Appendix to
"Evidence of Neighborhood Effects from Moving to Opportunity: LATEs of Neighborhood Quality"

Dionissi Aliprantis* Francisca G.-C. Richter⋆

February 15, 2018

A Proofs

A.1 The j to $j+1$ LATE Is Identified by the Wald Estimator

Proposition 1. *Assume the model from the text. Consider the set of observed and unobserved characteristics in the ordered choice model that would result in (i) selection into treatment level j when not receiving the instrument, (ii) treatment level j or $j+1$ with the instrument, and (iii) a positive probability of selection intro treatment level $j+1$ with the instrument. In the text we define this identification support set first using*

$$\Omega_j \equiv \left\{ \; i \in \Omega \;\middle|\; D_0(i) = j, \;\; D_1(i) \in \{j, j+1\}, \;\; Pr(D_1(i) = j+1) > 0 \; \right\},$$

and then

$$\mathcal{S}_j^M \equiv \left\{ \; \big(\mu(x_i), U_D(i)\big) \;\middle|\; i \in \Omega_j \; \right\}.$$

Applying the Wald estimator to the subsample of experimental and control households in $\mathcal{S}_j^M$ identifies the j to $j+1$ transition-specific LATE:

$$\frac{\mathbb{E}\big[\, Y \,\big|\, Z^M = 1, (\mu(X), U_D) \in \mathcal{S}_j^M \,\big] \; - \; \mathbb{E}\big[\, Y \,\big|\, Z^M = 0, (\mu(X), U_D) \in \mathcal{S}_j^M \,\big]}{\mathbb{E}\big[\, D \,\big|\, Z^M = 1, (\mu(X), U_D) \in \mathcal{S}_j^M \,\big] \; - \; \mathbb{E}\big[\, D \,\big|\, Z^M = 0, (\mu(X), U_D) \in \mathcal{S}_j^M \,\big]} = \triangle_{j,j+1}^{LATE}(Z^M).$$

*: Federal Reserve Bank of Cleveland, +1(216)579-3021, *dionissi.aliprantis@clev.frb.org*

⋆: Case Western Reserve University, +1(216)368-8686, *francisca.richter@case.edu*

Disclaimer: The views stated herein are those of the authors and are not necessarily those of the Federal Reserve Bank of Cleveland or the Board of Governors of the Federal Reserve System.

Proof. Recall that $W_Z^M(i)$ is an indicator for individual i's counterfactual voucher take up. When offered a voucher ($Z_i^M = 1$), a household in $\mathcal{S}_j^M$ can possibly respond with any of the following mutually exclusive options:

1 Not use the MTO voucher and remain in a neighborhood of quality j; we denote this set of households by $S_j^M(W_1^M(i) = 0, D(i) = j)$.

2 Move with the MTO voucher, but not to a higher quality neighborhood, denoted by $S_j^M(W_1^M(i) = 1, D(i) = j)$.[1]

3 Move with the MTO voucher to a higher quality neighborhood, denoted by $S_j^M(W_1^M(i) = 1, D(i) = j+1)$.

We can then classify households into non-compliers and compliers as follows:

$$\mathcal{NC}_j^M = S_j^M(W_1^M(i) = 0, D(i) = j) \ \sqcup \ S_j^M(W_1^M(i) = 1, D(i) = j)$$

$$\mathcal{C}_j^M = S_j^M(W_1^M(i) = 1, D(i) = j+1),$$

where $\sqcup$ represents a disjoint union. We denote the probability of being a complier as $\pi(\mathcal{C}_j^M)$.

The Wald estimator applied to the subsample of experimental and control households in $\mathcal{S}_j^M$ identifies the j to $j+1$ transition-specific LATE for compliers:

$$\begin{aligned}
&\frac{\mathbb{E}\left[\, Y \mid Z^M = 1, (\mu(X), U_D) \in \mathcal{S}_j^M \,\right] - \mathbb{E}\left[\, Y \mid Z^M = 0, (\mu(X), U_D) \in \mathcal{S}_j^M \,\right]}{\mathbb{E}\left[\, D \mid Z^M = 1, (\mu(X), U_D) \in \mathcal{S}_j^M \,\right] - \mathbb{E}\left[\, D \mid Z^M = 0, (\mu(X), U_D) \in \mathcal{S}_j^M \,\right]} \\
&\quad = \frac{\mathbb{E}\left(Y_{j+1} - Y_j \mid (\mu(X), U_D, W_1^M) \in \mathcal{C}_j^M \right) \pi(\mathcal{C}_j^M)}{\pi(\mathcal{C}_j^M)} \\
&\quad = \mathbb{E}\left[\, Y_{j+1} - Y_j \mid (\mu(X), U_D, W_1^M) \in \mathcal{C}_j^M \,\right] \\
&\quad \equiv \triangle_{j,j+1}^{LATE}(Z^M),
\end{aligned} \tag{1}$$

[1] One possibility generating this case is that the household moves to a low-poverty neighborhood of the same quality. Another possibility is that because the interim study was conducted four to seven years after randomization, households could have moved more than once, with their final move being to a neighborhood of quality level j.

