Online Appendix

A Defining Neighborhood Choices

The neighborhood choices are defined according to the eligibility criteria of MTO vouchers:

- Low poverty neighborhood (t_l) are the neighborhoods whose poverty level is below 10% according to the 1990 U.S. Census.
- High poverty neighborhood (t_h) are the housing projects targeted by the MTO experiment.
- Medium poverty neighborhood (t_m) are the remaining neighborhoods.

Each choice refers to the neighborhood decision at the beginning of the intervention. Thus, each neighborhood choice indicates the initial family decision of neighborhood relocation but also eventual subsequent moves made by the family.

Families using the vouchers were supposed to move from housing projects within six months of the voucher assignment. However, this rule was not strictly enforced: 17% of the families that used the Section 8 voucher and 36% of families that used the experimental voucher took more than 6 months to move. Thus the neighborhood choice depends on the voucher utilization, the neighborhood poverty level and also on the time that the family took to relocate.

It is useful to classify the families into three groups: stayers, compliers, and self-movers. *Stayers* are families that had not moved from their original housing projects since the intervention onset until the time of the interim evaluation in 2002. *Compliers* are families that use the experimental or Section 8 vouchers to relocate. *Self-movers* are families that had moved at the time of the interim evaluation without using the voucher. Table A.1 presents the distribution of these family types across sites. Around 20% of families that receive vouchers and 30% of the control families stayed in their original dwellings by the time of the interim evaluation. Self-movers totals 36% of experimental families and 24% of Section 8 families in 2002.

Table A.1: Relocation Rates by Site at the Time of the Interim Evaluation in 2002

Voucher Assignment	All Sites N	All Sites %	Relocation Decision	All Sites N	All Sites %	Baltimore N	Baltimore %	Boston N	Boston %	Chicago N	Chicago %	Los Angeles N	Los Angeles %	New York N	New York %
Experimental	1729	41%	Compliers	818	47%	146	58%	168	46%	155	34%	167	67%	182	45%
			Self-movers	618	36%	97	38%	149	41%	234	51%	53	21%	85	21%
			Stayers	293	17%	9	4%	49	13%	71	15%	30	12%	134	33%
Section 8	1209	28%	Compliers	716	59%	135	72%	129	48%	134	66%	130	77%	188	49%
			Self-movers	276	23%	45	24%	86	32%	55	27%	25	15%	65	17%
			Stayers	217	18%	7	4%	52	19%	13	6%	13	8%	132	34%
Control	1310	31%	Self-movers	917	70%	174	88%	240	74%	189	81%	172	66%	142	48%
			Stayers	393	30%	23	12%	86	26%	43	19%	88	34%	153	52%
Total	4248														

This tables describe the relocation of families by voucher assignment and site in 2002. MTO families are classified into three groups: (1) compliers – families that used the vouchers to relocate; (2) self-movers – families that had moved without the voucher at the time of the interim evaluation in 2002; (3) stayers – families that had not moved since intervention onset in 1994–1998 until the interim evaluation in 2002.

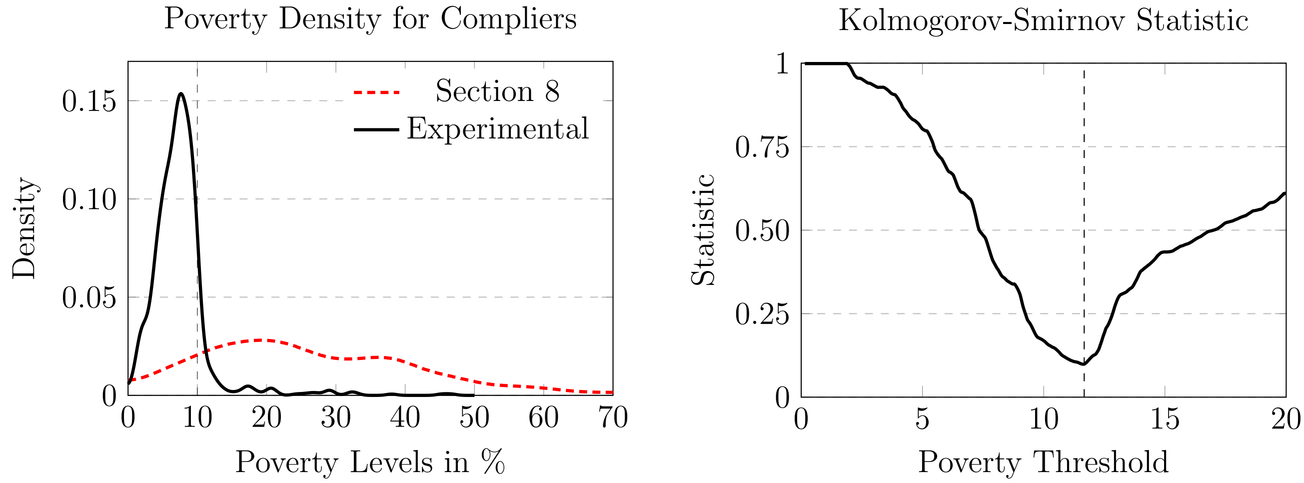
The neighborhood choices of stayers and compliers are easily characterized. The neighborhood choice of families who stay in their original dwellings is t_h. The experimental voucher can only be used to relocate to low poverty neighborhoods. Thus the neighborhood choice of experimental families that use the voucher is t_l. Families that decide to use the Section 8 voucher choose between low (t_l) or medium-poverty (t_m) neighborhoods. This ambiguity is resolved by assessing the poverty levels of the chosen neighborhoods.

The experimental voucher defines low-poverty neighborhoods as those whose poverty level is below a soft target of 10%.[46] In practice, 11% of neighborhoods classified as low-poverty were slightly above the nominal threshold (first graph of Figure 7). I employ a simple approach to address for this fact. I use the poverty distribution of Section 8 compliers to estimate a threshold that best conforms with the poverty distribution of low-poverty neighborhoods. Specifically, I estimate the threshold that minimizes the Kolmogorov-Smirnov statistic between the poverty distribution of Section 8 compliers and the poverty distribution of experimental compliers. The empirical threshold is 11.67% (second graph of Figure 7).

It remains to determine the neighborhood choice for the self-movers, which comprise all families that have relocated between surveys. The goal is to identify families who decided to move by the time of the onset of the intervention. To do so, I explore the available information on the time spell from voucher assignment until the first relocation. I account for this fact using the same procedure that yields the poverty threshold. I estimate the threshold the minimizes the difference on the distribution on relocation time between compliers and self-movers. The first graph of Figure 8 presents the distribution of relocation time for compliers while the second graph presents the Kolmogorov-Smirnov statistics for the difference on relocation time between compliers and self-movers. The corrected thresholds for relocation time are 8.6 months for medium-poverty neighborhoods and 10.6 months for low-poverty

[46] Using the 1990 US Census.

Figure 7: Poverty Densities and Threshold Investigation

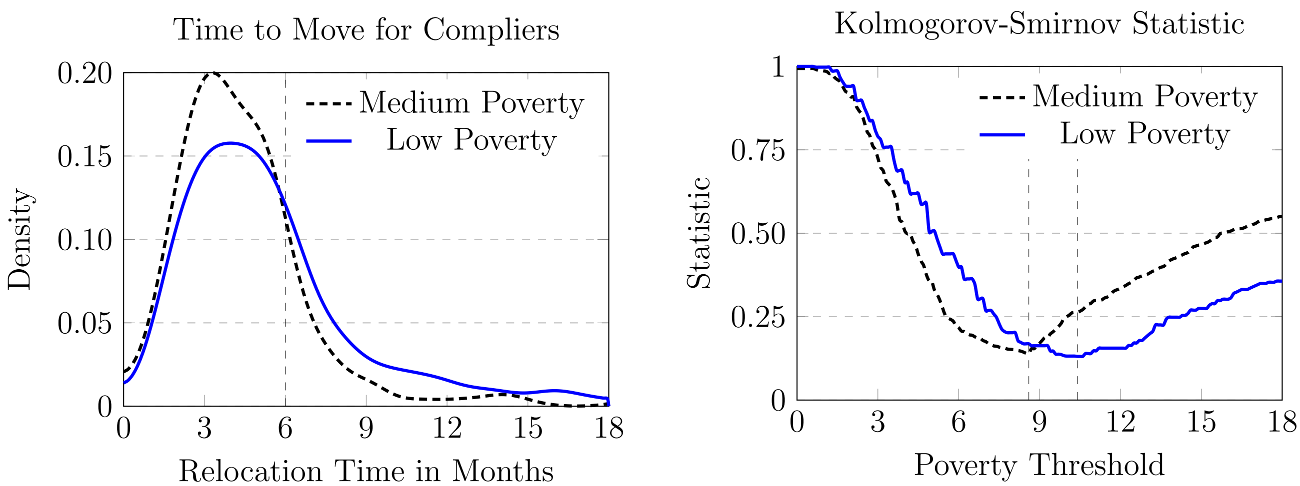


The first graph presents the poverty density of chosen neighborhoods for families who comply with the Experimental and Section 8 vouchers. The second graph presents the Kolmogorov-Smirnov statistics (y-axis) between the poverty distribution of experimental compliers and the poverty distribution of Section 8 compliers that is right-bounded by a threshold (x-axis).

neighborhoods. The neighborhood choice of self-movers that relocate before these thresholds is set at either low or medium-poverty neighborhoods.

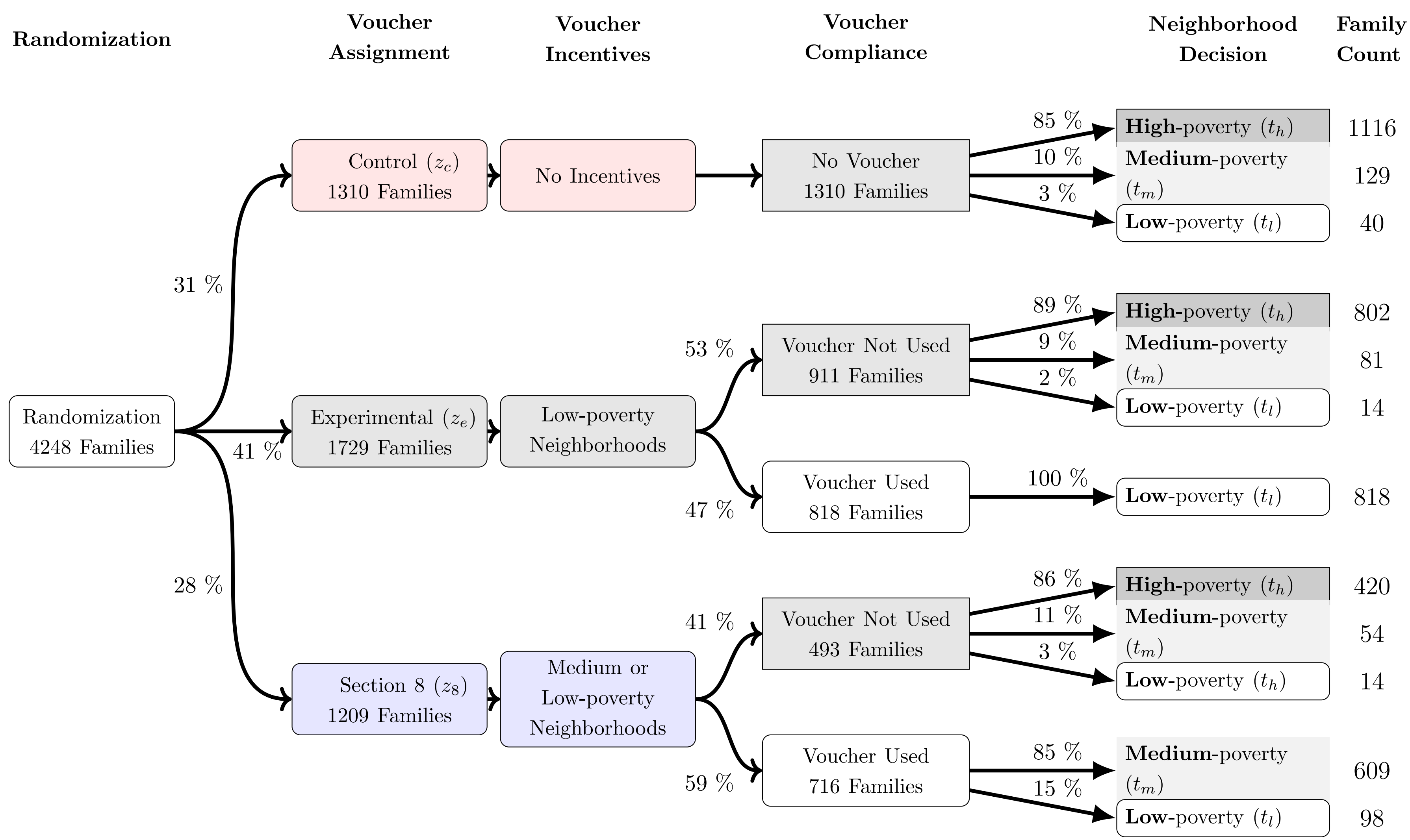
Figure 9 summarizes the neighborhood decision of the MTO families by voucher assignment. Nearly 85% the control families choose high-poverty neighborhoods, 10% choose medium poverty neighborhoods and 3% choose low-poverty neighborhoods. Families that do not use the voucher share a similar composition of neighborhood choices. Around 15% of families that use the Section 8 voucher decide for low-poverty neighborhoods while 85% of Section 8 compliers choose medium-poverty neighborhoods. Neighborhood choices are robust across variations of the assignment procedure. For instance, we can generate alternative values for the neighborhood choices by setting the poverty threshold to its nominal value of 10% and the relocation time to 6 months generates. These values agree with the neighborhood choices described by the procedure above in 97% of the cases.

