Online Appendix for
Community Colleges and Upward Mobility

Jack Mountjoy*

A Proof of Binary 2SLS Decomposition

This appendix section derives the binary two-stage least squares (2SLS) decomposition in equation (2) of the main text, showing that binary 2SLS estimates a weighted average of local average treatment effects along the 2←0 (2-year entry vs. no college) and 2←4 (2-year entry vs. 4-year entry) complier margins.[1] Recall the 2SLS specification:

$$
\begin{aligned}
Y &= \beta_0 + \beta_2 D_2 + \epsilon \\
D_2 &= \alpha_0 + \alpha_2 Z_2 + \eta,
\end{aligned}
$$

where Y is a student outcome, D_2 is an indicator for 2-year college entry, and Z_2 is an exogenous and excludable binary instrument that induces students into 2-year entry from the alternative treatments of no college (D_0) and 4-year entry (D_4). In this system, β_2 is the familiar Wald (1940) estimand:

$$
\beta_2 = \frac{E[Y|Z_2 = 1] - E[Y|Z_2 = 0]}{E[D_2|Z_2 = 1] - E[D_2|Z_2 = 0]}.
$$

Decompose $E[Y|Z_2 = 1]$ in the numerator using the fact that $Y = Y_0 D_0 + Y_2 D_2 + Y_4 D_4$, where Y_j is the potential outcome associated with treatment $j \in \{0, 2, 4\}$:

$$
\begin{aligned}
E[Y|Z_2 = 1] &= E[Y_0 D_0 + Y_2 D_2 + Y_4 D_4 | Z_2 = 1] \\
&= E[Y_0 | D_0 = 1, Z_2 = 1] Pr(D_0 = 1 | Z_2 = 1) \\
&+ E[Y_2 | D_2 = 1, Z_2 = 1] Pr(D_2 = 1 | Z_2 = 1) \\
&+ E[Y_4 | D_4 = 1, Z_2 = 1] Pr(D_4 = 1 | Z_2 = 1).
\end{aligned}
$$

Letting $D(z_2) \in \{0, 2, 4\}$ denote the potential choice an individual would make if exogenously assigned to $Z_2 = z_2 \in \{0, 1\}$, by instrument independence and exclusion this becomes

$$
\begin{aligned}
E[Y|Z_2 = 1] &= E[Y_0 | D(1) = 0] Pr(D(1) = 0) \\
&+ E[Y_2 | D(1) = 2] Pr(D(1) = 2) \\
&+ E[Y_4 | D(1) = 4] Pr(D(1) = 4).
\end{aligned}
$$

*University of Chicago Booth School of Business and NBER (email: jack.mountjoy@chicagobooth.edu).

[1] Heckman and Urzua (2010), Kline and Walters (2016), and Hull (2018) provide related derivations, as do Angrist and Imbens (1995) for the case of ordered multivalued treatments.

The monotonicity assumption that Z_2 induces students into D_2 from D_0 and D_4 permits the following five complier types: $\{D(0) = 0, D(1) = 0\}$, $\{D(0) = 0, D(1) = 2\}$, $\{D(0) = 2, D(1) = 2\}$, $\{D(0) = 4, D(1) = 4\}$, and $\{D(0) = 4, D(1) = 2\}$. Hence we can further decompose:

$$\begin{aligned}
E[Y|Z_2 = 1] = & \, E[Y_0|D(0) = 0, D(1) = 0]Pr(D(0) = 0, D(1) = 0) \\
& + E[Y_2|D(0) = 0, D(1) = 2]Pr(D(0) = 0, D(1) = 2) \\
& + E[Y_2|D(0) = 2, D(1) = 2]Pr(D(0) = 2, D(1) = 2) \\
& + E[Y_2|D(0) = 4, D(1) = 2]Pr(D(0) = 4, D(1) = 2) \\
& + E[Y_4|D(0) = 4, D(1) = 4]Pr(D(0) = 4, D(1) = 4).
\end{aligned}$$

These permitted complier types also decompose $E[Y|Z_2 = 0]$ into

$$\begin{aligned}
E[Y|Z_2 = 0] &= E[Y_0D_0|Z_2 = 0] + E[Y_2D_2|Z_2 = 0] + E[Y_4D_4|Z_2 = 0] \\
&= E[Y_0|D_0 = 1, Z_2 = 0]Pr(D_0 = 1|Z_2 = 0) \\
&+ E[Y_2|D_2 = 1, Z_2 = 0]Pr(D_2 = 1|Z_2 = 0) \\
&+ E[Y_4|D_4 = 1, Z_2 = 0]Pr(D_4 = 1|Z_2 = 0) \\
&= E[Y_0|D(0) = 0, D(1) = 0]Pr(D(0) = 0, D(1) = 0) \\
&+ E[Y_0|D(0) = 0, D(1) = 2]Pr(D(0) = 0, D(1) = 2) \\
&+ E[Y_2|D(0) = 2, D(1) = 2]Pr(D(0) = 2, D(1) = 2) \\
&+ E[Y_4|D(0) = 4, D(1) = 4]Pr(D(0) = 4, D(1) = 4) \\
&+ E[Y_4|D(0) = 4, D(1) = 2]Pr(D(0) = 4, D(1) = 2).
\end{aligned}$$

Subtracting $E[Y|Z_2 = 1] - E[Y|Z_2 = 0]$ eliminates the always-taker and never-taker groups, leaving only the instrument compliers:

$$\begin{aligned}
E[Y|Z_2 = 1] - E[Y|Z_2 = 0] &= E[Y_2|D(0) = 0, D(1) = 2]Pr(D(0) = 0, D(1) = 2) \\
&- E[Y_0|D(0) = 0, D(1) = 2]Pr(D(0) = 0, D(1) = 2) \\
&+ E[Y_2|D(0) = 4, D(1) = 2]Pr(D(0) = 4, D(1) = 2) \\
&- E[Y_4|D(0) = 4, D(1) = 2]Pr(D(0) = 4, D(1) = 2) \\
&= E[Y_2 - Y_0|D(0) = 0, D(1) = 2]Pr(D(0) = 0, D(1) = 2) \\
&+ E[Y_2 - Y_4|D(0) = 4, D(1) = 2]Pr(D(0) = 4, D(1) = 2).
\end{aligned}$$

To identify the two complier probabilities $Pr(D(0) = 0, D(1) = 2)$ and $Pr(D(0) = 4, D(1) = 2)$, recall from above that independence and monotonicity of the instrument imply

$$\begin{aligned}
& Pr(D_0|Z_2 = 0) = Pr(D(0) = 0) = Pr(D(0) = 0, D(1) = 0) + Pr(D(0) = 0, D(1) = 2) \\
& Pr(D_0|Z_2 = 1) = Pr(D(1) = 0) = Pr(D(0) = 0, D(1) = 0) \\
\implies & Pr(D_0|Z_2 = 0) - Pr(D_0|Z_2 = 1) = Pr(D(0) = 0, D(1) = 2)
\end{aligned}$$

$$\begin{aligned}
& Pr(D_4|Z_2 = 0) = Pr(D(0) = 4) = Pr(D(0) = 4, D(1) = 4) + Pr(D(0) = 4, D(1) = 2) \\
& Pr(D_4|Z_2 = 1) = Pr(D(1) = 4) = Pr(D(0) = 4, D(1) = 4) \\
\implies & Pr(D_4|Z_2 = 0) - Pr(D_4|Z_2 = 1) = Pr(D(0) = 4, D(1) = 2).
\end{aligned}$$

This yields

$$\begin{aligned}
E[Y|Z_2 = 1] - E[Y|Z_2 = 0] = & \, E[Y_2 - Y_0|D(0) = 0, D(1) = 2](E[D_0|Z_2 = 0] - E[D_0|Z_2 = 1]) \\
& + E[Y_2 - Y_4|D(0) = 4, D(1) = 2](E[D_4|Z_2 = 0] - E[D_4|Z_2 = 1]),
\end{aligned}$$

and plugging this back into the Wald expression yields the result:

$$\begin{aligned}\beta_2 &= \frac{E[Y|Z_2=1]-E[Y|Z_2=0]}{E[D_2|Z_2=1]-E[D_2|Z_2=0]}\\ &= \frac{E[Y_2-Y_0|D(0)=0,D(1)=2](E[D_0|Z_2=0]-E[D_0|Z_2=1])}{E[D_2|Z_2=1]-E[D_2|Z_2=0]}\\ &+ \frac{E[Y_2-Y_4|D(0)=4,D(1)=2](E[D_4|Z_2=0]-E[D_4|Z_2=1])}{E[D_2|Z_2=1]-E[D_2|Z_2=0]}\\ &= \omega E[Y_2-Y_0|D(0)=0,D(1)=2]+(1-\omega)E[Y_2-Y_4|D(0)=4,D(1)=2]\\ &= \omega LATE_{2\leftarrow 0}+(1-\omega)LATE_{2\leftarrow 4},\end{aligned}$$

where the weights

$$\omega \equiv -\frac{E[D_0|Z_2=1]-E[D_0|Z_2=0]}{E[D_2|Z_2=1]-E[D_2|Z_2=0]}, \quad (1-\omega) = -\frac{E[D_4|Z_2=1]-E[D_4|Z_2=0]}{E[D_2|Z_2=1]-E[D_2|Z_2=0]}$$

result from the fact that $D_0+D_2+D_4=1$.

B What Does Multivariate 2SLS Identify?