where the equality in Equation 1 is derived as follows,

$$
\begin{aligned}
&\mathbb{E}[\, Y \mid Z^M = 1, (\mu(X), U_D) \in \mathcal{S}_j^M \,] \;-\; \mathbb{E}[\, Y \mid Z^M = 0, (\mu(X), U_D) \in \mathcal{S}_j^M \,] \\
&\quad = \Big\{ \mathbb{E}[\, Y_{j+1} \mid (\mu(X), U_D, W_1^M) \in \mathcal{C}_j^M \,] \pi(\mathcal{C}_j^M) \;+\; \mathbb{E}[\, Y_j \mid (\mu(X), U_D, W_1^M) \in \mathcal{NC}_j^M \,](1 - \pi(\mathcal{C}_j^M)) \Big\} \\
&\qquad - \Big\{ E[\, Y_j \mid (\mu(X), U_D, W_1^M) \in \mathcal{C}_j^M \,] \pi(\mathcal{C}_j^M) \;+\; \mathbb{E}[\, Y_j \mid (\mu(X), U_D, W_1^M) \in \mathcal{NC}_j^M \,](1 - \pi(\mathcal{C}_j^M)) \Big\} \\
&\quad = \Big(\mathbb{E}[\, Y_{j+1} \mid (\mu(X), U_D, W_1^M) \in \mathcal{C}_j^M \,] \;-\; \mathbb{E}[\, Y_j \mid (\mu(X), U_D, W_1^M) \in \mathcal{C}_j^M \,] \Big) \pi(\mathcal{C}_j^M) \\
&\qquad + \Big(\mathbb{E}[\, Y_j \mid (\mu(X), U_D, W_1^M) \in \mathcal{NC}_j^M \,] \;-\; \mathbb{E}[\, Y_j \mid (\mu(X), U_D, W_1^M) \in \mathcal{NC}_j^M \,] \Big) (1 - \pi(\mathcal{C}_j^M)) \\
&\quad = \Big(\mathbb{E}[\, Y_{j+1} \mid (\mu(X), U_D, W_1^M) \in \mathcal{C}_j^M \,] \;-\; \mathbb{E}[\, Y_j \mid (\mu(X), U_D, W_1^M) \in \mathcal{C}_j^M \,] \Big) \pi(\mathcal{C}_j^M)
\end{aligned}
$$

and likewise,

$$
\begin{aligned}
&\mathbb{E}[\, D \mid Z^M = 1, (\mu(X), U_D) \in \mathcal{S}_j^M \,] \;-\; \mathbb{E}[\, D \mid Z^M = 0, (\mu(X), U_D) \in \mathcal{S}_j^M \,] \\
&\quad = \Big\{ (j+1)\pi(\mathcal{C}_j^M) + (j)(1 - \pi(\mathcal{C}_j^M)) \Big\} \;-\; \Big\{ (j)\pi(\mathcal{C}_j^M) + (j)(1 - \pi(\mathcal{C}_j^M)) \Big\} \\
&\quad = \Big[(j+1) - j \Big] \pi(\mathcal{C}_j^M) + \Big[j - j \Big] (1 - \pi(\mathcal{C}_j^M)) \\
&\quad = \pi(\mathcal{C}_j^M).
\end{aligned}
$$

□

Note that this entire identification strategy is predicated on identifying V_i, as it is required for identifying $\mathcal{S}_j^M$ and applying the Wald estimator to households in this set.

A.2 Equivalence of the Unobserved Component of Choice Specified Under Discrete and Continuous Models of Neighborhood Quality

Suppose that we observe a continuous measure of neighborhood quality $q \in [0, 1]$. If we partition quality to generate discrete treatment levels, there will be an associated ordered choice model. We can estimate the unobserved component V of the ordered choice model using the identification strategy in the text.

We interpret estimates obtained in this way to be estimates of the same random variable V regardless of the partition of quality used in estimation. We justify this interpretation by showing that a sequence of $\{V^n(i)\}_{n=1}^{\infty}$ derived from a sequence of refinements of quality converging in the norm will converge to the random variable $V(i)$ from a continuous model (Corollary 2).

In practice, this result allows us to move freely between different partitions of quality when estimating the ordered choice model and when estimating causal effects on outcomes. This means that we use one partition of quality when estimating the ordered choice model, using the sample population to determine where cutpoints are located so as to improve estimation. We use a different partition of quality when estimating causal effects on outcomes, placing cutpoints so as to characterize meaningful margins of neighborhood characteristics.

Before stating our results, we will first provide some relevant notation and definitions. Let

$$MB(q) = \mu(x_i) - C(q) - V(i) = 0 \tag{2}$$

be the First Order Condition determining neighborhood quality selection q in a continuous model of choice, where all variables are defined as in the text. Consider a partition of the continuous quality measure $q \in [\alpha, 1] \subset (0, 1]$ into K discrete levels:

$$\mathcal{P}^q(K) = \{\alpha = q_0, q_1, \ldots, q_{K-1}, q_K = 1\}.$$

Partition $\mathcal{P}^q(K)$ defines the discrete treatment

$$D^K = \begin{cases} 1 & \text{if } q \in [q_0, q_1]; \\ 2 & \text{if } q \in (q_1, q_2]; \\ \vdots & \vdots \quad \vdots \\ K & \text{if } q \in (q_{K-1}, q_K]. \end{cases}$$

Given a continuous cost function $C(q) : [\alpha, 1] \to B \subset \mathcal{R}$, partition $\mathcal{P}^q(K)$ implies a partition of the image of C

$$\mathcal{P}^C(K) = \{C_0, C_1, \ldots, C_{K-1}, C_K\},$$

where $C_k = C(q_k)$.