Figure 8: Time to Relocate Densities and Threshold Investigation



The first graph presents the density of the time to relocate into low and medium poverty neighborhoods since voucher assignment for families who comply with the vouchers. Density estimates use the Gaussian Kernel with optimal bandwidth. The second graph presents the Kolmogorov-Smirnov statistics (y-axis) between the distribution time to relocate of voucher compliers and the distribution of time to relocate for self-movers that are right-bounded by a threshold (x-axis).

Figure 9: **Neighborhood Relocation by Voucher Assignment and Compliance**



This figure describes the possible decision patterns of families in MTO that resulted from voucher assignment and family compliance. Low-poverty (t_l) neighborhoods are defined as those whose share of poor residents is below 10% according to the 1990 sensus (Orr et al., 2003). High poverty (t_h) neighborhoods are the housing projects originally targeted by the intervention. Medium poverty (t_m) neighborhoods are neither the high poverty nor the ones classified as low poverty. Families who stay in their baseline housing live in high poverty neighborhoods (t_h). Families who use the experimental voucher (z_e) relocate to low poverty neighborhoods (t_l). Families who use Section 8 voucher (z_8) can decide between low (t_l) or medium (t_m) poverty neighborhoods. Control families (z_c) and families that do not use the vouchers may choose freely among all three neighborhoods: high (t_h), medium (t_m) or low (t_l).

B Generating MTO Choice Restrictions

We seek to investigate the choice restrictions generated by the design of the MTO intervention when assuming that the Weak Axiom of Revealed Preference (WARP) hold. Choice restrictions are generated by applying the choice rule (55) to the MTO incentive matrix (56).

$$\text{If} \quad T_i(z) = t \quad \text{and} \quad \boldsymbol{L}[z', t'] - \boldsymbol{L}[z, t'] \leq 0 \leq \boldsymbol{L}[z', t] - \boldsymbol{L}[z, t] \quad \text{then} \quad T_i(z') \neq t'. \tag{55}$$

$$\boldsymbol{L} = \begin{array}{c} \\ \\ \\ \end{array} \begin{array}{ccc} t_h & t_m & t_l \end{array} \\ \boldsymbol{L} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{array}{c} z_c \\ z_8 \\ z_e \end{array} \tag{56}$$

MTO incentive matrix combined with WARP generates 20 choice restrictions as listed in Table A.2. The 20 choice restrictions in Table A.2 can be summarized into the eight choice restrictions described in Table A.3.

The two last choice restrictions of Table (A.3) are redundant given the first six restrictions. Normal choice assumption generates the choice restriction $T_i(z_c) \neq t_h \Rightarrow T_i(z_8) = T_i(z_c)$, which totalizes the seven choice restrictions presented in the main text of the paper.

Table A.2: Choice Restrictions Due to WARP

#	Revealed Choice	Incentive Inequalities					Choice Statement
	$T_i(z) = t$	$\boldsymbol{L}[z', t'] - \boldsymbol{L}[z, t'] \leq 0 \leq \boldsymbol{L}[z', t] - \boldsymbol{L}[z, t]$					$T(z') \neq t'$
1	$T_i(z_c) = t_h,$	$\boldsymbol{L}[z_e, t_m] - \boldsymbol{L}[z_c, t_m] = 0$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_e, t_h] - \boldsymbol{L}[z_c, t_h]$	$T_i(z_e) \neq t_m$
2	$T_i(z_c) = t_m,$	$\boldsymbol{L}[z_8, t_h] - \boldsymbol{L}[z_c, t_h] = 0$	$\leq$	0	$\leq$	$1 = \boldsymbol{L}[z_8, t_m] - \boldsymbol{L}[z_c, t_m]$	$T_i(z_8) \neq t_h$
3	$T_i(z_c) = t_m,$	$\boldsymbol{L}[z_e, t_h] - \boldsymbol{L}[z_c, t_h] = 0$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_e, t_m] - \boldsymbol{L}[z_c, t_m]$	$T_i(z_e) \neq t_h$
4	$T_i(z_c) = t_l,$	$\boldsymbol{L}[z_8, t_h] - \boldsymbol{L}[z_c, t_h] = 0$	$\leq$	0	$\leq$	$1 = \boldsymbol{L}[z_8, t_l] - \boldsymbol{L}[z_c, t_l]$	$T_i(z_8) \neq t_h$
5	$T_i(z_c) = t_l,$	$\boldsymbol{L}[z_e, t_h] - \boldsymbol{L}[z_c, t_h] = 0$	$\leq$	0	$\leq$	$1 = \boldsymbol{L}[z_e, t_l] - \boldsymbol{L}[z_c, t_l]$	$T_i(z_e) \neq t_h$
6	$T_i(z_c) = t_l,$	$\boldsymbol{L}[z_e, t_m] - \boldsymbol{L}[z_c, t_m] = 0$	$\leq$	0	$\leq$	$1 = \boldsymbol{L}[z_e, t_l] - \boldsymbol{L}[z_c, t_l]$	$T_i(z_e) \neq t_m$
7	$T_i(z_8) = t_h,$	$\boldsymbol{L}[z_c, t_m] - \boldsymbol{L}[z_8, t_m] = -1$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_c, t_h] - \boldsymbol{L}[z_8, t_h]$	$T_i(z_c) \neq t_m$
8	$T_i(z_8) = t_h,$	$\boldsymbol{L}[z_e, t_m] - \boldsymbol{L}[z_8, t_m] = -1$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_e, t_h] - \boldsymbol{L}[z_8, t_h]$	$T_i(z_e) \neq t_m$
9	$T_i(z_8) = t_h,$	$\boldsymbol{L}[z_c, t_l] - \boldsymbol{L}[z_8, t_l] = -1$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_c, t_h] - \boldsymbol{L}[z_8, t_h]$	$T_i(z_c) \neq t_l$
10	$T_i(z_8) = t_h,$	$\boldsymbol{L}[z_e, t_l] - \boldsymbol{L}[z_8, t_l] = 0$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_e, t_h] - \boldsymbol{L}[z_8, t_h]$	$T_i(z_e) \neq t_l$
11	$T_i(z_8) = t_l,$	$\boldsymbol{L}[z_e, t_h] - \boldsymbol{L}[z_8, t_h] = 0$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_e, t_l] - \boldsymbol{L}[z_8, t_l]$	$T_i(z_e) \neq t_h$
12	$T_i(z_8) = t_l,$	$\boldsymbol{L}[z_e, t_m] - \boldsymbol{L}[z_8, t_m] = -1$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_e, t_l] - \boldsymbol{L}[z_8, t_l]$	$T_i(z_e) \neq t_m$
13	$T_i(z_e) = t_h,$	$\boldsymbol{L}[z_c, t_m] - \boldsymbol{L}[z_e, t_m] = 0$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_c, t_h] - \boldsymbol{L}[z_e, t_h]$	$T_i(z_c) \neq t_m$
14	$T_i(z_e) = t_h,$	$\boldsymbol{L}[z_c, t_l] - \boldsymbol{L}[z_e, t_l] = -1$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_c, t_h] - \boldsymbol{L}[z_e, t_h]$	$T_i(z_c) \neq t_l$
15	$T_i(z_e) = t_h,$	$\boldsymbol{L}[z_8, t_l] - \boldsymbol{L}[z_e, t_l] = 0$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_8, t_h] - \boldsymbol{L}[z_e, t_h]$	$T_i(z_8) \neq t_l$
16	$T_i(z_e) = t_m,$	$\boldsymbol{L}[z_c, t_h] - \boldsymbol{L}[z_e, t_h] = 0$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_c, t_m] - \boldsymbol{L}[z_e, t_m]$	$T_i(z_c) \neq t_h$
17	$T_i(z_e) = t_m,$	$\boldsymbol{L}[z_8, t_h] - \boldsymbol{L}[z_e, t_h] = 0$	$\leq$	0	$\leq$	$1 = \boldsymbol{L}[z_8, t_m] - \boldsymbol{L}[z_e, t_m]$	$T_i(z_8) \neq t_h$
18	$T_i(z_e) = t_m,$	$\boldsymbol{L}[z_c, t_l] - \boldsymbol{L}[z_e, t_l] = -1$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_c, t_m] - \boldsymbol{L}[z_e, t_m]$	$T_i(z_c) \neq t_l$
19	$T_i(z_e) = t_m,$	$\boldsymbol{L}[z_8, t_l] - \boldsymbol{L}[z_e, t_l] = 0$	$\leq$	0	$\leq$	$1 = \boldsymbol{L}[z_8, t_m] - \boldsymbol{L}[z_e, t_m]$	$T_i(z_8) \neq t_l$
20	$T_i(z_e) = t_l,$	$\boldsymbol{L}[z_8, t_h] - \boldsymbol{L}[z_e, t_h] = 0$	$\leq$	0	$\leq$	$0 = \boldsymbol{L}[z_8, t_l] - \boldsymbol{L}[z_e, t_l]$	$T_i(z_8) \neq t_h$

Table A.3: Summary of Choice Restrictions generated by applying WARP to the MTO Incentive Matrix

#	Choice Restrictions		
4,5,6	$T_i(z_c) = t_l$	$\Rightarrow$	$T_i(z_e) = t_l$ and $T_i(z_8) \neq t_h$
2,3	$T_i(z_c) = t_m$	$\Rightarrow$	$T_i(z_e) \neq t_h$ and $T_i(z_8) \neq t_h$
16,17,18,19	$T_i(z_e) = t_m$	$\Rightarrow$	$T_i(z_c) = t_m$ and $T_i(z_8) = t_m$
13,14,15	$T_i(z_e) = t_h$	$\Rightarrow$	$T_i(z_c) = t_h$ and $T_i(z_8) \neq t_l$
7,8,9,10	$T_i(z_8) = t_h$	$\Rightarrow$	$T_i(z_c) = t_h$ and $T_i(z_e) = t_h$
11,12	$T_i(z_8) = t_l$	$\Rightarrow$	$T_i(z_e) = t_l$
1	$T_i(z_c) = t_l$	$\Rightarrow$	$T_i(z_e) \neq t_m$
20	$T_i(z_e) = t_l$	$\Rightarrow$	$T_i(z_8) \neq t_h$

C MTO Incentives do not Justify Ordered Choices

A natural inquiry is whether it is possible to model neighborhood choices in MTO as an ordered choice model. Viewing the treatment as ordered is appealing because it relates to the well-known monotonicity condition of Angrist and Imbens (1995) in (57), which has shown to be equivalent to assuming an ordered choice selection model which allows for random thresholds (Vytlacil, 2006).

$$\text{For any } z, z' \text{ }, T_i(z) \leq T_i(z') \forall\, i \text{ or } T_i(z) \geq T_i(z') \forall\, i. \tag{57}$$

Monotonicity condition (57) benefits from well-establish literature in policy evaluation. In particular, Angrist and Imbens (1995) has shown that the standard TSLS regression evaluates a causal parameter under (57). In this section, I show that monotonicity condition (57) is incompatible with MTO incentives. I then exemplify incentives that justify the condition.