This appendix section derives and decomposes the multivariate two-stage least squares (2SLS) estimands in equation (3) of the main text under Assumptions IE, UPM, and CC.[2] For efficient notation, write this specification as

$$\begin{aligned}Y &= \gamma+\beta_0 D_0+\beta_4 D_4+\epsilon\\ E[D_0|Z] &= \alpha_0^0+\alpha_0^2 Z_2+\alpha_0^4 Z_4\\ E[D_4|Z] &= \alpha_4^0+\alpha_4^2 Z_2+\alpha_4^4 Z_4.\end{aligned}$$

$D_2=1$ is the omitted treatment case in the outcome equation, making $-\beta_0$ comparable to $MTE_{2\leftarrow 0}$ and $-\beta_4$ comparable to $MTE_{2\leftarrow 4}$. Z_2 and Z_4 are continuous, and the entire specification is local to a given evaluation point (z_2, z_4) such that the linear first stages are arbitrarily close to exact for small partial shifts in Z_2 and Z_4. Plug these first stage conditional expectations into the reduced form:

$$\begin{aligned}E[Y|Z] &= \gamma+\beta_0(\alpha_0^0+\alpha_0^2 Z_2+\alpha_0^4 Z_4)+\beta_4(\alpha_4^0+\alpha_4^2 Z_2+\alpha_4^4 Z_4)+E[\epsilon|Z]\\ &= \underbrace{\gamma+\beta_0\alpha_0^0+\beta_4\alpha_4^0}_{\equiv\alpha_y^0}+\underbrace{(\beta_0\alpha_0^2+\beta_4\alpha_4^2)}_{\equiv\alpha_y^2}Z_2+\underbrace{(\beta_0\alpha_0^4+\beta_4\alpha_4^4)}_{\equiv\alpha_y^4}Z_4\\ &= \alpha_y^0+\alpha_y^2 Z_2+\alpha_y^4 Z_4\end{aligned}$$

where $E[\epsilon|Z]=0$ by Assumption IE. Note that

$$\begin{pmatrix}\alpha_y^2\\ \alpha_y^4\end{pmatrix} = \begin{pmatrix}\alpha_0^2 & \alpha_4^2\\ \alpha_0^4 & \alpha_4^4\end{pmatrix} \times \begin{pmatrix}\beta_0\\ \beta_4\end{pmatrix},$$

[2] See Kirkeboen, Leuven and Mogstad (2016) for a related derivation involving discrete instruments, a less restrictive monotonicity condition, and no comparable compliers assumption, which yields more complicated estimands due to additional margins of instrument compliance. See also Kline and Walters (2016) and Hull (2018) for related derivations involving one binary instrument interacted with a stratifying covariate.

so we can solve for β_0 and β_4 as

$$\begin{pmatrix}\beta_0\\ \beta_4\end{pmatrix} = \begin{pmatrix}\alpha_0^2 & \alpha_4^2\\ \alpha_0^4 & \alpha_4^4\end{pmatrix}^{-1} \times \begin{pmatrix}\alpha_y^2\\ \alpha_y^4\end{pmatrix} = \frac{1}{\alpha_0^2\alpha_4^4 - \alpha_4^2\alpha_0^4}\begin{pmatrix}\alpha_4^4 & -\alpha_4^2\\ -\alpha_0^4 & \alpha_0^2\end{pmatrix} \times \begin{pmatrix}\alpha_y^2\\ \alpha_y^4\end{pmatrix}$$

$$\beta_0 = \frac{\alpha_4^4\alpha_y^2 - \alpha_4^2\alpha_y^4}{\alpha_0^2\alpha_4^4 - \alpha_4^2\alpha_0^4}, \quad \beta_4 = \frac{\alpha_0^2\alpha_y^4 - \alpha_0^4\alpha_y^2}{\alpha_0^2\alpha_4^4 - \alpha_4^2\alpha_0^4}.$$

Using the complier mean potential outcome identification results of Section IIID, we can decompose the reduced form w.r.t. Z_2 into

$$\begin{aligned}
\alpha_y^2 &= \frac{\partial E[Y|Z]}{\partial Z_2} = \frac{\partial E[YD_0|Z]}{\partial Z_2} + \frac{\partial E[YD_2|Z]}{\partial Z_2} + \frac{\partial E[YD_4|Z]}{\partial Z_2}\\
&= E[Y_0|2\text{–}0]\frac{\partial E[D_0|Z]}{\partial Z_2} + E[Y_2|2\text{–}0]\left(-\frac{\partial E[D_0|Z]}{\partial Z_2}\right) + E[Y_2|2\text{–}4]\left(-\frac{\partial E[D_4|Z]}{\partial Z_2}\right) + E[Y_4|2\text{–}4]\frac{\partial E[D_4|Z]}{\partial Z_2}\\
&= -\frac{\partial E[D_0|Z]}{\partial Z_2}\left(E[Y_2|2\text{–}0] - E[Y_0|2\text{–}0]\right) - \frac{\partial E[D_4|Z]}{\partial Z_2}\left(E[Y_2|2\text{–}4] - E[Y_4|2\text{–}4]\right)\\
&= -\alpha_0^2 MTE_{2\leftarrow 0} - \alpha_4^2 MTE_{2\leftarrow 4}
\end{aligned}$$

where $E[Y_0|2\text{–}0]$, for example, is shorthand for

$$\lim_{z_2'\uparrow z_2} E[Y_0|D(z_2', z_4) = 2, D(z_2, z_4) = 0] = E[Y_0|\text{Marginal 2–0 complier w.r.t. } Z_2 \text{ at } (z_2, z_4)],$$

and dependence on the local evaluation point (z_2, z_4) is suppressed in the notation of each element. Likewise with respect to Z_4, we have

$$\begin{aligned}
\alpha_y^4 &= \frac{\partial E[Y|Z]}{\partial Z_4} = \frac{\partial E[YD_0|Z]}{\partial Z_4} + \frac{\partial E[YD_2|Z]}{\partial Z_4} + \frac{\partial E[YD_4|Z]}{\partial Z_4}\\
&= E[Y_0|4\text{–}0]\frac{\partial E[D_0|Z]}{\partial Z_4} + E[Y_2|2\text{–}4 \text{ w.r.t.} Z_4]\frac{\partial E[D_2|Z]}{\partial Z_4}\\
&\quad + E[Y_4|2\text{–}4 \text{ w.r.t. } Z_4]\left(-\frac{\partial E[D_2|Z]}{\partial Z_4}\right) + E[Y_4|4\text{–}0]\left(-\frac{\partial E[D_0|Z]}{\partial Z_4}\right)
\end{aligned}$$

where $E[Y_0|4\text{–}0]$, for example, is shorthand for

$$\lim_{z_4'\uparrow z_4} E[Y_0|D(z_2, z_4') = 4, D(z_2, z_4) = 0] = E[Y_0|\text{Marginal 4–0 complier w.r.t. } Z_4 \text{ at } (z_2, z_4)].$$

By Assumption CC, we can equate $E[Y_2|2\text{–}4 \text{ w.r.t. } Z_4] = E[Y_2|2\text{–}4 \text{ w.r.t. } Z_2]$ at a given evaluation point and thus write this mean complier potential outcome in shorthand as $E[Y_2|2\text{–}4]$. Assumption CC in the main text is silent about the relationship between $E[Y_4|2\text{–}4 \text{ w.r.t. } Z_4]$ and $E[Y_4|2\text{–}4 \text{ w.r.t. } Z_2]$, however, since no restrictions are needed on these Y_4 potential outcomes to secure identification of the desired treatment effects in the separate identification approach of this paper. To simplify the 2SLS decomposition, however, let us make a slightly stronger comparable compliers assumption and equate these mean Y_4 potential outcomes across 2–4 compliers w.r.t. Z_2 vs. Z_4, as with Y_2. Hence we equate $E[Y_4|2\text{–}4 \text{ w.r.t. } Z_4] = E[Y_4|2\text{–}4 \text{ w.r.t. } Z_2] \equiv E[Y_4|2\text{–}4]$, which simplifies the expression for α_y^4 to

$$\begin{aligned}
\alpha_y^4 &= -\frac{\partial E[D_0|Z]}{\partial Z_4}\left(E[Y_4|4\text{–}0] - E[Y_0|4\text{–}0]\right) + \frac{\partial E[D_2|Z]}{\partial Z_4}\left(E[Y_2|2\text{–}4] - E[Y_4|2\text{–}4]\right)\\
&= -\alpha_0^4 MTE_{4\leftarrow 0} - (\alpha_0^4 + \alpha_4^4) MTE_{2\leftarrow 4},
\end{aligned}$$

again suppressing dependence on the local evaluation point (z_2, z_4) and using the fact that

$$\frac{\partial E[D_2|Z]}{\partial Z_4} = \frac{\partial E[1 - D_0 - D_4|Z]}{\partial Z_4} = -\frac{\partial E[D_0|Z]}{\partial Z_4} - \frac{\partial E[D_4|Z]}{\partial Z_4} = -\alpha_0^4 - \alpha_4^4.$$

Plugging these results into the expressions above for β_0 and β_4 yields:

$$\begin{aligned}
\beta_0 &= \frac{\alpha_4^4(-\alpha_0^2 MTE_{2\leftarrow 0} - \alpha_4^2 MTE_{2\leftarrow 4}) - \alpha_4^2(-\alpha_0^4 MTE_{4\leftarrow 0} - (\alpha_0^4 + \alpha_4^4) MTE_{2\leftarrow 4})}{\alpha_0^2\alpha_4^4 - \alpha_4^2\alpha_0^4} \\
&= -\frac{\alpha_0^2\alpha_4^4 MTE_{2\leftarrow 0} - \alpha_4^2\alpha_0^4(MTE_{4\leftarrow 0} + MTE_{2\leftarrow 4})}{\alpha_0^2\alpha_4^4 - \alpha_4^2\alpha_0^4}, \\
\beta_4 &= \frac{\alpha_0^2(-\alpha_0^4 MTE_{4\leftarrow 0} - (\alpha_0^4 + \alpha_4^4) MTE_{2\leftarrow 4}) - \alpha_0^4(-\alpha_0^2 MTE_{2\leftarrow 0} - \alpha_4^2 MTE_{2\leftarrow 4})}{\alpha_0^2\alpha_4^4 - \alpha_4^2\alpha_0^4} \\
&= -\frac{(\alpha_0^2\alpha_4^4 - \alpha_4^2\alpha_0^4 + \alpha_0^2\alpha_0^4) MTE_{2\leftarrow 4} + (-\alpha_0^2\alpha_0^4)(MTE_{2\leftarrow 0} - MTE_{4\leftarrow 0})}{\alpha_0^2\alpha_4^4 - \alpha_4^2\alpha_0^4}.
\end{aligned}$$

Finally, defining the weights

$$\theta_0 \equiv \frac{\alpha_0^2\alpha_4^4}{\alpha_0^2\alpha_4^4 - \alpha_4^2\alpha_0^4}, \quad \theta_4 \equiv \frac{\alpha_0^2\alpha_4^4 - \alpha_4^2\alpha_0^4 + \alpha_0^2\alpha_0^4}{\alpha_0^2\alpha_4^4 - \alpha_4^2\alpha_0^4}$$

yields the main result of this appendix section:

$$\begin{aligned}
-\beta_0 &= \theta_0 MTE_{2\leftarrow 0} + (1 - \theta_0)(MTE_{4\leftarrow 0} + MTE_{2\leftarrow 4}) \\
-\beta_4 &= \theta_4 MTE_{2\leftarrow 4} + (1 - \theta_4)(MTE_{2\leftarrow 0} - MTE_{4\leftarrow 0}).
\end{aligned}$$

Each local multivariate 2SLS estimand in this setting is thus a linear combination of the marginal treatment effect of interest and a biasing term involving effects for compliers along the other two treatment margins. In the special case of constant treatment effects across all individuals, note that $MTE_{4\leftarrow 0} + MTE_{2\leftarrow 4} = (Y_4 - Y_0) + (Y_2 - Y_4) = Y_2 - Y_0$ and $MTE_{2\leftarrow 0} - MTE_{4\leftarrow 0} = (Y_2 - Y_0) - (Y_4 - Y_0) = Y_2 - Y_4$, which confirms that 2SLS identifies the effects of interest in the absence of effect heterogeneity. With heterogeneous effects, however, $MTE_{4\leftarrow 0} + MTE_{2\leftarrow 4} \neq MTE_{2\leftarrow 0}$ and $MTE_{2\leftarrow 0} - MTE_{4\leftarrow 0} \neq MTE_{2\leftarrow 4}$ in general, since each of these treatment effects conditions on a different complier subpopulation. Each multivariate 2SLS estimand therefore does not generally recover a well-defined treatment effect for any well-defined complier population.

