# Is Zip Code Destiny? <br> Re-visiting Long-run Neighborhood Effects 

Sadegh Eshaghnia

March 2, 2023

## Introduction

■ Children's incomes in adulthood vary remarkably by the local region where they grow up (Chetty et al. (2014)).

- Spatial variation in intergenerational mobility has been documented for the US and many other developed countries.

■ What is the causal status of the link between neighborhood of residence and longrun economic well-being?

- To what extent do the differences in income mobility across geographical areas reflect causal effects of place (Chetty \& Hendren (2018a,b); Chetty et al. (2020a,b); Chetty (2021))?
- This paper
- documents life cycle heterogeneity in the neighborhood sorting
- critically reviews the estimation procedures and underlying assumptions of the extant literature: causality or correlation?


## Motivation

- Chetty \& Hnedren (2018a) analyze data on families who moved across commuting zones (CZ) in the US and argue that neighborhoods shape various adulthood outcomes of children:
- Adult incomes of children who moved converge to the adult incomes of children of permanent residents in the destination at a rate of $4 \%$ per year of exposure
- They interpret their results as causal effects of neighborhoods
- Chetty et al. (2020a) repeat the analysis at the Census tracts
- Replicated using data from other countries
- Chetty \& Hendren (2018b): Causal effects of each county/CZ

■ Chetty et al. (2020a) construct an "Opportunity Atlas"

- Touted as "zip code destiny" or "power of place"


## Motivation- Cont'd

- Influence on the design of housing policies
- Relocation policies as a way to promote upward mobility
- Creating Moves to Opportunity Experiment (CMTO) in Seattle and King county (Bergman et al. (2019))
- Should we invest in families and local amenities, or whether should we relocate families across neighborhoods?


## This Paper

■ Replicates Chetty et al. (2018) using Danish registers

- Investigates the mechanisms behind the exposure estimates
- Can one interpret the results as causal effects of neighborhoods or "power of place"?
- The role of selection and sorting

■ Examines identifying assumptions in Chetty et al. (2018): Selection effects do not vary with the child's age when moving

- This requires children potential outcomes to be orthogonal to their age when families move across neighborhoods
- Documents life cycle heterogeneity in the nbhd sorting process that invalidates the assumption of constant selection effects
- Conducts a placebo test to examine the credibility of the estimation strategies for identifying long-run nbhd effects


## Preview of Results

- I find similar estimates to those of Chetty et al. (2018)
- Placebo tests suggest: exposure effect estimates in the literature reflect the correlational estimates of place effects
- I provide evidence for a violation of the main identifying assumption (constant selection effects) in previous studies
- Self-selection into "permanent residency" status and into timing of moves (wrt the age of children)
- Families sort into heterogenous areas and the age of child when parents move is not orthogonal to the extent to which there is a positive sorting between parents and neighborhoods:


## Chetty et al. (2018)

- Given birth cohort s and CZ c , let p be the parents' percentile in the national income distribution
- Let $y_{i}$ denote the child's national income rank in adulthood

$$
y_{i}=\alpha_{c s}+\psi_{c s} p_{i}+\epsilon_{i}
$$

then, estimate $y_{p c s}$, the mean rank of children with parents at percentile $p$ of the income distribution in CZ c in birth cohort s , using the fitted values:

$$
\bar{y}_{p c s}=\hat{\alpha}_{c s}+\hat{\psi}_{c s} p
$$

## Exposure Effects

Exposure effect at age $m$ : the impact of spending year $m$ of one's childhood in an area where PR's outcomes are 1 pp higher

Thought experiment: randomly assign children to new NBHD d starting at age $m$ for the rest of childhood. The best linear predictor of children's outcomes $y_{i}$ in the experimental sample, based on the PR's outcomes in CZ d ( $\bar{y}_{p d s}$ ):

$$
\begin{equation*}
y_{i}=\alpha_{m}+\beta_{m} \bar{y}_{p d s}+\theta_{i} \tag{3}
\end{equation*}
$$

Random assignment: $\theta \perp \bar{y}_{p d s}$
Exposure effect at $m: \gamma_{m}=\beta_{m}-\beta_{m+1}$, the effect on $y_{i}$ of spending the year from age $m$ to age $(m+1)$ in the destination