Angrist and Imbens (1995) monotonicity condition in (57) captures the essence of an ordered choice model. The condition is equivalent to state that there exist a sequence of instrumental variables $z_1, ..., z_J$ and such that the corresponding sequence of counterfactual choices is weakly increasing, that is, $T_i(z_1) < ... < T_i(z_J)$ for all agents $i \in \mathcal{I}$. In the case of MTO, it means that, if (57) were true, it would be possible to relabel the instrumental values and neighborhood choices of MTO, say $supp(Z) = \{z_1, z_2, z_3\}$, and $T \in \{1, 2, 3\}$, such that

$$T_i(z_1) \leq T_i(z_2) \leq T_i(z_3) \text{ holds for all } i \in \mathcal{I}. \tag{58}$$

Unfortunately, monotonicity condition (58) does not hold regardless of how we label instrumental values and neighborhood choices. To see this, let the instrumental values z_c, z_8, z_e be relabeled as z_1, z_2, z_3 and the neighborhood choices z_h, z_m, z_l as $1, 2, 3$. In this notation, the

MTO response matrix is given by:

$$\text{Relabeled MTO Response Matrix}: \quad \boldsymbol{R} = \begin{array}{c} \begin{matrix} \boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 \end{matrix} \\ \begin{bmatrix} 1 & 2 & 3 & 1 & 1 & 2 & 1 \\ 1 & 2 & 3 & 2 & 3 & 2 & 2 \\ 1 & 2 & 3 & 3 & 3 & 3 & 1 \end{bmatrix} \end{array} \begin{matrix} T_i(z_1) \\ T_i(z_2) \\ T_i(z_3) \end{matrix} \tag{59}$$

The response-types $\boldsymbol{s}_1$ until $\boldsymbol{s}_6$ are weakly increasing, which comply with the monotonicity condition $T_i(z_1) \leq T_i(z_2) \leq T_i(z_3)$. Response-type $\boldsymbol{s}_7$ however violates this condition as $T_i(z_2) > T_i(z_3)$. Switching the second and third rows of (59) would make $\boldsymbol{s}_7$ comply with the monotonicity criteria (58), but $\boldsymbol{s}_4$ would violate it. It is easy to see that the monotonicity condition would not be satisfied by relabeling the neighborhood choices either. In other words, it is not possible to model the MTO response matrix as an ordered choice model with random thresholds represented by monotonicity condition of Angrist and Imbens (1995).

In the paper called "economics of monotonicity," I investigate which incentive schemes justify the ordered choice models. The paper describe the necessary and sufficient condition that the incentive matrix must have to generate the monotonicity condition of Angrist and Imbens (1995) in (57) under WARP and normality of treatment choices.

The are several incentive schemes that justify the ordered monotonicity condition (58). Let the incentive matrix $\boldsymbol{L}$ be the $J \times K$ matrix that characterises the incentives induced by instrumental values in $Z \in \{z_1, ..., z_J\}$, toward choices in $T \in \{1, ..., K\}$. The matrix input $\boldsymbol{L}[z_j, k]$ denotes the incentive for choosing choice $k \in \{1, ..., K\}$ when assigned to instrumental variable $z_j\{z_1, ..., z_J\}$. One incentive scheme that generates the ordered choice models is the presence of increasing incentive increments, that is:

$$\boldsymbol{L}[z_{j+1}, t_k] - \boldsymbol{L}[z_j, t_k] < \boldsymbol{L}[z_{j+1}, t_{k+1}] - \boldsymbol{L}[z_j, t_{k+1}] \text{ for } j \in \{1, ..., J-1\} \text{ and } k \in \{1, ..., K-1\}. \tag{60}$$

One example of incentive matrix where (60) hold is the Vandermond Matrix:

$$\boldsymbol{L} = \begin{array}{c} \begin{matrix} {\scriptstyle 1} & {\scriptstyle 2} & \cdots & {\scriptstyle K} \end{matrix} \\ \begin{bmatrix} \alpha_1 & \alpha_1^2 & \cdots & \alpha_1^K \\ \alpha_2 & \alpha_2^2 & \cdots & \alpha_2^K \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_J & \alpha_J^2 & \cdots & \alpha_J^K \end{bmatrix} \end{array} \begin{matrix} {\scriptstyle z_1} \\ {\scriptstyle z_2} \\ \vdots \\ {\scriptstyle z_J} \end{matrix}, \text{ for } 1 \leq \alpha_1 < \alpha_2 < ... < \kappa_J.. \tag{61}$$

Examples of such incentives for three-choice model and a three-valued instrument of MTO are:

$$\boldsymbol{L}_1 = \begin{matrix} 1 & 2 & 3 & \\ \left[\begin{matrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \end{matrix}\right. & & \left.\right] & \begin{matrix} z_1 \\ z_2 \\ z_3 \end{matrix} \end{matrix}, \quad \boldsymbol{L}_2 = \overset{\begin{matrix}1 & 2 & 3\end{matrix}}{\begin{bmatrix} 1 & 2 & 3 \\ 1 & 4 & 9 \\ 1 & 8 & 27 \end{bmatrix}} \begin{matrix} z_1 \\ z_2 \\ z_3 \end{matrix}, or \quad \boldsymbol{L}_3 = \overset{\begin{matrix}1 & 2 & 3\end{matrix}}{\begin{bmatrix} 1 & 1 & 1 \\ 2 & 4 & 8 \\ 4 & 16 & 64 \end{bmatrix}} \begin{matrix} z_1 \\ z_2 \\ z_3 \end{matrix} \tag{62}$$

The method described in my MTO paper employ revealed preference analysis to convert incentive matrices into response matrices. Applying the method to any of the incentive matrices $\boldsymbol{L}_1$, $\boldsymbol{L}_2$ or $\boldsymbol{L}_3$ in (62) generates the following response matrix:

$$\text{MTO Response Matrix:} \quad \boldsymbol{R} = \overset{\begin{matrix}\boldsymbol{s}_1 & \boldsymbol{s}_2 & \boldsymbol{s}_3 & \boldsymbol{s}_4 & \boldsymbol{s}_5 & \boldsymbol{s}_6 & \boldsymbol{s}_7 & \boldsymbol{s}_8 & \boldsymbol{s}_9 & \boldsymbol{s}_{10}\end{matrix}}{\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 3 \\ 1 & 1 & 1 & 2 & 2 & 3 & 2 & 2 & 3 & 3 \\ 1 & 2 & 3 & 2 & 3 & 3 & 2 & 3 & 3 & 3 \end{bmatrix}} \begin{matrix} T_i(z_1) \\ T_i(z_2) \\ T_i(z_3) \end{matrix} \tag{63}$$

The choices in each of the response-types of (63) are weakly increasing. This response matrix is a case of an ordered choice model. It contains all the ten admissible response-types that satisfy the monotonicity condition $T_i(z_1) \leq T_i(z_2) \leq T_i(z_3)$. The only response-type probabilities that are point identified are $P(\boldsymbol{S} = \boldsymbol{s}_1)$ and $P(\boldsymbol{S} = \boldsymbol{s}_{10})$. In contrast, the MTO response matrix (59) has seven response-types identify all response-type probabilities.

The incentive scheme in (60) is a sufficient condition to generate an ordered choice model. The incentive matrices $\boldsymbol{L}_4 - -\boldsymbol{L}_6$ (64) also generate response matrices where $T_i(z_1) \leq T_i(z_2) \leq T_i(z_3)$ holds. However, these incentives generate response matrices that have less than ten response-types. Specifically, $\boldsymbol{L}_4$ and $\boldsymbol{L}_5$ generate the same response matrix that has 8 response-types including the response-types in (63) except $\boldsymbol{s}_4, \boldsymbol{s}_9$. Incentive matrix $\boldsymbol{L}_6$ generates a response matrix that contains 7 response-types including those in (63) except $\boldsymbol{s}_4, \boldsymbol{s}_5, \boldsymbol{s}_9$.

$$\boldsymbol{L}_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \boldsymbol{L}_5 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \boldsymbol{L}_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}. \tag{64}$$

D Identification of Counterfactual Outcomes and Response-type Probabilities

The MTO model relies on the following assumptions:

Assumption A. *(i)* $Z \perp\!\!\!\perp \big(Y(t), T(z)\big)|\boldsymbol{X}$ for $(z,t) \in \{z_c, z_8, z_e\} \times \{t_h, t_m, t_l\}$; *(ii)* $supp(\boldsymbol{S}) = \{\boldsymbol{s}_{ah}, ..., \boldsymbol{s}_{ph}\}$ as described in (59); *(iii)* $P(T = t|Z = z, \boldsymbol{X}) > 0$ for all $(z,t) \in supp(Z) \times$

supp(T); *(iv)* $E(|Y|) < \infty$.

A *(i)* is the IV exogeneity condition. **A** *(ii)* states that the support of $\boldsymbol{S}$ consists of the seven response-types generated by the revealed preference analysis in Section 4. **A** *(iii)* means that each neighborhood type is chosen by some families for each voucher assignment. **A** *(iv)* assures that $E(Y)$ exists. All variables lie in the probability space $(\mathcal{I}, \mathcal{C}, P)$ and $\boldsymbol{X}$ is suppressed henceforward for notational simplicity.

Equation (65) describes the outcome equation of the IV model. We seek to identify the counterfactual outcomes $\boldsymbol{Q}_S(t)$.

$$\boldsymbol{Q}_Z(t) \odot \boldsymbol{P}_Z(t) = \boldsymbol{B}_t \cdot \Big(\boldsymbol{Q}_S(t) \odot \boldsymbol{P}_S\Big); \quad t \in \{t_h, t_m, t_l\}, \tag{65}$$

where
$$\begin{aligned}
\boldsymbol{P}_Z(t) &= [P(T=t|Z=z_c), P(T=t|Z=z_8), P(T=t|Z=z_e)]',\\
\boldsymbol{Q}_Z(t) &= [E(Y|T=t,Z=z_c), E(Y|T=t,Z=z_8), E(Y|T=t,Z=z_e)]',\\
\boldsymbol{P}_S &= [P(\boldsymbol{S}=\boldsymbol{s}_{ah}), P(\boldsymbol{S}=\boldsymbol{s}_{am}), P(\boldsymbol{S}=\boldsymbol{s}_{al}), P(\boldsymbol{S}=\boldsymbol{s}_{fc}), P(\boldsymbol{S}=\boldsymbol{s}_{pl}), P(\boldsymbol{S}=\boldsymbol{s}_{pm}), P(\boldsymbol{S}=\boldsymbol{s}_{ph})]',\\
\boldsymbol{Q}_S(t) &= [E(Y(t)|\boldsymbol{S}=\boldsymbol{s}_{ah}), ..., E(Y(t)|\boldsymbol{S}=\boldsymbol{s}_{ph})]',\\
\boldsymbol{B}_t &= \mathbf{1}[\boldsymbol{R}=t]; t \in \{t_l, t_m, t_h\}.
\end{aligned}$$

The the binary matrices $\boldsymbol{B}_t = \mathbf{1}[\boldsymbol{R}=t]$ for t_l, t_m, and t_h are displayed in equations (67),(68) and (69) respectively. It is useful to decompose each binary matrix $\boldsymbol{B}_t$ into $\boldsymbol{B}_t = \boldsymbol{C}_t \cdot \boldsymbol{A}_t$, where $\boldsymbol{C}_t$ is the array the non-zero columns of $\boldsymbol{B}_t$ and $\boldsymbol{A}_t$ is a mapping between the vectors in $\boldsymbol{C}_t$ and $\boldsymbol{B}_t$. Specifically, the response matrix $\boldsymbol{R}$ is decomposed as

$$\boldsymbol{R} \equiv \sum_{t \in supp(T)} t \cdot \boldsymbol{B}_t = \sum_{t \in supp(T)} t \cdot \boldsymbol{C}_t \boldsymbol{A}_t, \tag{66}$$

where $\boldsymbol{B}_t, \boldsymbol{C}_t, \boldsymbol{A}_t$ for $t \in \{t_h, t_m, t_h\}$ are given by:

$$\boldsymbol{B}_{t_h} = \overset{\begin{matrix}\boldsymbol{s}_{ah} & \boldsymbol{s}_{am} & \boldsymbol{s}_{al} & \boldsymbol{s}_{fc} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{pm} & \boldsymbol{s}_{ph}\end{matrix}}{\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}} = \underbrace{\overset{\begin{matrix}\boldsymbol{s}_{fc},\boldsymbol{s}_{pl} & \boldsymbol{s}_{ph} & \boldsymbol{s}_{ah}\end{matrix}}{\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}}}_{\boldsymbol{C}_{t_h}} \cdot \underbrace{\overset{\begin{matrix}\boldsymbol{s}_{ah} & \boldsymbol{s}_{am} & \boldsymbol{s}_{al} & \boldsymbol{s}_{fc} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{pm} & \boldsymbol{s}_{ph}\end{matrix}}{\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}}_{\boldsymbol{A}_{t_h}} \tag{67}$$

$$\boldsymbol{B}_{t_m} = \overset{\begin{matrix}\boldsymbol{s}_{ah} & \boldsymbol{s}_{am} & \boldsymbol{s}_{al} & \boldsymbol{s}_{fc} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{pm} & \boldsymbol{s}_{ph}\end{matrix}}{\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}} = \underbrace{\overset{\begin{matrix}\boldsymbol{s}_{fc},\boldsymbol{s}_{ph} & \boldsymbol{s}_{pm} & \boldsymbol{s}_{am}\end{matrix}}{\begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}}}_{\boldsymbol{C}_{t_m}} \cdot \underbrace{\overset{\begin{matrix}\boldsymbol{s}_{ah} & \boldsymbol{s}_{am} & \boldsymbol{s}_{al} & \boldsymbol{s}_{fc} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{pm} & \boldsymbol{s}_{ph}\end{matrix}}{\begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}}}_{\boldsymbol{A}_{t_m}} \tag{68}$$

$$\boldsymbol{B}_{t_l} = \begin{array}{c} \begin{matrix} \boldsymbol{s}_{ah} & \boldsymbol{s}_{am} & \boldsymbol{s}_{al} & \boldsymbol{s}_{fc} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{pm} & \boldsymbol{s}_{ph} \end{matrix} \\ \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{bmatrix} \end{array} = \underbrace{\begin{array}{c} \begin{matrix} \boldsymbol{s}_{fc}, \boldsymbol{s}_{pm} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{al} \end{matrix} \\ \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \end{array}}_{\boldsymbol{C}_{t_l}} \cdot \underbrace{\begin{array}{c} \begin{matrix} \boldsymbol{s}_{ah} & \boldsymbol{s}_{am} & \boldsymbol{s}_{al} & \boldsymbol{s}_{fc} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{pm} & \boldsymbol{s}_{ph} \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}}_{\boldsymbol{A}_{t_l}} \tag{69}$$

Matrices $\boldsymbol{C}_t$ and $\boldsymbol{A}_t$ are used to generate a closed-form solution for the nonparametric identification of counterfactual outcomes.