$\mathcal{P}^q(K)$ and $\mathcal{P}^C(K)$ define a partition of the closed interval $[\mu(x_i) - C_K,\ \mu(x_i) - C_0] \subset \mathbb{R}$:

$$\mathcal{P}^V(K) = \{\mu(x_i) - C_K, \ldots, \mu(x_i) - C_0\}.$$

The selection conditions for the discrete model of neighborhood quality choice associated with the continuous model 2 and partition $\mathcal{P}^q(K)$ are:

$$D^K(i) = k \iff \mu(x_i) - C(q_k) < V(i) \leq \mu(x_i) - C(q_{k-1}) \qquad \text{for} \qquad k = 1, \ldots, K. \quad (3)$$

Some definitions that we use in our proof are as follows:

Partition-Specific Discrete Model: We say that the ordered choice model given by Condition 3 is the $\mathcal{P}^q(K)$-discrete model associated with the continuous choice model in Equation 2.

Refinement of a Partition: $\mathcal{P}^q_n$ is a refinement of $\mathcal{P}^q(K_0)$ if $\mathcal{P}^q(K_0) \subset \mathcal{P}^q_n$, with $\subset$ representing strict inclusion.

Sequence of Refinements of a Partition: $\{\mathcal{P}^q_n\}_{n=1}^{\infty}$ is a sequence of refinements of $\mathcal{P}^q(K_0)$ if $\mathcal{P}^q(K_0) \subset \mathcal{P}^q_1$ and $\mathcal{P}^q_{n-1} \subset \mathcal{P}^q_n$ for $n = 2, \ldots, \infty$. The n^{th} refinement in the sequence is denoted by $\mathcal{P}^q_n = \{\alpha = q_{0,n}, q_{1,n}, \ldots, q_{N_n,n} = 1\}$, where $N_n + 1$ is the cardinality of the refinement.

Norm of a Partition: The norm of $\mathcal{P}^q_n$ is $\max_{k \in \{1,\ldots,N_n\}} |q_{k,n} - q_{k-1,n}|$.

Convergence in the Norm: The norm of the refinements in the sequence $\{\mathcal{P}^q_n\}_{n=1}^{\infty}$ converges to zero if

$$\lim_{n\to\infty} \max_{k \in \{1,\ldots,N_n\}} |q_{k,n} - q_{k-1,n}| = 0. \quad (4)$$

Proposition 2. *Let the $\mathcal{P}^q(K_0)$-discrete model be associated with the continuous choice model in Equation 2. Define a sequence of refinements of $\mathcal{P}^q(K_0)$ indexed by n, $\{\mathcal{P}_n^q\}_{n=1}^{\infty}$, such that the norm of the refinements converges to zero. The sequence of refinements of $\mathcal{P}^C(K_0)$ generated by $\{\mathcal{P}_n^q\}_{n=1}^{\infty}$ also converges to zero in the norm.*

Proof. We want to show that

$$\lim_{n\to\infty} \max_{k\in\{1,\ldots,N_n\}} |C(q_{k,n}) - C(q_{k-1,n})| = 0. \tag{5}$$

This is equivalent to:

$$\forall\ \epsilon > 0, \exists\, m^* \text{ such that } |C(q_{k,n}) - C(q_{k-1,n})| < \epsilon,\ \ \forall\, k \in \{1,\ldots,N_n\},\ \text{ when } n > m^*. \tag{6}$$

We know that every continuous function on a closed and bounded interval is uniformly continuous (Heine-Cantor Theorem, Aliprantis and Border (2006) Corollary 3.31). Thus, $C : [\alpha, 1] \to B \subset \mathcal{R}$ satisfies:

$$\forall\ \epsilon > 0, \exists\, \delta > 0 \text{ such that } |q - p| < \delta \Rightarrow |C(q) - C(p)| < \epsilon,\ \ \forall\ p, q \in [\alpha, 1]. \tag{7}$$

The condition that the norm of refinements in the series $\{\mathcal{P}_n^q\}_{n=1}^{\infty}$ converges to zero (condition 4) can be restated as

$$\forall\ \delta > 0, \exists\ n^* \text{ such that } |q_{k,n} - q_{k-1,n}| < \delta,\ \ \forall\, k \in \{1,\ldots,N_n\},\ \text{ when } n > n^*. \tag{8}$$

Take $\epsilon_0 > 0$. By uniform continuity of $C(q)$ there exists $\delta(\epsilon_0)$ that satisfies

$$|C(q) - C(p)| < \epsilon_0,\ \ \forall\ p, q \in [\alpha, 1] \text{ whenever } |q - p| < \delta(\epsilon_0).$$

Given Condition 8, we can find $n^*(\delta(\epsilon_0))$ that satisfies

$$|q_{k,n} - q_{k-1,n}| < \delta(\epsilon_0),\ \ \forall\, k \in \{1,\ldots,N_n\},\ \text{ when } n > n^*(\delta(\epsilon_0)).$$

Thus, given ϵ_0, we can find $m^* = n^*(\delta(\epsilon_0))$ such that

$$|C(q_{k,n}) - C(q_{k-1,n})| < \epsilon_0,\ \ \forall\, k \in \{1,\ldots,N_n\},\ \text{ when } n > m^* = n^*(\delta(\epsilon_0)),$$

which satisfies Condition 6. □

Corollary 1. *The sequence of refinements of $\mathcal{P}^V(K_0)$ converges to zero in the norm if it is generated by a sequence of partitions $\{\mathcal{P}^q_n\}_{n=1}^{\infty}$ that converges to zero in the norm.*

