborhood such students tend to experience, but are instead randomly assigning a neighborhood from the full population distribution for both the low ($E[Y|T^{10}]$) and high ($E[Y|T^{90}]$) school/commuting zone treatments. Specifically, the expected outcome of a randomly chosen student at a particular $\boldsymbol{X}_i\boldsymbol{B}$ percentile q' (denoted $(\boldsymbol{X}_i\boldsymbol{B})^{q'}$ below) who is assigned a school-commuting zone combination at the q-th percentile in the "School and CZ" counterfactual is estimated via:

$$E[\hat{Y}^q] = \frac{1}{P}\sum_p \frac{1}{I}\sum_i \Phi((\boldsymbol{X}_i\boldsymbol{B})^{q'} + \boldsymbol{X}_n\hat{\boldsymbol{G}}_1^N + \boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S + T^q$$
$$+ (\boldsymbol{X}_i\boldsymbol{B})^{q'}(\boldsymbol{X}_n^N\hat{\boldsymbol{G}}_1^N)\hat{r}_1^N + (\boldsymbol{X}_i\boldsymbol{B})^{q'}(\boldsymbol{X}_s\hat{\boldsymbol{G}}_1^S)\hat{r}_1^S + \boldsymbol{M}_i^{q'} \otimes (\boldsymbol{Z}_{2s}^S\boldsymbol{G}_2^S)_p\hat{\boldsymbol{r}}_2^S$$
$$+ \boldsymbol{M}_i^{q'} \otimes (\boldsymbol{Z}_{2c}^C\boldsymbol{G}_2^C)_p\hat{\boldsymbol{r}}_2^C)/(1 + Var(v_n - v_s)), \quad (45)$$

where $\boldsymbol{M}_i^{q'} = [(\boldsymbol{X}_i\boldsymbol{B})^{q'}, 1(Female), 1(URM), 1(Low\_Income)]$.

Note are averaging over the empirical distribution of $1(Female), 1(URM), 1(Low\_Income)$, not the distribution conditionalon $\boldsymbol{X}_i\boldsymbol{B} = (\boldsymbol{X}_i\boldsymbol{B})^{q'}$

B4 Monte Carlo Simulations

This section describes the methodology and summarizes the results from a set of monte carlo simulations of our multilevel mixed effects estimator. The simulation results in AM already established that our control function $\boldsymbol{X}_g$ can absorb nearly all of the variation in $\boldsymbol{X}_g^U$ even when small samples of individuals in group g are used to construct $\boldsymbol{X}_g$, and even when the spanning condition A5 only approximately holds. Thus, the set of simulations described here are designed instead to highlight properties of the estimator that relate to the addition of interactions between individual and group inputs in the production function.

Specifically, the aim of these simulations is threefold. First, we wish to provide a particular (plausible) data generating process in which the key assumptions A6 and A7 that underlie Propositions 2-4 approximately hold, in addition to assumptions A1-A5. Second, we wish to verify that our estimates of the key coefficients used to construct our lower bound estimates of treatment effects from shifts in group membership ($\hat{\boldsymbol{r}}_2$ and $\hat{\boldsymbol{G}}_2$) closely match the formulas presented in equations (18) and (19) that are derived under assumptions A1-A7. Third, we wish to examine how the MME estimator of $\hat{\boldsymbol{G}}_2$ and particularly the interaction coefficients $\hat{\boldsymbol{r}}_2$ performs when small samples of individuals in group g are used to construct $\boldsymbol{X}_g$, since we rely on such small samples in our empirical work.

B4.1 Description of the Data Generating Process

The data generating process we consider closely mirrors the one presented in AM. Agents choose groups by solving the problem described in Section 2. Since the market for locations is assumed to be perfectly competitive, it maximizes social surplus, and the equilibrium allocation is found by solving a large scale linear programming problem. There are 25,000 individuals in the location market. Individuals choose among 100 groups, and each group has a capacity of 250 individuals, so that the equilibrium allocation places each individual in a group.

The parameters governing the choice problem are chosen as follows:

1. The elements of $[\boldsymbol{X}_i, \boldsymbol{X}_i^U, \boldsymbol{Q}_i]$ are jointly normally distributed, with each element featuring a unit variance; the elements of $\boldsymbol{Q}_i$ are independent of each other and $[\boldsymbol{X}_i, \boldsymbol{X}_i^U]$, and each pair of characteristics in $[\boldsymbol{X}_i, \boldsymbol{X}_i^U]$ features a .25 correlation.[40]
2. The latent amenity factors $\boldsymbol{A}_g$ are normally distributed with a .25 correlation between any pair of amenities across groups. Each amenity factor in $\boldsymbol{A}_g$ features a unit variance.
3. The matrices of taste parameters $\boldsymbol{\Theta}$ and $\boldsymbol{\Theta}^U$ represent draws from a multivariate normal distribution in which (a) $corr(\Theta_{k\ell}, \Theta_{jm}) \equiv .25$ if $j = k$ or $\ell = m$, and 0 otherwise, (b) $corr(\Theta^U_{k\ell}, \Theta^U_{jm}) = .25$ if $j = k$ or $\ell = m$, and 0 otherwise, and (c) $corr(\Theta_{k\ell}, \Theta^U_{jm}) = .25$ if $\ell = m$, and 0 otherwise.
4. There are 5 elements of $\boldsymbol{X}_i$ and of $\boldsymbol{X}_i^U$, so that $L = 5$ and $L^U = 5$. There are $K = 3$ amenity factors in $\boldsymbol{A}_g$.
5. The number of elements of $\boldsymbol{Q}_i$ is equal to the number of elements of $\boldsymbol{A}_g$ ($K = 3$). $\boldsymbol{\Theta}^Q$ is the identity matrix.
6. $\varepsilon_{i,g}$ are drawn i.i.d from a normal distribution with a standard deviation of 15, which was chosen to create interclass correlations for $\boldsymbol{X}_i$ and $\boldsymbol{X}_i^U$ of between .1 and .25 across specifications. These values are in line with the range observed across the datasets used in the empirical analysis.

Note that the description above (notably points 1-4) implies that assumptions A1-A5 are satisfied, so that the results of Proposition 1 hold. Next, we describe the parameters governing the production function.

1. All the observable and unobservable characteristics in $\boldsymbol{X}_i$ and $\boldsymbol{X}_i^U$ are equally important in determining the outcome, so that each characteristic features the same (unit) variance, $\beta_\ell = 1 \; \forall \; \ell$, and $\beta^U_\ell = 1 \; \forall \; \ell$.
2. There are no peer effects, so that $\boldsymbol{\Gamma}_1 = \boldsymbol{\Gamma}_1^U = \boldsymbol{0}$.
3. There is a single observed non-average group characteristic Z_{2g} and a single unobserved non-average group characteristic Z^U_{2g}. Each features a unit variance across groups, and each enters the production function with a coefficient of 1: $\Gamma_2 = 1$ and $\Gamma_2^U = 1$.
4. The correlation between Z_{2g} and Z^U_{2g} is .25. The correlation between Z_{2g} and each of the 3 amenity factors in $\boldsymbol{A}_g$ is denoted $corr_{AZ}$. $corr_{AZ}$ also governs the correlation between Z^U_{2g} and each amenity factor in $\boldsymbol{A}_g$. $corr_{AZ}$ determines the degree to which student sorting is related to the average causal treatment effect associated with group g. We consider three alternative specifications featuring different values of $corr_{AZ}$: 0, .125, and .25.
5. In the production function, M_i is a scalar equal to $\boldsymbol{X}_i\boldsymbol{B}$, where $\boldsymbol{B}$ adheres to the formula provided in (16). This is one of the M_i variables that we use in our empirical work.
6. The scalar interaction $M_i Z_{2g} G_2$ enters the production function with a coefficient of either $\rho_2 = 0.25$ or $\rho_2 = 0.5$, depending on specification. There are no interactions between M_i and either X_g or X_g^U, so that $\boldsymbol{\rho}_1 = \boldsymbol{\rho}_1^U = \boldsymbol{0}$. There are also no interactions between M_i and Z^U_{2g}, so

[40] This is the average correlation between observed continuous student-level characteristics in ELS2002.

that $\rho_2^U = 0$.

7. We set η_{gi} and ξ_{gi} equal to zero $\forall\ (g,i)$.

Because our goal is to evaluate the role of interactions and the additional assumptions necessary to accommodate them in determining the performance of our estimator, we restrict the characteristics generating group treatment effects in our simulations Z_{2g} and Z_{2g}^U to only operate at a single group level g, rather than allowing for separate sets of productive characteristics at the neighborhood, school, and commuting zone levels.

The formulas (17) and (18) reveal that setting $\boldsymbol{\rho}_1 = 0$, $\boldsymbol{\rho}_1^U = 0$ and $\rho_2^U = 0$ implies that $\boldsymbol{r}_1 = \mathbf{0}$ and $r_2 = \rho_2$ when assumptions A1-A7 hold. However, the spanning assumption A5 need not hold when sample averages $\hat{\boldsymbol{X}}_g$ are used in place of the population expectations $\boldsymbol{X}_g$. Furthermore, assumptions A6 and A7 need not hold with the finite number of schools and number of students per school considered and with a non-zero variance of ε_{ig}. Thus, the degree to which this DGP generates violations of A6 and A7 is one of the objects of interest.

The restricted specification laid out above yields the following simplified production function:

$$Y_{ig} = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{X}_i^U\boldsymbol{\beta}^U + Z_{2g}\Gamma_2 + (\boldsymbol{X}_i(\boldsymbol{\beta} + \boldsymbol{\Pi}_{X^UX}\boldsymbol{\beta}^U))(Z_{2g})\rho_2 \tag{46}$$

We estimate the following restricted version of our estimating equation via a mixed effects estimator:

$$Y_{ig} = \boldsymbol{X}_i\boldsymbol{B} + \boldsymbol{X}_g\boldsymbol{G}_1 + Z_{2g}G_2 + (\boldsymbol{X}_i\boldsymbol{B})(Z_{2g}G_2)r_2 + v_g + (v_i - v_g) \tag{47}$$

Note that, as in our empirical work, the index $M_i = \boldsymbol{X}_i(\boldsymbol{\beta} + \boldsymbol{\Pi}_{X^UX}\boldsymbol{\beta}^U)$ depends on the unknown parameters $\boldsymbol{\beta} + \boldsymbol{\Pi}_{X^UX}\boldsymbol{\beta}^U$ that must be estimated (via the parameter vector $\boldsymbol{B}$) simultaneously along with the other parameters of the model. This departs slightly from the set-up of the model in Section 3, where M_i is assumed to be a known function of X_i.

When assessing the impact of observing only small subsamples of the population of individuals in each group, we replace the population means $\boldsymbol{X}_g$ in (47) with their sample mean counterparts $\hat{\boldsymbol{X}}_g$.

Finally, we calculate the sorting equilibrium for twenty economy-wide draws of all of the random variables described above for each of the six combinations of ρ_2 and $Corr_{AZ}$ we consider, and report averages of each reported coefficient or statistic across these twenty draws. Furthermore, when considering estimates from specifications featuring the control function based on small sample means $\hat{\boldsymbol{X}}_g$, for each of the twenty draws we collect 50 random samples of 10, 20, or 40 students in each group and re-estimate the model using $\hat{\boldsymbol{X}}_g$ constructed from the chosen samples. This allows us to abstract from the additional volatility in estimates caused by the reliance on such small samples and instead focus on the bias it generates in the coefficients of interest.