The main paper shows that response matrix of MTO generates nestes sets of response-types. Formally, a response matrix is nested regarding a choice t if:

$$\text{for any} z, z' \in \textit{supp}(Z), \text{ we have that } \boldsymbol{B}_t[z, \boldsymbol{s}] \leq \boldsymbol{B}_t[z', \boldsymbol{s}] \forall\, \boldsymbol{s} \text{ or } \boldsymbol{B}_t[z, \boldsymbol{s}] \geq \boldsymbol{B}_t[z', \boldsymbol{s}] \forall\, \boldsymbol{s}. \tag{70}$$

The nested definition (70) is equivalent to state that there exists a sequence of of IV-values $z_1^{(t)}, ... z_N^{(t)}$ of the values in $\textit{supp}(Z) \equiv \{z_1, ..., z_N\}$ such that:

$$\boldsymbol{B}_t[z_k^{(t)}, \boldsymbol{s}] \leq \boldsymbol{B}_t[z_{k+1}^{(t)}, \boldsymbol{s}] \;\forall\; \boldsymbol{s} \in \textit{supp}(\boldsymbol{S}); k = 1, .., N-1. \tag{71}$$

Pinto (2016) shows that if a response matrix $\boldsymbol{R}$ is nested for a choice $t \in \textit{supp}(t)$, then all the identified counterfactual outcomes are identified by the following equation:

$$\Big(\boldsymbol{A}_t\big(\boldsymbol{Q}_S(t) \odot \boldsymbol{P}_S\big) \div \Big(\boldsymbol{A}_t \boldsymbol{P}_S\Big) = \Big((\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}\boldsymbol{C}_t'\big(\boldsymbol{Q}_Z(t) \odot \boldsymbol{P}_Z(t)\big) \div \Big(((\boldsymbol{C}_t'\boldsymbol{C}_t)^{-1}\boldsymbol{C}_t'\boldsymbol{P}_Z(t)\Big) \tag{72}$$

where $\div$ denotes element-wise division,[47] and $\boldsymbol{A}_t$ stems from the decomposition $\boldsymbol{B}_t = \boldsymbol{C}_t\boldsymbol{A}_t$ as in (67)–(69) Moreover, if the response matrix is generated by monotonic incentives, then $\boldsymbol{C}_t$ is invertible and equation (72) can be simplified as:

$$\underbrace{\Big(\boldsymbol{A}_t\big(\boldsymbol{Q}_S(t) \odot \boldsymbol{P}_S\big) \div \Big(\boldsymbol{A}_t \boldsymbol{P}_S\Big)}_{\text{Identified Counterfactual Outcomes}} = \underbrace{\Big(\boldsymbol{C}_t^{-1}\big(\boldsymbol{Q}_Z(t) \odot \boldsymbol{P}_Z(t)\big) \div \Big((\boldsymbol{C}_t^{-1}\boldsymbol{P}_Z(t)\Big)}_{\text{Identification Formulas}} \tag{73}$$

The right-hand side of (73) summarizes all identified counterfactual outcomes. The left-hand side of (73) generates identification formulas. Equations (74)–(76) exemplify the left-hand side of (73) for t_h.

$$\boldsymbol{A}_{t_h}\big(\boldsymbol{Q}_S(t_h) \odot \boldsymbol{P}_S\big) = \begin{bmatrix} E(Y(t_h)|\boldsymbol{S} = \boldsymbol{s}_{fc})\, P(\boldsymbol{S} = \boldsymbol{s}_{fc}) + E(Y(t_h)|\boldsymbol{S} = \boldsymbol{s}_{pl})\, P(\boldsymbol{S} = \boldsymbol{s}_{pl}) \\ E(Y(t_h)|\boldsymbol{S} = \boldsymbol{s}_{ph})\, P(\boldsymbol{S} = \boldsymbol{s}_{ph}) \\ E(Y(t_h)|\boldsymbol{S} = \boldsymbol{s}_{ah})\, P(\boldsymbol{S} = \boldsymbol{s}_{ah}) \end{bmatrix} \tag{74}$$

$$\boldsymbol{A}_{t_h}\boldsymbol{P}_S = \begin{bmatrix} P(\boldsymbol{S} = \boldsymbol{s}_{fc}) + P(\boldsymbol{S} = \boldsymbol{s}_{pl}) \\ P(\boldsymbol{S} = \boldsymbol{s}_{ph}) \\ P(\boldsymbol{S} = \boldsymbol{s}_{ah}) \end{bmatrix} \tag{75}$$

$$\therefore \Big(\boldsymbol{A}_{t_h}\big(\boldsymbol{Q}_S(t_h) \odot \boldsymbol{P}_S\big)\Big) \div \Big(\boldsymbol{A}_{t_h}\boldsymbol{P}_S\Big) = \begin{bmatrix} E(Y(t_h)|\boldsymbol{S} \in \boldsymbol{s}_{fc}, \boldsymbol{s}_{pl}) \\ E(Y(t_h)|\boldsymbol{S} = \boldsymbol{s}_{ph}) \\ E(Y(t_h)|\boldsymbol{S} = \boldsymbol{s}_{ah}) \end{bmatrix} \tag{76}$$

[47] Let $\boldsymbol{A}, \boldsymbol{B}$ be two vectors of same length, then $\boldsymbol{A} \div \boldsymbol{B} \equiv \text{diag}(\boldsymbol{B})^{-1}\boldsymbol{A}$, where $\text{diag}(\cdot)$ is the operator that transform a vector into a diagonal matrix.

The right-hand side of (73) for t_h employs the matrix $\boldsymbol{C}_{t_h}^{-1}$ displayed in equation (77):

$$\boldsymbol{C}_{t_h} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \Rightarrow \boldsymbol{C}_{t_h}^{-1} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \tag{77}$$

Equations (78) and (79) exemplify the numerator and the denominators of the right-hand side of (73) for t_h :

$$\begin{aligned} \boldsymbol{C}_{t_h}^{-1}\boldsymbol{Q}_Z(t_h) \odot \boldsymbol{P}_Z(t_h) &= \begin{bmatrix} E(Y|T=t_h, Z=z_c)\,P(T=t_h|Z=z_c) - E(Y|T=t_h, Z=z_e)\,P(T=t_h|Z=z_e) \\ E(Y|T=t_h, Z=z_e)\,P(T=t_h|Z=z_e) - E(Y|T=t_h, Z=z_8)\,P(T=t_h|Z=z_8) \\ E(Y|T=t_h, Z=z_8)\,P(T=t_h|Z=z_8) \end{bmatrix}, \\ &= \begin{bmatrix} E(Y \cdot D_{t_h}|Z=z_c) - E(Y \cdot \mathbf{1}[T=t_h]|Z=z_e) \\ E(Y \cdot D_{t_h}|Z=z_e) - E(Y \cdot \mathbf{1}[T=t_h]|Z=z_8) \\ E(Y \cdot D_{t_h}|Z=z_8) \end{bmatrix}, \end{aligned} \tag{78}$$

$$\text{and } \boldsymbol{C}_{t_h}^{-1}\boldsymbol{P}_Z(t_h) = \begin{bmatrix} P(T=t_h|Z=z_c) - P(T=t_h|Z=z_e) \\ P(T=t_h|Z=z_e) - P(T=t_h|Z=z_8) \\ P(T=t_h|Z=z_8) \end{bmatrix}, \tag{79}$$

The final equation for t_h is presented in (80). The left-hand side of (80) lists all the identified counterfactual outcome means of $Y(t_h)$. The right-hand side provides the identification formulas that can be evaluated from observed data.

$$\therefore \underbrace{\begin{bmatrix} E(Y(t_h)|\boldsymbol{S} \in \boldsymbol{s}_{fc}, \boldsymbol{s}_{pl}) \\ E(Y(t_h)|\boldsymbol{S} = \boldsymbol{s}_{ph}) \\ E(Y(t_h)|\boldsymbol{S} = \boldsymbol{s}_{ah}) \end{bmatrix}}_{\boldsymbol{A}_{t_h}(\boldsymbol{Q}_S(t_h) \odot \boldsymbol{P}_S) \div \boldsymbol{A}_{t_h}\boldsymbol{P}_S} = \underbrace{\begin{bmatrix} \frac{E(Y \cdot D_{t_h}|Z=z_c) - E(Y \cdot D_{t_h}|Z=z_e)}{P(T=t_h|Z=z_c) - P(T=t_h|Z=z_e)} \\ \frac{E(Y \cdot D_{t_h}|Z=z_e) - E(Y \cdot D_{t_h}|Z=z_8)}{P(T=t_h|Z=z_e) - P(T=t_h|Z=z_8)} \\ \frac{E(Y \cdot D_{t_h}|Z=z_8)}{P(T=t_h|Z=z_8)} \end{bmatrix}}_{(\boldsymbol{C}_{t_h}^{-1}\boldsymbol{Q}_Z(t_h) \odot \boldsymbol{P}_Z(t_h)) \div \boldsymbol{C}_{t_h}^{-1}\boldsymbol{P}_Z(t_h)}. \tag{80}$$

Equations (81)–(82) arise from applying formula (73) to t_m and t_l :

$$\underbrace{\begin{bmatrix} E(Y(t_m)|\boldsymbol{S} \in \{\boldsymbol{s}_{fc}, \boldsymbol{s}_{ph}\}) \\ E(Y(t_m)|\boldsymbol{S} = \boldsymbol{s}_{pm}) \\ E(Y(t_m)|\boldsymbol{S} = \boldsymbol{s}_{am}) \end{bmatrix}}_{\boldsymbol{A}_{t_m}(\boldsymbol{Q}_S(t_m) \odot \boldsymbol{P}_S) \div \boldsymbol{A}_{t_m}\boldsymbol{P}_S} = \underbrace{\begin{bmatrix} \frac{E(Y \cdot D_{t_m}|Z=z_8) - E(Y \cdot D_{t_m}|Z=z_c)}{P(T=t_m|Z=z_8) - P(T=t_m|Z=z_c)} \\ \frac{E(Y \cdot D_{t_m}|Z=z_c) - E(Y \cdot D_{t_m}|Z=z_e)}{P(T=t_m|Z=z_c) - P(T=t_m|Z=z_e)} \\ \frac{E(Y \cdot D_{t_m}|Z=z_e)}{P(T=t_m|Z=z_e)} \end{bmatrix}}_{\boldsymbol{C}_{t_m}^{-1}(\boldsymbol{Q}_Z(t_m) \odot \boldsymbol{P}_Z(t_m)) \div \boldsymbol{C}_{t_m}^{-1}\boldsymbol{P}_Z(t_m)}, \tag{81}$$

$$\underbrace{\begin{bmatrix} E(Y(t_l)|\boldsymbol{S} \in \{\boldsymbol{s}_{fc}, \boldsymbol{s}_{pm}\}) \\ E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{pl}) \\ E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{al}) \end{bmatrix}}_{\boldsymbol{A}_{t_l}(\boldsymbol{Q}_S(t_l) \odot \boldsymbol{P}_S) \div \boldsymbol{A}_{t_l}\boldsymbol{P}_S} = \underbrace{\begin{bmatrix} \frac{E(Y \cdot D_{t_l}|Z=z_e) - E(Y \cdot D_{t_l}|Z=z_8)}{P(T=t_l|Z=z_e) - P(T=t_l|Z=z_8)} \\ \frac{E(Y \cdot D_{t_l}|Z=z_8) - E(Y \cdot D_{t_l}|Z=z_c)}{P(T=t_l|Z=z_8) - P(T=t_l|Z=z_c)} \\ \frac{E(Y \cdot D_{t_l}|Z=z_c)}{P(T=t_l|Z=z_c)} \end{bmatrix}}_{\boldsymbol{C}_{t_l}^{-1}(\boldsymbol{Q}_Z(t_l) \odot \boldsymbol{P}_Z(t_l)) \div \boldsymbol{C}_{t_l}^{-1}\boldsymbol{P}_Z(t_l)}. \tag{82}$$