C Proofs for Equations (5), (6), (7), and (8)

This appendix section proves the mean potential outcome identification results in equations (5), (6), (7), and (8) of the main text. Consider a decrease in Z_2 from z_2 to z_2' while holding Z_4 fixed at z_4. By Assumption UPM, this induces 2←0 and 2←4 compliers. Changes in D_0 with respect to this shift therefore must be driven by 2←0 compliers:

$$\begin{aligned}
&Pr[D = 0|z_2', z_4] - Pr[D = 0|z_2, z_4] \\
=&Pr[D(z_2', z_4) = 0] - Pr[D(z_2, z_4) = 0] \\
=&Pr[D(z_2', z_4) = 0, D(z_2, z_4) = 0] - \big(Pr[D(z_2', z_4) = 0, D(z_2, z_4) = 0] + Pr[D(z_2', z_4) = 2, D(z_2, z_4) = 0]\big) \\
=& - Pr[D(z_2', z_4) = 2, D(z_2, z_4) = 0] \\
\equiv& - Pr[2\leftarrow 0 \text{ complier w.r.t. } (z_2', z_4) \leftarrow (z_2, z_4)]. \qquad (16)
\end{aligned}$$

The first equation is due to Assumption IE: conditioning on a given instrument value is as good as exogenously assigning that instrument value. The second equation is due to Assumption UPM: the group of individuals who choose $D(z_2, z_4) = 0$ includes 2←0 compliers, who would switch to $D = 2$ in response to the reduction in Z_2, as well as non-responders, who would continue to choose $D = 0$. Meanwhile, the group of individuals who choose $D(z_2', z_4) = 0$ can only include non-responders w.r.t. $(z_2', z_4) \leftarrow (z_2, z_4)$, since if $D = 2$ still is not attractive to them with a lower Z_2 cost, it would not have been more attractive at a higher Z_2 cost.

To prove (5), next note that

$$\begin{aligned}
E[YD_0|z_2, z_4] =& E[Y_0|D_0 = 1, z_2, z_4]Pr[D_0 = 1|z_2, z_4] \\
=& E[Y_0|D(z_2, z_4) = 0]Pr[D(z_2, z_4) = 0] \\
=& \big[E[Y_0|D(z_2', z_4) = 0, D(z_2, z_4) = 0]Pr[D(z_2', z_4) = 0|D(z_2, z_4) = 0] \\
&+ E[Y_0|D(z_2', z_4) = 2, D(z_2, z_4) = 0]Pr[D(z_2', z_4) = 2|D(z_2, z_4) = 0]\big]Pr[D(z_2, z_4) = 0] \\
=& E[Y_0|D(z_2', z_4) = 0, D(z_2, z_4) = 0]Pr[D(z_2', z_4) = 0, D(z_2, z_4) = 0] \\
&+ E[Y_0|D(z_2', z_4) = 2, D(z_2, z_4) = 0]Pr[D(z_2', z_4) = 2, D(z_2, z_4) = 0],
\end{aligned}$$

where the second equation is due to Assumption IE and the third equation again decomposes the mass of individuals with $D(z_2, z_4) = 0$ into the complier groups w.r.t. $(z_2', z_4) \leftarrow (z_2, z_4)$ allowed by Assumption UPM. Likewise, Assumptions IE and UPM imply

$$\begin{aligned}
E[YD_0|z_2', z_4] =& E[Y_0|D(z_2', z_4) = 0]Pr[D(z_2', z_4) = 0] \\
=& E[Y_0|D(z_2', z_4) = 0, D(z_2, z_4) = 0]Pr[D(z_2', z_4) = 0, D(z_2, z_4) = 0].
\end{aligned}$$

Hence

$$\begin{aligned}
& E[YD_0|z_2', z_4] - E[YD_0|z_2, z_4] \\
=& - E[Y_0|D(z_2', z_4) = 2, D(z_2, z_4) = 0]Pr[D(z_2', z_4) = 2, D(z_2, z_4) = 0].
\end{aligned}$$

Dividing by (16) above yields equation (5) in the main text:[3]

$$\begin{aligned}
\frac{E[YD_0|z_2', z_4] - E[YD_0|z_2, z_4]}{E[D_0|z_2', z_4] - E[D_0|z_2, z_4]} &= E[Y_0|D(z_2', z_4) = 2, D(z_2, z_4) = 0] \\
&\equiv E[Y_0|2\leftarrow 0 \text{ complier w.r.t. } (z_2', z_4) \leftarrow (z_2, z_4)].
\end{aligned}$$

[3] Instead of working with YD_0, one could alternatively work with selected outcomes; a rewriting of (5) yields

$$E[Y|D = 0, z_2', z_4] = E[Y_0|2\leftarrow 0 \text{ complier w.r.t. } (z_2', z_4) \leftarrow (z_2, z_4)] - \frac{E[Y|D = 0, z_2', z_4] - E[Y|D = 0, z_2, z_4]}{\big(E[D_0|z_2', z_4] - E[D_0|z_2, z_4]\big)/E[D_0|z_2, z_4]}.$$

In words, the mean selected outcome among the $D = 0$ treatment group at (z_2', z_4) is equal to the unselected complier potential outcome mean of interest adjusted by a selection term, which is proportional to the instrument-induced compositional change in the observed outcome within the selected treatment group. This formulation has the flavor of a nonparametric control function (e.g. Heckman and Robb, 1985; Blundell and Powell, 2003; Wooldridge, 2015; Brinch, Mogstad and Wiswall, 2017; Kline and Walters, 2019), and as such suggests a simple test of selection: if the instrument induces no compositional change in the mean selected outcome, i.e. if the selection term is zero, then the mean complier potential outcome of interest is identified directly from the conditional mean $E[Y|D = 0, z_2, z_4]$ with no selection adjustment. Otherwise, the sign of the selection term helps inform whether the average $D = 0$ treatment group member at (z_2', z_4) tends to be positively or negatively selected on their potential outcome level Y_0 relative to the 2←0 compliers of interest.

We can proceed analogously for D_4:

$$\begin{aligned}
E[YD_4|z_2,z_4] =& E[Y_4|D(z_2,z_4)=4]Pr[D(z_2,z_4)=4]\\
=&\big(E[Y_4|D(z_2',z_4)=4,D(z_2,z_4)=4]Pr[D(z_2',z_4)=4|D(z_2,z_4)=4]\\
&+E[Y_4|D(z_2',z_4)=2,D(z_2,z_4)=4]Pr[D(z_2',z_4)=2|D(z_2,z_4)=4]\big)Pr[D(z_2,z_4]=4)\\
=&E[Y_4|D(z_2',z_4)=4,D(z_2,z_4)=4]Pr[D(z_2',z_4)=4,D(z_2,z_4)=4]\\
&+E[Y_4|D(z_2',z_4)=2,D(z_2,z_4)=4]Pr[D(z_2',z_4)=2,D(z_2,z_4)=4]
\end{aligned}$$

$$\begin{aligned}
E[YD_4|z_2',z_4] =& E[Y_4|D(z_2',z_4)=4]Pr[D(z_2',z_4)=4]\\
=&E[Y_4|D(z_2',z_4)=4,D(z_2,z_4)=4]Pr[D(z_2',z_4)=4,D(z_2,z_4)=4]
\end{aligned}$$

$$\begin{aligned}
E[YD_4|z_2',z_4]-E[YD_4|z_2,z_4] =& -E[Y_4|D(z_2',z_4)=2,D(z_2,z_4)=4]Pr[D(z_2',z_4)=2,D(z_2,z_4)=4]\\
=&E[Y_4|D(z_2',z_4)=2,D(z_2,z_4)=4]\left(Pr[D(z_2',z_4)=4]-Pr[D(z_2,z_4)=4]\right),
\end{aligned}$$

which yields (6):

$$\begin{aligned}
\frac{E[YD_4|z_2',z_4]-E[YD_4|z_2,z_4]}{E[D_4|z_2',z_4]-E[D_4|z_2,z_4]} &= E[Y_4|D(z_2',z_4)=2,D(z_2,z_4)=4]\\
&\equiv E[Y_4|2\leftarrow4 \text{ complier w.r.t. } (z_2',z_4)\leftarrow(z_2,z_4)].
\end{aligned}$$

Turning to D_2,

$$\begin{aligned}
E[YD_2|z_2,z_4] =& E[Y_2|D(z_2,z_4)=2]Pr[D(z_2,z_4)=2]\\
=&E[Y_2|D(z_2',z_4)=2,D(z_2,z_4)=2]Pr[D(z_2',z_4)=2,D(z_2,z_4)=2]
\end{aligned}$$

$$\begin{aligned}
E[YD_2|z_2',z_4] =& E[Y_2|D(z_2',z_4)=2]Pr[D(z_2',z_4)=2]\\
=&E[Y_2|D(z_2',z_4)=2,D(z_2,z_4)=2]Pr[D(z_2',z_4)=2,D(z_2,z_4)=2]\\
&+E[Y_2|D(z_2',z_4)=2,D(z_2,z_4)=0]Pr[D(z_2',z_4)=2,D(z_2,z_4)=0]\\
&+E[Y_2|D(z_2',z_4)=2,D(z_2,z_4)=4]Pr[D(z_2',z_4)=2,D(z_2,z_4)=4],
\end{aligned}$$

which yields the pooled expression in (7):

$$\begin{aligned}
E[YD_2|z_2',z_4]-E[YD_2|z_2,z_4] =& E[Y_2|D(z_2',z_4)=2,D(z_2,z_4)=0]Pr[D(z_2',z_4)=2,D(z_2,z_4)=0]\\
&+E[Y_2|D(z_2',z_4)=2,D(z_2,z_4)=4]Pr[D(z_2',z_4)=2,D(z_2,z_4)=4].
\end{aligned}$$