Observational data: $b_{m}=\beta_{m}+\delta_{m}$
Bias $=\delta_{m}=\frac{\operatorname{cov}\left(\theta_{i}, \bar{y}_{p d s}\right)}{\operatorname{var}\left(\bar{y}_{p d s}\right)}:$ parent inputs \& unobserved det. of children's outcomes covary with PR's outcomes

## Exposure Effects- Constant-in-Age Selection Assumption

$$
\operatorname{Bias}=\delta_{m}=\frac{\operatorname{cov}\left(\theta_{i}, \bar{y}_{p d s}\right)}{\operatorname{var}\left(\bar{y}_{p d s}\right)}
$$

ASSUMPTION 1 (A.1): Selection effects do not vary with the child's age at move: $\delta_{m}=\delta$ for all m.

Under A.1, we obtain consistent estimates of exposure effects:

$$
\gamma_{m}=\left(\beta_{m}+\delta_{m}\right)-\left(\beta_{m+1}+\delta_{m+1}\right)=b_{m}-b_{m+1}
$$

- Selection effects $\delta$ cancel out when estimating the exposure effect.
- Rules out differential preferences among parents by age of child for local amenities (schools) not captured by income

■ Even an stronger assumption when identifying county level estimates (Chetty \& Hendren (2018b))

## What if Assumption A. 1 Is violated?

Under A.1:

$$
\gamma_{m}=\left(\beta_{m}-\beta_{m+1}\right)+\left(\delta_{m}-\delta_{m+1}\right)=b_{m}-b_{m+1}
$$

If A. 1 is violated:
11 If sorting decreases in child's age:
$\delta_{m}>\delta_{m+1} \quad \forall m \in\{\underline{m}, \ldots, \bar{m}\} \Rightarrow$ Equ (3) overestimates the exposure effect, $\gamma_{m}$

2 If sorting becomes stronger as age increases:
$\delta_{m}<\delta_{m+1} \quad \forall m \in\{\underline{m}, \ldots, \bar{m}\} \Rightarrow$ Equ (3) underestimates the exposure effect, $\gamma_{m}$.
3 Unclear if sorting not monotonically changes over the age support exploited for the estimation.

## Exposure Effects- Estimation Strategy

Consider the set of children whose families moved when they were exactly $m$ years old.

We can analyze how these children's incomes in adulthood are related to those of PR in their destination CZ as below:

$$
\begin{equation*}
y_{i}=\alpha_{q o s}+b_{m} \Delta_{o d p s}+\epsilon_{1 i} \tag{4}
\end{equation*}
$$

- $y_{i}$ : child's income rank at age 24 ,
- $\alpha_{\text {qos }}$ : FE for the origin o by parent income decile $q$ by birth cohort s,
- $\Delta_{o d p s}=\bar{y}_{p d s}-\bar{y}_{p o s}$ : difference in predicted income rank (at age 24) of permanent residents in the destination versus origin for the relevant parent income rank $p$ and birth cohort $s$.


## Childhood Exposure Effects on Inc. Ranks in Adulthood

$$
y_{i}=\underbrace{\alpha_{\text {qosm }}}_{\mathrm{FE}}
$$

## Childhood Exposure Effects on Inc. Ranks in Adulthood

$$
y_{i}=\underbrace{\alpha_{\text {qosm }}}_{\text {FE }}+\underbrace{\sum_{m=9}^{30} b_{m} \mathbb{I}_{m_{i}=m} \Delta_{\text {odps }}}_{\text {by-age exposure effects }}
$$

## Childhood Exposure Effects on Inc. Ranks in Adulthood

$$
y_{i}=\underbrace{\alpha_{q o s m}}_{\text {FE }}+\underbrace{\sum_{m=9}^{30} b_{m} \mathbb{I}_{m_{i}=m} \Delta_{o d p s}}_{\text {by-age exposure effects }}+\underbrace{\sum_{s=1980}^{1987} \kappa_{s} \mathbb{I}_{s_{i}=s} \Delta_{o d p s}}_{\text {cohort-specific selection effects }}+\varepsilon_{2 i}
$$