Response-type probabilities can be identified by equations $\boldsymbol{A}_t\boldsymbol{P}_S = \boldsymbol{C}_{t_h}^{-1}\boldsymbol{P}_Z(t_h)$ for $t = t_h, t_l, t_m$:

$$\therefore \underbrace{\begin{bmatrix} P(\boldsymbol{S} \in \boldsymbol{s}_{fc}, \boldsymbol{s}_{pl}) \\ P(\boldsymbol{S} = \boldsymbol{s}_{ph}) \\ P(\boldsymbol{S} = \boldsymbol{s}_{ah}) \end{bmatrix}}_{\boldsymbol{A}_{t_h} \boldsymbol{P}_S} = \underbrace{\begin{bmatrix} P(T = t_h | Z = z_c) - P(T = t_h | Z = z_e) \\ P(T = t_h | Z = z_e) - P(T = t_h | Z = z_8) \\ P(T = t_h | Z = z_8) \end{bmatrix}}_{\boldsymbol{C}_{t_h}^{-1} \boldsymbol{P}_{Z(t_h)}}. \tag{83}$$

$$\underbrace{\begin{bmatrix} P(\boldsymbol{S} \in \{\boldsymbol{s}_{fc}, \boldsymbol{s}_{ph}\}) \\ P(\boldsymbol{S} = \boldsymbol{s}_{pm}) \\ P(\boldsymbol{S} = \boldsymbol{s}_{am}) \end{bmatrix}}_{\boldsymbol{A}_{t_m} \boldsymbol{P}_S} = \underbrace{\begin{bmatrix} P(T = t_m | Z = z_8) - P(T = t_m | Z = z_c) \\ P(T = t_m | Z = z_c) - P(T = t_m | Z = z_e) \\ P(T = t_m | Z = z_e) \end{bmatrix}}_{\boldsymbol{C}_{t_m}^{-1} \boldsymbol{P}_{Z(t_m)}}, \tag{84}$$

$$\underbrace{\begin{bmatrix} P(\boldsymbol{S} \in \{\boldsymbol{s}_{fc}, \boldsymbol{s}_{pm}\}) \\ P(\boldsymbol{S} = \boldsymbol{s}_{pl}) \\ P(\boldsymbol{S} = \boldsymbol{s}_{al}) \end{bmatrix}}_{\boldsymbol{A}_{t_l} \boldsymbol{P}_S} = \underbrace{\begin{bmatrix} P(T = t_l | Z = z_e) - P(T = t_l | Z = z_8) \\ P(T = t_l | Z = z_8) - P(T = t_l | Z = z_c) \\ P(T = t_l | Z = z_c) \end{bmatrix}}_{\boldsymbol{C}_{t_l}^{-1} \boldsymbol{P}_{Z(t_l)}}. \tag{85}$$

E Understanding the Problem of Partial Identification

The partial identification problem of discrete instruments stems from the lack of variability of the instrumental variable. The matrices in (86) help to illustrate this fact. The first matrix lists the response-types in the response-matrix $\boldsymbol{R}$ in which t_l appears. Those are $\boldsymbol{s}_{al}, \boldsymbol{s}_{fc}, \boldsymbol{s}_{pl}, \boldsymbol{s}_{pm}$. The second matrix indicates the t_l-choices. The matrix consists of the columns $\boldsymbol{s}_{al}, \boldsymbol{s}_{fc}, \boldsymbol{s}_{pl}, \boldsymbol{s}_{pm}$ of $\boldsymbol{B}_{t_l}$. The third matrix reorders these columns into a triangular matrix.

$$\begin{array}{c} \\ z_c \\ z_8 \\ z_e \end{array} \underbrace{\begin{array}{cccc} \boldsymbol{s}_{al} & \boldsymbol{s}_{fc} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{pm} \\ \hline t_l & t_h & t_h & t_m \\ t_l & t_m & t_l & t_m \\ t_l & t_l & t_l & t_l \end{array}}_{\text{Selection of } \boldsymbol{R}} \Rightarrow \underbrace{\begin{array}{cccc} \boldsymbol{s}_{al} & \boldsymbol{s}_{fc} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{pm} \\ \hline 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{array}}_{\text{Columns of } \boldsymbol{B}_{t_l}} \Rightarrow \underbrace{\begin{array}{cccc} \boldsymbol{s}_{al} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{fc} & \boldsymbol{s}_{pm} \\ \hline 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{array}}_{\text{Reordered } \boldsymbol{B}_{t_l} \text{ Columns}} \therefore \begin{array}{l} \\ \boldsymbol{s}_{al} \\ \boldsymbol{s}_{al}, \boldsymbol{s}_{pl} \\ \boldsymbol{s}_{al}, \boldsymbol{s}_{pl}, \boldsymbol{s}_{fc}, \boldsymbol{s}_{pm} \end{array} \tag{86}$$

The first row of the matrices in (86) correspond to the instrumental value z_c. It shows that $\boldsymbol{s}_{al}$ is the response-type that takes value t_l. Under z_8 (second row) response-types $\boldsymbol{s}_{al}$ and $\boldsymbol{s}_{pl}$ take value t_l. The difference of response-type sets between z_8 and z_c is $\boldsymbol{s}_{pl}$ which renders $E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{pl})$ identified. Partial identification arises because the difference between the sets of response-types that take value t_l when the instrument shifts from z_8 to z_e consists of two response-types $\{\boldsymbol{s}_{fc}, \boldsymbol{s}_{pm}\}$. Thereby $E(Y(t_l)|\boldsymbol{S} \in \{\boldsymbol{s}_{fc}, \boldsymbol{s}_{pm}\})$ is identified but cannot be disentangled into $E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{fc})$ and $E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{pm})$.

The partial identification problem would be solved if there were an hypothetical instrumental value z^* that bridges the response-types that take value t_l between the IV values z_8 and z_e. This means that the set of response-types that take value t_l under the hypothetical IV value z^* is either $\{\boldsymbol{s}_{al}, \boldsymbol{s}_{pl}, \boldsymbol{s}_{fc}\}$ or $\{\boldsymbol{s}_{al}, \boldsymbol{s}_{pl}, \boldsymbol{s}_{pm}\}$.

It is useful to characterise the hypothetical IV value z^* in terms of the neighborhood choices under z_8. That is to say that IV value z^* that enables the identification of $E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{fc})$ and $E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{pm})$ has the following properties:

1. z^* induces the same neighborhood choices of z_8 for the response-types $\boldsymbol{s}_{al}, \boldsymbol{s}_{pl}$, that is t_l

2. Moreover, z^* shifts the neighborhood choice of z_8 from t_m to t_l for either $\boldsymbol{s}_{pm}$ or $\boldsymbol{s}_{fc}$.

In summary, we seek to investigate the properties of the response-matrix when we shift the neighborhood choices of response-types $\boldsymbol{s}_{fc}$ or $\boldsymbol{s}_{pm}$ from t_m to t_l as the IV value change from z_8 to z^*. The matrices in (87) illustrate these two possibilities by inserting a z^*-row between the rows z_8 and z_e. The second matrix displays the $\boldsymbol{s}_{fc}$-shift from t_m to t_l while the third matrix presents the $\boldsymbol{s}_{pm}$-shift.

$$\underbrace{\begin{array}{c} \\ z_c \\ z_8 \\ z^* \\ z_e \end{array} \begin{array}{c} \begin{array}{cccc} \boldsymbol{s}_{al} & \boldsymbol{s}_{fc} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{pm} \end{array} \\ \begin{bmatrix} t_l & t_h & t_h & t_m \\ t_l & t_m & t_l & t_m \\ t_l & ? & t_l & ? \\ t_l & t_l & t_l & t_l \end{bmatrix} \end{array}}_{\text{New Instrument Value } z^*} \Rightarrow \underbrace{\overbrace{\begin{array}{c} \begin{array}{cccc} \boldsymbol{s}_{al} & \boldsymbol{s}_{fc} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{pm} \end{array} \\ \begin{bmatrix} t_l & t_h & t_h & t_m \\ t_l & t_m & t_l & t_m \\ t_l & \boxed{t_l} & t_l & \boxed{t_m} \\ t_l & t_l & t_l & t_l \end{bmatrix} \end{array}}^{\boldsymbol{s}_{fc}\text{-shift}}}_{\substack{\Sigma_{t_l}(z^*) = \{\boldsymbol{s}_{al}, \boldsymbol{s}_{pl}, \boldsymbol{s}_{fc}\} \\ \text{Unordered monotonicity holds}}} \text{ or } \underbrace{\overbrace{\begin{array}{c} \begin{array}{cccc} \boldsymbol{s}_{al} & \boldsymbol{s}_{fc} & \boldsymbol{s}_{pl} & \boldsymbol{s}_{pm} \end{array} \\ \begin{bmatrix} t_l & \boxed{t_h} & t_h & \boxed{t_m} \\ t_l & t_m & t_l & t_m \\ t_l & \boxed{t_m} & t_l & \boxed{t_l} \\ t_l & t_l & t_l & t_l \end{bmatrix} \end{array}}^{\boldsymbol{s}_{pm}\text{-shift}}}_{\substack{\Sigma_{t_l}(z^*) = \{\boldsymbol{s}_{al}, \boldsymbol{s}_{pl}, \boldsymbol{s}_{pm}\} \\ \text{Unordered monotonicity does \textbf{not} hold}}} \quad (87)$$

We can now examine if the $\boldsymbol{s}_{fc}$-shift and the $\boldsymbol{s}_{pm}$-shift in (87) preserve the nested property of the neighborhood choices. In the case of choice t_l, the matrices in (87) show that both the $\boldsymbol{s}_{fc}$-shift and the $\boldsymbol{s}_{pm}$-shift preserve the nested property of neighborhood choice t_l. If we investigate choice t_h, we can also conclude that both the $\boldsymbol{s}_{fc}$-shift and the $\boldsymbol{s}_{pm}$-shift in preserve the nested property of neighborhood choice t_h. If we focus on the second matrix in (87), we can conclude that $\boldsymbol{s}_{fc}$-shift preserves the nested property of of neighborhood choice t_m. The $\boldsymbol{s}_{pm}$-shift however violates the nested property of neighborhood choice t_m. Note that response-type $\boldsymbol{s}_{pm}$ takes value t_m under z_c while response-type $\boldsymbol{s}_{fc}$ takes value t_m under z_*. These sets are not nested and thereby unordered monotonicity does not hold. We conclude that only the $\boldsymbol{s}_{fc}$-shift preserves the nested property for all neighborhood choices. Thereby the sequence of nested sets of response-types for choice t_l is given by $\{\boldsymbol{s}_{al}\} \to \{\boldsymbol{s}_{al}, \boldsymbol{s}_{pl}\} \to \{\boldsymbol{s}_{al}, \boldsymbol{s}_{pl}, \boldsymbol{s}_{fc}\} \to \{\boldsymbol{s}_{al}, \boldsymbol{s}_{pl}, \boldsymbol{s}_{fc}, \boldsymbol{s}_{pm}\}$. The identification equations for $E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{fc})$ and $E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{pm})$ generated by this sequence are displayed in Table A.4.