Proof. By stating the definition and using the fact that taking the limit is compatible with algebraic operations, we know that

$$\lim_{n\to\infty}\left(\max_{k=1,\ldots,K_n} \Big|[\mu(x_i)-C(q_{k-1,n})]-[\mu(x_i)-C(q_{k,n})]\Big|\right) = \lim_{n\to\infty}\left(\max_{k=2,\ldots,K-1} \big|C(q_{k,n})-C(q_{k-1,n})\big|\right).$$

By Proposition 2, we know that the right hand side of this equation is equal to 0. $\square$

Corollary 2. *For a sequence of partitions $\{\mathcal{P}^q_n\}_{n=1}^{\infty}$ that converges in the norm, a sequence of $\{V^n(i)\}_{n=1}^{\infty}$ satisfying Equation 3 converges to $V(i)$ in Equation 2.*

Proof. By Equation 3 we know that for the $V(i)$ in Equation 2,

$$V(i) \in \big(\mu(x_i)-C(q_{k,n}), \mu(x_i)-C(q_{k-1,n})\big] \text{ when } D^K(i)=k.$$

By construction, the $V^n(i)$ in our identification strategy is also in this same interval:

$$V^n(i) \in \big(\mu(x_i)-C(q_{k,n}), \mu(x_i)-C(q_{k-1,n})\big] \text{ when } D^K(i)=k.$$

It follows that

$$\lim_{n\to\infty}\left(\big|V^n(i)-V(i)\big|\right) \leq \lim_{n\to\infty}\left(\max_{k=1,\ldots,K_n} \Big|[\mu(x_i)-C(q_{k-1,n})]-[\mu(x_i)-C(q_{k,n})]\Big|\right).$$

We know from Corollary 1 that the right hand side of this inequality is 0. $\square$

A.2.1 Discussion of Intuition

Corollary 2 can be interpreted as the discrete choice conditions of Equation 3 converging to the continuous first order condition of Equation 2. The implication for interpreting the estimates in our model comes from the fact that in practice, $V(i)$ is estimated from a given partition of quality. Corollary 2 assures us that a sequence of $V^n(i)$'s derived from a series of refinements of that partition satisfying Condition 4 will converge to $V(i)$ from the same continuous model, regardless of the initial partition.

We now seek to add intuition to the link between the continuous and discrete models, and to discuss how a distributional assumption on V can be seen as a normalization due to the flexibility of the cost function. Suppose there is a continuous measure of neighborhood quality $q \in [\alpha, 50]$ for arbitrary $\alpha > 0$, and that there are two partitions into discrete levels of quality, where under the first partition

$$Q_i^3 = \begin{cases} 1 & \text{if } q_i \in [q_0, q_1] = [\alpha, 10]; \\ 2 & \text{if } q_i \in (q_1, q_2] = (10, 40]; \\ 3 & \text{if } q_i \in (q_2, q_3] = (40, 50], \end{cases}$$

and under the second partition (a refinement of the first partition)

$$Q_i^5 = \begin{cases} I & \text{if } q_i \in [q_0, q_I] = [\alpha, 10]; \\ II & \text{if } q_i \in (q_I, q_{II}] = (10, 20]; \\ III & \text{if } q_i \in (q_{II}, q_{III}] = (20, 30]; \\ IV & \text{if } q_i \in (q_{III}, q_{IV}] = (30, 40]; \\ V & \text{if } q_i \in (q_{IV}, q_V] = (40, 50]. \end{cases}$$

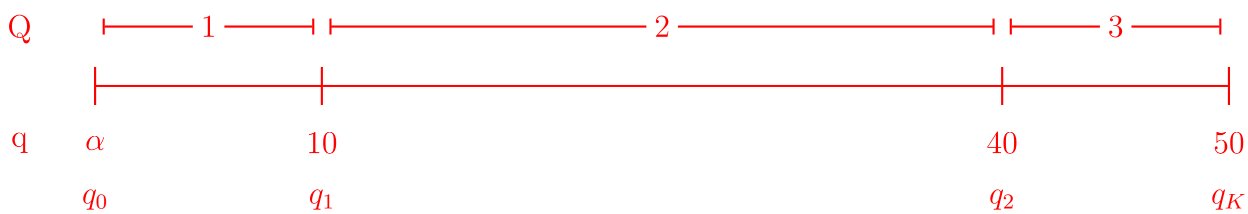
Focusing on observed characteristics for a particular realization $X(i) = x_i$, we have that

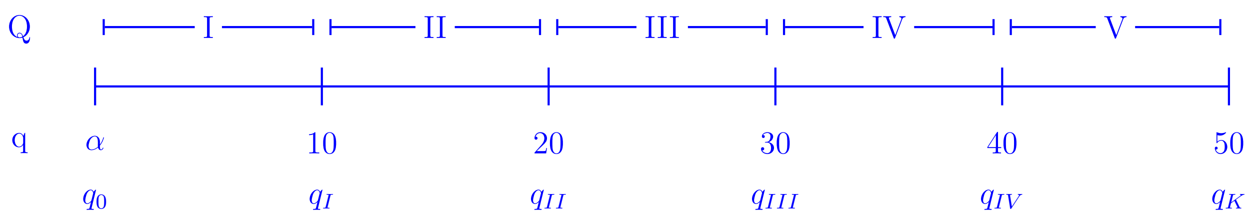
$$\begin{aligned} Pr(q_i^* > 40) = Pr(Q_i^3 = 3) = Pr(q_i^* > q_2) = \Phi(\mu(x_i) - C_2) \\ = Pr(Q_i^5 = V) = Pr(q_i^* > q_{IV}) = \Phi(\mu(x_i) - C_{IV}). \end{aligned}$$