- $\Delta_{\text {qosm }}$ : (origin $\times$ parent income decile $\times$ cohort $\times$ age) FE
- $\hat{b}_{m}$ : the average effect on age- 24 income rank $y_{i}$, conditional on moving from $o$ to $d$ at age $m$, of a 1 pctile $\uparrow$ in $\Delta_{\text {odps }}$


## Childhood Exposure Effects on Inc. Ranks in Adulthood

$$
y_{i}=\underbrace{\alpha_{\text {qosm }}}_{\mathrm{FE}}+\underbrace{\sum_{m=9}^{30} b_{m} \mathbb{I}_{m_{i}=m} \Delta_{o d p s}}_{\text {by-age exposure effects }}+\underbrace{\sum_{s=1980}^{1987} \kappa_{s} \mathbb{I}_{s_{i}=s} \Delta_{o d p s}}_{\text {cohort-specific selection effects }}+\varepsilon_{2 i}
$$

- $\Delta_{\text {qosm }}$ : (origin $\times$ parent income decile $\times$ cohort $\times$ age) FE
- $\hat{b}_{m}$ : the average effect on age- 24 income rank $y_{i}$, conditional on moving from o to $d$ at age $m$, of a 1 pctile $\uparrow$ in $\Delta_{\text {odps }}$

If we had only one cohort and one parent income percentile:

$$
y_{i}=\underbrace{\alpha_{o m}}_{\text {FE }}+\underbrace{\sum_{m=9}^{30} b_{m} \mathbb{I}_{m_{i}=m} \Delta_{o d}}_{\text {by-age exposure effects }}+\underbrace{\kappa \Delta_{o d}}_{\text {selection effects }}+\varepsilon_{2 i}
$$

## Parametric model w. cohort- and age-specific slopes

If we had only one cohort and one parent income percentile:

$$
\mathrm{y}_{i}=\alpha^{1}+\underbrace{\alpha^{2} \bar{y}_{o}}_{\text {origin FE }}+\underbrace{\sum_{m=9}^{30} \mathbb{I}_{m_{i}=m} \zeta_{m}}_{\text {age FE }}
$$

## Parametric model w. cohort- and age-specific slopes

If we had only one cohort and one parent income percentile:

$$
\begin{aligned}
& \mathrm{y}_{i}=\alpha^{1}+\underbrace{\alpha^{2} \bar{y}_{o}}_{\text {origin FE }}+\underbrace{\sum_{m=9}^{30} \mathbb{I}_{m_{i}=m} \zeta_{m}}_{\text {age FE }} \\
& +\sum_{m=9}^{30} b_{m} \mathbb{I}_{m_{i}=m} \Delta_{o d}+\kappa \Delta_{o d}+\varepsilon_{3 i}
\end{aligned}
$$

## Parametric model w. cohort- and age-specific slopes

If we had only one cohort and one parent income percentile:

$$
\begin{aligned}
& \mathrm{y}_{i}=\alpha^{1}+\underbrace{\alpha^{2} \bar{y}_{o}}_{\text {origin FE }}+\underbrace{\sum_{m=9}^{30} \mathbb{I}_{m_{i}=m} \zeta_{m}}_{\text {age FE }} \\
& +\sum_{m=9}^{30} b_{m} \mathbb{I}_{m_{i}=m} \Delta_{o d}+\kappa \Delta_{o d}+\varepsilon_{3 i}
\end{aligned}
$$

Generalizing to various cohorts and parental income:
$\mathrm{y}_{i}=\underbrace{\sum_{s=1980}^{1988} \mathbb{I}_{s_{i}=s}\left(\alpha_{s}^{1}+\alpha_{s}^{2} \bar{y}_{\text {pos }}\right)}$
origin effects by cohort

## Parametric model w. cohort- and age-specific slopes

If we had only one cohort and one parent income percentile:

$$
\mathrm{y}_{i}=\alpha^{1}+\underbrace{\alpha^{2} \bar{y}_{o}}_{\text {origin FE }}+\underbrace{\sum_{m=9}^{30} \mathbb{I}_{m_{i}=m} \zeta_{m}}_{\text {age FE }}
$$

$$
+\sum_{m=9}^{30} b_{m} \mathbb{I}_{m_{i}=m} \Delta_{o d}+\kappa \Delta_{o d}+\varepsilon_{3 i}
$$