Table A.4: Identification Formulas for $E(Y(t_l)|\boldsymbol{S}=\boldsymbol{s}_{fc})$ and $E(Y(t_l)|\boldsymbol{S}=\boldsymbol{s}_{pm})$

Counterfactual Outcome	Integral Representation	Function of Propensity Scores
$E(Y(t_l)\|\boldsymbol{S}\in\{\boldsymbol{s}_{fc},\boldsymbol{s}_{pm}\}) =$	$\frac{\int_{P_{t_l}(z_8)}^{P_{t_l}(z_e)} E(Y(t_l)\|U_{t_l}=u)du}{P_{t_l}(z_e)-P_{t_l}(z_8)}$	$\equiv g_{t_l}\Big(P_{t_l}(z_e),P_{t_l}(z_8)\Big)$
$E(Y(t_h)\|\boldsymbol{S}=\boldsymbol{s}_{fc}) =$	$\frac{\int_{P_{t_h}(z_8)}^{p^*_{t_l}} E(Y(t_l)\|U_{t_l}=u)du}{p^*_{t_l}-P_{t_l}(z_8)}$	$\equiv g_{t_h}\Big(p^*_{t_l},P_{t_l}(z_8)\Big)$
$E(Y(t_h)\|\boldsymbol{S}=\boldsymbol{s}_{pm}) =$	$\frac{\int_{p^*_{t_l}}^{P_{t_l}(z_e)} E(Y(t_h)\|U_{t_l}=u)du}{P_{t_h}(z_e)-p^*_{t_l}}$	$\equiv g_{t_l}\Big(P_{t_l}(z_e),p^*_{t_l}\Big)$
where $p^*_{t_l}$	$=P(\boldsymbol{S}\in\{\boldsymbol{s}_{al},\boldsymbol{s}_{pl},\boldsymbol{s}_{fc}\})\in\Big(P_{t_l}(z_8),P_{t_l}(z_e)\Big)$	
because $P_{t_l}(z_8)$	$=P(\boldsymbol{S}\in\{\boldsymbol{s}_{al},\boldsymbol{s}_{pl}\})$ and $P_{t_l}(z_e)=P(\boldsymbol{S}\in\{\boldsymbol{s}_{al},\boldsymbol{s}_{pl},\boldsymbol{s}_{fc},\boldsymbol{s}_{pm}\})$	

A parallel analysis hold for choices t_h. We seek to disentangle $E(Y(t_h)|\boldsymbol{S}\in\{\boldsymbol{s}_{fc},\boldsymbol{s}_{pl}\})$ into $E(Y(t_h)|\boldsymbol{S}=\boldsymbol{s}_{fc})$ and $E(Y(t_h)|\boldsymbol{S}=\boldsymbol{s}_{pl})$. The sequence of IV values for increasing values of the propensity score $P(T=t_h|Z=z)$ is $z_m\rightarrow z_e\rightarrow z_c$. The associated sequence of response-types is given by $\{\boldsymbol{s}_{ah}\}\rightarrow\{\boldsymbol{s}_{ah},\boldsymbol{s}_{ph}\}\rightarrow\{\boldsymbol{s}_{ah},\boldsymbol{s}_{ph},\boldsymbol{s}_{fc},\boldsymbol{s}_{pl}\}$. Following the same rationale applied to choice t_l we can verify that $\boldsymbol{s}_{fc}$ must antecede $\boldsymbol{s}_{pl}$, otherwise the nested property for t_l. Thereby the sequence of nested sets of response-types for choice t_h is given by $\{\boldsymbol{s}_{ah}\}\rightarrow\{\boldsymbol{s}_{ah},\boldsymbol{s}_{ph}\}\rightarrow\{\boldsymbol{s}_{ah},\boldsymbol{s}_{ph},\boldsymbol{s}_{fc},\boldsymbol{s}_{pl}\}\rightarrow\{\boldsymbol{s}_{ah},\boldsymbol{s}_{ph},\boldsymbol{s}_{fc},\boldsymbol{s}_{pl}\}$. The identification equations for $E(Y(t_h)|\boldsymbol{S}=\boldsymbol{s}_{fc})$ and $E(Y(t_h)|\boldsymbol{S}=\boldsymbol{s}_{pl})$ generated by this sequence are displayed in Table A.5.

Table A.5: Identification Formulas for $E(Y(t_h)|\boldsymbol{S}=\boldsymbol{s}_{fc})$ and $E(Y(t_h)|\boldsymbol{S}=\boldsymbol{s}_{pl})$

Counterfactual Outcome	Integral Representation	Function of Propensity Scores
$E(Y(t_h)\|\boldsymbol{S}\in\{\boldsymbol{s}_{fc},\boldsymbol{s}_{pl}\}) =$	$\frac{\int_{P_{t_h}(z_e)}^{P_{t_h}(z_c)} E(Y(t_h)\|U_{t_h}=u)du}{P_{t_h}(z_c)-P_{t_h}(z_e)}$	$\equiv g_{t_h}\Big(P_{t_h}(z_c),P_{t_h}(z_e)\Big)$
$E(Y(t_h)\|\boldsymbol{S}=\boldsymbol{s}_{fc}) =$	$\frac{\int_{P_{t_h}(z_e)}^{p^*_{t_h}} E(Y(t_h)\|U_{t_h}=u)du}{p^*_{t_h}-P_{t_h}(z_e)}$	$\equiv g_{t_h}\Big(p^*_{t_h},P_{t_h}(z_e)\Big)$
$E(Y(t_h)\|\boldsymbol{S}=\boldsymbol{s}_{pl}) =$	$\frac{\int_{p^*_{t_h}}^{P_{t_h}(z_c)} E(Y(t_h)\|U_{t_h}=u)du}{P_{t_h}(z_c)-p^*_{t_h}}$	$\equiv g_{t_h}\Big(P_{t_h}(z_c),p^*_{t_h}\Big)$
where $p^*_{t_h}$	$=P(\boldsymbol{S}\in\{\boldsymbol{s}_{ah},\boldsymbol{s}_{ph},\boldsymbol{s}_{fc}\})\in\Big(P_{t_h}(z_e),P_{t_h}(z_c)\Big)$	
because $P_{t_h}(z_e)$	$=P(\boldsymbol{S}\in\{\boldsymbol{s}_{ah},\boldsymbol{s}_{ph}\})$ and $P_{t_h}(z_c)=P(\boldsymbol{S}\in\{\boldsymbol{s}_{ah},\boldsymbol{s}_{ph},\boldsymbol{s}_{fc},\boldsymbol{s}_{pl}\})$	

A parallel analysis hold for choices t_m. We seek to disentangle $E(Y(t_m)|\boldsymbol{S}\in\{\boldsymbol{s}_{fc},\boldsymbol{s}_{ph}\})$ into $E(Y(t_m)|\boldsymbol{S}=\boldsymbol{s}_{fc})$ and $E(Y(t_m)|\boldsymbol{S}=\boldsymbol{s}_{ph})$. The sequence of IV values for increasing values of the propensity score $P(T=t_m|Z=z)$ is $z_m\rightarrow z_e\rightarrow z_c$. The associated sequence of response-types is given by $\{\boldsymbol{s}_{am}\}\rightarrow\{\boldsymbol{s}_{am},\boldsymbol{s}_{pm}\}\rightarrow\{\boldsymbol{s}_{am},\boldsymbol{s}_{pm},\boldsymbol{s}_{fc},\boldsymbol{s}_{ph}\}$. Following the same rationale applied to choice t_l we can verify that $\boldsymbol{s}_{fc}$ must antecede $\boldsymbol{s}_{ph}$, otherwise the nested property for t_h. Thereby the sequence of nested sets of response-types for choice t_m is given by $\{\boldsymbol{s}_{am}\} \rightarrow \{\boldsymbol{s}_{am}, \boldsymbol{s}_{pm}\} \rightarrow \{\boldsymbol{s}_{am}, \boldsymbol{s}_{pm}, \boldsymbol{s}_{fc}, \boldsymbol{s}_{ph}\} \rightarrow \{\boldsymbol{s}_{am}, \boldsymbol{s}_{pm}, \boldsymbol{s}_{fc}, \boldsymbol{s}_{ph}\}$. The identification equations for $E(Y(t_m)|\boldsymbol{S} = \boldsymbol{s}_{fc})$ and $E(Y(t_m)|\boldsymbol{S} = \boldsymbol{s}_{ph})$ generated by this sequence are displayed in Table A.6.

Table A.6: Identification Formulas for $E(Y(t_m)|\boldsymbol{S} = \boldsymbol{s}_{fc})$ and $E(Y(t_m)|\boldsymbol{S} = \boldsymbol{s}_{ph})$

Counterfactual Outcome	Integral Representation	Function of Propensity Scores
$E(Y(t_m)\|\boldsymbol{S} \in \{\boldsymbol{s}_{fc}, \boldsymbol{s}_{ph}\}) =$	$\frac{\int_{P_{t_m}(z_8)}^{P_{t_m}(z_c)} E(Y(t_m)\|U_{t_m}=u)du}{P_{t_m}(z_8)-P_{t_m}(z_c)}$	$\equiv g_{t_m}\Big(P_{t_m}(z_8), P_{t_m}(z_c)\Big)$
$E(Y(t_m)\|\boldsymbol{S} = \boldsymbol{s}_{fc}) =$	$\frac{\int_{P_{t_m}(z_8)}^{p^*_{t_m}} E(Y(t_m)\|U_{t_m}=u)du}{p^*_{t_m}-P_{t_m}(z_c)}$	$\equiv g_{t_m}\Big(p^*_{t_m}, P_{t_m}(z_8)\Big)$
$E(Y(t_m)\|\boldsymbol{S} = \boldsymbol{s}_{ph}) =$	$\frac{\int_{p^*_{t_m}}^{P_{t_m}(z_8)} E(Y(t_m)\|U_{t_m}=u)du}{P_{t_m}(z_8)-p^*_{t_m}}$	$\equiv g_{t_m}\Big(P_{t_m}(z_8), p^*_{t_m}\Big)$
where $p^*_{t_m}$	$= P(\boldsymbol{S} \in \{\boldsymbol{s}_{am}, \boldsymbol{s}_{pm}, \boldsymbol{s}_{fc}\}) \in \Big(P_{t_m}(z_c), P_{t_m}(z_8)\Big)$	
because $P_{t_m}(z_e)$	$= P(\boldsymbol{S} \in \{\boldsymbol{s}_{am}, \boldsymbol{s}_{pm}\})$ and $P_{t_m}(z_c) = P(\boldsymbol{S} \in \{\boldsymbol{s}_{am}, \boldsymbol{s}_{pm}, \boldsymbol{s}_{fc}, \boldsymbol{s}_{ph}\})$	

F Nonparametric Propensity Score Estimator with Covariates

As mentioned in the main paper, the response matrix of MTO implies the unordered monotonicity conditions, which in turn is equivalent to state that the choice indicator can be expressed by the following inequality $D_t = \mathbf{1}[P_t(Z) \geq U_t]$ such that $U_t \sim Unif[0,1]$ and for $t \in \{t_h, t_m, t_l\}$ (Heckman and Pinto, 2018).

The main paper shows that the counterfactual outcome $E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{pl})$ can be expressed as:

$$E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{pl}) = \frac{E(YD_{t_l}|Z = z_8) - E(YD_{t_l}|Z = z_c)}{E(D_{t_l}|Z = z_8) - E(D_{t_l}|Z = z_c)} = \frac{\int_{P_{t_l}(z_c)}^{P_{t_l}(z_8)} E(Y(t_l)|U_{t_l} = u)du}{P_{t_l}(z_8) - P_{t_l}(z_c)}, \tag{88}$$

where $P_t(z) \equiv P(T = t|Z = z)$ is the unconditional propensity score for $t \in \{t_h, t_m, t_l\}$ and $z \in \{z_c, z_8, z_e\}$.