Thus $C_2 = C_{IV}$, so the values at the common cut point/knot will be the same under both partitions, with $C(40) = C(40) = C_2 = C_{IV}$. The same logic applies to see that

$$\begin{aligned} Pr(q_i^* > 10) = Pr(Q_i^3 \geq 2) = Pr(q_i^* > q_1) = \Phi(\mu(x_i) - C_1) \\ = Pr(Q_i^5 \geq II) = Pr(q_i^* > q_I) = \Phi(\mu(x_i) - C_I). \end{aligned}$$

Thus we will likewise have $C(10) = C(10) = C_1 = C_I$. The partitions and cut points/knots from this example are illustrated below.





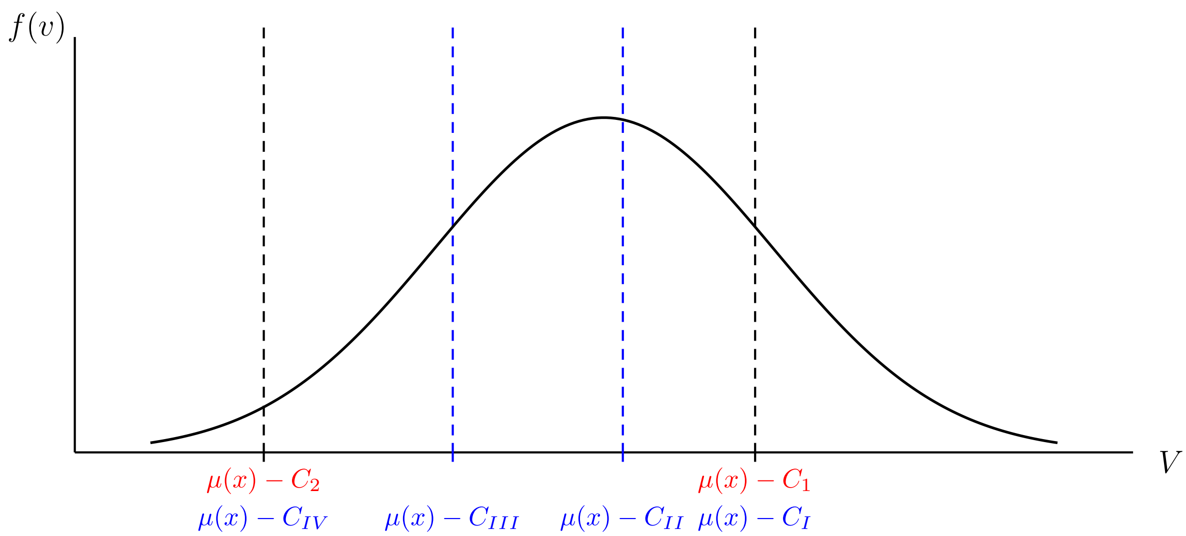


Figure 1 illustrates the relationship between the linearly interpolation implied by the partitions $\mathcal{P}_3^q$ and $\mathcal{P}_V^q$ along with $C(q)$. The linear interpolation $C_n(q)$ between $C(q_{k,n})$ and $C(q_{k-1,n})$ for $q \in (q_{k-1}, q_k)$ can be made to approximate the true continuous cost function $C(q)$ to an arbitrary degree of accuracy. That is, for all $\epsilon > 0$, there exists some $m^* \in \mathbb{N}$ such that $n > m^*$ implies that the norm of the the partition $\mathcal{P}_n^q$ is less than $\delta_\epsilon > 0$. By the uniform continuity of $C(q)$, if C is increasing, then linear interpolation of $C_n(q)$ implies that $|C_n(q) - C(q)| \leq \max_{k\in\{1,\ldots,K\}} |C(q_{k,n}) - C(q_{k-1,n})|\,] < \epsilon$ for all $q \in [\alpha, 50]$ when $C_n(q)$ is estimated using the partition $\mathcal{P}_n^q$.