Generalizing to various cohorts and parental income:

$$
\mathrm{y}_{i}=\underbrace{\sum_{s=1980}^{1988} \mathbb{I}_{s_{i}=s}\left(\alpha_{s}^{1}+\alpha_{s}^{2} \bar{y}_{p o s}\right)}_{\text {origin effects by cohort }}+\underbrace{\sum_{m=9}^{30} \mathbb{I}_{m_{i}=m}\left(\zeta_{m}^{1}+\zeta_{m}^{2} p_{i}\right)}_{\text {age-specific disruption effect }}
$$

## Parametric model w. cohort- and age-specific slopes

If we had only one cohort and one parent income percentile:

$$
\begin{aligned}
\mathrm{y}_{i} & =\alpha^{1}+\underbrace{\alpha^{2} \bar{y}_{o}}_{\text {origin FE }}+\underbrace{\sum_{m=9}^{30} \mathbb{I}_{m_{i}=m} \zeta_{m}}_{\text {age FE }} \\
& +\sum_{m=9}^{30} b_{m} \mathbb{I}_{m_{i}=m} \Delta_{o d}+\kappa \Delta_{o d}+\varepsilon_{3 i}
\end{aligned}
$$

Generalizing to various cohorts and parental income:

$$
\begin{aligned}
\mathrm{y}_{i}= & \underbrace{\sum_{s=1980}^{1988} \mathbb{I}_{s_{i}=s}\left(\alpha_{s}^{1}+\alpha_{s}^{2} \bar{y}_{p o s}\right)}_{\text {origin effects by cohort }}+\underbrace{\sum_{m=9}^{30} \mathbb{I}_{m_{i}=m}\left(\zeta_{m}^{1}+\zeta_{m}^{2} p_{i}\right)}_{\text {age-specific disruption effect }} \\
& +\sum_{\text {selection effect by cohort }}^{30} b_{m} \mathbb{I}_{m_{i}=m} \Delta_{o d p s}+\underbrace{\sum_{s} \kappa_{s} \mathbb{I}_{s_{i}=s} \Delta_{\text {odps }}}_{s=1980}+\varepsilon_{3 i},
\end{aligned}
$$

## Results: $\hat{b}_{m}$ as Function of Age $m$



Figure IV
Childhood Exposure Effects on Income Ranks in Adulthood

## Childhood Exposure Effect Estimates- Linear Specification

$$
\begin{aligned}
\mathrm{y}_{i} & =\underbrace{\sum_{s=1980}^{1988} \mathbb{I}_{s_{i}=s}\left(\alpha_{s}^{1}+\alpha_{s}^{2} \bar{y}_{p o s}\right)}_{\text {origin effects by cohort }}+\underbrace{\sum_{m=9}^{30} \mathbb{I}_{m_{i}=m}\left(\zeta_{m}^{1}+\zeta_{m}^{2} p_{i}\right)}_{\text {age-specific disruption effect }} \\
& +\underbrace{\kappa_{s} \mathbb{I}_{s_{i}=s} \Delta_{o d p s}}_{\text {selection effect by cohort }}+\mathbb{I}_{m_{i} \leq 23}\left(b_{0}+\left(23-m_{i}\right) \gamma\right) \Delta_{o d p s} \\
& +I_{m_{i}>23}\left(\delta+\left(23-m_{i}\right) \delta^{\prime}\right) \Delta_{o d p s}+\varepsilon_{3 i}
\end{aligned}
$$

## Childhood Exposure Effect Estimates- Linear Specification

$$
\begin{aligned}
\mathrm{y}_{i} & =\underbrace{\sum_{s=1980}^{1988} \mathbb{I}_{s_{i}=s}\left(\alpha_{s}^{1}+\alpha_{s}^{2} \bar{y}_{p o s}\right)}_{\text {origin effects by cohort }}+\underbrace{\sum_{m=9}^{30} \mathbb{I}_{m_{i}=m}\left(\zeta_{m}^{1}+\zeta_{m}^{2} p_{i}\right)}_{\text {age-specific disruption effect }} \\
& +\underbrace{\kappa_{s} \mathbb{I}_{s_{i}=s} \Delta_{o d p s}}_{\text {selection effect by cohort }}+\mathbb{I}_{m_{i} \leq 23}\left(b_{0}+\left(23-m_{i}\right) \gamma\right) \Delta_{o d p s} \\
& +I_{m_{i}>23}\left(\delta+\left(23-m_{i}\right) \delta^{\prime}\right) \Delta_{o d p s}+\varepsilon_{3 i}
\end{aligned}
$$