Consider a more general terminology in which the values $z, z' \in \{z_c, z_8, z_e\}$ and the choice value $t \in \{t_h, t_m, t_l\}$ are such that such that $P_t(z') > P_t(z)$ are associated with a response-type $\boldsymbol{s} \in supp(\boldsymbol{S})$ for which (89) holds.

$$E(Y(t)|\boldsymbol{S} = \boldsymbol{s}) = \frac{\int_{P_t(z)}^{P_t(z')} E(Y(t)|U_t = u)du}{P_t(z) - P_t(z')}, \tag{89}$$

In summary, equation (88) simply describes a connection between instrumental values z, z', neighborhood choice t and their associated with response-type $\boldsymbol{s} \in supp(\boldsymbol{S})$. The numerator in (89) is identified by:

$$\begin{aligned}
&\int_{P_t(z)}^{P_t(z')} E(Y(t)|U_t = u)du = \\
&\qquad = E(Y(t)\mathbf{1}[P_t(z) \leq U_t \leq P_t(z')]) \\
&\qquad = E(Y(t)\mathbf{1}[U_t \leq P_t(z')] - \mathbf{1}[U_t \geq P_t(z)]) \\
&\qquad = E(Y(t)\mathbf{1}[U_t \leq P_t(Z)]|P_t(Z) = P_t(z')) - E(Y(t)\mathbf{1}[U_t \geq P_t(Z)]|P_t(Z) = P_t(z)) \\
&\qquad = E(YD_t|P_t(Z) = P_t(z')) - E(YD_t|P_t(Z) = P_t(z)) \qquad (90)
\end{aligned}$$

This section seeks to identify (89) as a function of propensity scores conditional of $\boldsymbol{X}$, that is $P_t(z, \boldsymbol{x}) \equiv P(T = t|Z = z, \boldsymbol{X} = \boldsymbol{x})$. The conditional version of (90) is given by:

$$\begin{aligned}
E(Y(t)|\boldsymbol{S} = \boldsymbol{s}, \boldsymbol{X} = \boldsymbol{x}) &= \frac{\int_{P_t(z',\boldsymbol{x})}^{P_t(z,\boldsymbol{x})} E(Y(t)|U_t = u, \boldsymbol{X} = \boldsymbol{x})du}{P_t(z, \boldsymbol{x}) - P_t(z', \boldsymbol{x})} \\
&= \frac{E(YD_t|P_t(Z) = P_t(z', \boldsymbol{x}), \boldsymbol{X} = \boldsymbol{x}) - E(YD_t|P_t(Z) = P_t(z, \boldsymbol{x}), \boldsymbol{X} = \boldsymbol{x})}{P_t(z', \boldsymbol{x}) - P_t(z, \boldsymbol{x})} \qquad (91)
\end{aligned}$$

Integrating $E(Y(t)|\boldsymbol{S} = \boldsymbol{s}, \boldsymbol{X} = \boldsymbol{x})$ over $\boldsymbol{X}$ generates the following equation:

$$\begin{aligned}
\int E(Y(t)|\boldsymbol{S} = \boldsymbol{s}, \boldsymbol{X} = \boldsymbol{x})dF_{\boldsymbol{X}|\boldsymbol{S}=\boldsymbol{s}}(\boldsymbol{x}) &= \int E(Y(t)|\boldsymbol{S} = \boldsymbol{s}, \boldsymbol{X} = \boldsymbol{x})dF_{\boldsymbol{X}|\boldsymbol{S}=\boldsymbol{s}}(\boldsymbol{x}) \\
&= \int E(Y(t)|\boldsymbol{S} = \boldsymbol{s}, \boldsymbol{X} = \boldsymbol{x})\frac{P(\boldsymbol{S} = \boldsymbol{s}|\boldsymbol{X} = \boldsymbol{x})}{P(\boldsymbol{S} = \boldsymbol{s})}dF_{\boldsymbol{X}}(\boldsymbol{x}), \quad (92)
\end{aligned}$$

where the second equality is due to Bayes' theorem. Recall that $P(\boldsymbol{S} = \boldsymbol{s}|\boldsymbol{X} = \boldsymbol{x}) = P_t(z', \boldsymbol{x}) - P_t(z, \boldsymbol{x})$ and thereby $P(\boldsymbol{S} = \boldsymbol{s}) = \int P_t(z', \boldsymbol{x}) - P_t(z, \boldsymbol{x})dF_{\boldsymbol{X}}(\boldsymbol{x})$. Inserting (91) into (92) and using the above results generates:

$$\begin{aligned}
&E(Y(t)|\boldsymbol{S} = \boldsymbol{s}) = \\
&\quad = \frac{\int \Big(E(YD_t|P_t(Z) = P_t(z', \boldsymbol{x}), \boldsymbol{X} = \boldsymbol{x}) - E(YD_t|P_t(Z) = P_t(z, \boldsymbol{x}), \boldsymbol{X} = \boldsymbol{x})\Big)dF_{\boldsymbol{X}}(\boldsymbol{x})}{\int P_t(z', \boldsymbol{x}) - P_t(z, \boldsymbol{x})dF_{\boldsymbol{X}}(\boldsymbol{x})} \qquad (93)
\end{aligned}$$

G Connection between TSLS and Propensity Score Estimations

$E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{fc})$ and $E(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{pm})$ are estimated by interpolation.[48] A simple approach is to estimate $P_t(z) = P(T = t|Z = z)$ and evaluate D_{t_l} and YD_{t_l} as a polynomial of propensity scores:

$$D_{t_l,i} = \theta_0 + \theta_1 P_{t_l,i} + \theta_2 P_{t_l,i}^2 + \epsilon_{i,D} = \boldsymbol{\Lambda}(P_{t_l,i})\boldsymbol{\theta} + \epsilon_{i,D}, \tag{94}$$

$$Y_i D_{t_l,i} = \beta_0 + \beta_1 P_{t_l,i} + \beta_2 P_{t_l,i}^2 + \epsilon_{i,Y} = \boldsymbol{\Lambda}(P_{t_l,i})\boldsymbol{\beta} + \epsilon_{i,D}, \tag{95}$$

where $P_{t_l,i} = P_{t_l}(Z_i)$ denotes the propensity score of t_l for family i, and $\boldsymbol{\Lambda}(p) = [1, p, p^2]$ simply stacks the polynomial as a vector. The estimator for $E(Y(t_l)|\boldsymbol{S} \in \{\boldsymbol{s}_{fc}, \boldsymbol{s}_{pm}\})$ in (96) is numerically the same as the TSLS estimator of Table 6.[49]

$$\hat{E}(Y(t_l)|\boldsymbol{S} \in \{\boldsymbol{s}_{fc}, \boldsymbol{s}_{pm}\}) = \frac{\Big(\boldsymbol{\Lambda}(P_{t_l}(z_e)) - \boldsymbol{\Lambda}(P_{t_l}(z_8))\Big)'\widehat{\boldsymbol{\beta}}_t}{\Big(\boldsymbol{\Lambda}(P_{t_l}(z_e)) - \boldsymbol{\Lambda}(P_{t_l}(z_8))\Big)'\widehat{\boldsymbol{\theta}}_t}, \tag{96}$$

We can evaluate the probability $P^* = P_{t_l}(z_8) + P(\boldsymbol{S} = \boldsymbol{s}_{fc})$ where $P(\boldsymbol{S} = \boldsymbol{s}_{fc})$ is estimated by (**??**). We can disentangle $E(Y(t_l)|\boldsymbol{S} \in \{\boldsymbol{s}_{fc}, \boldsymbol{s}_{pm}\})$ in (96) via:

$$\hat{E}(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{fc}) = \frac{\Big(\boldsymbol{\Lambda}(P^*) - \boldsymbol{\Lambda}(P_{t_l}(z_8))\Big)'\widehat{\boldsymbol{\beta}}_t}{\Big(\boldsymbol{\Lambda}(P^*) - \boldsymbol{\Lambda}(P_{t_l}(z_8))\Big)'\widehat{\boldsymbol{\theta}}_t}, \quad \hat{E}(Y(t_l)|\boldsymbol{S} = \boldsymbol{s}_{pm}) = \frac{\Big(\boldsymbol{\Lambda}(P_{t_l}(z_e)) - \boldsymbol{\Lambda}(P^*)\Big)'\widehat{\boldsymbol{\beta}}_t}{\Big(\boldsymbol{\Lambda}(P_{t_l}(z_e)) - \boldsymbol{\Lambda}(P^*)\Big)'\widehat{\boldsymbol{\theta}}_t}. \tag{97}$$

[48] Recently, Brinch, Mogstad, and Wiswall (2017); Kline and Walters (2017); Mogstad, Andres, and Torgovitsky (2017); Mogstad and Torgovitsky (2018) have studied the problem of assessing ATE in binary choice models with discrete instruments by extrapolating the Marginal Treatment Effect MTE parameter of Heckman and Vytlacil (2005). The method described here differs from this literature in three instances: it investigates a multiple choice model instead of the binary case, it seeks to identify a LATE parameter instead of ATE and it employs interpolation instead of extrapolation.

[49] This equivalence holds for any choice of l.i. functions in $\boldsymbol{\Lambda}(z)$ and for any method that estimates the propensity scores. See Kline and Walters (2019) for a recent discussion on numerical equivalence among estimators for the LATE models.

Table A.7: Elimination of MTO Response-types by the Unordered Monotonicity Condition

Counterfactual Choices	All 27 Possible Response-types																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
$T_i(z_c)$	t_h	t_h	t_h	t_h	t_h	t_h	t_h	t_h	t_h	t_m	t_m	t_m	t_m	t_m	t_m	t_m	t_m	t_m	t_l	t_l	t_l	t_l	t_l	t_l	t_l	t_l	t_l
$T_i(z_8)$	t_h	t_h	t_h	t_m	t_m	t_m	t_l	t_l	t_l	t_h	t_h	t_h	t_m	t_m	t_m	t_l	t_l	t_l	t_h	t_h	t_h	t_m	t_m	t_m	t_l	t_l	t_l
$T_i(z_e)$	t_h	t_m	t_l	t_h	t_m	t_l	t_h	t_m	t_l	t_h	t_m	t_l	t_h	t_m	t_l	t_h	t_m	t_l	t_h	t_m	t_l	t_h	t_m	t_l	t_h	t_m	t_l
Monotonicity 1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓
Monotonicity 2	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓
Monotonicity 3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✓	✗	✓	✓	✗	✓	✓	✗	✓	✓	✗	✓	✓
Monotonicity 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✓	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
Monotonicity 5	✓	✗	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✓	✓	✓	✓	✗	✓
Monotonicity 6	✓	✗	✓	✓	✗	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✓	✗	✓
Monotonicity 7	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗	✓	✓	✓
Monotonicity 8	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓
Monotonicity 9	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗	✗	✓	✗	✗	✓
Not Eliminated	1			4		6			9					14	15												27

The top section of this table lists all the 27 possible response-types that the response variable $S_i = [T_i(z_c)\ ,\ T_i(z_8)\ ,\ T_i(z_e)]$ can take. Rows present the counterfactual neighborhood choices that would arise if a family were assigned to control group, Section 8 and experimental group, that is $T_i(z_c), T_i(z_8)$ and $T_i(z_e)$ respectively. Columns present all the values of response-type as choices range over $supp(T) = \{t_h, t_m, t_l\}$. The second section of this table indicate whether the response-type in the column of the first panel violates any of the following monotonicity relations:

	Z-pairs	Values of T	Unordered Monotonicity Relations		
Monotonicity Relation 1	(z_c, z_8)	t_h	$\mathbf{1}[T_i(z_c) = t_h]$	$\geq$	$\mathbf{1}[T_i(z_8) = t_h]$
Monotonicity Relation 2	(z_8, z_e)	t_h	$\mathbf{1}[T_i(z_8) = t_h]$	$\leq$	$\mathbf{1}[T_i(z_e) = t_h]$
Monotonicity Relation 3	(z_e, z_c)	t_h	$\mathbf{1}[T_i(z_c) = t_h]$	$\geq$	$\mathbf{1}[T_i(z_e) = t_h]$
Monotonicity Relation 4	(z_c, z_8)	t_m	$\mathbf{1}[T_i(z_c) = t_m]$	$\leq$	$\mathbf{1}[T_i(z_8) = t_m]$
Monotonicity Relation 5	(z_8, z_e)	t_m	$\mathbf{1}[T_i(z_8) = t_m]$	$\geq$	$\mathbf{1}[T_i(z_e) = t_m]$
Monotonicity Relation 6	(z_e, z_c)	t_m	$\mathbf{1}[T_i(z_e) = t_m]$	$\leq$	$\mathbf{1}[T_i(z_c) = t_m]$
Monotonicity Relation 7	(z_c, z_8)	t_l	$\mathbf{1}[T_i(z_c) = t_l]$	$\leq$	$\mathbf{1}[T_i(z_8) = t_l]$
Monotonicity Relation 8	(z_8, z_e)	t_l	$\mathbf{1}[T_i(z_8) = t_l]$	$\leq$	$\mathbf{1}[T_i(z_e) = t_l]$
Monotonicity Relation 9	(z_e, z_c)	t_l	$\mathbf{1}[T_i(z_e) = t_l]$	$\geq$	$\mathbf{1}[T_i(z_c) = t_l]$

A check mark sign indicates that the response-type indicated by the column in the top of the table does not violate the choice restriction indicated by the row. A cross sign indicates that the associated response-type violates the relation. The last row of the panel indicates the response-types that are not eliminated by any of the monotonicity relatinons.

H Sensitivity Analyses

This section present additional evaluations that check the robustness of these findings under modifications of the baseline model. I am presenting only the estimations regarding the causal effects for full compliers $\boldsymbol{s}_{fc}$ in order to satisfy the maximum limit of 25 pages for this Appendix. The estimates for counterfactual outcomes, TOT decomposition, additional empirical analyses and specification tests can be obtained by request (rodrig@econ.ucla.edu).

Tables A.8–A.10 presents results based on variations of the original model that generates Table 10 of the main text.

Table A.8 suppresses the interaction of site fixed effects and the propensity scores. Suppressing the interaction between propensity scores and site fixed effects forces cities to shift the the mean potential outcomes for a given neighborhood choice in parallel across all response-types.