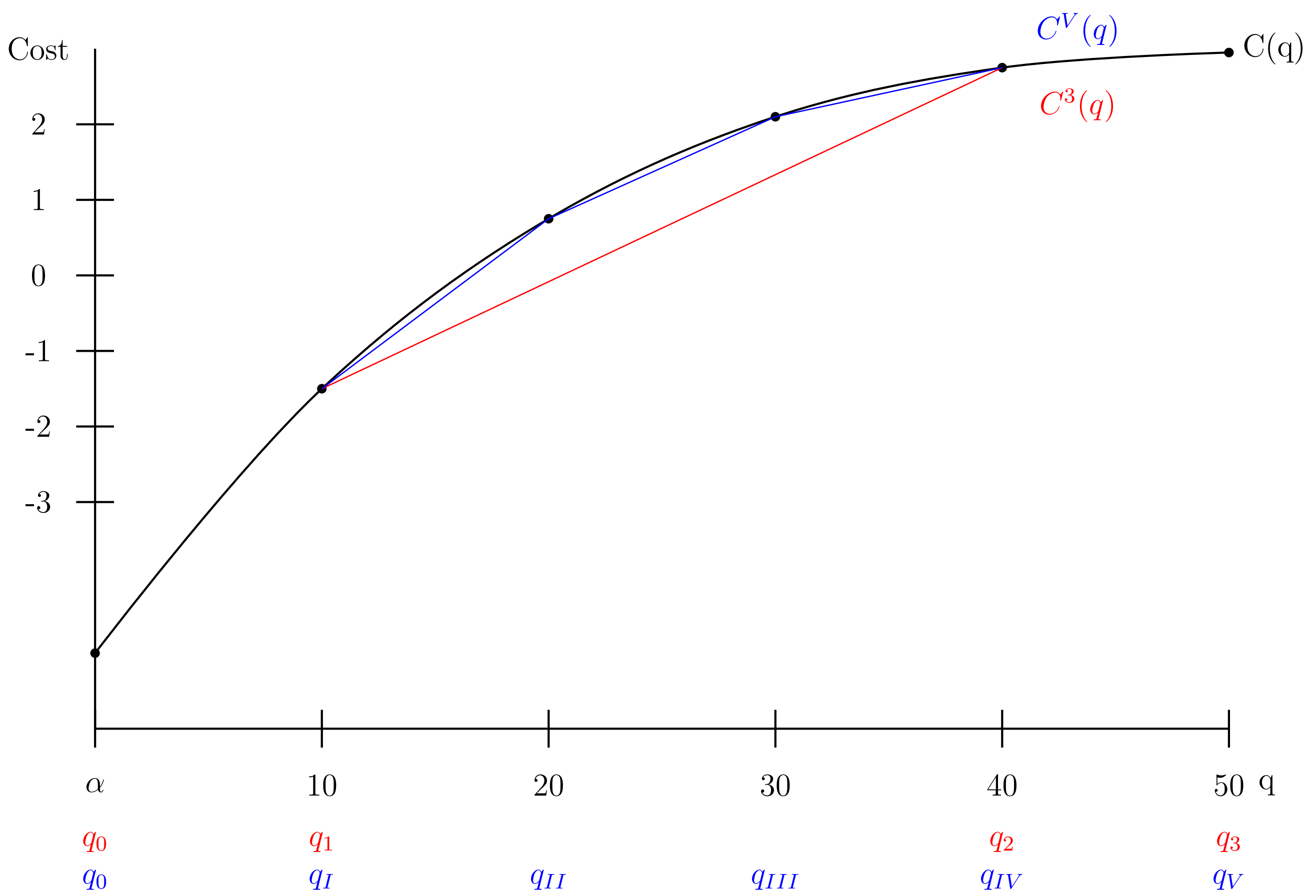


Figure 1: Approximating $C(q)$ to Arbitrary Accuracy

B Estimation Algorithm

The general estimation algorithm is as follows:

Step 1 Estimate the ordered choice model to obtain $\widehat{\mu}(x_i)$, $\{\widehat{C}_k\}$, $\{\widehat{\gamma}_k^S\}$, and $\{\widehat{\gamma}_k^M\}$.

Step 2 Linearly interpolate to obtain $\widehat{C}(q)$, $\widehat{\gamma}^S(q)$, and $\widehat{\gamma}^M(q)$. In the case of $\widehat{C}(q)$,

$$\widehat{C}(q) = \widehat{C}_k + (q - \overline{q}_k)\left(\frac{\widehat{C}_{k+1} - \widehat{C}_k}{\overline{q}_{k+1} - \overline{q}_k}\right) \quad \text{for } q \in (\overline{q}_k, \overline{q}_{k+1}).$$

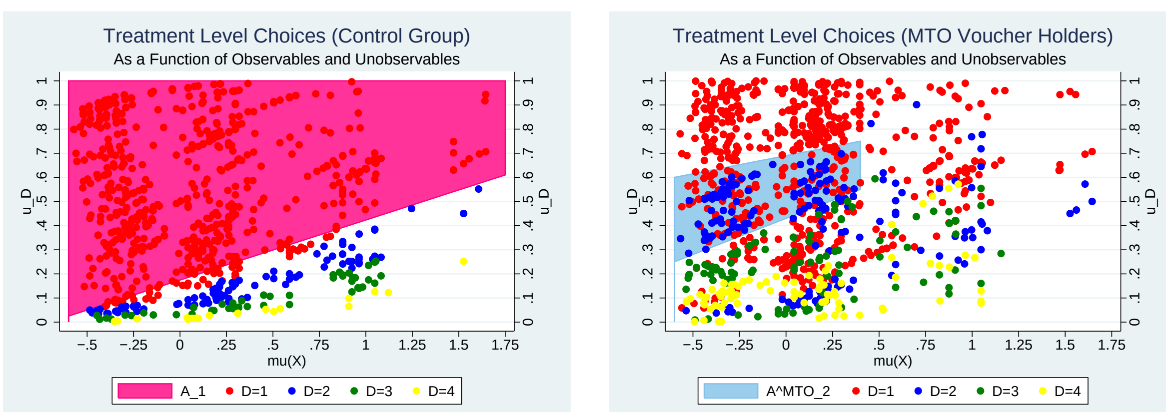
Step 3 Estimate $\widehat{V}$ using the FOC