$\hat{\gamma} \approx 0.04$

# Neighborhood Exposure Effects in Denmark 

## Distribution of Child's Age when Family Moves

Figure: Distribution of the Child's Age when Parents Move


# Neighborhood Exposure Effects 

## Childhood Exposure Effects on Inc. Ranks

Figure: Childhood Exposure Effects on Income Ranks in Adulthood


## Childhood Exposure Effect Estimates

| Dependent Variable: Child's Income Rank in Adulthood (Age 30) |  |  |  |  |  |  | Family FE |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  |  |  |  |  |
| Specification: | Pooled <br> (1) | $\text { Age }<=23$ <br> (2) | $\text { Age }<18$ <br> (3) | No cohort controls <br> (4) | Family Income <br> (5) | Child nbhd FE (6) | Baseline <br> (7) | No cohort controls (8) | Timevarying controls (9) |
| US: Exposure Effect ( $\gamma$ ) | $\begin{gathered} 0.040 \\ (0.002) \end{gathered}$ | $\begin{gathered} 0.040 \\ (0.002) \end{gathered}$ | $\begin{gathered} 0.037 \\ (0.005) \end{gathered}$ | $\begin{gathered} 0.036 \\ (0.002) \end{gathered}$ | $\begin{gathered} 0.041 \\ (0.002) \end{gathered}$ | $\begin{gathered} 0.031 \\ (0.002) \end{gathered}$ | $\begin{gathered} 0.044 \\ (0.008) \end{gathered}$ | $\begin{gathered} 0.031 \\ (0.005) \end{gathered}$ | $\begin{gathered} 0.043 \\ (0.008) \end{gathered}$ |
| Denmark: Exposure Effect ( $\gamma$ ) | $\begin{gathered} 0.023 \\ (0.003) \end{gathered}$ | $\begin{gathered} 0.023 \\ (0.003) \end{gathered}$ | $\begin{gathered} 0.019 \\ (0.005) \end{gathered}$ | $\begin{gathered} 0.016 \\ (0.003) \end{gathered}$ | $\begin{gathered} 0.016 \\ (0.003) \end{gathered}$ | $\begin{gathered} 0.021 \\ (0.003) \end{gathered}$ | $\begin{gathered} \hline 0.020 \\ (0.013) \end{gathered}$ | $\begin{gathered} \hline 0.017 \\ (0.009) \end{gathered}$ | $\begin{gathered} 0.023 \\ (0.015)) \end{gathered}$ |
| Number of Obs.: | 107,289 | 102,521 | 80,237 | 107,289 | 107,123 | 107,252 | 107,289 | 107,289 | 107,289 |

## Placebo Tests Using Birth Characteristics

## Placebo Tests

- Examine the credibility of the estimation strategies for identifying long-run neighborhood effects
- The extent to which nbhd exposure estimates are driven by the sorting of heterogeneous families across nbhd with different amenities rather than by causal impacts of nbhd on children
- Data on birth characteristics of children born between 1997-2005 in Denmark
- Chetty \& Hnedren (2018a) investigate how children's earnings in adulthood are related to the quality of the destination neighborhood and the child's age when moving
- I examine how a child's birth length is related to such factors
- One expects to find insignificant estimates. Otherwise, the effect would be preceding the cause