Table A.9 suppresses the interaction of baseline variables and propensity scores. This forces that the family baseline characteristics to shift the mean potential outcomes for a given neighborhood choice in parallel.

Table A.10 estimates the same outcome equation displayed in the main text. The model however uses a multinomial logit model to estimate propensity scores, instead of the linear probability model in (50).

Estimates in Tables A.8–A.10 are very close to those presented in Table 10.

Table A.11 presents the estimates of the neighborhood causal effects for the three major cities of the experiment. Most of the estimates align with the those using full sample. However, the sample reduction generates imprecise estimates. Unfortunately, none of the neighborhood effects by city are statistically significant.

Table A.8: Causal Effects for Full Compliers $\boldsymbol{S} = \boldsymbol{s}_4$ (No Site Interaction)

	$E(Y(t_l) - Y(t_h)\|\boldsymbol{s}_4)$		$E(Y(t_l) - Y(t_m)\|\boldsymbol{s}_4)$		$E(Y(t_m) - Y(t_h)\|\boldsymbol{s}_4)$	
Income of Family Head	2.030	***	0.110		1.919	**
(s.e.)	0.761		0.896		0.902	
(p-value)	0.005		0.905		0.032	
Income of Head and Spouse	0.710		0.486		0.224	
(s.e.)	0.826		0.899		1.019	
(p-value)	0.393		0.598		0.800	
Total household income	1.421		1.083		0.337	
(s.e.)	0.871		0.981		1.060	
(p-value)	0.117		0.277		0.752	
Above Poverty Line	0.089	**	0.038		0.051	
(s.e.)	0.039		0.050		0.053	
(p-value)	0.018		0.453		0.348	
Employed without welfare	0.102	**	0.034		0.068	
(s.e.)	0.044		0.057		0.059	
(p-value)	0.027		0.570		0.245	
Currently on welfare	−0.111	***	0.061		−0.172	***
(s.e.)	0.041		0.055		0.057	
(p-value)	0.008		0.270		0.010	
Job tenure	0.073		0.028		0.044	
(s.e.)	0.044		0.052		0.053	
(p-value)	0.102		0.607		0.393	
Economic self-sufficiency	0.062	*	−0.023		0.085	*
(s.e.)	0.033		0.045		0.045	
(p-value)	0.062		0.592		0.055	
Neighborhood Poverty	−32.843	***	−20.999	***	−11.844	***
(s.e.)	0.996		1.690		1.914	
(p-value)	0.000		0.000		0.000	

This table evaluates the neighborhood effects for full compliers $\boldsymbol{s}_4$ across several outcomes. The first column lists the outcome variables. The second column evaluates the causal effect between the neighborhood types of low and high poverty. The third column compares low versus medium poverty neighborhoods and the last column evaluates the neighborhood effects between medium versus high poverty types. The results are based on a semi-parametric method that evaluates propensity scores and response-type probabilities using a linear probability model. All estimates are conditioned on the site of intervention and account for the person-level weight for adult survey of the interim analyses (Interim Impacts Evaluation manual, 2003, Appendix B). Inference is obtained by a bootstrap method that employs a weighted sampling scheme. The p-values are associated with the double-tailed inference that tests if the estimates are equal to zero. Asterisks indicate the typical p-value thresholds: *** for p-value < 0.01, ** for $0.01 \leq p$-value < 0.05, * for $0.05 \leq p$-value < 0.1.

Table A.9: Causal Effects for Full Compliers $\boldsymbol{S} = \boldsymbol{s}_4$ (No Covariate Interaction)

	$E(Y(t_l) - Y(t_h) \mid \boldsymbol{s}_4)$		$E(Y(t_l) - Y(t_m) \mid \boldsymbol{s}_4)$		$E(Y(t_m) - Y(t_h) \mid \boldsymbol{s}_4)$	
Income of Family Head	2.188	***	−0.289		2.477	
(s.e.)	0.822		1.516		1.469	
(p-value)	0.005		0.847		0.112	
Income of Head and Spouse	0.865		0.176		0.689	
(s.e.)	0.862		1.642		1.711	
(p-value)	0.335		0.910		0.663	
Total household income	2.040	**	0.693		1.347	
(s.e.)	0.908		1.621		1.623	
(p-value)	0.035		0.647		0.420	
Above Poverty Line	0.122	***	−0.050		0.173	*
(s.e.)	0.042		0.086		0.084	
(p-value)	0.008		0.607		0.082	
Employed without welfare	0.114	**	0.140		−0.026	
(s.e.)	0.046		0.095		0.097	
(p-value)	0.022		0.157		0.803	
Currently on welfare	−0.130	***	−0.048		−0.081	
(s.e.)	0.044		0.090		0.089	
(p-value)	0.007		0.565		0.343	
Job tenure	0.094	**	0.061		0.033	
(s.e.)	0.047		0.096		0.094	
(p-value)	0.048		0.533		0.723	
Economic self-sufficiency	0.077	**	−0.133		0.210	**
(s.e.)	0.034		0.087		0.084	
(p-value)	0.027		0.177		0.030	
Neighborhood Poverty	−33.283	***	−21.631	***	−11.652	***
(s.e.)	1.008		2.216		2.279	
(p-value)	0.000		0.000		0.000	

This table evaluates the neighborhood effects for full compliers $\boldsymbol{s}_4$ across several outcomes. The first column lists the outcome variables. The second column evaluates the causal effect between the neighborhood types of low and high poverty. The third column compares low versus medium poverty neighborhoods and the last column evaluates the neighborhood effects between medium versus high poverty types. The results are based on a semi-parametric method that evaluates propensity scores and response-type probabilities using a linear probability model. All estimates are conditioned on the site of intervention and account for the person-level weight for adult survey of the interim analyses (Interim Impacts Evaluation manual, 2003, Appendix B). Inference is obtained by a bootstrap method that employs a weighted sampling scheme. The p-values are associated with the double-tailed inference that tests if the estimates are equal to zero. Asterisks indicate the typical p-value thresholds: *** for p-value < 0.01, ** for $0.01 \leq p$-value < 0.05, * for $0.05 \leq p$-value < 0.1.

Table A.10: Causal Effects for Full Compliers $\boldsymbol{S} = \boldsymbol{s}_4$ (Using Multinomial Logit)

	$E(Y(t_l) - Y(t_h)\|\boldsymbol{s}_4)$		$E(Y(t_l) - Y(t_m)\|\boldsymbol{s}_4)$		$E(Y(t_m) - Y(t_h)\|\boldsymbol{s}_4)$	
Income of Family Head	2.471	***	0.887		1.585	
(s.e.)	0.821		1.179		1.225	
(p-value)	0.005		0.453		0.220	
Income of Head and Spouse	1.298		1.278		0.021	
(s.e.)	0.923		1.371		1.399	
(p-value)	0.195		0.350		0.993	
Total household income	2.161	**	2.780	**	−0.619	
(s.e.)	0.954		1.379		1.423	
(p-value)	0.032		0.048		0.668	
Above Poverty Line	0.133	***	0.077		0.057	
(s.e.)	0.048		0.069		0.065	
(p-value)	0.010		0.275		0.408	
Employed without welfare	0.117	**	0.096		0.021	
(s.e.)	0.051		0.079		0.079	
(p-value)	0.023		0.223		0.808	
Currently on welfare	−0.108	**	−0.051		−0.057	
(s.e.)	0.046		0.069		0.068	
(p-value)	0.028		0.462		0.372	
Job tenure	0.120	**	0.053		0.067	
(s.e.)	0.050		0.077		0.079	
(p-value)	0.028		0.457		0.398	
Economic self-sufficiency	0.076	*	−0.044		0.119	*
(s.e.)	0.042		0.059		0.057	
(p-value)	0.087		0.485		0.067	
Neighborhood Poverty	−32.893	***	−22.737	***	−10.157	***
(s.e.)	1.150		1.725		1.995	
(p-value)	0.000		0.000		0.000	

This table evaluates the neighborhood effects for full compliers $\boldsymbol{s}_4$ across several outcomes. The first column lists the outcome variables. The second column evaluates the causal effect between the neighborhood types of low and high poverty. The third column compares low versus medium poverty neighborhoods and the last column evaluates the neighborhood effects between medium versus high poverty types. The results are based on a semi-parametric method that evaluates propensity scores and response-type probabilities using a multinomial logit model. All estimates are conditioned on the site of intervention and account for the person-level weight for adult survey of the interim analyses (Interim Impacts Evaluation manual, 2003, Appendix B). Inference is obtained by a bootstrap method that employs a weighted sampling scheme. The p-values are associated with the double-tailed inference that tests if the estimates are equal to zero. Asterisks indicate the typical p-value thresholds: *** for p-value < 0.01, ** for $0.01 \leq p$-value < 0.05, * for $0.05 \leq p$-value < 0.1.

Table A.11: Causal Effects for Full Compliers $\boldsymbol{S}_{fc}$ (Chicago)

	New York			Los Angeles			Chicago		
	$t_l v.s. t_h$	$t_l v.s. t_m$	$t_m v.s. t_h$	$t_l v.s. t_h$	$t_l v.s. t_m$	$t_m v.s. t_h$	$t_l v.s. t_h$	$t_l v.s. t_m$	$t_m v.s. t_h$
Income of Family Head	5.340	2.445	2.895	1.203	1.431	−0.228	3.245	0.024	3.221
(s.e.)	1.599	2.838	2.967	1.909	2.586	2.717	1.691	2.536	2.587
(p-value)	0.115	0.432	0.382	0.534	0.521	0.934	0.184	0.991	0.360
Income of Head and Spouse	2.709	2.458	0.251	1.170	1.135	0.034	1.789	1.261	0.528
(s.e.)	1.775	2.996	3.045	2.081	3.156	3.275	1.801	2.789	2.938
(p-value)	0.392	0.498	0.919	0.559	0.662	0.983	0.537	0.635	0.830
Total household income	4.380	2.944	1.436	3.267	3.980	−0.712	0.361	−2.669	3.030
(s.e.)	1.893	65.589	65.570	2.233	3.169	3.437	1.896	3.073	3.339
(p-value)	0.244	0.416	0.672	0.233	0.353	0.828	0.859	0.620	0.516
Above Poverty Line	0.238	0.139	0.099	0.203	0.173	0.031	0.079	−0.428	0.506
(s.e.)	0.092	0.155	0.157	0.104	0.143	0.155	0.092	0.161	0.161
(p-value)	0.161	0.449	0.501	0.153	0.358	0.827	0.410	0.613	0.380
Employed without welfare	0.255	0.120	0.135	0.078	0.075	0.003	0.130	−0.109	0.238
(s.e.)	0.099	0.180	0.179	0.114	0.158	0.170	0.100	0.171	0.177
(p-value)	0.263	0.485	0.570	0.481	0.611	0.987	0.312	0.702	0.516
Currently on welfare	−0.171	−0.035	−0.135	0.040	−0.169	0.209	−0.143	0.189	−0.332
(s.e.)	0.089	0.152	0.161	0.105	0.142	0.156	0.086	0.133	0.141
(p-value)	0.172	0.785	0.452	0.857	0.434	0.690	0.181	0.663	0.386
Job tenure	0.085	−0.040	0.125	0.144	0.160	−0.016	0.163	−0.120	0.284
(s.e.)	0.099	0.184	0.188	0.108	0.169	0.181	0.099	0.156	0.160
(p-value)	0.384	0.812	0.451	0.305	0.505	0.916	0.373	0.564	0.425
Economic self-sufficiency	0.117	0.008	0.109	0.111	0.059	0.052	0.085	−0.395	0.481
(s.e.)	0.074	0.133	0.133	0.088	0.135	0.141	0.075	0.133	0.140
(p-value)	0.226	0.936	0.353	0.286	0.694	0.666	0.303	0.383	0.240
Neighborhood Poverty	−37.85	−26.19	−11.66	−33.48	−9.343	−24.14	−33.24	−24.87	−8.37
(s.e.)	2.289	3.939	4.514	2.759	3.801	4.651	2.565	3.408	4.187
(p-value)	0.000	0.000	0.018	0.000	0.893	0.023	0.000	0.000	0.053

This table evaluates the neighborhood effects for full compliers $\boldsymbol{s}_{fc}$ across several outcomes for the three main sites of MTO. The first column lists the outcome variables. The second column evaluates the causal effect between the neighborhood types of low and high poverty. The third column compares low versus medium poverty neighborhoods and the last column evaluates the neighborhood effects between medium versus high poverty types. The results are based on a propensity score estimator using a linear probability model. Outcome equation are estimated using a polynomial of third degree. The estimates are conditioned on baseline variables described in Section 2 of the main paper. All estimates account MTO weights. Inference is obtained by a bootstrap method that employs a weighted sampling scheme. The p-values are associated with the double-tailed inference that tests if the estimates are equal to zero.