Finally, we turn to Z_4. From the same initial evaluation point (z_2, z_4), consider an increase in Z_4 from z_4 to z_4', while holding Z_2 fixed at z_2. By Assumption UPM, this induces $2\leftarrow4$ and $0\leftarrow4$ compliers. Changes in D_2 with respect to this shift therefore must only involve $2\leftarrow4$ compliers:

$$\begin{aligned}
E[YD_2|z_2,z_4] =& E[Y_2|D(z_2,z_4)=2]Pr[D(z_2,z_4)=2]\\
=&E[Y_2|D(z_2,z_4')=2,D(z_2,z_4)=2]Pr[D(z_2,z_4')=2,D(z_2,z_4)=2]
\end{aligned}$$

$$\begin{aligned}
E[YD_2|z_2,z_4'] =& E[Y_2|D(z_2,z_4')=2]Pr[D(z_2,z_4')=2]\\
=&E[Y_2|D(z_2,z_4')=2,D(z_2,z_4)=2]Pr[D(z_2,z_4')=2,D(z_2,z_4)=2]\\
&+E[Y_2|D(z_2,z_4')=2,D(z_2,z_4)=4]Pr[D(z_2,z_4')=2,D(z_2,z_4)=4]
\end{aligned}$$

$$
\begin{aligned}
E[YD_2|z_2,z_4'] - E[YD_2|z_2,z_4] =& E[Y_2|D(z_2,z_4')=2, D(z_2,z_4)=4] Pr[D(z_2,z_4')=2, D(z_2,z_4)=4] \\
=& E[Y_2|D(z_2,z_4')=2, D(z_2,z_4)=4] \left(Pr[D(z_2,z_4')=2] - Pr[D(z_2,z_4)=2] \right),
\end{aligned}
$$

which yields (8):

$$
\begin{aligned}
\frac{E[YD_2|z_2,z_4'] - E[YD_2|z_2,z_4]}{E[D_2|z_2,z_4'] - E[D_2|z_2,z_4]} &= E[Y_2|D(z_2,z_4')=2, D(z_2,z_4)=4] \\
&\equiv E[Y_2|2\leftarrow 4 \text{ complier w.r.t. } (z_2,z_4') \leftarrow (z_2,z_4)].
\end{aligned}
$$

D The Index Model Is Sufficient for Assumptions UPM and CC

This appendix section shows that the index model in equation (4) of the main text satisfies the more general Assumptions UPM and CC as a special case. Recall the choice equations from (4):

$$
\begin{aligned}
D_0(z_2,z_4) &= \mathbb{1}[U_2 < \mu_2(z_2), U_4 < \mu_4(z_4)] \\
D_2(z_2,z_4) &= \mathbb{1}[U_2 > \mu_2(z_2), U_4 - U_2 < \mu_4(z_4) - \mu_2(z_2)] \\
D_4(z_2,z_4) &= \mathbb{1}[U_4 > \mu_4(z_4), U_4 - U_2 > \mu_4(z_4) - \mu_2(z_2)].
\end{aligned}
$$

D.1 Assumption UPM

To prove that the first part of Assumption UPM holds in this model, fix an arbitrary base point (z_2,z_4) and consider a decrease in Z_2 to $z_2' < z_2$ while holding Z_4 fixed at z_4. We must show that $D_0(z_2',z_4) \le D_0(z_2,z_4)$, $D_2(z_2',z_4) \ge D_2(z_2,z_4)$, and $D_4(z_2',z_4) \le D_4(z_2,z_4)$ for all individuals, with each inequality holding strictly for at least some individuals.

By (4), whether an individual would choose a given treatment at a given instrument value depends entirely on her values of (U_2,U_4). We can therefore completely characterize the set of individuals with $D_0(z_2,z_4)=1$ as $\mathcal{I}_0(z_2,z_4) = \{(U_2,U_4) : U_2 < \mu_2(z_2), U_4 < \mu_4(z_4)\}$, and the set of individuals with $D_0(z_2',z_4)=1$ as $\mathcal{I}_0(z_2',z_4) = \{(U_2,U_4) : U_2 < \mu_2(z_2'), U_4 < \mu_4(z_4)\}$. Since $\mu_2(\cdot)$ is strictly increasing, any individual satisfying $U_2 < \mu_2(z_2')$ also satisfies $U_2 < \mu_2(z_2)$, which implies $\mathcal{I}_0(z_2',z_4) \subset \mathcal{I}_0(z_2,z_4)$ and thus $D_0(z_2',z_4) \le D_0(z_2,z_4)$ for all individuals, with the inequality holding strictly for $2\leftarrow 0$ compliers with $\{(U_2,U_4) : U_2 \in (\mu_2(z_2'), \mu_2(z_2)), U_4 < \mu_4(z_4)\}$.

Those choosing $D_2(z_2,z_4)=1$ are $\mathcal{I}_2(z_2,z_4) = \{(U_2,U_4) : U_2 > \mu_2(z_2), U_4 - U_2 < \mu_4(z_4) - \mu_2(z_2)\}$. Likewise $\mathcal{I}_2(z_2',z_4) = \{(U_2,U_4) : U_2 > \mu_2(z_2'), U_4 - U_2 < \mu_4(z_4) - \mu_2(z_2')\}$. Since $\mu_2(\cdot)$ is strictly increasing, any individual satisfying $U_2 > \mu_2(z_2)$ also satisfies $U_2 > \mu_2(z_2')$, and any individual satisfying $U_4 - U_2 < \mu_4(z_4) - \mu_2(z_2)$ also satisfies $U_4 - U_2 < \mu_4(z_4) - \mu_2(z_2')$, which implies $\mathcal{I}_2(z_2,z_4) \subset \mathcal{I}_2(z_2',z_4)$ and thus $D_2(z_2',z_4) \ge D_2(z_2,z_4)$ for all individuals, with the inequality holding strictly for $2\leftarrow 0$ compliers with $\{(U_2,U_4) : U_2 \in (\mu_2(z_2'), \mu_2(z_2)), U_4 < \mu_4(z_4)\}$ and $2\leftarrow 4$ compliers with $\{(U_2,U_4) : U_4 - U_2 \in (\mu_4(z_4) - \mu_2(z_2), \mu_4(z_4) - \mu_2(z_2')), U_4 > \mu_4(z_4)\}$.

Those choosing $D_4(z_2,z_4)=1$ are $\mathcal{I}_4(z_2,z_4) = \{(U_2,U_4) : U_4 > \mu_4(z_4), U_4 - U_2 > \mu_4(z_4) - \mu_2(z_2)\}$. Likewise $\mathcal{I}_4(z_2',z_4) = \{(U_2,U_4) : U_4 > \mu_4(z_4), U_4 - U_2 > \mu_4(z_4) - \mu_2(z_2')\}$. Since $\mu_2(\cdot)$ is strictly increasing, any individual satisfying $U_4 - U_2 > \mu_4(z_4) - \mu_2(z_2')$ also satisfies $U_4 - U_2 > \mu_4(z_4) - \mu_2(z_2)$, which implies $\mathcal{I}_4(z_2',z_4) \subset \mathcal{I}_4(z_2,z_4)$ and thus $D_4(z_2',z_4) \le D_4(z_2,z_4)$ for all individuals, with the inequality holding strictly for $2\leftarrow 4$ compliers with $\{(U_2,U_4) : U_4 - U_2 \in (\mu_4(z_4) - \mu_2(z_2), \mu_4(z_4) - \mu_2(z_2')), U_4 > \mu_4(z_4)\}$.

To prove that the second part of Assumption UPM holds in this model, fix an arbitrary base point (z_2,z_4) and consider an increase in Z_4 to $z_4' > z_4$ (to match the direction of the visualized shift in Figure 4) while holding Z_2 fixed at z_2. We must show that $D_0(z_2,z_4') \ge D_0(z_2,z_4)$,

$D_2(z_2, z_4') \geq D_2(z_2, z_4)$, and $D_4(z_2, z_4') \leq D_4(z_2, z_4)$ for all individuals, with each inequality holding strictly for at least some individuals.

Those choosing $D_0(z_2, z_4') = 1$ are $\mathcal{I}_0(z_2, z_4') = \{(U_2, U_4) : U_2 < \mu_2(z_2), U_4 < \mu_4(z_4')\}$. Since $\mu_4(\cdot)$ is strictly increasing, any individual satisfying $U_4 < \mu_4(z_4)$ also satisfies $U_4 < \mu_4(z_4')$, which implies $\mathcal{I}_0(z_2, z_4) \subset \mathcal{I}_0(z_2, z_4')$ and thus $D_0(z_2, z_4') \geq D_0(z_2, z_4)$ for all individuals, with the inequality holding strictly for 0←4 compliers with $\{(U_2, U_4) : U_2 < \mu_2(z_2), U_4 \in (\mu_4(z_4), \mu_4(z_4'))\}$.

Those choosing $D_2(z_2, z_4') = 1$ are $\mathcal{I}_2(z_2, z_4') = \{(U_2, U_4) : U_2 > \mu_2(z_2), U_4 - U_2 < \mu_4(z_4') - \mu_2(z_2)\}$. Since $\mu_4(\cdot)$ is strictly increasing, any individual satisfying $U_4 - U_2 < \mu_4(z_4) - \mu_2(z_2)$ also satisfies $U_4 - U_2 < \mu_4(z_4') - \mu_2(z_2)$, which implies $\mathcal{I}_2(z_2, z_4) \subset \mathcal{I}_2(z_2, z_4')$ and thus $D_2(z_2, z_4') \geq D_2(z_2, z_4)$ for all individuals, with the inequality holding strictly for 2←4 compliers with $\{(U_2, U_4) : U_2 > \mu_2(z_2), U_4 - U_2 \in (\mu_4(z_4) - \mu_2(z_2), \mu_4(z_4') - \mu_2(z_2))\}$.

Those choosing $D_4(z_2, z_4') = 1$ are $\mathcal{I}_4(z_2, z_4') = \{(U_2, U_4) : U_4 > \mu_4(z_4'), U_4 - U_2 > \mu_4(z_4') - \mu_2(z_2)\}$. Since $\mu_4(\cdot)$ is strictly increasing, any individual satisfying $U_4 > \mu_4(z_4')$ also satisfies $U_4 > \mu_4(z_4)$, and any individual satisfying $U_4 - U_2 > \mu_4(z_4') - \mu_2(z_2)$ also satisfies $U_4 - U_2 > \mu_4(z_4) - \mu_2(z_2)$, which implies $\mathcal{I}_4(z_2, z_4') \subset \mathcal{I}_4(z_2, z_4)$ and thus $D_4(z_2, z_4') \leq D_4(z_2, z_4)$ for all individuals, with the inequality holding strictly for 0←4 compliers with $\{(U_2, U_4) : U_2 < \mu_2(z_2), U_4 \in (\mu_4(z_4), \mu_4(z_4'))\}$ and 2←4 compliers with $\{(U_2, U_4) : U_2 > \mu_2(z_2), U_4 - U_2 \in (\mu_4(z_4) - \mu_2(z_2), \mu_4(z_4') - \mu_2(z_2))\}$. This proves that the index model in (4) is sufficient for Assumption UPM.