$$\widehat{V}(i) = \widehat{\mu}(x_i) - \widehat{C}(q_i^*) + \widehat{\gamma}^S(q_i^*)z_i^S w_i^S + \widehat{\gamma}^M(q_i^*)z_i^M w_i^M$$

and $\widehat{U}_D(i)$ via

$$\widehat{U}_D(i) = \Phi(\widehat{V}(i)).$$

Step 4 Using the control group, find an area $\widehat{\mathcal{A}}_j \subset \mathcal{M} \times [0,1]$ such that households with $(\widehat{\mu}(x_i), \widehat{u}_{Di}) \in \widehat{\mathcal{A}}_j$ would select into neighborhood quality $D_i = j$ without any voucher. Using the MTO voucher group, find the subset $\widehat{\mathcal{A}}_{j,j+1}^M$ for which some households would select into neighborhood quality $D_i = j+1$ with an MTO voucher. The identification support set is $\widehat{\mathcal{S}}_j^M \equiv \widehat{\mathcal{A}}_j \cap \widehat{\mathcal{A}}_{j,j+1}^M$.



(a) Control Group and $\widehat{\mathcal{A}}_1$

(b) MTO Voucher Holders and $\widehat{\mathcal{A}}_{1,2}^M$

Figure 2: Selection into Treatment and Counterfactual Areas $\widehat{\mathcal{A}}_1$ and $\widehat{\mathcal{A}}_{1,2}^M$

Step 5 Estimate the j to $j+1$ transition-specific LATE over $\widehat{\mathcal{S}}_j^M$ using the Wald estimator from Equation 1 applied to $\widehat{\mathcal{S}}_j^M$.

Step 6 Bootstrap by repeating the following steps T times:

Step 6a Sample with replacement

Step 6b Repeat Step 1: Estimate the ordered choice model on the new sample

Step 6c Repeat Step 3: Calculate $\mathbb{E}[\widehat{\triangle}_{j,j+1}^{LATE}(Z^M)|(\widehat{\mu}(x_i), \widehat{u}_{Di}) \in \widehat{\mathcal{S}}_j^M]$ on the new sample where the set $\widehat{\mathcal{S}}_j^M$ maintains the definition determined in Step 2 for the original sample

Construct standard errors using the T parameter estimates.

C Specification of the Full Likelihood Function

Recall that $V(i)$ represents the unobserved cost for household i of moving up in the absence of a voucher program. We will use $V^S(i)$ and $V^M(i)$ to denote unobserved variables influencing the decision of household i to take up a Section 8 voucher and an MTO voucher when these are offered. We allow for these variables to be correlated in an arbitrary way, possibly exhibiting patterns of correlation anywhere between being exactly identical variables to being independently distributed variables to being negatively correlated variables. This is the maintained assumption required to identify the j to $j+1$ LATE parameter.

Because we would also like to understand the relationship between $V(i)$, $V^S(i)$, and $V^M(i)$, we adopt additional assumptions. We further assume that whether a household moves with a voucher is determined by the latent index models:

$$W^S(i) = \mathbf{1}\{\ \mu^S(x_i) - V_i^S \ \geq\ 0\ \} \quad = \mathbf{1}\{\ \beta_1^S x_1 + \cdots + \beta_8^S x_8 - v_i^S \ \geq\ 0\ \},$$
$$W^M(i) = \mathbf{1}\{\ \mu^M(x_i) - V_i^M \ \geq\ 0\ \} \quad = \mathbf{1}\{\ \beta_1^M x_1 + \cdots + \beta_8^M x_8 - v_i^M \ \geq\ 0\ \}.$$

under the distributional assumption:

$$\mathbf{V}(i) \equiv \big(V(i), V^S(i), V^M(i)\big) \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho^S & \rho^M \\ \rho^S & 1 & \rho^{SM} \\ \rho^M & \rho^{SM} & 1 \end{bmatrix}\right).$$

We stress that the role of assuming this joint normal distribution in aiding identification is entirely for understanding the relationship between $V(i)$, $V^S(i)$, and $V^M(i)$. No assumption on the joint distribution of these variables is necessary to identify the j to $j+1$ LATE identified in the paper, beyond the standard normal distribution assumed of $V(i)$.

The marginal distributions implied by this joint distribution are as follows:

$$V(i) \sim \mathcal{N}(0,1), \tag{9}$$
$$(V(i), V^S(i)) \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_S \\ \rho_S & 1 \end{bmatrix}\right), \quad \text{and} \tag{10}$$
$$(V(i), V^M(i)) \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_M \\ \rho_M & 1 \end{bmatrix}\right). \tag{11}$$

We also know from the latent index models that the probabilities of moving when offered a Section 8 and experimental MTO voucher are $P(V^S(i) \leq \mu^S(x_i))$ and $P(V^M(i) \leq \mu^M(x_i))$, respectively.