B4.2 Simulation Results

The results of our simulations are presented in Online Appendix Table B1. Our first objective is to demonstrate that the data generating process described above satisfies or nearly satisfies assumptions A6 and A7 that underlie Propositions 2-4.

B4.2.1 Evaluating Assumptions A6 and A7

Recall that assumption A6 requires that $Cov(D\tilde{x}_i^U, Z_{2g}) = 0$ and $\boldsymbol{Cov}(D\tilde{x}_i^U DM_i, \boldsymbol{X}_g) = \mathbf{0}$. Since the scale of Y_i is only implicitly determined in the simulation, rather than directly reporting the sample counterparts to these covariances, we report instead the sample correlation $Corr(D\tilde{x}_i^U, Z_{2g})$ and the mean absolute correlation $\frac{1}{L}|Corr(D\tilde{x}_i^U DM_i, \boldsymbol{X}_{gl})|$ among the $L = 5$ elements of $\boldsymbol{X}_g$.

In Panel A of Online Appendix Table B1, Row 1-3 of Column 1 reports $Corr(D\tilde{x}_i^U, Z_{2g})$ for specifications in which the correlation between each amenity A_{gk} and Z_{2g} is set at 0, .125, and .25, respectively. Note that setting $Corr(A_{gk}, Z_{2g}) = .25$ constitutes a fairly extreme scenario in which half of the variance in Z_{2g} is predictable based on $\boldsymbol{A}_g$. In each specification, $Corr(D\tilde{x}_i^U, Z_{2g})$ almost exactly zero. Similarly, Row 1 of Column 2 shows that the mean absolute correlation $\frac{1}{L}Corr(D\tilde{x}_i^U DM_i, \boldsymbol{X}_{gl})$ is near zero as well (.008). Since this set of correlations does not depend on the relationship between $\boldsymbol{A}_g$ and Z_g, Rows 2 and 3 report the exact same value.[41] Thus, the results confirm that assumption A6 is satisfied by this sorting process.

Assumption A7 requires that $\tilde{Z}_{2g}^U$ is independent of $\boldsymbol{X}_g$, Z_{2g}, and DX_i. Independence is a difficult property to verify. However, note that our proofs of Propositions 2 and 3 use assumption A7 only to argue that the projection coefficients from four projection equations are zero. Specifically, A7 is used to zero out the coefficients from the following projections:

$$D\tilde{x}_i^U = \boldsymbol{DX}_i\boldsymbol{\Pi}_1 + DM_i\boldsymbol{X}_g\boldsymbol{\Pi}_2 + DM_i\boldsymbol{Z}_{2g}\boldsymbol{\Pi}_3 + \psi_{D\tilde{x}_i^U} \tag{48}$$

$$\boldsymbol{DM}_i\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\rho}^U = \boldsymbol{DX}_i\boldsymbol{\Pi}_4 + DM_i\boldsymbol{X}_g\boldsymbol{\Pi}_5 + DM_i\boldsymbol{Z}_{2g}\boldsymbol{\Pi}_6 + \psi_{\boldsymbol{DM}_i\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\rho}^U} \tag{49}$$

$$\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\Gamma}_2^U = \boldsymbol{X}_g\boldsymbol{\Pi}_7 + \boldsymbol{Z}_{2g}\boldsymbol{\Pi}_8 + M_g\boldsymbol{X}_g\boldsymbol{\Pi}_9 + M_g\boldsymbol{Z}_{2g}\boldsymbol{\Pi}_{10} + \psi_{\tilde{\boldsymbol{Z}}_{2g}^U\boldsymbol{\Gamma}_2^U} \tag{50}$$

$$M_g\widetilde{\boldsymbol{Z}_{2g}^U}\boldsymbol{\rho}_2^U = \boldsymbol{X}_g\boldsymbol{\Pi}_{11} + \boldsymbol{Z}_{2g}\boldsymbol{\Pi}_{12} + M_g\boldsymbol{X}_g\boldsymbol{\Pi}_{13} + M_g\boldsymbol{Z}_{2g}\boldsymbol{\Pi}_{14} + \psi_{M_g\widetilde{\boldsymbol{Z}_{2g}^U}\boldsymbol{\rho}_2^U} \tag{51}$$

Thus, we test the key implication of A7 that is used in the proofs by examining whether the coefficients are jointly zero in each projection equation separately. Zero coefficients in the first two equations are required for Proposition 2, while Proposition 3 also requires zero coefficients in the last two equations. Row 1, columns 3 and 4 report the adjusted R-squared (denoted R_{adj}^2 in the table) from the first two projections. We see that the adjusted R-squared values for both of these regressions are almost exactly zero (-.0002 and .0002), suggesting that the formulas in Proposition 2 are likely to be quite accurate here, at least when a large population is used to construct $\hat{\boldsymbol{X}}_g$. Columns 5 and 6 report the corresponding adjusted R-squared values from the third and fourth projections. While Column 5 reports a negative value in all three specifications, indicating that the regressors have no predictive power whatsoever, Column 6 displays small positive values between .044 and .046 in each specification. Taken together, the results suggest that Assumption A7

[41] Note that the validity of A6 and A8 do not depend on the true value of the interaction parameters $\boldsymbol{\rho}$. Thus, we only report results pertaining to A6 and A8 for the specification in which $\rho_2 = 0.25$.

is well approximated by the chosen DGP. However, to gauge whether the minor departures from A7 observed in Column 6 might cause the estimated coefficients to meaningfully diverge from the formulas in Proposition 3, we turn to Panel B, which compares the estimated coefficients with the true coefficients implied by the parameters of the sorting process and production function.

B4.2.2 Evaluating the Accuracy of $\hat{G}_2$ and $\hat{r}_2$ as Estimators of G_2 and r_2

Since G_2 and r_2 are the key parameters used to construct our lower bound estimates of group treatment effects, we are particularly interested in whether the MME estimator can produce accurate estimates of these parameters. To this end, Column 1 of Panel B of Online Appendix Table B1 reports the true value of r_2 used in the production function, while Column 2 provides the full-sample MME estimate, $\hat{r}_2$. The first three rows consider specifications in which $r_2 = 0.25$ and the correlation between Z_{2g} and each element of $\boldsymbol{A}_g$ is 0, .125, and .25, respectively. As one would expect given that A6 and A7 are essentially satisfied, $\hat{r}_2$ very closely matches r_2; even in Row 3, where amenity-driven student sorting is closely related to Z_{2g}, the estimate of $\hat{r}_2$ is still .249. Rows 4-6 repeat the specifications from Rows 1-3, except that the true interaction coefficient is set at $r_2 = 0.5$, a large value given that the standard deviation of $M_i = X_i B$ is around 5, so that a student with a value of $\boldsymbol{X}_i\boldsymbol{B}$ one standard deviation above the mean would be 3.5 times as sensitive to a one unit change in $Z_{2s}G_2$ as a student at the mean of $\boldsymbol{X}_i\boldsymbol{B}$. The estimate of $\hat{r}_2$ again closely matches r_2 for all values of $Corr_{AZ}$.

Columns 4 and 5 provide the corresponding true and full-sample estimated values of the group-level coefficient (G_2 and $\hat{G}_2$). Note that since G_2 depends on the partial projection matrix $\Pi_{Z^U_{2g}Z_{2g}}$, it varies with the sorting equilibrium and thus the exact draws of $\{\varepsilon_{ig}\}$, $[\boldsymbol{X}_i, \boldsymbol{X}^U_i, \boldsymbol{Q}_i]$, $[\Theta, \Theta^U, \Theta^Q]$, $\boldsymbol{A}_g$, and $[Z_{2g}, Z^U_{2g}]$. As mentioned above, we calculate the sorting equilibrium for twenty economy-wide draws of all of these random variables for each of the six specifications represented by rows 1-6, and report averages of G_2 and $\hat{G}_2$ across these twenty draws.

As with the interaction coefficient, $\hat{G}_2$ almost exactly matches G_2. Even when $Corr_{AZ} = .25$, $\hat{G}_2$ only very slightly overstates G_2 (1.138 vs. 1.134). Doubling the magnitude of the true interaction coefficient r_2 has no effect on the estimates of G_2. Thus, a plausible if quite stylized DGP featuring jointly normally distributed amenities and student characteristics leads to quite accurate MME estimates of both the main effects of group characteristics G_2 as well as their interactions with student characteristics r_2.

B4.2.3 Assessing the Impact of Using Small Subsamples to Construct $\hat{\boldsymbol{X}}_g$

The final issue we consider is the performance of our MME estimator when sample means $\hat{\boldsymbol{X}}_g$ based on small samples of individuals in each group are used in place of the population mean $\boldsymbol{X}_g$ to serve as the control function. Recall that we use samples of ~ 20 per school to construct $\boldsymbol{X}_s$ in our empirical work. Online Appendix Table B1, Panel B, column 3 reports the estimates $\hat{r}_2$ generated when 10, 20, and 40 students are used to construct $\hat{\boldsymbol{X}}_g$, while column 6 reports the corresponding estimates of $\hat{G}_2$ for these specifications.

Row 1 and Row 4 show that when Z_{2g} and Z_{2g}^U are uncorrelated with the amenities, even samples of 10 or 20 suffice. This is because Z_{2g} and Z_{2g}^U are essentially uncorrelated with sorting-driven mean differences in $\boldsymbol{X}_i$ or $\boldsymbol{X}_i^U$ across groups. Thus, student sorting, while significant, does not generate any bias that needs to be controlled for. Row 2 shows that the estimator of r_2 still performs quite well for a moderate correlation of 0.125 between Z_{2g} (or Z_{2g}^U) and each amenity factor: relative to a true r_2 of .250, the estimates $\hat{r}_2$ are .242, .245, and .249 when 10, 20, and 40 students, respectively, are used to construct the sample means $\hat{\boldsymbol{X}}_g$. The corresponding estimates $\hat{G}_2$ show a bit of upward bias: relative to a truth of $G_2 = 1.214$, are 1.263, 1.248, and 1.228.

However, when a high value of $Corr_{AZ} = .25$ is considered (Row 3), very small samples of students per school do lead to moderate underestimates of the magnitude of r_2 and moderate overestimates of the magnitude of G_2. Relative to a true r_2 of .250, the estimates $\hat{r}_2$ are .231, .238, and .244 when 10, 20, and 40 students, respectively, are used to construct the sample means $\hat{\boldsymbol{X}}_g$. These represent percentage understatements of 7.6%, 4.8%, and 2.4%, respectively. The corresponding estimates $\hat{G}_2$, relative to a truth of $G_2 = 1.134$, are 1.238, 1.202, and 1.169. These represent percentage understatements of 9.2%, 6.0%, and 3.1%, respectively. Column 6 of Rows 4-6 shows that doubling the true r_2 doubles the absolute bias in $\hat{r}_2$ but maintains the same percentage of the true value, and does not affect $\hat{G}_2$ at all.

Overall, we see that using small samples of students to construct the control function $\hat{X}_g$ generates a small but non-negligible bias in estimates of the key parameters r_2 and G_2 only when the causal group characteristics are closely related to the amenities that drive student sorting.