D.2 Assumption CC

To prove that the index model is sufficient for Assumption CC, first consider the left side of Assumption CC, which in the index model translates to

$$\begin{aligned}\lim_{z_2' \uparrow z_2} E[Y_2 | D(z_2', z_4) = 2, D(z_2, z_4) = 4] &= \lim_{z_2' \uparrow z_2} E[Y_2 | U_4 - U_2 \in \left(\mu_4(z_4) - \mu_2(z_2), \mu_4(z_4) - \mu_2(z_2')\right), U_4 > \mu_4(z_4)] \\ &= E[Y_2 | U_4 - U_2 = \mu_4(z_4) - \mu_2(z_2), U_4 > \mu_4(z_4)], \end{aligned} \tag{17}$$

where exact indifference with $U_4 - U_2 = \mu_4(z_4) - \mu_2(z_2)$ is assumed to be decided in favor of $D(z_2, z_4) = 4$. Now consider the right side of Assumption CC, which in the index model translates to

$$\begin{aligned}\lim_{z_4' \downarrow z_4} E[Y_2 | D(z_2, z_4') = 2, D(z_2, z_4) = 4] &= \lim_{z_4' \downarrow z_4} E[Y_2 | U_4 - U_2 \in \left(\mu_4(z_4) - \mu_2(z_2), \mu_4(z_4') - \mu_2(z_2)\right), U_2 > \mu_2(z_2)] \\ &= E[Y_2 | U_4 - U_2 = \mu_4(z_4) - \mu_2(z_2), U_2 > \mu_2(z_2)]. \end{aligned} \tag{18}$$

The conditioning set in (17) can be written as $U_4 = U_2 + \mu_4(z_4) - \mu_2(z_2) > \mu_4(z_4)$, which implies $U_2 > \mu_2(z_2)$. Hence a verbose version of (17) is $E[Y_2 | U_4 - U_2 = \mu_4(z_4) - \mu_2(z_2), U_2 > \mu_2(z_2), U_4 > \mu_4(z_4)]$. The conditioning set in (18) can be written as $U_2 = U_4 - \mu_4(z_4) + \mu_2(z_2) > \mu_2(z_2)$, which implies $U_4 > \mu_4(z_4)$. Hence a verbose version of (18) is $E[Y_2 | U_4 - U_2 = \mu_4(z_4) - \mu_2(z_2), U_2 > \mu_2(z_2), U_4 > \mu_4(z_4)]$. Therefore

$$\begin{aligned}\lim_{z_2' \uparrow z_2} E[Y_2 | D(z_2', z_4) = 2, D(z_2, z_4) = 4] &= [Y_2 | U_4 - U_2 = \mu_4(z_4) - \mu_2(z_2), U_2 > \mu_2(z_2), U_4 > \mu_4(z_4)] \\ &= \lim_{z_4' \downarrow z_4} E[Y_2 | D(z_2, z_4') = 2, D(z_2, z_4) = 4], \end{aligned}$$

i.e. the index model is sufficient for Assumption CC.

E A Nonseparable Model Satisfying Assumptions UPM and CC

This appendix section shows that the converse of the previous appendix section does not hold: the index model in equation (4) of the main text is not necessary for Assumptions UPM and CC. I consider a nonseparable model that strictly nests (4) and no longer satisfies the two-dimensional visualization in Figure 4, but still satisfies Assumptions UPM and CC.

The key generalization from the separable index model in (4) will be to allow for unobservable individual-level heterogeneity in the cost (instrument response) functions, i.e. nonseparability. As a useful preamble, however, note that (4) can accommodate such heterogeneity if it affects both $\mu_2(\cdot)$ and $\mu_4(\cdot)$ with equal sign and magnitude. To see this, consider the following variation on the index model,

$$\begin{aligned} I_0 &= 0 \\ I_2 &= U_2 - V\mu_2(Z_2) \\ I_4 &= U_4 - V\mu_4(Z_4), \end{aligned} \tag{19}$$

where $V > 0$ is a random variable that varies unobservably across individuals. This implies the choice equations

$$\begin{aligned} D_0(z_2, z_4) &= \mathbb{1}[U_2 < V\mu_2(z_2), U_4 < V\mu_4(z_4)] \\ D_2(z_2, z_4) &= \mathbb{1}[U_2 > V\mu_2(z_2), U_4 - U_2 < V\mu_4(z_4) - V\mu_2(z_2)] \\ D_4(z_2, z_4) &= \mathbb{1}[U_4 > V\mu_4(z_4), U_4 - U_2 > V\mu_4(z_4) - V\mu_2(z_2)]. \end{aligned}$$

By dividing through by $V > 0$, we can see that this model is isomorphic to a separable index model given by

$$\begin{aligned} \tilde{I}_0 &= 0 \\ \tilde{I}_2 &= \tilde{U}_2 - \mu_2(Z_2) \\ \tilde{I}_4 &= \tilde{U}_4 - \mu_4(Z_4), \end{aligned}$$

where each tilde'd variable is its original value in (19) divided by V. Since dividing by $V > 0$ preserves each individual's relative ranking of I_0, I_2, and I_4 for every given instrument value (z_2, z_4), the model in (19) is weakly separable, i.e. it can be renormalized as a separable model yielding identical choice behavior.

What is important in strictly generalizing from (4), then, is that the unobserved heterogeneity in instrument responses be differential across $\mu_2(\cdot)$ and $\mu_4(\cdot)$. Thus consider the following model,

$$\begin{aligned} I_0 &= 0 \\ I_2 &= U_2 - V_2\mu_2(Z_2) \\ I_4 &= U_4 - V_4\mu_4(Z_4), \end{aligned}$$

where $V_2 > 0$ and $V_4 > 0$ are random variables that vary unobservably across individuals. This model nests the weakly separable case above as $V_2 = V_4 = V$, and it nests (4) as $V_2 = V_4 = 1$. Unlike those models, however, this one does not generally admit a separable representation when $V_2 \neq V_4$. Dividing through by $V_4 > 0$, for example, still leaves the following nonseparable representation that we will work with for the remainder of this section, with some abuse of notation that redefines the quantites divided by V_4:

$$\begin{aligned} I_0 &= 0 \\ I_2 &= U_2 - V\mu_2(Z_2) \\ I_4 &= U_4 - \mu_4(Z_4). \end{aligned} \tag{20}$$

$V \equiv V_2/V_4 > 0$ thus captures individual-level heterogeneity in relative responsiveness to Z_2 vs. Z_4. That is, individuals with high values of V are relatively more sensitive to changes in Z_2 than changes in Z_4, compared to individuals with low values of V. This third dimension of choice heterogeneity is shut down by weakly separable models like (4) and (19): those models allow all individuals to respond differently to Z_2 relative to Z_4, since $\mu_2(\cdot)$ and $\mu_4(\cdot)$ can differ from each other, but this relative responsiveness must be homogeneous across individuals, since $\mu_2(\cdot)$ and $\mu_4(\cdot)$ do not differ across individuals. The nonseparable model in (20) thus strictly generalizes those separable models by allowing for such heterogeneity.

E.1 Assumption UPM

This subsection shows that the nonseparable model in (20) still satisfies Assumption UPM. As long as $V > 0$ for all individuals, the logic of the proof in Appendix D.1 still goes through, since all arguments using the fact that $\mu_2(\cdot)$ is strictly increasing for each individual still apply to $V\mu_2(\cdot)$. Thus we proceed with a proof with nearly identical structure to that in D.1. The choice equations implied by (20) are

$$\begin{aligned}
D_0(z_2, z_4) &= \mathbb{1}[U_2 < V\mu_2(z_2), U_4 < \mu_4(z_4)] \\
D_2(z_2, z_4) &= \mathbb{1}[U_2 > V\mu_2(z_2), U_4 - U_2 < \mu_4(z_4) - V\mu_2(z_2)] \\
D_4(z_2, z_4) &= \mathbb{1}[U_4 > \mu_4(z_4), U_4 - U_2 > \mu_4(z_4) - V\mu_2(z_2)].
\end{aligned}$$

To prove that the first part of Assumption UPM holds in this model, fix an arbitrary base point (z_2, z_4) and consider a decrease in Z_2 to $z_2' < z_2$ while holding Z_4 fixed at z_4. We must show that $D_0(z_2', z_4) \leq D_0(z_2, z_4)$, $D_2(z_2', z_4) \geq D_2(z_2, z_4)$, and $D_4(z_2', z_4) \leq D_4(z_2, z_4)$ for all individuals, with each inequality holding strictly for at least some individuals.

Whether an individual would choose a given treatment at a given instrument value depends entirely on her values of (U_2, U_4, V). We can therefore completely characterize the set of individuals with $D_0(z_2, z_4) = 1$ as $\mathcal{I}_0(z_2, z_4) = \{(U_2, U_4, V) :\ U_2 < V\mu_2(z_2), U_4 < \mu_4(z_4)\}$, and the set of individuals with $D_0(z_2', z_4) = 1$ as $\mathcal{I}_0(z_2', z_4) = \{(U_2, U_4, V) :\ U_2 < V\mu_2(z_2'), U_4 < \mu_4(z_4)\}$. Since $V\mu_2(\cdot)$ is strictly increasing, any individual satisfying $U_2 < V\mu_2(z_2')$ also satisfies $U_2 < V\mu_2(z_2)$, which implies $\mathcal{I}_0(z_2', z_4) \subset \mathcal{I}_0(z_2, z_4)$ and thus $D_0(z_2', z_4) \leq D_0(z_2, z_4)$ for all individuals, with the inequality holding strictly for $2 \leftarrow 0$ compliers with $\{(U_2, U_4, V) : U_2 \in (V\mu_2(z_2'), V\mu_2(z_2)), U_4 < \mu_4(z_4)\}$.

