Online Appendix

Neighborhoods Matter: Assessing the Evidence for Place Effects

Eric Chyn and Lawrence F. Katz

June 16, 2021

A. Econometric Models of Neighborhood Effects

To aid our review of the empirical literature on neighborhood effects, we present formal econometric models of individual outcomes that allow for contemporaneous and developmental neighborhood effects and discuss identification issues. Let y_{it}^{j} denote outcome j (e.g., earnings or health status) of individual i in year t. Define the index $n(i)t$ to denote the neighborhood where individual i lives in year t, and let $c(i,a)$ represent the neighborhood in which individual i grew up at ages $a \in \{1,\ldots,A\}$. We assume $A < t$ and denote the years associated with childhood ages using the index $t(a)$. Let the vector $W_{n(i)t}$ have entries that contain indicators of neighborhood quality and other neighborhood characteristics measured in a given year. Finally, for current outcome j, let the term θ_i be the impact of family or individual background factors, such as family inputs or genetic endowments, and let the term ϵ_{it} represent time-varying idiosyncratic influences, such as household-level shocks.

A simple model assumes y_{it}^{j} is an additive function of neighborhood effects and other factors:

$$y_{it}^{j} = W_{n(i)t}\lambda' + \sum_{a=1}^{A} W_{c(i,a)t(a)}\mu_a' + \theta_i + \epsilon_{it}. \tag{1}$$

The contemporaneous effects of current neighborhood characteristics are captured by the coefficients contained in the vector λ. The possibility that neighborhoods have lasting exposure effects due to impacts on child development is captured by the coefficients contained in the age-specific vector μ_a. These effects may vary (i.e., $\mu_a \neq \mu_{a-1}$) which embodies the "critical age effects" hypothesis that some childhood ages may be more important than others. Note that Equation 1 assumes that there are no lingering effects from an individual's previous adult neighborhoods (residential locations in the years between childhood and t)—an assumption that is often tested in the empirical literature (e.g., Chetty and Hendren 2018).

The production function for current outcomes embodied in Equation 1 encompasses a range of models from the neighborhood effects literature. Theoretically, much attention has focused on a canonical linear-in-means model of social interactions that assumes the presence of only contemporaneous neighborhood effects (Manski 1993; Brock and Durlauf 2001). In this model, there are three sources of neighborhoods effects. First, endogenous peer effects arise due to the propensity for individual behavior to depend on the expected (mean) behavior of their neighborhood peers. Second, exogenous effects represent the possibility that individual behavior is shaped by a vector of average characteristics (e.g., socioeconomic background) of neighbor peers. Third, correlated effects refer to the fact that individuals within a neighborhood face the same institutional and physical environments (e.g., access to schools, law enforcement practices, temperature, or air pollution). As discussed in Manski (1993), it is typically not possible to separately identify endogenous effects from exogenous effects (or from unmeasured correlated effects) in the canonical linear-in-means model. Rather, a reduced form can be examined and estimated to test for evidence of some form of contemporaneous neighborhood effects:

$$y_i^j = \alpha + x_i\gamma' + \overline{x}_n\beta' + z_n\gamma' + \epsilon_i, \tag{2}$$

where x_i is a vector for individual characteristics (e.g., socioeconomic background), $\overline{x}_n$ is a vector of the averages of the individual background characteristics for the residents of neighborhood n, z_n is a vector of other neighborhood characteristics (e.g., school resources), and ϵ_i is an error term. In this framework, the main coefficient of interest is the vector β as a measure of the reduced form impacts of both endogenous and exogenous neighborhood effects.

There has been much recent attention on models that focus solely on childhood neighborhood effects. Chetty and Hendren (2018) study the effects of moving a child to a new area where other children do well. They characterize neighborhoods in terms of the mean adult outcomes of children who spend their entire childhood in an area (those who are "permanent residents"). Since place effects may vary based on parent income and the child's birth cohort, Chetty and Hendren examine the impact of the mean outcome of children who are permanent residents of place n with parents at the percentile p of the income distribution in birth cohort s denoted as $\overline{y}_{nps}^j$. Formally, they focus on a restricted version of Equation 1 for children who moved across areas:

$$y_i^j = \alpha_m + \beta_m\overline{y}_{nps}^j + \theta_i, \tag{3}$$

where y_i^j is an adult-age outcome (e.g., income at age 24) for child i who moved to commuting zone n at age $m \in \{1, \ldots, A\}$ and stayed for the rest of their childhood, and θ_i is an error term. The coefficient of interest β_m represents the mean impact of spending year m of one's childhood onward in an area where permanent residents have better outcomes. In this model, the exposure effect at age m is defined as $\gamma_m = \beta_m - \beta_{m+1}$.

Identification of neighborhood effect parameters is empirically challenging due to the non-random sorting of families into neighborhoods. Formally, the concern is that $cov(\overline{x}_n, \epsilon_i) \neq 0$ and $cov(\overline{y}^j_{nps}, \theta_i) \neq 0$ in Equations 2 and 3, respectively. In line with broader trends in economics and social science, the recent neighborhood effects literature has addressed self-selection using experimental and quasi-experimental approaches. As discussed in the main text, several studies rely on experimental data from the MTO demonstration, which provided housing vouchers to a randomly selected group of low-income households living in distressed public housing. For example, Kling, Liebman, and Katz (2007) use the MTO treatment group variable as an instrument for the neighborhood poverty rate and estimate a variant of Equation 2. The quasi-experimental approach in Chetty and Hendren (2018) studies movers and estimates exposure effects by assuming that selection effects for movers to different areas do not vary with the child's age at move. This allows for the possibility that families that move to better areas may differ from those that move to worse areas. This assumption implies that selection effects in the estimates from Equation 3 will cancel out when estimating the exposure effect.