Appendix Table B1: Monte Carlo Simulation Results Analyzing the Assumptions and Finite Sample Properties of the Multilevel Mixed Effects Estimator

Panel A: Evaluating Assumptions A6 and A8

		A6		A8 - Prop. 2		A8 - Prop. 3	
		(1)	(2)	(3)	(4)	(5)	(6)
Row	Specification	$Corr\ of$ $DM_i\tilde{x}_i^U$ and Z_{2s}	$Avg.\ \lvert Corr \rvert\ of$ $DM_i\tilde{x}_i^U$ and X_{gl}	R^2_{Adj}: $D\tilde{x}_i^U$	R^2_{Adj}: $DM_iZ^U_{2g}$	R^2_{Adj}: $\tilde{z}_g^U$	R^2_{Adj}: $M_gZ^U_{2g}$
(1)	$Corr_{AZ} = 0$	.001	.008	-.0002	.0002	-.062	.046
(2)	$Corr_{AZ} = 0.125$	.000	.008	-.0002	.0002	-.063	.045
(3)	$Corr_{AZ} = 0.25$	.000	.008	-.0002	.0002	-.064	.044

Panel B: Evaluating Estimator Bias with Population and Sample Data

		Interaction Coef.			Group Effect Coef.		
		(1)	(2)	(3)	(4)	(5)	(6)
Row	Specification	r_2	$\hat{r}_2$	$\hat{r}_2$ (10/20/40)	G_2	$\hat{G}_2$	$\hat{G}_2$ (10/20/40)
(1)	$Corr_{AZ} = 0$	0.250	0.251	0.249 0.250 0.252	1.238	1.237	1.251 1.250 1.239
(2)	$Corr_{AZ} = 0.125$	0.250	0.250	0.242 0.245 0.249	1.214	1.215	1.263 1.248 1.228
(3)	$Corr_{AZ} = 0.25$	0.250	0.249	0.231 0.238 0.244	1.134	1.138	1.238 1.202 1.169
(4)	$Corr_{AZ} = 0$	0.500	0.502	0.498 0.499 0.502	1.238	1.237	1.251 1.250 1.239
(5)	$Corr_{AZ} = 0.125$	0.500	0.500	0.485 0.490 0.497	1.214	1.215	1.262 1.248 1.228
(6)	$Corr_{AZ} = 0.25$	0.500	0.499	0.462 0.475 0.488	1.134	1.138	1.238 1.202 1.169

Notes: See Appendix B4 for a full description of the data generating process used to generate the simulation results.
$Corr_{AZ}$: Correlation between Z_{2g} and each amenity factor A_{gk}.
$Avg.\ \lvert Corr \rvert\ of DM_i\tilde{x}_i^U\ and\ X_{gl}$: Mean absolute value of the correlation across all L elements of $\boldsymbol{X}_g$ of $DM_i\tilde{x}_i^U$ and X_{gl}.
R^2_{Adj} : "Y" : Adjusted R-squared from the projection of "Y" on the corresponding within-group or between-group observable characteristics.
$\hat{r}_2(10/20/40)$: Estimated interaction coefficient when samples of 10, 20 or 40 individuals are used to compute the sample means $\hat{\boldsymbol{X}}_g$.

Appendix Table B2: Descriptive Statistics for Explanatory Variables in Basic and Full Specifications (by Data Set)

	A. NELS				B. ELS			
Variable		**Mean**	**Std. Dev.**	**% Imputed**		**Mean**	**Std. Dev.**	**% Imputed**
Student Characteristics (X_i)								
1(Female)		0.514	0.500	0.00		0.508	0.500	0.00
1(Black)		0.106	0.308	0.77		0.129	0.336	0.00
1(Hispanic)		0.132	0.338	0.77		0.136	0.343	0.00
1(Asian)		0.069	0.253	0.77		0.093	0.290	0.00
1(White)	*	0.681	0.466	0.77	*	0.574	0.495	0.00
1(Other race)						0.059	0.236	0.00
1(Immigrant)		0.073	0.249	7.22		0.102	0.282	15.00
1(Native English speaker)		0.872	0.334	0.54		0.836	0.367	2.35
1(Athletic)		2.215	1.356	0.00		0.360	0.456	10.72
1(Black Male)		0.049	0.217	0.77		0.064	0.244	0.00
1(Hispanic Male)		0.063	0.242	0.77		0.067	0.250	0.00
# Weekly homework hours	x	6.060	5.157	5.94	x	10.880	8.696	7.47
1(Parent checks HW)	x	0.439	0.495	0.64	x	0.347	0.433	17.60
# Weekly reading hours	x	2.218	2.615	4.54	x	2.751	3.881	7.82
1(Often missing pencil)	x	0.221	0.405	4.44	x	0.168	0.356	9.21
1(Fought at school)	x	0.203	0.399	1.71	x	0.126	0.318	8.35
Std. math score	x	0.065	1.001	0.00	x	0.109	0.984	0.00
Std. reading score	x	0.051	0.993	0.00	x	0.091	0.979	0.00
Parent and Family Characteristics (X_i)								
SES Index (standardized)		-0.011	0.710	0.00		0.102	1.040	0.00
Number of siblings		2.293	1.581	0.60		2.295	1.373	21.40
family composition:								
1(Does not live with both M and F)		0.319	0.464	1.07		0.400	0.466	10.80
Father's years education		13.275	5.254	6.53		13.881	2.687	9.01
Mother's years education		12.846	2.405	0.00		13.665	2.352	0.00
1(Mother's ed missing)		0.023	0.151	0.00		0.034	0.182	0.00
Log(family income)		10.876	2.169	10.09		10.979	0.889	24.82
1(Mother or father is immigrant)		0.176	0.367	9.57		0.254	0.414	16.81
1(Protestant)	*	0.459	0.491	3.68	*	0.341	0.426	22.64
1(Catholic)		0.318	0.458	3.68		0.347	0.432	22.64
1(Other Christian)		0.069	0.248	3.68		0.182	0.341	22.64
1(Religion other)		0.090	0.280	3.68		0.130	0.296	22.64
1(Religion missing)		0.037	0.188	0.00		0.226	0.419	22.64
mother's occupation:								
1(Manager, accountant, nurse, business owner, teacher)		0.272	0.429	11.56		0.367	0.442	0.00
1(Missing)		0.115	0.319	0.00		0.250	0.433	25.05
1(Sales, service)		0.197	0.376	11.56		0.205	0.351	0.00
1(Clerical)		0.215	0.390	11.56		0.175	0.330	25.05
1(Other, homemaker)	*	0.306	0.434	11.56	*	0.257	0.376	25.05
father's occupation:								
1(Accountant, nurse, teacher, manager, dentist, lawyer, business owner, etc)		0.333	0.447	24.69		0.378	0.447	34.28
1(Service, clerical, sales, missing, other, homemaker)		0.107	0.265	24.69		0.127	0.265	34.28
1(Military, security, craftsman, technician)		0.250	0.374	24.69		0.233	0.340	34.28
1(Farmer, laborer, operative)	*	0.300	0.419	24.69	*	0.272	0.383	34.28

Appendix Table B2 Continued:						
	A. NELS			B. ELS		
Variable	Mean	Std. Dev.	% Imputed	Mean	Std. Dev.	% Imputed
Home environment index	-0.025	1.658	6.76	-0.018	1.319	16.58
Parental sch engagement idx	-0.082	1.475	11.22	-0.028	1.382	24.73
Parents yrs ed desired for child	16.218	1.826	18.34	16.677	1.910	31.76
Neighborhood Characteristics (ZIP Code, treated as X_n)						
% Black	0.103	0.184	2.52	0.125	0.194	1.26
% Hispanic	0.106	0.194	2.52	0.116	0.187	1.26
% White and other	* 0.791	0.275	2.52	* 0.759	0.301	1.26
% Non-married household	0.522	0.079	2.53	0.259	0.128	1.26
% Married household	* 0.478	0.079	2.53	* 0.741	0.128	1.26
% Foreign born	0.076	0.111	2.52	0.101	0.122	1.26
% Native born	* 0.924	0.111	2.52	* 0.899	0.122	1.26
% High school or less	* 0.570	0.145	2.52	* 0.491	0.157	1.26
% Some college or assoc deg	0.242	0.070	2.52	0.276	0.061	1.26
% Four-year col deg or higher	0.187	0.127	2.52	0.233	0.146	1.26
Log(median income)	10.424	0.365	2.53	10.654	0.347	1.26
Gini coefficient				0.399	0.046	1.26
% SSI or welfare recipients	0.160	0.207	2.53	0.080	0.059	1.26
% Not SSI or welfare recipients	* 0.840	0.207	2.53	* 0.920	0.059	1.26
Log(median house value)				11.646	0.514	1.30
% Housing properties occupied				0.922	0.069	1.26
Neighborhood Characteristics (Block Group, treated as X_n)						
Proportion of jobs in:						
Agriculture, mining, oil, utility, construction, manufacturing				0.147	0.089	0.00
Information, finance, insurance, real estate, professional, science				0.124	0.073	0.00
Management, admin, waste mgmt				0.062	0.032	0.00
Education, other services and public administration				0.161	0.075	0.00
Transportation and warehousing				0.124	0.058	0.00
Health care, arts, entertainment, recreation, accommodation, food				* 0.381	0.153	0.00
% White				* 0.693	0.308	0.00
% Black				0.126	0.223	0.00
% Hispanic				0.116	0.200	0.00
% Other				0.068	0.119	0.00
% Married household				* 0.752	0.157	0.00
% Non-married household				0.248	0.157	0.00
% Native born				* 0.898	0.136	0.00
% Foreign born				0.102	0.136	0.00
% High school or less				* 0.481	0.195	0.00
% Some college or assoc deg				0.277	0.081	0.00
% Four-year col deg or higher				0.241	0.178	0.01
Log(median income)				10.691	0.472	0.00
Gini coefficient				0.373	0.064	0.00
% SSI or welfare recipients				0.079	0.082	0.00
% Not SSI or welfare recipients				* 0.921	0.082	0.00
Log(median house value)				11.654	0.576	0.99
% Housing properties occupied				0.931	0.069	0.00