Those choosing $D_2(z_2, z_4) = 1$ are $\mathcal{I}_2(z_2, z_4) = \{(U_2, U_4, V) :\ U_2 > V\mu_2(z_2), U_4 - U_2 < \mu_4(z_4) - V\mu_2(z_2)\}$. Likewise $\mathcal{I}_2(z_2', z_4) = \{(U_2, U_4, V) :\ U_2 > V\mu_2(z_2'), U_4 - U_2 < \mu_4(z_4) - V\mu_2(z_2')\}$. Since $V\mu_2(\cdot)$ is strictly increasing, any individual satisfying $U_2 > V\mu_2(z_2)$ also satisfies $U_2 > V\mu_2(z_2')$, and any individual satisfying $U_4 - U_2 < \mu_4(z_4) - V\mu_2(z_2)$ also satisfies $U_4 - U_2 < \mu_4(z_4) - V\mu_2(z_2')$, which implies $\mathcal{I}_2(z_2, z_4) \subset \mathcal{I}_2(z_2', z_4)$ and thus $D_2(z_2', z_4) \geq D_2(z_2, z_4)$ for all individuals, with the inequality holding strictly for $2 \leftarrow 0$ compliers with $\{(U_2, U_4, V) : U_2 \in (V\mu_2(z_2'), V\mu_2(z_2)), U_4 < \mu_4(z_4)\}$ and $2 \leftarrow 4$ compliers with $\{(U_2, U_4, V) : U_4 - U_2 \in (\mu_4(z_4) - V\mu_2(z_2), \mu_4(z_4) - V\mu_2(z_2')), U_4 > \mu_4(z_4)\}$.

Those choosing $D_4(z_2, z_4) = 1$ are $\mathcal{I}_4(z_2, z_4) = \{(U_2, U_4, V) :\ U_4 > \mu_4(z_4), U_4 - U_2 > \mu_4(z_4) - V\mu_2(z_2)\}$. Likewise $\mathcal{I}_4(z_2', z_4) = \{(U_2, U_4, V) :\ U_4 > \mu_4(z_4), U_4 - U_2 > \mu_4(z_4) - V\mu_2(z_2')\}$. Since $V\mu_2(\cdot)$ is strictly increasing, any individual satisfying $U_4 - U_2 > \mu_4(z_4) - V\mu_2(z_2')$ also satisfies $U_4 - U_2 > \mu_4(z_4) - V\mu_2(z_2)$, which implies $\mathcal{I}_4(z_2', z_4) \subset \mathcal{I}_4(z_2, z_4)$ and thus $D_4(z_2', z_4) \leq D_4(z_2, z_4)$ for all individuals, with the inequality holding strictly for $2 \leftarrow 4$ compliers with $\{(U_2, U_4, V) : U_4 - U_2 \in (\mu_4(z_4) - V\mu_2(z_2), \mu_4(z_4) - V\mu_2(z_2')), U_4 > \mu_4(z_4)\}$.

To prove that the second part of Assumption UPM holds in this model, fix an arbitrary base point (z_2, z_4) and consider an increase in Z_4 to $z_4' > z_4$ (to match the direction of the visualized shift in Figure 4) while holding Z_2 fixed at z_2. We must show that $D_0(z_2, z_4') \geq D_0(z_2, z_4)$, $D_2(z_2, z_4') \geq D_2(z_2, z_4)$, and $D_4(z_2, z_4') \leq D_4(z_2, z_4)$ for all individuals, with each inequality holding strictly for at least some individuals.

Those choosing $D_0(z_2, z_4') = 1$ are $\mathcal{I}_0(z_2, z_4') = \{(U_2, U_4, V) : U_2 < V\mu_2(z_2), U_4 < \mu_4(z_4')\}$. Since $\mu_4(\cdot)$ is strictly increasing, any individual satisfying $U_4 < \mu_4(z_4)$ also satisfies $U_4 < \mu_4(z_4')$, which implies $\mathcal{I}_0(z_2, z_4) \subset \mathcal{I}_0(z_2, z_4')$ and thus $D_0(z_2, z_4') \geq D_0(z_2, z_4)$ for all individuals, with the inequality holding strictly for 0←4 compliers with $\{(U_2, U_4, V) : U_2 < V\mu_2(z_2), U_4 \in (\mu_4(z_4), \mu_4(z_4'))\}$.

Those choosing $D_2(z_2, z_4') = 1$ are $\mathcal{I}_2(z_2, z_4') = \{(U_2, U_4, V) : U_2 > V\mu_2(z_2), U_4 - U_2 < \mu_4(z_4') - V\mu_2(z_2)\}$. Since $\mu_4(\cdot)$ is strictly increasing, any individual satisfying $U_4 - U_2 < \mu_4(z_4) - V\mu_2(z_2)$ also satisfies $U_4 - U_2 < \mu_4(z_4') - V\mu_2(z_2)$, which implies $\mathcal{I}_2(z_2, z_4) \subset \mathcal{I}_2(z_2, z_4')$ and thus $D_2(z_2, z_4') \geq D_2(z_2, z_4)$ for all individuals, with the inequality holding strictly for 2←4 compliers with $\{(U_2, U_4, V) : U_2 > V\mu_2(z_2), U_4 - U_2 \in (\mu_4(z_4) - V\mu_2(z_2), \mu_4(z_4') - V\mu_2(z_2))\}$.

Those choosing $D_4(z_2, z_4') = 1$ are $\mathcal{I}_4(z_2, z_4') = \{(U_2, U_4, V) : U_4 > \mu_4(z_4'), U_4 - U_2 > \mu_4(z_4') - V\mu_2(z_2)\}$. Since $\mu_4(\cdot)$ is strictly increasing, any individual satisfying $U_4 > \mu_4(z_4')$ also satisfies $U_4 > \mu_4(z_4)$, and any individual satisfying $U_4 - U_2 > \mu_4(z_4') - V\mu_2(z_2)$ also satisfies $U_4 - U_2 > \mu_4(z_4) - V\mu_2(z_2)$, which implies $\mathcal{I}_4(z_2, z_4') \subset \mathcal{I}_4(z_2, z_4)$ and thus $D_4(z_2, z_4') \leq D_4(z_2, z_4)$ for all individuals, with the inequality holding strictly for 0←4 compliers with $\{(U_2, U_4, V) : U_2 < V\mu_2(z_2), U_4 \in (\mu_4(z_4), \mu_4(z_4'))\}$ and 2←4 compliers with $\{(U_2, U_4, V) : U_2 > V\mu_2(z_2), U_4 - U_2 \in (\mu_4(z_4) - V\mu_2(z_2), \mu_4(z_4') - V\mu_2(z_2))\}$. This proves that the nonseparable model in (20) satisfies Assumption UPM.

E.2 Assumption CC

To prove that the nonseparable model in (20) satisfies Assumption CC, first consider the left side of Assumption CC, which in the nonseparable model translates to

$$\begin{aligned}
&\lim_{z_2' \uparrow z_2} E[Y_2 | D(z_2', z_4) = 2, D(z_2, z_4) = 4] \\
=&\lim_{z_2' \uparrow z_2} E[Y_2 | U_4 - U_2 \in \left(\mu_4(z_4) - V\mu_2(z_2), \mu_4(z_4) - V\mu_2(z_2')\right), U_4 > \mu_4(z_4)] \\
=&E[Y_2 | U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_4 > \mu_4(z_4)], \qquad (21)
\end{aligned}$$

where exact indifference with $U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2)$ is assumed to be decided in favor of $D(z_2, z_4) = 4$. Now consider the right side of Assumption CC, which in the nonseparable model translates to

$$\begin{aligned}
&\lim_{z_4' \downarrow z_4} E[Y_2 | D(z_2, z_4') = 2, D(z_2, z_4) = 4] \\
=&\lim_{z_4' \downarrow z_4} E[Y_2 | U_4 - U_2 \in \left(\mu_4(z_4) - V\mu_2(z_2), \mu_4(z_4') - V\mu_2(z_2)\right), U_2 > V\mu_2(z_2)] \\
=&E[Y_2 | U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2)]. \qquad (22)
\end{aligned}$$

The conditioning set in (21) can be written as $U_4 = U_2 + \mu_4(z_4) - V\mu_2(z_2) > \mu_4(z_4)$, which implies $U_2 > V\mu_2(z_2)$. Hence a verbose version of (21) is $E[Y_2 | U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2), U_4 > \mu_4(z_4)]$. The conditioning set in (22) can be written as $U_2 = U_4 - \mu_4(z_4) + V\mu_2(z_2) > V\mu_2(z_2)$, which implies $U_4 > \mu_4(z_4)$. Hence a verbose version of (22) is $E[Y_2 | U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2), U_4 > \mu_4(z_4)]$. Therefore

$$\begin{aligned}
\lim_{z_2' \uparrow z_2} E[Y_2 | D(z_2', z_4) = 2, D(z_2, z_4) = 4] &= [Y_2 | U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2), U_4 > \mu_4(z_4)] \\
&= \lim_{z_4' \downarrow z_4} E[Y_2 | D(z_2, z_4') = 2, D(z_2, z_4) = 4],
\end{aligned}$$

i.e. the nonseparable model in (20) satisfies Assumption CC.

F A Nonseparable Model Satisfying UPM but Not CC

To show that Assumption CC does not quite come for free in the general framework of Section IIID of the main text, consider a generalization of the nonseparable model in (20) with the same structure,

$$
\begin{aligned}
I_0 &= 0 \\
I_2 &= U_2 - V\mu_2(Z_2) \\
I_4 &= U_4 - \mu_4(Z_4),
\end{aligned}
$$

but now allow the instrument response heterogeneity variable $V \geq 0$ to take on the value of zero for some positive mass of students in the population, with $0 < Pr[V = 0|U_2, U_4] < 1$ for all (U_2, U_4).

F.1 Assumption UPM

Such a model still satisfies Assumption UPM. To see this, note first that for the subpopulation of individuals with V strictly positive, the proof in E.1 goes through exactly, so they satisfy Assumption UPM. For the complementary subpopulation of individuals with $V = 0$, their implied choice equations are

$$
\begin{aligned}
D_0(z_2, z_4) &= D_0(z_4) = \mathbb{1}[U_2 < 0, U_4 < \mu_4(z_4)] \\
D_2(z_2, z_4) &= D_2(z_4) = \mathbb{1}[U_2 > 0, U_4 - U_2 < \mu_4(z_4)] \\
D_4(z_2, z_4) &= D_4(z_4) = \mathbb{1}[U_4 > \mu_4(z_4), U_4 - U_2 > \mu_4(z_4)].
\end{aligned}
$$

That is, their potential treatment functions do not depend on Z_2. This means they satisfy the weak inequalities in first part of Assumption UPM by holding with equality—$D_0(z_2', z_4) = D_0(z_2, z_4)$, $D_2(z_2', z_4) = D_2(z_2, z_4)$, and $D_4(z_2', z_4) = D_4(z_2, z_4)$—but they do not contribute any compliers to satisfy the requirement that each inequality hold strictly for at least some individuals. This does not lead to a violation of Assumption UPM, however, since other students with $V > 0$ are available to fulfill this role for all possible instrument shifts given $Pr[V > 0|U_2, U_4] > 0$ for all (U_2, U_4).