## Placebo Exposure Effect Estimates

$$
\begin{aligned}
b l_{i}= & \sum_{s=1997}^{2005} \kappa_{s} I\left(s_{i}=s\right)\left(\alpha_{s}^{1}+\alpha_{s}^{2} \overline{b_{p o s}}\right)+\sum_{m=1}^{20} I\left(m_{i}=m\right)\left(\zeta_{m}^{1}+\zeta_{m}^{2} p_{i}\right) \\
& +\sum_{m=1}^{20} \beta_{m} I\left(m_{i}=m\right) \Delta_{\text {odps }}^{b l}+\sum_{s=1997}^{2004} \kappa_{s}^{d} I\left(s_{i}=s\right) \Delta_{\text {odps }}^{b l}+\epsilon_{3 i},
\end{aligned}
$$

where $b l_{i}$ denotes the child's percentile rank on her position in the national birth length distribution relative to all others in her birth cohort, and $\Delta_{o d p s}^{b l}=\bar{b} l_{p d s}-\bar{b} l_{p o s}$ is the mean difference in permanent residents' birth length ranks between the destination and origin for the relevant parent income rank $p$ and birth cohort $s$.

## Placebo Tests

Figure: Placebo Effects Using Birth Length


## Placebo Exposure Effect Estimates- Parametric Estimates

$$
\begin{aligned}
b l_{i}= & \sum_{s=1997}^{2005} \kappa_{s} I\left(s_{i}=s\right)\left(\alpha_{s}^{1}+\alpha_{s}^{2} \bar{b} l_{\text {pos }}\right)+\sum_{m=1}^{20} I\left(m_{i}=m\right)\left(\zeta_{m}^{1}+\zeta_{m}^{2} p_{i}\right) \\
& +\sum_{s=1997}^{2004} \kappa_{s}^{d} I\left(s_{i}=s\right) \Delta_{\text {odps }}^{b l}+I\left(m_{i} \geq 0\right)\left(b_{0}+m_{i} \gamma\right) \Delta_{\text {odps }}^{b l} \\
& +I\left(m_{i}<0\right)\left(\delta_{0}+m_{i} \delta^{\prime}\right) \Delta_{\text {odps }}^{b l}+\epsilon_{3 i},
\end{aligned}
$$

## Placebo Estimates

Dependent Variable: Child's Birth Length Rank

| Specification: | Pooled <br> (1) | $\text { Age }>=0$ <br> (2) | $\text { Age }<22$ <br> (3) | No cohort controls <br> (4) | Family Level (5) | Child nbhd FE (6) | Family FE |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  |  | Baseline | No cohort controls | Timevarying controls (9) |
| US: Exposure Effect ( $\gamma$ ) | $\begin{gathered} \hline 0.040 \\ (0.002) \end{gathered}$ | $\begin{gathered} 0.040 \\ (0.002) \end{gathered}$ | $\begin{gathered} 0.037 \\ (0.005) \end{gathered}$ | $\begin{gathered} 0.036 \\ (0.002) \end{gathered}$ | $\begin{gathered} 0.041 \\ (0.002) \end{gathered}$ | $\begin{gathered} 0.031 \\ (0.002) \end{gathered}$ | $\begin{gathered} 0.044 \\ (0.008) \end{gathered}$ | $\begin{gathered} 0.031 \\ (0.005) \end{gathered}$ | $\begin{gathered} 0.043 \\ (0.008) \end{gathered}$ |
| Denmark: Placebo Effect ( $\gamma$ ) | $\begin{gathered} 0.044 \\ (0.006) \end{gathered}$ | $\begin{gathered} 0.045 \\ (0.006) \end{gathered}$ | $\begin{gathered} \hline 0.031 \\ (0.006) \end{gathered}$ | $\begin{gathered} 0.044 \\ (0.006) \end{gathered}$ | - | $\begin{gathered} 0.043 \\ (0.006) \end{gathered}$ | $\begin{gathered} \hline 0.028 \\ (0.014) \end{gathered}$ | $\begin{gathered} 0.033 \\ (0.014) \end{gathered}$ | $\begin{gathered} 0.029 \\ (0.014) \end{gathered}$ |
| Number of Obs.: | 127,536 | 73,746 | 133,159 | 127,536 | - | 127,536 | 127,536 | 127,536 | 127,536 |

## Age of Child at Move and Child Potential Outcomes

Figure: Birth Length Rank and the Age of the Child at the Time of the Move


## Parental Selection based on Education

Chetty (2018) estimates:

$$
\begin{equation*}
y_{i}=\alpha+\beta_{m} \Delta_{o d p s}+\epsilon_{i} \tag{4}
\end{equation*}
$$

Parent's education level is one of the omitted variables affecting both child's outcome and quality of the move across NBHDs.