Appendix Table B2 Continued:								
	A. NELS				B. ELS			
Variable		Mean	Std. Dev.	% Imputed		Mean	Std. Dev.	% Imputed
School Characteristics (Treated as part of X_s)								
% Minority		0.226	0.289	0.43		0.333	0.306	1.81
% Limited English proficient		0.070	0.084	0.25		0.042	0.082	3.98
% Free/reduced lunch		0.233	0.231	0.45		0.231	0.241	7.35
% in special ed	x	0.063	0.050	0.25	x	0.092	0.087	5.62
% in remedial reading	x	0.099	0.127	0.14	x	0.042	0.065	17.76
% in remedial math	x	0.073	0.101	0.14	x	0.059	0.085	19.11
School Characteristics (Z_{2s})								
1(Catholic school)		0.094	0.292	0.00		0.129	0.334	0.47
1(Private school)		0.071	0.256	0.00		0.092	0.287	1.05
Teacher turnover rate						0.060	0.061	27.85
School enrollment		665.093	372.846	0.00		1262.200	820.095	0.38
Student-teacher ratio		17.768	5.044	0.00		16.568	4.130	3.00
% Teachers minority		0.111	0.185	1.82		0.131	0.184	38.45
% Teachers with certification						91.984	17.723	2.87
% Teachers w masters deg		0.466	0.244	2.65				
Collective bargaining		0.562	0.495	0.45				
Log(min teacher salary)		9.760	0.180	1.34				
Teacher evaluation policy index						0.008	1.114	14.57
Teacher incentive pay index (1)						0.003	1.359	13.54
Teacher technology access index (1)						0.010	1.547	15.95
School's physical environment index (1)						-0.096	1.753	30.02
Admin's security policies index (1)		0.109	1.129	0.31		0.027	1.468	15.77
Admin's security policies index (2)		-0.034	1.061	0.31		0.010	1.208	15.77
Admin's school facility index (1)						0.034	2.135	19.73
1(Rural, not in MSA)						0.102	0.303	0.25
1(Rural, in MSA)						0.098	0.297	0.25
1(Town)						0.104	0.306	0.25
1(Suburb of medium city)						0.084	0.278	0.25
1(Medium city)						0.165	0.370	0.25
1(Large city)						0.162	0.368	0.25
Admin's Crime in Neighborhood						2.931	0.607	12.15
1(Urban)		0.257	0.437	0.00				
1(Suburban)		0.425	0.494	0.00				
1(Rural)	*	0.319	0.466	0.00				
1(Gifted program)		0.658	0.475	0.00				
JH school minutes / year		70918.510	5800.581	0.85				
JH course assignment index (1)		0.019	1.242	0.43				
JH student retainment policy index (1)		-0.025	1.998	6.51				
JH student activities index (1)		-0.010	2.009	1.57				
JH school environment index (1)		-0.055	2.231	1.14				
JH movement in sch policy index (1)		-0.143	1.716	0.43				
JH counseling policy index		-0.056	1.265	0.48				
JH sch uniforms index		0.066	0.975	0.00				
JH student problems index (1)		-0.032	2.201	0.81				
JH student punishment policy index (1)		-0.052	1.992	2.55				

Appendix Table B2 Continued:							
	A. NELS			B. ELS			
Variable	Mean	Std. Dev.	% Imputed	Mean	Std. Dev.	% Imputed	
Commuting Zone Characteristics (Z_{2c})							
Household income per capita	38240.630	7097.921	0.00	38768.060	6962.543	0.00	
Theil racial segregation index	0.231	0.110	0.00	0.237	0.109	0.00	
Log population density	5.413	1.419	0.00	5.535	1.369	0.00	
% Black	0.118	0.105	0.00	0.127	0.105	0.00	
Income segregation	0.083	0.035	0.00	0.086	0.035	0.00	
Social capital index	-0.342	0.966	0.73	-0.387	0.952	0.97	
Poverty rate	0.128	0.049	0.00	0.126	0.044	0.00	
Unemployment rate	0.049	0.012	0.00	0.049	0.015	0.00	
Fraction of Children with Single Mothers	0.219	0.039	0.00	0.221	0.037	0.00	
Gini coefficient	0.466	0.084	0.00	0.473	0.080	0.00	
HS dropout rate (income adj)	0.005	0.017	26.22	0.006	0.018	27.41	
College grad rate (income adj)	-0.020	0.102	1.41	-0.019	0.104	1.27	
Number of Colleges per Capita	0.014	0.008	1.57	0.013	0.007	1.71	
CZ causal effect on college attendance from 18-23	0.050	0.362	1.26	0.014	0.358	0.32	
CZ causal effect on rank in national income distribution at age 26	0.018	0.270	1.26	0.000	0.242	0.32	
1(Northeast)	0.186	0.389	0.00	0.179	0.384	0.00	
1(Midwest)	0.266	0.442	0.00	0.255	0.436	0.00	
1(West)	0.195	0.396	0.00	0.202	0.401	0.00	
1(South)	* 0.353	0.478	0.00	* 0.364	0.481	0.00	

(*) indicates that variable is the left-out group. Variables marked with (x) are excluded from the basic specification. Indices are constructed using principal components. (i) indicates the i-th principal component. Zip code characteristics (Xn) are measured from longform Census. Year 1990 Census used for NELS, year 2000 Census used for ELS. Block group characteristics are measured from both longform Census and LODES. Commuting zone characteristics (Z2c) are measured in year 2000. School characteristics treated as elements of Xs are used to avoid measurement error in school sample averages of the corresponding student characteristics. They do not contribute to the estimated lower bound on contributions of schools or commuting zones. These summary statistics refer to the high school graduation samples. They are slightly different for other outcomes, depending on response rates in further follow-up surveys which observations are lost due to missing data. All statistics and calculations are equal-weighted.

Appendix Table B3-1. Fraction of the Variance of Education and Wages at the Individual, Neighborhood, School, and Commuting Zone Levels (Base Set of Student Variables)

		A. HS Grad			B. College Enroll			C. College Grad			D. Log Wage		
		NELS	ELSbg	ELSz	NELS	ELSbg	ELSz	NELS	ELSbg	ELSz	NELS	ELSbg	ELSz
	Row	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Total Individual	(1)	0.815	0.804	0.832	0.738	0.708	0.699	0.745	0.806	0.813	0.879	0.827	0.844
$Var(Y_i-Y_n)/Var(Y_i)$		[.77, .844]	[.76, .832]	[.77, .857]	[.704, .757]	[.668, .739]	[.68, .746]	[.698, .757]	[.746, .815]	[.771, .83]	[.809, .906]	[.759, .83]	[.763, .85]
Total neighborhood	(2)	0.013	0.031	0.018	0.016	0.028	0.010	0.012	0.026	0.011	0.008	0.028	0.012
$Var(Y_n-Y_s)/Var(Y_i)$		[.005, .016]	[.015, .038]	[.009, .019]	[.009, .019]	[.023, .047]	[.008, .015]	[.008, .015]	[.025, .045]	[.009, .016]	[.005, .01]	[.021, .043]	[.007, .015]
Total School	(3)	0.106	0.102	0.102	0.147	0.161	0.175	0.152	0.102	0.108	0.036	0.081	0.081
$Var(Y_s-Y_c)/Var(Y_i)$		[.082, .136]	[.076, .134]	[.074, .135]	[.13, .171]	[.122, .188]	[.133, .185]	[.134, .182]	[.08, .144]	[.086, .139]	[.011, .07]	[.07, .125]	[.074, .123]
Total CZ	(4)	0.067	0.063	0.048	0.099	0.102	0.117	0.091	0.066	0.068	0.077	0.064	0.063
$Var(Y_c)/Var(Y_i)$		[.055, .097]	[.048, .096]	[.045, .094]	[.084, .124]	[.084, .123]	[.089, .132]	[.079, .124]	[.057, .087]	[.06, .088]	[.058, .13]	[.051, .099]	[.05, .116]

Note: ELSbg represents the ELS data set with the neighborhood specification of block group. ELSz represents the ELS data set with the neighborhood specification of zip code. This table reports the fractions of the variance of log wage and of the latent variables for high school graduation, enrollment at a 4-year college, and a 4-year college degree that is person-specific within a neighborhood (total individual), neighborhood-specific within a school (total neighborhood), school-specific within a commuting zone (total school) and commuting-zone-specific (total CZ). They are computed based on estimates of equation (26), for the base set of X_i variables (not reported). The 5th and 95th percentile values of the bootstrap replications are reported in brackets.

Appendix Table B3-2. Fraction of the Variance of XiB at the Individual, Neighborhood, School, and Commuting Zone Levels (Base Set of Student Variables)

		A. HS Grad			B. College Enroll			C. College Grad			D. Log Wage		
		NELS	ELSbg	ELSz	NELS	ELSbg	ELSz	NELS	ELSbg	ELSz	NELS	ELSbg	ELSz
	Row	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Individual Share of var(X_iB)	(1)	0.694	0.689	0.718	0.655	0.619	0.657	0.605	0.608	0.650	0.727	0.671	0.716
		[.652, .884]	[.663, .729]	[.695, .756]	[.613, .701]	[.6, .64]	[.635, .675]	[.58, .637]	[.58, .641]	[.625, .675]	[.69, .781]	[.637, .723]	[.679, .756]
Neighborhood Share of var(X_iB)	(2)	0.015	0.043	0.012	0.016	0.057	0.017	0.017	0.064	0.019	0.012	0.061	0.015
		[.000, .023]	[.023, .058]	[.005, .021]	[.008, .026]	[.042, .073]	[.010, .028]	[.009, .027]	[.040, .081]	[.012, .028]	[.001, .023]	[.032, .08]	[.003, .026]
School Share of var(X_iB)	(3)	0.158	0.171	0.172	0.194	0.203	0.205	0.228	0.208	0.209	0.160	0.174	0.175
		[.088, .173]	[.142, .188]	[.141, .188]	[.162, .213]	[.181, .223]	[.182, .223]	[.196, .246]	[.185, .225]	[.188, .225]	[.124, .183]	[.146, .194]	[.146, .197]
CZ Share of var(X_iB)	(4)	0.133	0.097	0.098	0.135	0.121	0.122	0.150	0.120	0.122	0.100	0.093	0.094
		[.071, .163]	[.081, .115]	[.082, .114]	[.105, .168]	[.107, .141]	[.108, .143]	[.12, .183]	[.105, .14]	[.107, .143]	[.079, .122]	[.075, .114]	[.076, .115]
X_iB Share of Var(Y_i)	(5)	0.176	0.151	0.153	0.239	0.237	0.233	0.288	0.229	0.234	0.118	0.148	0.151
		[.157, .339]	[.138, .177]	[.136, .174]	[.212, .267]	[.212, .262]	[.218, .262]	[.264, .322]	[.208, .25]	[.216, .258]	[.106, .15]	[.132, .179]	[.134, .182]

Note: ELSbg represents the ELS data set with the neighborhood specification of block group. ELSz represents the ELS data set with the neighborhood specification of zip code. The rows of Appendix Table B3-2 report the fractions of the variance of X_iB that is within a neighborhood (total individual), neighborhood-specific within a school (total neighborhood), school-specific within a commuting zone (total school), and commuting-zone-specific (total CZ). X_iB is the index of student characteristics that affect the outcome indicated in the column headings. B is estimated separately for each outcome and sample as part of the estimation of equation (26), using the base set of X_i variables (not reported). The 5th and 95th percentile values of the bootstrap replications are reported in brackets.