To prove that the second part of Assumption UPM holds for the $V = 0$ subpopulation, fix an arbitrary base point (z_2, z_4) and consider an increase in Z_4 to $z_4' > z_4$ while holding Z_2 fixed at z_2. We must show that $D_0(z_2, z_4') \geq D_0(z_2, z_4)$, $D_2(z_2, z_4') \geq D_2(z_2, z_4)$, and $D_4(z_2, z_4') \leq D_4(z_2, z_4)$ for all individuals with $V = 0$. For this subpopulation, V is fixed and does not enter the choice equations, so whether an individual would choose a given treatment at a given instrument value depends entirely on her values of (U_2, U_4).

Those choosing $D_0(z_2, z_4') = 1$ are $\mathcal{I}_0(z_2, z_4') = \{(U_2, U_4) : U_2 < 0, U_4 < \mu_4(z_4')\}$. Since $\mu_4(\cdot)$ is strictly increasing, any individual satisfying $U_4 < \mu_4(z_4)$ also satisfies $U_4 < \mu_4(z_4')$, which implies $\mathcal{I}_0(z_2, z_4) \subset \mathcal{I}_0(z_2, z_4')$ and thus $D_0(z_2, z_4') \geq D_0(z_2, z_4)$ for all individuals in the $V = 0$ subpopulation, with the inequality holding strictly for $0 \leftarrow 4$ compliers with $\{(U_2, U_4) : U_2 < 0, U_4 \in (\mu_4(z_4), \mu_4(z_4'))\}$.

Those choosing $D_2(z_2, z_4') = 1$ are $\mathcal{I}_2(z_2, z_4') = \{(U_2, U_4) : U_2 > 0, U_4 - U_2 < \mu_4(z_4')\}$. Since $\mu_4(\cdot)$ is strictly increasing, any individual satisfying $U_4 - U_2 < \mu_4(z_4)$ also satisfies $U_4 - U_2 < \mu_4(z_4')$, which implies $\mathcal{I}_2(z_2, z_4) \subset \mathcal{I}_2(z_2, z_4')$ and thus $D_2(z_2, z_4') \geq D_2(z_2, z_4)$ for all individuals in the $V = 0$ subpopulation, with the inequality holding strictly for $2 \leftarrow 4$ compliers with $\{(U_2, U_4) : U_2 > 0, U_4 - U_2 \in (\mu_4(z_4), \mu_4(z_4'))\}$.

Those choosing $D_4(z_2, z_4') = 1$ are $\mathcal{I}_4(z_2, z_4') = \{(U_2, U_4) : U_4 > \mu_4(z_4'), U_4 - U_2 > \mu_4(z_4')\}$. Since $\mu_4(\cdot)$ is strictly increasing, any individual satisfying $U_4 > \mu_4(z_4')$ also satisfies $U_4 > \mu_4(z_4)$, and any individual satisfying $U_4 - U_2 > \mu_4(z_4')$ also satisfies $U_4 - U_2 > \mu_4(z_4)$, which implies $\mathcal{I}_4(z_2, z_4') \subset \mathcal{I}_4(z_2, z_4)$ and thus $D_4(z_2, z_4') \leq D_4(z_2, z_4)$ for all individuals in the $V = 0$ subpopulation, with the inequality holding strictly for 0←4 compliers with $\{(U_2, U_4) : U_2 < 0, U_4 \in (\mu_4(z_4), \mu_4(z_4'))\}$ and 2←4 compliers with $\{(U_2, U_4) : U_2 > 0, U_4 - U_2 \in (\mu_4(z_4), \mu_4(z_4'))\}$. This proves that the $V = 0$ subpopulation satisfies Assumption UPM.

F.2 Assumption CC

This model does not satisfy Assumption CC, however. Since individuals with $V = 0$ are responsive to Z_4 but not Z_2, their existence can break the exact overlap of marginal 2←4 compliers w.r.t. Z_2 vs. Z_4 featured in the separable index model in (4) and the nonseparable model in (20). To see this, first note that the left side of Assumption CC in this model can only involve individuals with $V > 0$, since those with $V = 0$ are never compliers w.r.t. Z_2. Thus

$$\begin{aligned}
&\lim_{z_2' \uparrow z_2} E[Y_2 | D(z_2', z_4) = 2, D(z_2, z_4) = 4] \\
=& \lim_{z_2' \uparrow z_2} E[Y_2 | V > 0, U_4 - U_2 \in \left(\mu_4(z_4) - V\mu_2(z_2), \mu_4(z_4) - V\mu_2(z_2')\right), U_4 > \mu_4(z_4)] \\
=& E[Y_2 | V > 0, U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_4 > \mu_4(z_4)], \qquad (23)
\end{aligned}$$

where exact indifference with $U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2)$ is assumed to be decided in favor of $D(z_2, z_4) = 4$. Now consider the right side of Assumption CC, which does involve individuals with $V = 0$,

$$\begin{aligned}
&\lim_{z_4' \downarrow z_4} E[Y_2 | D(z_2, z_4') = 2, D(z_2, z_4) = 4] \\
=& \lim_{z_4' \downarrow z_4} E[Y_2 | U_4 - U_2 \in \left(\mu_4(z_4) - V\mu_2(z_2), \mu_4(z_4') - V\mu_2(z_2)\right), U_2 > V\mu_2(z_2)] \\
=& E[Y_2 | U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2)],
\end{aligned}$$

which we can decompose as a weighted average across the subpopulations with $V > 0$ vs. $V = 0$,

$$\begin{aligned}
&E[Y_2 | V > 0, U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2)] Pr[V > 0 | U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2)] \\
+& E[Y_2 | V = 0, U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2)] Pr[V = 0 | U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2)].
\end{aligned}$$

As in E.2, the conditioning set $V > 0, U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_4 > \mu_4(z_4)$ is equivalent to $V > 0, U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2)$. Putting these pieces together, the difference between the left side and right side of Assumption CC in this model is

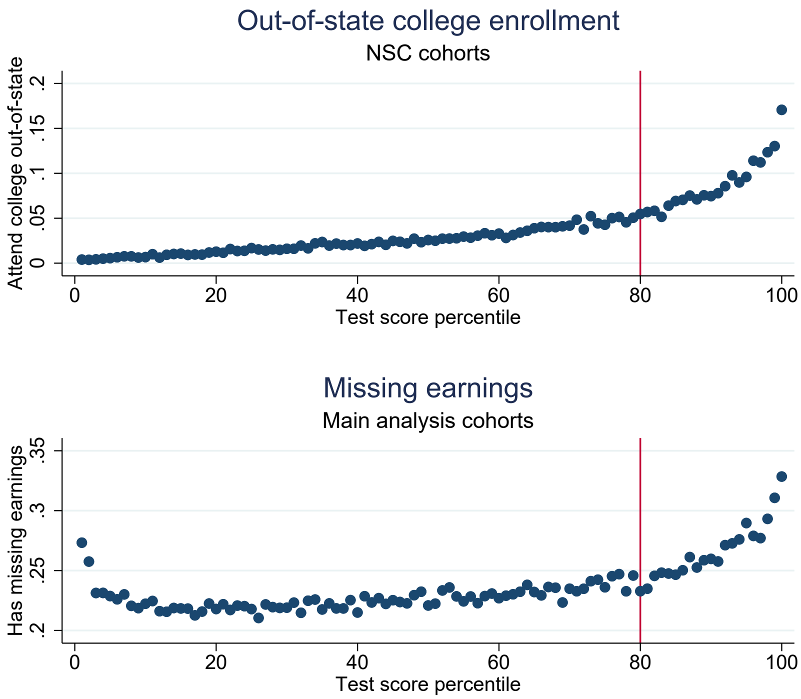
$$\begin{aligned}
&E[Y_2 | V > 0, U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_4 > \mu_4(z_4)] \\
-& E[Y_2 | V > 0, U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2)]\left(1 - Pr[V = 0 | U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2)]\right) \\
-& E[Y_2 | V = 0, U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2)] Pr[V = 0 | U_4 - U_2 = \mu_4(z_4) - V\mu_2(z_2), U_2 > V\mu_2(z_2)],
\end{aligned}$$

which simplifies to

$$\left\{E[Y_2 | V > 0, D(z_2, z_4') = 2, D(z_2, z_4) = 4] - E[Y_2 | V = 0, D(z_2, z_4') = 2, D(z_2, z_4) = 4]\right\} Pr[V = 0 | D(z_2, z_4') = 2, D(z_2, z_4) = 4].$$

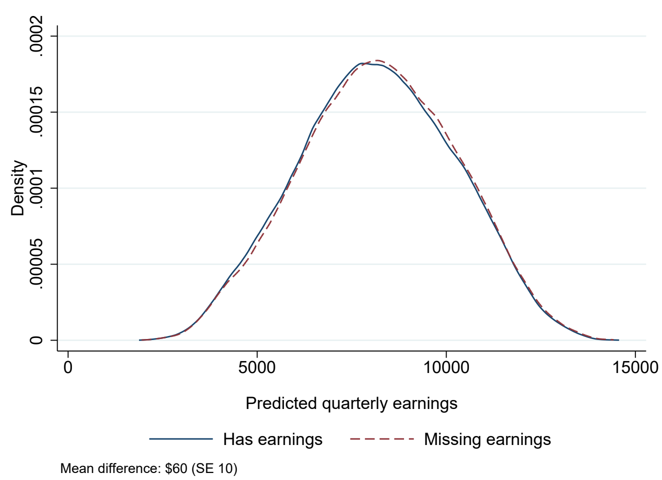
Thus, Assumption CC will fail to hold in this model if two conditions are both met. First, individuals with $V = 0$ must be among 2←4 compliers w.r.t. Z_4 at (z_2, z_4), i.e. $Pr[V = 0 | D(z_2, z_4') = 2, D(z_2, z_4) = 4] > 0$. Second, those compliers with $V = 0$ must differ in their mean Y_2 from their counterparts with $V > 0$, i.e. $E[Y_2 | V > 0, D(z_2, z_4') = 2, D(z_2, z_4) = 4] \neq E[Y_2 | V = 0, D(z_2, z_4') = 2, D(z_2, z_4) = 4]$.