Let's assume that the true model is as follows:

$$
\begin{equation*}
y_{i}=\alpha+\beta_{m} \Delta_{o d p s}+\beta_{e} e d u_{i}^{p}+u_{i} \tag{5}
\end{equation*}
$$

Then,

$$
\begin{aligned}
& \operatorname{Plim} \hat{\beta}_{m}= \beta_{m}+\beta_{e} \frac{\operatorname{cov}\left(e d u_{i}^{p}, \Delta_{p d s}\right)}{\operatorname{var}\left(\Delta_{p d s}\right)} \\
&=\beta_{m}+\beta_{e} \delta_{m}
\end{aligned}
$$

$$
\operatorname{Plim} \hat{\gamma}_{m}=\left(\beta_{m}-\beta_{m+1}\right)+\beta_{e}\left(\delta_{m}-\delta_{m+1}\right)
$$

## Intensity of Sorting by Age of Child at Move

Figure: Intensity of Sorting b/w Parent's Education and Quality of Move
(a) Parental Education

(b) Selection Bias by Age


## Back-of-the-envelope Calculation of the Bias

To evaluate the size of the bias, $\beta_{e}\left(\delta_{m}-\delta_{m+1}\right)$ :
1 Using equ (5), obtain some estimates for $\beta_{e}: \hat{\beta}_{e} \in[0.82,1.15]$
2 Using the slope of covariance term (between parents' education level and quality of the move) over age of child, obtain an estimate for $\left(\delta_{m}-\delta_{m+1}\right):\left(\delta_{m}-\delta_{m+1}\right) \approx 0.005$

Life Cycle Heterogeneity in the Neighborhood Sorting Process

# Selection and Age of Child at Move: 

(A) Parental Characteristics

## Education of Parents and Age of Child when Parents Move

Figure: Age of Child at Move and Parental Edu. by Ownership Status
(a) Owners
(b) Renters



## Income of Parents and Age of Child when Parents Move

Figure: Parental Income Rank and Age of Child when Parents Move
(a) Owners
(b) Renters



## Family Structure and Age of Child at Move

Figure: Fraction of Intact Families and Age of Child when Parents Move
(a) Owners

(b) Renters


## Selection and Age of Child at Move:

(B) Quality of Moves

# (B.1) Difference in Mean Income Ranks of Children of PR's in Orig. vs Dest. 

## The Quality of Moves and Age of Child at Move

Figure: The Quality of Moves by Ownership Status
(a) Owners

(b) Renters


## The Quality of Origin and Age of Child at Move

Figure: The Quality of Moves by Ownership Status
(a) Owners
(b) Renters



## The Quality of Destination and Age of Child at Move

Figure: The Quality of Moves by Ownership Status
(a) Owners
(b) Renters


(B.2) NBHD Avg Inc Rank at Orig. vs Dest.

## NBHD Income Rank and Age of Child at Move

Figure: Change in NBHD Inc Rank and Age of Child
(a) Parish Level

(b) Cluster Level

(B.3) School Quality Rank at Orig. vs Dest.

## NBHD (Parish) School Quality Rank and Age at Move

Figure: Change in nbhd School Rank (Math Grades) and Age of Child


## (B.4) Average Neighborhood House Price Rank

## Neighborhood (Parish) House Price Rank and Age of Child at Move

Figure: Change in NBHD House Price Rank. and Age of Child
(a) Owners
(b) Renters


(C) Timing of Moves and Lifecycle Shocks

## Divorce and Age of Child at Move

Figure: Age of Child at Move \& Frac. of Parents Separated when Moving
(a) Owners
(b) Renters



## Conclusion

- Recent studies have exploited quasi-experimental strategies to identify the causal impact of NBHDs on children.
- One of the main challenges in estimating the causal impact of NBHDs on child is the endogeneity of NBHD quality.
- I investigate the methodology and main identifying assumptions of the influential studies in the literature.
- Parental sorting into NBHDs has an important lifecycle gradient; it is not orthogonal to children's age at the move.
- The constant selection effects assumption in recent empirical works is violated $\rightarrow$ overestimating NBHD impacts on children
- The placebo tests clearly showcase the methodological problems of the popular studies in the literature.