Appendix Table B4. Estimates of the Education and Wage Model Parameters (Base Set of Student Variables)

	A. High School Graduation				B. College Enrollment			
	NELS	**ELSbg**	**ELSz**	**N+E**	**NELS**	**ELSbg**	**ELSz**	**N+E**
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
sd(X_iB)	0.506 ***	0.454 ***	0.449 ***	0.478 ***	0.645 ***	0.647 ***	0.657 ***	0.651 ***
	(.107)	(.028)	(.028)	(.055)	(.033)	(.02)	(.02)	(.019)
sd($X_{1n}G_1^N$)	0.000	0.000	0.089	0.045	0.000	0.066 *	0.000	0.000
	(.039)	(.035)	(.058)	(.035)	(.034)	(.034)	(.039)	(.026)
sd($X_{1s}G_1^S$)	0.173 ***	0.153 ***	0.133 ***	0.153 ***	0.172 ***	0.227 ***	0.247 ***	0.209 ***
	(.042)	(.038)	(.038)	(.029)	(.034)	(.04)	(.036)	(.025)
sd($Z_{2s}G_2^S$)	0.104 ***	0.116 ***	0.139 ***	0.122 ***	0.144 ***	0.074 *	0.101 **	0.122 ***
	(.036)	(.042)	(.044)	(.029)	(.028)	(.044)	(.045)	(.026)
sd($Z_{2c}G_2^C$)	0.162 ***	0.111 *	0.087	0.125 ***	0.154 ***	0.135 **	0.144 **	0.149 ***
	(.04)	(.066)	(.075)	(.043)	(.026)	(.056)	(.056)	(.031)
X_iB x $X_{1n}G_1^N$ (r_1^N)	-0.031	-0.036	0.028	-0.002	0.001	-0.018	-0.083 **	-0.041 *
	(.036)	(.034)	(.023)	(.022)	(.029)	(.031)	(.034)	(.022)
X_iB x $X_{1S}G_1^S$ (r_1^S)	0.020	0.031	0.031	0.025	0.015	-0.005	0.043	0.029
	(.031)	(.027)	(.027)	(.02)	(.025)	(.038)	(.032)	(.02)
X_iB x $Z_{2S}G_2^S$ (r_{21}^S)	-0.016	-0.043	-0.034	-0.025	-0.073 ***	-0.048	-0.061 **	-0.067 ***
	(.039)	(.038)	(.037)	(.027)	(.027)	(.033)	(.03)	(.02)
X_iB x $Z_{2C}G_2^C$ (r_{21}^C)	0.034	-0.011	-0.024	0.005	-0.004	0.028	0.018	0.007
	(.027)	(.031)	(.031)	(.021)	(.026)	(.031)	(.029)	(.019)
Female x $Z_{2S}G_2^S$ (r_{22}^S)	0.021	-0.050	-0.054	-0.017	-0.002	-0.015	-0.005	-0.004
	(.047)	(.052)	(.051)	(.035)	(.029)	(.029)	(.028)	(.02)
Minority x $Z_{2S}G_2^S$ (r_{23}^S)	-0.059	0.110 *	0.104 *	0.022	0.027	-0.027	-0.019	0.004
	(.048)	(.06)	(.054)	(.036)	(.05)	(.039)	(.036)	(.031)
LowInc x $Z_{2S}G_2^S$ (r_{24}^S)	0.079	0.011	-0.003	0.038	0.030	0.067	0.066	0.048
	(.067)	(.063)	(.067)	(.047)	(.054)	(.045)	(.044)	(.035)
Female x $Z_{2C}G_2^C$ (r_{22}^C)	-0.007	0.000	0.032	0.012	-0.036	-0.027	-0.032	-0.034 *
	(.039)	(.039)	(.041)	(.028)	(.027)	(.027)	(.028)	(.019)
Minority x $Z_{2C}G_2^C$ (r_{23}^C)	0.064	-0.021	-0.021	0.021	-0.042	0.034	0.016	-0.013
	(.053)	(.059)	(.06)	(.04)	(.045)	(.038)	(.036)	(.029)
LowInc x $Z_{2C}G_2^C$ (r_{24}^C)	0.012	-0.035	-0.037	-0.012	-0.079 **	0.024	0.034	-0.022
	(.045)	(.045)	(.044)	(.031)	(.034)	(.047)	(.041)	(.026)
(Intercept)	1.295 ***	1.698 ***	1.689 ***	1.492 ***	-0.572 ***	-0.249 ***	-0.256 ***	-0.414 ***
	(.036)	(.046)	(.046)	(.029)	(.039)	(.037)	(.037)	(.027)
RANDOM EFFECTS								
sd(v_n - v_s)	0.118 ***	0.180 ***	0.134 ***	0.126 ***	0.138 ***	0.159 ***	0.105 ***	0.122 ***
	(.016)	(.026)	(.017)	(.012)	(.022)	(.026)	(.015)	(.013)
sd(v_s - v_c)	0.100 ***	0.125 ***	0.121 ***	0.110 ***	0.146 ***	0.228 ***	0.218 ***	0.182 ***
	(.015)	(.019)	(.017)	(.011)	(.023)	(.037)	(.036)	(.021)
sd(v_c)	0.064 ***	0.068 ***	0.060 ***	0.062 ***	0.049 ***	0.080 ***	0.090 ***	0.070 ***
	(.011)	(.013)	(.011)	(.008)	(.01)	(.024)	(.026)	(.014)

Notes: *, **, *** indicate significance at 10%, 5%, and 1% levels respectively. Bootstrap standard errors are in parentheses. The dependent variables are high school graduation (HSGRAD) in Panel A, enrollment in a 4 year college within 2 years after expected high school graduation (ENROLL) in Panel B, attainment of a BA degree (COLLBA) in Panel C and the log hourly wage rate (ln(wage)) at about age 25 in Panel D. Panels A, B, and C refer to the latent index of an MME probit specification. Panel D is based on an MME regression specification. The model is equation (23). The column heading NELS refers to the NELS data. ELSbg represents the ELS data set with the neighborhood specification of block group. ELSz represents the ELS data set with the neighborhood specification of zip code. Column 4 (8) report the average and standard error of the average of the NELS and ELSz estimates in columns 1 and 3 (5 and 7). The parameter vectors B, G1, G2S, and G2C that define the explanatory index variables XiB, X1nG1N, Z2SG2S and Z2CG2C are estimated in a first step using a nonlinear probit model. The model includes the "baseline" set of Xi variables (student level) and the corresponding "baseline" set of Xs variables (school means). See Appendix A1 for a list of Xi, X1n, Xs, and Z2c variables and Appendix table B2 for summary statistics. The index variables in the interaction terms are standardized to be mean 0 and sd 1. Standard deviation of vi is 1 in probit specifications. Names of interaction coefficients are next to the variables. See Section 6.2 for details about the estimation and bootstrap standard error procedures.

Appendix Table B4, continued. Estimates of the Education and Wage Model Parameters (Base Set of Student Variables)								
	C. College Graduation				D. Log Wage			
	NELS	ELSbg	ELSz	N+E	NELS	ELSbg	ELSz	N+E
VARIABLES	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
sd(X_iB)	0.714 ***	0.593 ***	0.603 ***	0.659 ***	0.123 ***	0.130 ***	0.132 ***	0.127 ***
	(.028)	(.022)	(.022)	(.018)	(.007)	(.007)	(.006)	(.005)
sd($X_{1n}G_1^N$)	0.000	0.088 ***	0.000	0.000	0.040 ***	0.015	0.024	0.032 ***
	(.042)	(.032)	(.039)	(.029)	(.014)	(.011)	(.015)	(.01)
sd($X_{1s}G_1^S$)	0.153 ***	0.134 ***	0.139 ***	0.146 ***	0.056 ***	0.051 ***	0.043 ***	0.049 ***
	(.036)	(.034)	(.034)	(.025)	(.013)	(.009)	(.011)	(.009)
sd($Z_{2s}G_2^S$)	0.082 ***	0.058 *	0.064 **	0.073 ***	0.010	0.010	0.015	0.013
	(.028)	(.033)	(.033)	(.021)	(.013)	(.012)	(.012)	(.009)
sd($Z_{2c}G_2^C$)	0.073 ***	0.091 ***	0.080 **	0.076 ***	0.035 ***	0.025 **	0.023 **	0.029 ***
	(.024)	(.033)	(.034)	(.021)	(.009)	(.01)	(.012)	(.008)
X_iB x $X_{1n}G_1^N$ (r_1^N)	-0.053	-0.051 **	-0.034	-0.043 *	-0.014	0.003	0.005	-0.005
	(.037)	(.026)	(.035)	(.026)	(.009)	(.008)	(.006)	(.006)
X_iB x $X_{1S}G_1^S$ (r_1^S)	0.050	-0.041	-0.038	0.006	-0.011	0.002	-0.001	-0.006
	(.034)	(.03)	(.03)	(.023)	(.01)	(.008)	(.009)	(.007)
X_iB x $Z_{2S}G_2^S$ (r_{21}^S)	-0.041	-0.053 *	-0.055 *	-0.048 **	0.005	0.000	-0.005	0.000
	(.032)	(.03)	(.031)	(.022)	(.009)	(.008)	(.008)	(.006)
X_iB x $Z_{2C}G_2^C$ (r_{21}^C)	0.042	-0.028	-0.052	-0.005	-0.016	-0.006	-0.005	-0.010 *
	(.028)	(.028)	(.033)	(.022)	(.01)	(.007)	(.007)	(.006)
Female x $Z_{2S}G_2^S$ (r_{22}^S)	-0.029	0.029	0.028	-0.001	0.009	0.012	0.007	0.008
	(.038)	(.027)	(.03)	(.024)	(.015)	(.014)	(.015)	(.011)
Minority x $Z_{2S}G_2^S$ (r_{23}^S)	0.010	0.035	0.033	0.022	-0.021	0.006	-0.003	-0.012
	(.052)	(.038)	(.039)	(.033)	(.016)	(.012)	(.012)	(.01)
LowInc x $Z_{2S}G_2^S$ (r_{24}^S)	-0.003	0.049	0.019	0.008	0.009	-0.002	-0.009	0.000
	(.042)	(.065)	(.063)	(.038)	(.015)	(.016)	(.016)	(.011)
Female x $Z_{2C}G_2^C$ (r_{22}^C)	-0.022	0.015	0.031	0.004	0.010	0.017	0.012	0.011
	(.035)	(.027)	(.027)	(.022)	(.017)	(.017)	(.019)	(.013)
Minority x $Z_{2C}G_2^C$ (r_{23}^C)	-0.054	-0.020	-0.038	-0.046	-0.024	-0.019	-0.010	-0.017
	(.054)	(.044)	(.046)	(.035)	(.021)	(.013)	(.013)	(.012)
LowInc x $Z_{2C}G_2^C$ (r_{24}^C)	-0.031	0.092	0.061	0.015	0.017	0.004	-0.002	0.007
	(.046)	(.058)	(.055)	(.036)	(.015)	(.014)	(.015)	(.011)
(Intercept)	-0.496 ***	-0.373 ***	-0.383 ***	-0.440 ***	2.547 ***	2.662 ***	2.664 ***	2.605 ***
	(.042)	(.032)	(.032)	(.026)	(.008)	(.009)	(.009)	(.006)
RANDOM EFFECTS								
sd(v_i)					0.319 ***	0.289 ***	0.290 ***	0.304 ***
					(.004)	(.006)	(.004)	(.003)
sd(v_n - v_s)	0.108 ***	0.128 ***	0.092 ***	0.100 ***	0.025 ***	0.047 ***	0.033 ***	0.029 ***
	(.016)	(.024)	(.013)	(.01)	(.003)	(.003)	(.003)	(.002)
sd(v_s - v_c)	0.079 ***	0.074 ***	0.086 ***	0.082 ***	0.022 ***	0.018 ***	0.022 ***	0.022 ***
	(.01)	(.013)	(.011)	(.007)	(.002)	(.002)	(.002)	(.001)
std(v_c)	0.034 ***	0.045 ***	0.040 ***	0.037 ***	0.013	0.012 **	0.013 *	0.013 **
	(.005)	(.007)	(.008)	(.004)	(.009)	(.005)	(.007)	(.006)