For the *Control* Group we do not observe whether the household would move with either type of voucher, while for both the *Section 8* and *Experimental MTO* voucher groups we observe whether the household is a "mover" or a "never-mover" with respect to the voucher they received (ie, whether they moved when offered that type of voucher). Together with actually observing the household's W^S when $z^S = 1$ and W^M when $z^M = 1$, the ordered choice condition in Equation 1 in the text and the marginal distributions in Equation 9 above allow us to express the probability of observing $D(i) = k$ for households in each observed group. Where $\Phi_2(a, b; \rho)$ is the cumulative distribution function of the standardized bivariate normal distribution with correlation coefficient ρ, these probabilities are:

Control Group

$$\Pr\big(D(i) = k \mid x_i, z_i^S = 0, z_i^M = 0\big) = \Phi\big(\ \mu(x_i) - C_{k-1}\ \big) - \Phi\big(\ \mu(x_i) - C_k\ \big) \tag{12}$$

Section 8 Voucher Movers and Non-Movers

$$\begin{aligned}\Pr\big(D(i) = k \mid x_i, z_i^S = 1, x_i^M = 0, w_i^S = 1\big) = {} & \Phi_2\big(\ \mu(x_i) + \gamma_{k-1}^S - C_{k-1},\quad \mu^S(x_i);\quad \rho_S\ \big) \\ & - \Phi_2\big(\ \mu(x_i) + \gamma_k^S - C_k,\quad \mu^S(x_i);\quad \rho_S\ \big)\end{aligned}$$

$$\begin{aligned}\Pr\big(D(i) = k \mid x_i, z_i^S = 1, z_i^M = 0, w_i^S = 0\big) = {} & \Phi_2\big(\ \mu(x_i) - C_{k-1},\quad -\mu^S(x_i);\quad \rho_S\ \big) \\ & - \Phi_2\big(\ \mu(x_i) - C_k,\quad -\mu^S(x_i);\quad \rho_S\ \big)\end{aligned}$$

MTO Voucher Movers and Non-Movers

$$\begin{aligned}\Pr\big(D(i) = k \mid x_i, z_i^S = 0, z_i^M = 1, w_i^M = 1\big) = {} & \Phi_2\big(\ \mu(x_i) + \gamma_{k-1}^M - C_{k-1},\quad \mu^M(x_i);\quad \rho_M\ \big) \\ & - \Phi_2\big(\ \mu(x_i) + \gamma_k^M - C_k,\quad \mu^M(x_i);\quad \rho_M\ \big)\end{aligned}$$

$$\begin{aligned}\Pr\big(D(i) = k \mid x_i, z_i^S = 0, z_i^M = 1, w_i^M = 0\big) = {} & \Phi_2\big(\ \mu(x_i) - C_{k-1},\quad -\mu^M(x_i);\quad \rho_M\ \big) \\ & - \Phi_2\big(\ \mu(x_i) - C_k,\quad -\mu^M(x_i);\quad \rho_M\ \big).\end{aligned}$$

These probabilities allow us to identify the parameters of the ordered choice model by expressing its log-likelihood function as:

$$
\begin{aligned}
\mathcal{LL}(\theta|\mathbf{X},\mathbf{Z},\mathbf{D},\mathbf{W}) &= \sum_{i=1}^{N}\sum_{k=1}^{K}\mathbf{1}\{d_i=k\}\,ln\Big(\Pr\big(D(i)=k\big|x_i,z_i^S,z_i^M,w_i^S,w_i^M\big)\Big)\\
&= \sum_{i=1}^{N^0}\sum_{k=1}^{K}\mathbf{1}\{d_i=k\}\,ln\Big(\Pr\big(D(i)=k\big|x_i,z_i^S=0,z_i^M=0\big)\Big) \qquad (13)\\
&\quad+\sum_{i=1}^{N^S}\sum_{k=1}^{K}\mathbf{1}\{d_i=k\}\,ln\Big(\sum_{t=0}^{1}\Pr\big(D(i)=k\big|x_i,z_i^S=1,w_i^M=0,w_i^S=t\big)1(w_i^S=t)\Big)\\
&\quad+\sum_{i=1}^{N^M}\sum_{k=1}^{K}\mathbf{1}\{d_i=k\}\,ln\Big(\sum_{t=0}^{1}\Pr\big(D(i)=k\big|x_i,z_i^S=0,z_i^M=1,w_i^M=t\big)1(w_i^M=t)\Big).
\end{aligned}
$$

The parameter estimates of the full model are presented below.

Table 1: Ordered Choice Model Parameter Estimates

X_k and $\mathbf{V}$	$\widehat{\beta}_k$	$\widehat{\beta}_k^S$	$\widehat{\beta}_k^M$
Baseline Characteristics			
Teens in HH	–0.08	–0.48	–0.39
	(0.05)	(0.10)	(0.09)
Family in Nbd	–0.14	–0.16	0.00
	(0.05)	(0.12)	(0.06)
HH Member Victim	0.03	0.10	0.10
	(0.05)	(0.10)	(0.10)
Baseline Nbd Quality	0.13	–0.02	–0.10
	(0.01)	(0.03)	(0.02)
Site Fixed Effects/Constant			
Baltimore	0	0.03	–0.25
	–	(0.13)	(0.13)
Boston	0.31	–0.49	0.02
	(0.10)	(0.21)	(0.02)
Chicago	–0.04	0.04	–0.50
	(0.09)	(0.13)	(0.13)
Los Angeles	–0.52	0.39	0.47
	(0.10)	(0.17)	(0.13)
New York City	–0.58	0.60	–0.13
	(0.09)	(0.15)	(0.15)
Unobserved Factors			
ρ^S and ρ^M	–	0.07	–0.17
	–	(0.10)	(0.11)

References

Aliprantis, C. D. and K. C. Border (2006). *Infinite Dimensional Analysis* (Third ed.). Berlin: Springer.