Thanks!

## Appendix

## Data

■ Data source: Danish registers
■ Data span: 1980-2017
■ Sample: Children who were born between 1970-1982

- permanent residents (stayers/PR): subset of parents who reside in a single municipality (parish) c in 1982-2000
- movers: individuals in the main sample who are not PR

■ Income type: Disposable income

- averaged over 1982-2000 to get parental income

■ Unit of Analysis: Family income for parents and individual income for children

## Summary Statistics

Table 1: Summary Statistics for Municifality Permanent Residents and Movers

| Variable | Mean <br> $(1)$ | Std. dev. <br> $(2)$ | Median <br> $(3)$ | Num. of obs. <br> $(4)$ |
| :--- | :--- | :--- | :--- | :--- |
| Panel A: Permanent residents: Families |  |  |  |  |
| who do not move across municipalities |  |  |  |  |
| Child individual income at 30 | 25,495 | 9,710 | 25,415 | 536,993 |
| Child family income at 30 | 43,090 | 19,368 | 44,476 | 536,072 |
| Child cohabiting at 30 | 0.67 | 0.47 | 1.00 | 537,801 |
| Child years of schooling by 30 | 14.68 | 2.37 | 14.50 | 524,959 |
| Child individual property value at 30 | 81,794 | 99,120 | 69,070 | 529,849 |
| Parent family income | 43,832 | 13,272 | 42,660 | 527,670 |
| Parent property value | 109,882 | 79,499 | 106,692 | 525,677 |
| Nuclear (intact) Family | 0.62 | 0.49 | 1.00 | 484,164 |
|  |  |  |  |  |
| Panel B: Families who move $1-3$ | times |  |  |  |
| across municipalities | 24,880 | 10,007 | 24,846 | 258,295 |
| Child individual income at 30 | 41,732 | 19,911 | 42,257 | 257,744 |
| Child family income at 30 | 0.65 | 0.48 | 1.00 | 258,592 |
| Child cohabiting at 30 | 14.50 | 2.55 | 14.50 | 251,296 |
| Child years of schooling by 30 | 69,105 | 92,740 | 47,726 | 255,337 |
| Child individual property value at 30 | 43,586 | 13,549 | 41,948 | 252,652 |
| Parent family income | 94,273 | 77,781 | 86,069 | 251,903 |
| Parent property value | 0.39 | 0.49 | 0.00 | 234,262 |
| Nuclear (intact) Family |  |  |  |  |
| Panel C: Families who move exactly once |  |  |  |  |
| across municipalities |  |  |  |  |
| Child individual income at 30 | 25,197 | 10,066 | 25,146 | 157,428 |
| Child family income at 30 | 42,313 | 19,955 | 42,968 | 157,119 |
| Child cohabiting at 30 | 0.65 | 0.48 | 1.00 | 157,633 |
| Child years of schooling by 30 | 14.63 | 2.51 | 14.50 | 153,221 |
| Child individual property value at 30 | 72,892 | 94,934 | 54,975 | 155,601 |
| Parent family income | 44,180 | 13,879 | 42,528 | 154,143 |
| Parent property value | 100,761 | 78,964 | 94,480 | 153,667 |
| Nuclear (intact) Family | 0.45 | 0,50 | 0.00 | 143,172 |

## Number of Moves

Figure: Number of Moves by Education Level


## Birth Length and Academic Achievement

Figure: Birth Length Rank and the Age of the Child at the Time of the Move

(a) Mathematics Knowledge

(c) Reading

(b) problem-solving

(d) Writing

## Birth Weight and Academic Achievement

Figure: Birth Weight Rank and the Age of the Child at the Time of the Move

(a) Mathematics Knowledge

(c) Reading

(b) problem-solving

(d) Writing

## Academic Achievement and Income Ranks

Figure: Test Scores and Adulthood Income Rank

(a) Mathematics Knowledge

(c) Reading

(b) problem-solving

(d) Writing