Appendix Table B5. Education and Wage Model Parameters, No Interactions (Full Set of Student Variables)

	A. High School Graduation				B. College Enrollment			
	NELS	ELSbg	ELSz	N+E	NELS	ELSbg	ELSz	N+E
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
sd(X_iB)	0.613 ***	0.592 ***	0.587 ***	0.600 ***	0.854 ***	0.951 ***	0.946 ***	0.900 ***
	(.07)	(.023)	(.025)	(.037)	(.025)	(.025)	(.025)	(.018)
sd($X_{1n}G_1^N$)	0.000	0.000	0.077 *	0.038	0.000	0.049	0.000	0.000
	(.052)	(.032)	(.04)	(.032)	(.039)	(.037)	(.06)	(.036)
sd($X_{1s}G_1^S$)	0.150 ***	0.113 ***	0.089 ***	0.120 ***	0.190 ***	0.163 ***	0.179 ***	0.185 ***
	(.04)	(.037)	(.033)	(.026)	(.03)	(.033)	(.037)	(.024)
sd($Z_{2s}G_2^S$)	0.083 *	0.126 ***	0.134 ***	0.109 ***	0.095 ***	0.081 **	0.107 ***	0.101 ***
	(.044)	(.046)	(.042)	(.03)	(.03)	(.041)	(.038)	(.024)
sd($Z_{2c}G_2^C$)	0.115 ***	0.122 ***	0.118 ***	0.116 ***	0.157 ***	0.125 ***	0.137 ***	0.147 ***
	(.039)	(.032)	(.043)	(.029)	(.028)	(.038)	(.047)	(.028)
(Intercept)	1.335 ***	1.776 ***	1.763 ***	1.549 ***	-0.630 ***	-0.300 ***	-0.297 ***	-0.464 ***
	(.036)	(.046)	(.045)	(.029)	(.04)	(.037)	(.037)	(.027)
RANDOM EFFECTS								
sd(v_n - v_s)	0.121 ***	0.159 ***	0.127 ***	0.124 ***	0.123 ***	0.192 ***	0.128 ***	0.125 ***
	(.016)	(.027)	(.017)	(.012)	(.02)	(.028)	(.016)	(.013)
sd(v_s - v_c)	0.091 ***	0.115 ***	0.109 ***	0.100 ***	0.140 ***	0.147 ***	0.138 ***	0.139 ***
	(.011)	(.016)	(.021)	(.012)	(.019)	(.023)	(.023)	(.015)
sd(v_c)	0.062 ***	0.059 ***	0.062 **	0.062 ***	0.058 ***	0.098 ***	0.096 ***	0.077 ***
	(.01)	(.01)	(.027)	(.014)	(.008)	(.026)	(.035)	(.018)

	C. College Graduation				D. Log Wage			
	NELS	ELSbg	ELSz	N+E	NELS	ELSbg	ELSz	N+E
VARIABLES	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
sd(X_iB)	0.741 ***	0.820 ***	0.827 ***	0.784 ***	0.134 ***	0.163 ***	0.163 ***	0.149 ***
	(.026)	(.025)	(.026)	(.018)	(.027)	(.036)	(.035)	(.022)
sd($X_{1n}G_1^N$)	0.000	0.064 *	0.000	0.000	0.037	0.012	0.032	0.034
	(.035)	(.033)	(.039)	(.026)	(.047)	(.043)	(.062)	(.039)
sd($X_{1s}G_1^S$)	0.160 ***	0.057 **	0.065 **	0.113 ***	0.050	0.037	0.029	0.039
	(.034)	(.027)	(.028)	(.022)	(.042)	(.038)	(.042)	(.029)
sd($Z_{2s}G_2^S$)	0.079 **	0.045	0.044	0.062 **	0.000	0.013	0.018	0.009
	(.035)	(.033)	(.035)	(.024)	(.042)	(.042)	(.045)	(.031)
sd($Z_{2c}G_2^C$)	0.097 ***	0.092 ***	0.068 ***	0.082 ***	0.034	0.028	0.025	0.030
	(.026)	(.029)	(.025)	(.018)	(.031)	(.038)	(.042)	(.026)
(Intercept)	-0.536 ***	-0.451 ***	-0.450 ***	-0.493 ***	2.544 ***	2.666 ***	2.666 ***	2.605 ***
	(.04)	(.034)	(.034)	(.026)	(.008)	(.009)	(.009)	(.006)
RANDOM EFFECTS								
sd(v_i)					0.319 ***	0.274 ***	0.277 ***	0.298 ***
					(.007)	(.007)	(.009)	(.006)
sd(v_n - v_s)	0.110 ***	0.131 ***	0.098 ***	0.104 ***	0.028 ***	0.050 **	0.030 *	0.029 ***
	(.012)	(.024)	(.014)	(.009)	(.011)	(.024)	(.017)	(.01)
sd(v_s - v_c)	0.073 ***	0.068 ***	0.071 ***	0.072 ***	0.021 **	0.019 **	0.020 **	0.020 ***
	(.009)	(.01)	(.01)	(.007)	(.009)	(.008)	(.01)	(.006)
std(v_c)	0.035 ***	0.039 ***	0.039 ***	0.037 ***	0.014 **	0.012 *	0.013 *	0.013 ***
	(.005)	(.007)	(.006)	(.004)	(.007)	(.007)	(.008)	(.005)

Notes: *, **, *** indicate significance at 10%, 5%, and 1% levels respectively. Bootstrap standard errors are in parentheses. The dependent variables are high school graduation (HSGRAD) in Panel A, enrollment in a 4 year college within 2 years after expected high school graduation (ENROLL) in Panel B, attainment of a BA degree (COLLBA) in Panel C and the log hourly wage rate (ln(wage)) at about age 25 in Panel D. Panels A, B, and C refer to the latent index of an MME probit specification. Panel D is based on an MME regression specification. The model is equation (26). The column heading NELS refers to the NELS data. The neighborhood is ZIP code ELSbg represents the ELS data set with the neighborhood specification of block group. ELSz represents the ELS data set with the neighborhood specification of ZIP code. The parameter vectors B, G_1, G_2^S, and G_2^C that define the explanatory index variables X_iB, $X_{1n}G_1^N$, $Z_{2S}G_2^S$ and $Z_{2C}G_2^C$ are estimated in a first step using a nonlinear probit model. The model includes the "full" set of X_i variables (student level) and the corresponding "full" set of X_s variables (school means). See Appendix A1 for a list of X_i, X_n, X_s, and Z_{2c} variables. See Sections 6.2 and 6.3 for details about the estimation and bootstrap strap standard error procedures. The index variables in the interaction terms are standardized to be mean 0 and sd 1. Standard deviation of v_i is 1 in probit specifications. ELSbg represents the ELS data set with the neighborhood specification of block group.

Appendix Table B6. Education and Wage Model Parameters, No Interactions (Base Set of Student Variables)

	A. High School Graduation				**B. College Enrollment**			
	NELS	**ELSbg**	**ELSz**	**N+E**	**NELS**	**ELSbg**	**ELSz**	**N+E**
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
sd(X_iB)	0.505 ***	0.465 ***	0.460 ***	0.482 ***	0.641 ***	0.650 ***	0.654 ***	0.647 ***
	(.103)	(.021)	(.02)	(.052)	(.033)	(.02)	(.02)	(.019)
sd($X_{1n}G_1^N$)	0.000	0.000	0.072	0.036	0.000	0.053	0.000	0.000
	(.052)	(.037)	(.052)	(.037)	(.034)	(.035)	(.039)	(.026)
sd($X_{1s}G_1^S$)	0.166 ***	0.114 ***	0.110 ***	0.138 ***	0.173 ***	0.217 ***	0.244 ***	0.208 ***
	(.041)	(.035)	(.04)	(.028)	(.032)	(.039)	(.034)	(.023)
sd($Z_{2s}G_2^S$)	0.100 **	0.139 ***	0.153 ***	0.127 ***	0.100 ***	0.057	0.086 **	0.093 ***
	(.041)	(.04)	(.043)	(.03)	(.03)	(.043)	(.043)	(.026)
sd($Z_{2c}G_2^C$)	0.132 ***	0.127 ***	0.099 ***	0.116 ***	0.160 ***	0.131 ***	0.134 ***	0.147 ***
	(.035)	(.028)	(.038)	(.026)	(.028)	(.041)	(.038)	(.024)
(Intercept)	1.286 ***	1.696 ***	1.682 ***	1.484 ***	-0.573 ***	-0.411 ***	-0.410 ***	-0.491 ***
	(.034)	(.041)	(.043)	(.028)	(.036)	(.034)	(.034)	(.025)
RANDOM EFFECTS								
sd(v_n - v_s)	0.124 ***	0.185 ***	0.148 ***	0.136 ***	0.142 ***	0.162 ***	0.104 ***	0.123 ***
	(.016)	(.029)	(.019)	(.012)	(.022)	(.026)	(.014)	(.013)
sd(v_s - v_c)	0.091 ***	0.122 ***	0.106 ***	0.099 ***	0.131 ***	0.218 ***	0.210 ***	0.170 ***
	(.012)	(.016)	(.014)	(.009)	(.021)	(.036)	(.033)	(.02)
sd(v_c)	0.052 ***	0.048 ***	0.061 ***	0.057 ****	0.054 ***	0.090 ***	0.096 ***	0.075 ***
	(.009)	(.011)	(.01)	(.007)	(.01)	(.025)	(.028)	(.015)

	C. College Graduation				**D. Log Wage**			
	NELS	**ELSbg**	**ELSz**	**N+E**	**NELS**	**ELSbg**	**ELSz**	**N+E**
VARIABLES	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
sd(X_iB)	0.710 ***	0.586 ***	0.596 ***	0.653 ***	0.124 ***	0.130 ***	0.131 ***	0.128 ***
	(.028)	(.02)	(.02)	(.017)	(.007)	(.007)	(.006)	(.005)
sd($X_{1n}G_1^N$)	0.000	0.062 **	0.000	0.000	0.037 **	0.014	0.022	0.030 ***
	(.033)	(.03)	(.031)	(.023)	(.015)	(.011)	(.015)	(.011)
sd($X_{1s}G_1^S$)	0.171 ***	0.111 ***	0.125 ***	0.148 ***	0.052 ***	0.052 ***	0.044 ***	0.048 ***
	(.036)	(.031)	(.03)	(.024)	(.013)	(.009)	(.011)	(.008)
sd($Z_{2s}G_2^S$)	0.059 **	0.030	0.040	0.050 **	0.003	0.014	0.016	0.009
	(.029)	(.031)	(.031)	(.021)	(.012)	(.011)	(.011)	(.008)
sd($Z_{2c}G_2^C$)	0.095 ***	0.086 ***	0.071 ***	0.083 ***	0.033 ***	0.025 **	0.022 *	0.027 ***
	(.027)	(.024)	(.02)	(.017)	(.009)	(.01)	(.012)	(.007)
(Intercept)	### ***	### ***	### ***	### ***	2.544 ***	2.664 ***	2.664 ***	2.604 ***
	(.039)	(.031)	(.03)	(.025)	(.008)	(.009)	(.009)	(.006)
RANDOM EFFECTS								
sd(v_i)					0.321 ***	0.288 ***	0.290 ***	0.305 ***
					(.004)	(.007)	(.004)	(.003)
sd(v_n - v_s)	0.107 ***	0.131 ***	0.099 ***	0.103 ***	0.029 ***	0.046 ***	0.033 ***	0.031 ***
	(.012)	(.022)	(.012)	(.009)	(.003)	(.003)	(.003)	(.002)
sd(v_s - v_c)	0.068 ***	0.094 ***	0.082 ***	0.075 ***	0.020 ***	0.020 ***	0.020 ***	0.020 ***
	(.009)	(.015)	(.012)	(.008)	(.002)	(.002)	(.002)	(.002)
std(v_c)	0.029 ***	0.046 ***	0.042 ***	0.036 ***	0.014 *	0.014 ***	0.013 *	0.013 **
	(.004)	(.007)	(.008)	(.004)	(.008)	(.005)	(.007)	(.005)