Figure A.1: Out-of-State Enrollment and Missing Earnings among Top Scorers



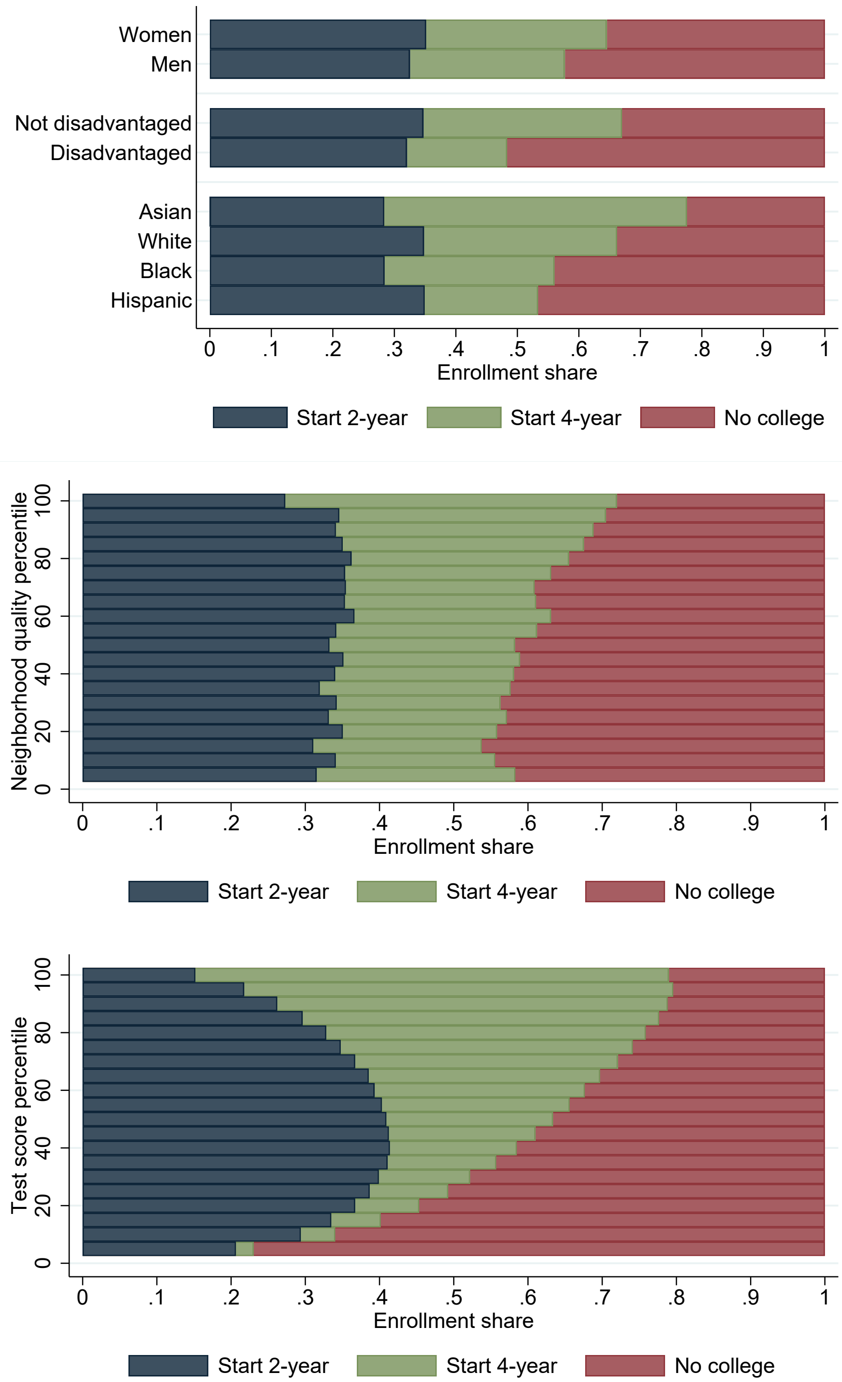
Notes: The top panel of this figure plots the share of students within each 10th grade test score percentile who enroll in college outside of Texas using the 2008-2009 cohorts with National Student Clearinghouse college enrollment coverage. The bottom panel plots the share of students within each test score percentile who have no Texas quarterly earnings records over ages 28-30 using the 2000-2004 main analysis cohorts.

Figure A.2: Predicted Earnings Are Similar for Students with Observed and Missing Earnings



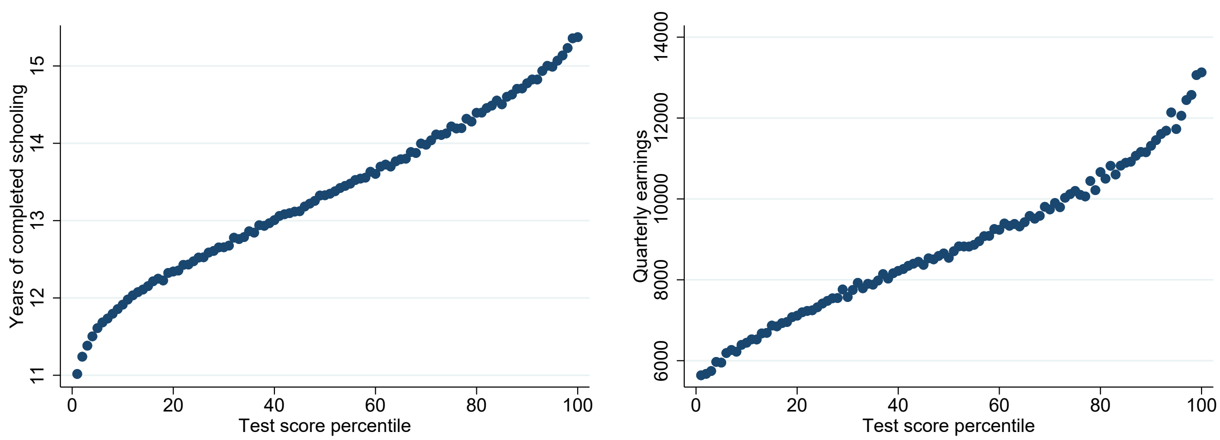
Notes: This figure plots the distributions of predicted mean quarterly earnings over ages 28-30 for students with and without observed earnings. Earnings are first projected on all covariates and instruments in Table 1 in the sample with valid earnings, then predicted in the full sample and plotted by earnings status.

Figure A.3: Sorting into College Enrollment Choices by Observables, 2000-2004 Analysis Cohorts



Notes: Disadvantaged is an indicator for free or reduced price lunch eligibility in 10th grade. Neighborhood quality and test score percentiles are grouped into 5-unit bins.

Figure A.4: Predictive Power of 10th Grade Test Scores on Long-Run Outcomes



Notes: Years of completed schooling are measured at age 28. Quarterly earnings are measured in real 2010 U.S. dollars and averaged within person over ages 28-30.

Table A.1: High School Graduation and Out-of-State College Enrollment

	Graduate from high school	Enroll in college out-of-state
Z_2: 2-year distance (miles/10)	0.0013	0.0001
	(0.0012)	(0.0006)
Z_4: 4-year distance (miles/10)	0.0002	-0.0011
	(0.0009)	(0.0007)
R^2	.018	.016
N	590,862	362,064
Sample	Main analysis cohorts	NSC cohorts
Baseline controls	✓	✓

Notes: NSC cohorts are those with National Student Clearinghouse college enrollment data. Standard errors in parentheses are clustered at the high school campus by cohort level. High school graduation is measured cumulatively through eight years after 10th grade. Out-of-state college enrollment is measured within two years of projected high school graduation due to NSC data availability.

Table A.2: Causal Effect Estimates: Time-to-Degree and Field of Study

	MTE_2 Net effect	=	ω Democratization share	$MTE_{2\leftarrow 0}$ Democratization effect	+	$(1-\omega)$ Diversion share	$MTE_{2\leftarrow 4}$ Diversion effect
BA completion within:							
4 years	0.046		0.657	0.041		0.343	0.055
	(0.018)		(0.049)	(0.012)		(0.049)	(0.051)
5 years	0.058		0.657	0.105		0.343	-0.033
	(0.032)		(0.049)	(0.020)		(0.049)	(0.076)
6 years	0.065		0.657	0.159		0.343	-0.116
	(0.037)		(0.049)	(0.025)		(0.049)	(0.078)
7 years	0.083		0.657	0.206		0.343	-0.154
	(0.041)		(0.049)	(0.030)		(0.049)	(0.080)
8 years	0.087		0.657	0.220		0.343	-0.167
	(0.045)		(0.049)	(0.032)		(0.049)	(0.081)
9 years	0.088		0.657	0.234		0.343	-0.191
	(0.045)		(0.049)	(0.034)		(0.049)	(0.081)
10 years	0.104		0.657	0.265		0.343	-0.204
	(0.048)		(0.049)	(0.036)		(0.049)	(0.082)
10-year BA completion in:							
STEM major	0.005		0.657	0.020		0.343	-0.024
	(0.012)		(0.049)	(0.010)		(0.049)	(0.027)
Non-STEM major	0.099		0.657	0.245		0.343	-0.180
	(0.042)		(0.049)	(0.034)		(0.049)	(0.076)

Notes: Locally weighted observations: 565,687. All estimates are evaluated at the mean values of the instruments. Standard errors in parentheses are block bootstrapped at the high school campus by cohort level. Complier shares are the same across outcomes due to common first stage equations.

References

Angrist, Joshua D., and Guido W. Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association*, 90(430): 431–442.

Blundell, Richard, and James Powell. 2003. "Endogeneity in Nonparametric and Semiparametric Regression Models." In *Advances in Economics and Econometrics: Theory and Applications.* , ed. Mathias Dewatripont, Lars Hansen and Stephen Turnovsky, 312–57. Cambridge:Cambridge University Press.

Brinch, Christian N., Magne Mogstad, and Matthew Wiswall. 2017. "Beyond LATE with a Discrete Instrument." *Journal of Political Economy*, 125(4): 985–1039.

Heckman, James J, and Richard Robb. 1985. "Alternative methods for evaluating the impact of interventions: An overview." *Journal of Econometrics*, 30(1): 239–267.

Heckman, James J., and Sergio Urzua. 2010. "Comparing IV with Structural Models: What Simple IV Can and Cannot Identify." *Journal of Econometrics*, 156(1): 27–37.

Hull, Peter. 2018. "IsoLATEing: Identifying Counterfactual-Specific Treatment Effects with Cross-Stratum Comparisons." Working paper.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad. 2016. "Field of Study, Earnings, and Self-Selection." *The Quarterly Journal of Economics*, 131(3): 1057–1111.

Kline, Patrick, and Christopher R. Walters. 2016. "Evaluating Public Programs with Close Substitutes: The Case of Head Start." *The Quarterly Journal of Economics*, 131(4): 1795–1848.

Kline, Patrick, and Christopher R. Walters. 2019. "On Heckits, LATE, and Numerical Equivalence." *Econometrica*, 87(2): 677–696.

Wald, Abraham. 1940. "The Fitting of Straight Lines if Both Variables are Subject to Error." *The Annals of Mathematical Statistics*, 11(3): 284–300.

Wooldridge, Jeffrey M. 2015. "Control Function Methods in Applied Econometrics." *Journal of Human Resources*, 50(2): 420–445.