Notes: *, **, *** indicate significance at 10%, 5%, and 1% levels respectively. Bootstrap standard errors are in parentheses. The dependent variables are high school graduation (HSGRAD) in Panel A, enrollment in a 4 year college within 2 years after expected high school graduation (ENROLL) in Panel B, attainment of a BA degree (COLLBA) in Panel C and the log hourly wage rate (ln(wage)) at about age 25 in Panel D. Panels A, B, and C refer to the latent index of an MME probit specification. Panel D is based on an MME regression specification. The model is equation (26). The column heading NELS refers to the NELS data. The neighborhood is ZIP code ELSbg represents the ELS data set with the neighborhood specification of block group. ELSz represents the ELS data set with the neighborhood specification of ZIP code. The parameter vectors B, G_1, G_2^S, and G_2^C that define the explanatory index variables X_iB, $X_{1n}G_1^N$, $Z_{2S}G_2^S$ and $Z_{2C}G_2^C$ are estimated in a first step using a nonlinear probit model. The model includes the "full" set of X_i variables (student level) and the corresponding "full" set of X_s variables (school means). See Appendix A1 for a list of X_i, X_n, X_s, and Z_{2c} variables. See Sections 6.2 and 6.3 for details about the estimation and bootstrap strap standard error procedures. The index variables in the interaction terms are standardized to be mean 0 and sd 1. Standard deviation of v_i is 1 in probit specifications. ELSbg represents the ELS data set with the neighborhood specification of block group.

Appendix Table B7: Sample Averages of the Treatment Effects on Education and Wages of a 10th to 50th Percentile Shift in School Quality and Commuting Zone Quality (Full Sample)

	A. HS Grad			B. College Enrollment			C. College Grad			D. Log Wage		
	NELS	**ELSbg**	**ELSz**	**NELS**	**ELSbg**	**ELSz**	**NELS**	**ELSbg**	**ELSz**	**NELS**	**ELSbg**	**ELSz**
Sch+CZ	0.049	0.036	0.035	0.080	0.085	0.090	0.042	0.044	0.041	0.055	0.052	0.051
	(.009)	(.007)	(.008)	(.008)	(.009)	(.009)	(.008)	(.006)	(.007)	(.013)	(.012)	(.013)
Sch Only	0.032	0.025	0.024	0.060	0.059	0.065	0.038	0.035	0.034	0.035	0.030	0.034
	(.008)	(.007)	(.007)	(.008)	(.009)	(.009)	(.008)	(.007)	(.008)	(.013)	(.012)	(.012)
CZ Only	0.037	0.023	0.020	0.054	0.055	0.059	0.023	0.036	0.029	0.051	0.039	0.038
	(.01)	(.006)	(.006)	(.008)	(.011)	(.01)	(.009)	(.008)	(.008)	(.012)	(.012)	(.013)

Note: The row Sch+Cz reports the average effect of moving students from a school and commuting zone at the 10th percentile of the distribution of the sum of school and commuting zone quality to the 90th percentile value. The row Sch Only (CZ Only) reports the average effect of a shift from the 10th percentile of school (commuting zone) quality to the 50th. The estimates are based on the model estimates in Table 4, which are for the full set of Xi variables. See Section 6.4 and online Appendix B3.3 for the details of the treatment effect calculations. Column heading indicate the outcome, the data set, and whether the neighborhood definition is block group or zip code. ELSbg represents the ELS data set with the neighborhood specification of block group. ELSz represents the ELS data set with the neighborhood specification of zip code. Bootstrap standard errors are in parentheses.

Appendix Table B8: Treatment Effects on Education and Wages of a 10th to 90th Percentile Shift in School Quality and Commuting Zone Quality for Students at the 10th, 50th, and 90th Percentile of the XiB Distribution (Full Sample)

	A. HS Grad				B. College Enrollment				C. College Grad				D. Log Wage			
	i. NELS / Zip Code				i. NELS / Zip Code				i. NELS / Zip Code				i. NELS / Zip Code			
	Mean	**10**	**50**	**90**	**Mean**	**10**	**50**	**90**	**Mean**	**10**	**50**	**90**	**Mean**	**10**	**50**	**90**
Sch+CZ	0.088	0.170	0.081	0.021	0.171	0.062	0.208	0.218	0.085	0.031	0.110	0.098	0.111	0.113	0.111	0.109
	(.014)	(.026)	(.013)	(.004)	(.018)	(.01)	(.023)	(.021)	(.016)	(.007)	(.021)	(.019)	(.025)	(.026)	(.025)	(.024)
Sch Only	0.060	0.115	0.055	0.014	0.125	0.047	0.151	0.157	0.077	0.029	0.099	0.087	0.069	0.069	0.069	0.069
	(.014)	(.027)	(.014)	(.004)	(.017)	(.009)	(.021)	(.019)	(.018)	(.008)	(.023)	(.019)	(.027)	(.027)	(.027)	(.027)
CZ Only	0.067	0.129	0.062	0.016	0.114	0.041	0.136	0.145	0.046	0.015	0.058	0.056	0.102	0.104	0.102	0.100
	(.017)	(.032)	(.016)	(.005)	(.017)	(.008)	(.021)	(.021)	(.019)	(.007)	(.024)	(.024)	(.024)	(.025)	(.024)	(.023)
	ii. ELS / Block Group				ii. ELS / Block Group				ii. ELS / Block Group				ii. ELS / Block Group			
Sch+CZ	0.062	0.138	0.047	0.010	0.173	0.077	0.235	0.159	0.090	0.039	0.115	0.098	0.104	0.104	0.104	0.105
	(.011)	(.023)	(.009)	(.002)	(.019)	(.011)	(.026)	(.019)	(.012)	(.007)	(.016)	(.013)	(.024)	(.024)	(.024)	(.024)
Sch Only	0.045	0.101	0.035	0.007	0.120	0.054	0.163	0.110	0.071	0.032	0.091	0.075	0.060	0.060	0.060	0.060
	(.01)	(.023)	(.008)	(.002)	(.019)	(.011)	(.026)	(.017)	(.015)	(.009)	(.02)	(.014)	(.023)	(.023)	(.023)	(.023)
CZ Only	0.041	0.091	0.032	0.007	0.112	0.049	0.152	0.104	0.073	0.031	0.092	0.079	0.078	0.078	0.078	0.079
	(.009)	(.02)	(.008)	(.002)	(.022)	(.011)	(.029)	(.02)	(.017)	(.008)	(.021)	(.019)	(.024)	(.024)	(.024)	(.024)
	iii. ELS / Zip Code				iii. ELS / Zip Code				iii. ELS / Zip Code				iii. ELS / Zip Code			
Sch+CZ	0.060	0.134	0.045	0.009	0.183	0.082	0.249	0.167	0.083	0.035	0.107	0.089	0.103	0.103	0.103	0.103
	(.011)	(.023)	(.009)	(.002)	(.019)	(.011)	(.026)	(.018)	(.015)	(.009)	(.019)	(.015)	(.027)	(.027)	(.027)	(.027)
Sch Only	0.043	0.098	0.033	0.007	0.133	0.061	0.180	0.120	0.068	0.030	0.088	0.072	0.069	0.069	0.069	0.069
	(.011)	(.024)	(.009)	(.002)	(.019)	(.011)	(.026)	(.016)	(.016)	(.009)	(.021)	(.015)	(.024)	(.024)	(.024)	(.024)
CZ Only	0.037	0.082	0.028	0.006	0.118	0.053	0.161	0.110	0.060	0.025	0.077	0.065	0.075	0.075	0.075	0.075
	(.01)	(.022)	(.008)	(.002)	(.021)	(.01)	(.028)	(.02)	(.016)	(.007)	(.021)	(.017)	(.026)	(.026)	(.026)	(.026)

Note: The row Sch+Cz reports effect of moving students from a school and commuting zone at the 10th percentile of the distribution of the sum of school and commuting zone quality to the 90th percentile value. The row Sch Only (CZ Only) reports corresponding values of the treatment effect of a shift from the 10th percentile of school (commuting zone) quality to the 90th. The columns labeled "mean", "10th", "50th" and "90th" report (respectively) the average effect and the effects for students that the 10th, 50th, and 90th quantile of the distribution of XiB. The estimates are based on the model estimates in Table 4, which are for the full set of Xi variables. See Section 6.4 and online Appendix B3.3 for the details of the treatment effect calculations. Column headings indicate the outcome, the data set, and whether the neighborhood definition is block group or ZIP code. Bootstrap standard errors are in parentheses.

Appendix Table B9: Sample Averages of Treatment Effects of a 10th to 90th Percentile Shift in School Quality and Commuting Zone Quality by Population Subgroup (Basic Set of Student Characteristics)

	A. HS Grad			B. College Enrollment			C. College Grad			D. Log Wage		
	NELS	**ELSbg**	**ELSz**	**NELS**	**ELSbg**	**ELSz**	**NELS**	**ELSbg**	**ELSz**	**NELS**	**ELSbg**	**ELSz**
Sch+CZ	0.100	0.066	0.07	0.192	0.233	0.226	0.095	0.103	0.11	0.108	0.102	0.104
	(.014)	(.017)	(.02)	(.018)	(.037)	(.033)	(.016)	(.02)	(.02)	(.022)	(.026)	(.028)
Sch Only	0.070	0.052	0.06	0.151	0.192	0.182	0.085	0.077	0.09	0.063	0.053	0.068
	(.015)	(.011)	(.012)	(.018)	(.033)	(.032)	(.018)	(.023)	(.023)	(.025)	(.023)	(.025)
CZ Only	0.082	0.040	0.03	0.119	0.125	0.137	0.061	0.082	0.07	0.096	0.072	0.068
	(.017)	(.019)	(.021)	(.018)	(.04)	(.039)	(.017)	(.024)	(.024)	(.023)	(.025)	(.028)

Note: The row Sch+Cz reports the average effect of moving students from a school and commuting zone at the 10th percentile of the distribution of the sum of school and commuting zone quality to the 90th percentile value. The row Sch Only (CZ Only) reports the average effect of a shift from the 10th percentile of school (commuting zone) quality to the 90th. The estimates are based on the model estimates using the basic set of Xi variables. See Section 6.4.4 for the details of the treatment effect calculations. The panel headings indicate the outcome, the data set and whether the neighborhood definition is block group or zip code. ELSbg represents the ELS data set with the neighborhood specification of block group. ELSz represents the ELS data set with the neighborhood specification of zip code.

Appendix Table B10: The Treatment Effects of a 10th to 90th Percentile Shift in School Quality and Commuting Zone Quality for Students at the 10th, 50th, and 90th Percentile of the XiB Distribution (Basic Set of Student Characteristics)

	A. HS Grad				**B. College Enrollment**				**C. College Grad**				**D. Log Wage**			
	i. NELS / Zip Code				**i. NELS / Zip Code**				**i. NELS / Zip Code**				**i. NELS / Zip Code**			
	Mean	**10**	**50**	**90**	**Mean**	**10**	**50**	**90**	**Mean**	**10**	**50**	**90**	**Mean**	**10**	**50**	**90**
Sch+CZ	0.100	0.173	0.097	0.034	0.192	0.109	0.217	0.236	0.095	0.046	0.112	0.112	0.108	0.109	0.107	0.106
	(.014)	(.022)	(.014)	(.007)	(.018)	(.014)	(.021)	(.021)	(.016)	(.01)	(.02)	(.019)	(.022)	(.023)	(.022)	(.022)
Sch Only	0.070	0.122	0.067	0.023	0.151	0.087	0.170	0.182	0.085	0.043	0.101	0.099	0.063	0.063	0.063	0.063
	(.015)	(.024)	(.015)	(.007)	(.018)	(.013)	(.021)	(.02)	(.018)	(.011)	(.022)	(.019)	(.025)	(.025)	(.025)	(.025)
CZ Only	0.082	0.139	0.080	0.029	0.119	0.067	0.134	0.150	0.061	0.028	0.071	0.075	0.096	0.098	0.096	0.095
	(.017)	(.029)	(.017)	(.007)	(.018)	(.012)	(.021)	(.022)	(.017)	(.008)	(.02)	(.022)	(.023)	(.024)	(.023)	(.022)
	ii. ELS / Block Group				**ii. ELS / Block Group**				**ii. ELS / Block Group**				**ii. ELS / Block Group**			
Sch+CZ	0.07	0.130	0.057	0.018	0.233	0.161	0.273	0.237	0.103	0.068	0.119	0.114	0.102	0.102	0.102	0.101
	(.017)	(.033)	(.015)	(.005)	(.037)	(.03)	(.042)	(.036)	(.02)	(.015)	(.023)	(.02)	(.026)	(.027)	(.027)	(.026)
Sch Only	0.05	0.103	0.045	0.014	0.192	0.136	0.226	0.195	0.077	0.051	0.088	0.085	0.053	0.053	0.053	0.053
	(.011)	(.022)	(.01)	(.003)	(.033)	(.026)	(.038)	(.031)	(.023)	(.017)	(.026)	(.023)	(.023)	(.023)	(.023)	(.023)
CZ Only	0.04	0.077	0.034	0.011	0.125	0.085	0.146	0.130	0.082	0.054	0.094	0.090	0.072	0.072	0.072	0.071
	(.019)	(.036)	(.016)	(.006)	(.04)	(.03)	(.046)	(.042)	(.024)	(.017)	(.028)	(.025)	(.025)	(.025)	(.025)	(.025)
	iii. ELS / Zip Code				**iii. ELS / Zip Code**				**iii. ELS / Zip Code**				**iii. ELS / Zip Code**			
Sch+CZ	0.07	0.132	0.058	0.018	0.226	0.156	0.267	0.230	0.107	0.069	0.123	0.117	0.104	0.105	0.104	0.104
	(.02)	(.037)	(.018)	(.006)	(.033)	(.026)	(.038)	(.032)	(.02)	(.015)	(.023)	(.02)	(.028)	(.029)	(.028)	(.028)
Sch Only	0.06	0.111	0.048	0.015	0.182	0.128	0.215	0.183	0.087	0.056	0.101	0.097	0.068	0.068	0.068	0.068
	(.012)	(.024)	(.011)	(.004)	(.032)	(.026)	(.038)	(.029)	(.023)	(.017)	(.027)	(.023)	(.025)	(.025)	(.025)	(.025)
CZ Only	0.03	0.063	0.028	0.009	0.137	0.093	0.161	0.142	0.072	0.048	0.083	0.078	0.068	0.069	0.068	0.068
	(.021)	(.039)	(.019)	(.006)	(.039)	(.027)	(.045)	(.041)	(.024)	(.017)	(.028)	(.024)	(.028)	(.028)	(.028)	(.028)

Note: The row "Sch+Cz" reports the effect of moving students from a school and commuting zone at the 10th percentile of the distribution of the sum of school and commuting zone quality to the 90th percentile value. The columns labeled "mean", "10th", "50th" and "90th" report (respectively) the average effect and the effects for students at the 10th, 50th, and 90th quantile of the distribution of XiB. The row "Sch Only" ("CZ Only") reports corresponding values of the treatment effect of a shift from the 10th percentile of school (commuting zone) quality to the 90th. The estimates are based on model estimates for the basic set of Xi variables (not reported.) See Section 6.4.1-6.4.4 for the details of the treatment effect calculations. Column heading indicate the outcome, the data set, and whether the neighborhood definition is block group or zip code.

Appendix Table B11: The Treatment Effects on Education and Wages of a 10th to 90th Percentile Shift in School Quality and Commuting Zone Quality, by Population Subgroup (Basic Set of Student Characteristics)

	A. HS Grad					B. College Enrollment				
	i. NELS / Zip Code					i. NELS / Zip Code				
	white	black	Hispanic	sg mother, hs deg	both par wh, col deg	white	black	Hispanic	sg mother, hs deg	both par wh, col deg
Sch+CZ	0.099	0.106	0.127	0.154	0.052	0.199	0.188	0.156	0.143	0.231
	(.014)	(.016)	(.017)	(.02)	(.008)	(.019)	(.019)	(.017)	(.017)	(.02)
Sch Only	0.069	0.074	0.090	0.108	0.036	0.156	0.148	0.122	0.116	0.178
	(.014)	(.016)	(.02)	(.021)	(.008)	(.019)	(.019)	(.016)	(.016)	(.02)
CZ Only	0.081	0.086	0.102	0.122	0.043	0.123	0.117	0.096	0.086	0.145
	(.017)	(.019)	(.021)	(.027)	(.009)	(.019)	(.018)	(.017)	(.014)	(.021)
	ii. ELS / Block Group					ii. ELS / Block Group				
Sch+CZ	0.057	0.078	0.092	0.101	0.027	0.242	0.228	0.199	0.207	0.247
	(.015)	(.021)	(.024)	(.022)	(.007)	(.038)	(.037)	(.034)	(.038)	(.036)
Sch Only	0.046	0.061	0.073	0.080	0.021	0.200	0.188	0.165	0.172	0.202
	(.01)	(.014)	(.016)	(.015)	(.004)	(.034)	(.033)	(.03)	(.033)	(.031)
CZ Only	0.035	0.047	0.056	0.057	0.016	0.131	0.120	0.106	0.110	0.134
	(.017)	(.023)	(.026)	(.024)	(.007)	(.042)	(.04)	(.035)	(.041)	(.041)
	iii. ELS / Zip Code					iii. ELS / Zip Code				
Sch+CZ	0.059	0.080	0.093	0.105	0.027	0.244	0.228	0.195	0.207	0.247
	(.018)	(.024)	(.027)	(.026)	(.008)	(.034)	(.033)	(.03)	(.033)	(.032)
Sch Only	0.049	0.066	0.079	0.087	0.023	0.200	0.190	0.163	0.171	0.202
	(.012)	(.014)	(.017)	(.016)	(.005)	(.033)	(.032)	(.029)	(.033)	(.028)
CZ Only	0.028	0.038	0.045	0.049	0.013	0.141	0.131	0.113	0.117	0.144
	(.019)	(.025)	(.028)	(.026)	(.008)	(.04)	(.037)	(.033)	(.039)	(.039)

	C. College Grad					D. Log Wage				
	i. NELS / Zip Code					i. NELS / Zip Code				
	white	black	Hispanic	sg mother, hs deg	both par wh, col deg	white	black	Hispanic	sg mother, hs deg	both par wh, col deg
Sch+CZ	0.100	0.083	0.073	0.063	0.111	0.107	0.109	0.108	0.108	0.107
	(.017)	(.016)	(.014)	(.012)	(.018)	(.022)	(.023)	(.022)	(.023)	(.022)
Sch Only	0.090	0.075	0.066	0.059	0.099	0.063	0.064	0.063	0.064	0.062
	(.019)	(.018)	(.016)	(.014)	(.019)	(.025)	(.025)	(.025)	(.025)	(.025)
CZ Only	0.064	0.053	0.046	0.039	0.074	0.096	0.098	0.096	0.098	0.095
	(.018)	(.016)	(.013)	(.011)	(.021)	(.023)	(.024)	(.023)	(.024)	(.023)
	ii. ELS / Block Group					ii. ELS / Block Group				
Sch+CZ	0.110	0.093	0.086	0.093	0.115	0.101	0.102	0.102	0.102	0.101
	(.021)	(.019)	(.017)	(.02)	(.019)	(.026)	(.026)	(.027)	(.027)	(.027)
Sch Only	0.082	0.068	0.064	0.070	0.084	0.053	0.054	0.052	0.051	0.054
	(.024)	(.021)	(.021)	(.023)	(.022)	(.023)	(.022)	(.022)	(.022)	(.023)
CZ Only	0.087	0.074	0.067	0.075	0.090	0.072	0.073	0.072	0.073	0.072
	(.025)	(.022)	(.02)	(.024)	(.024)	(.025)	(.025)	(.025)	(.025)	(.025)
	iii. ELS / Zip Code					iii. ELS / Zip Code				
Sch+CZ	0.114	0.096	0.088	0.095	0.118	0.104	0.104	0.104	0.104	0.104
	(.021)	(.018)	(.017)	(.02)	(.019)	(.028)	(.028)	(.028)	(.029)	(.028)
Sch Only	0.093	0.078	0.071	0.077	0.097	0.068	0.068	0.069	0.067	0.069
	(.024)	(.021)	(.021)	(.023)	(.022)	(.025)	(.025)	(.025)	(.025)	(.025)
CZ Only	0.077	0.066	0.059	0.065	0.079	0.068	0.068	0.069	0.067	0.068
	(.025)	(.022)	(.02)	(.025)	(.023)	(.028)	(.028)	(.028)	(.028)	(.028)

Note: The row Sch+CZ reports the average effect of moving students from a school and commuting zone at the 10th percentile of the distribution of the sum of school and commuting zone quality to the 90th percentile value. The row Sch Only (CZ Only) reports the average effect of a shift from the 10th percentile of school (commuting zone) quality to the 90th. The estimates are based on the model estimates in online Appendix Table B5, which are for the full set of X_i variables. See Section 6.4 and online Appendix B3.3 for the details of the treatment effect calculations. The panel headings indicate the outcome, the data set and whether the neighborhood definition is block group or zip code. The column heading identify the subgroup. "white" are white non-Hispanic students. "black" and "Hispanic" are non-Hispanic black and Hispanic students. "sg mother, hs deg" are students with a single mother who has a high school degree or less. "both par wh, col deg" are white students with two resident parents with college degrees. Bootstrap standard errors are in parentheses.