Is Zip Code Destiny? 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"

- 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

- 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:

- 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

$$y_i = \alpha_{cs} + \psi_{cs} p_i + \epsilon_i$$

then, estimate y_{pcs} , 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:

$$\bar{y}_{pcs} = \hat{\alpha}_{cs} + \hat{\psi}_{cs}p$$

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}_{pds}):

$$y_i = \alpha_m + \beta_m \bar{y}_{pds} + \theta_i \quad (3)$$

Random assignment: $\theta \perp \bar{y}_{pds}$

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 = $\delta_m = \frac{cov(\theta_i, \bar{y}_{pds})}{var(\bar{y}_{pds})}$: parent inputs & unobserved det. of children's outcomes covary with PR's outcomes

$$\mathsf{Bias} = \delta_m = rac{\mathsf{cov}(heta_i, ar{y}_{\mathsf{pds}})}{\mathsf{var}(ar{y}_{\mathsf{pds}})}$$

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:

$$\gamma_m = (\beta_m + \delta_m) - (\beta_{m+1} + \delta_{m+1}) = b_m - b_{m+1}$$

- \blacksquare Selection effects δ 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:

$$\gamma_m = (\beta_m - \beta_{m+1}) + (\delta_m - \delta_{m+1}) = b_m - b_{m+1}$$

If A.1 is violated:

- I If sorting decreases in child's age: $\delta_m > \delta_{m+1} \quad \forall m \in \{\underline{m}, ..., \overline{m}\} \Rightarrow \text{Equ } (3) \text{ overestimates the exposure effect, } \gamma_m$
- 2 If sorting becomes stronger as age increases: $\delta_m < \delta_{m+1} \quad \forall m \in \{\underline{m}, ..., \overline{m}\} \Rightarrow \text{Equ } (3) \text{ underestimates the exposure effect, } \gamma_m.$
- 3 Unclear if sorting not monotonically changes over the age support exploited for the estimation.

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

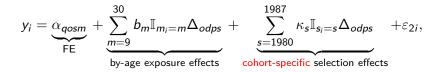
We can analyze how these children's incomes in adulthood are related to those of PR in their destination CZ as below:

$$y_i = \alpha_{qos} + b_m \Delta_{odps} + \epsilon_{1i}, \quad (4)$$

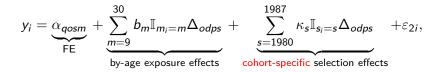
- y_i: child's income rank at age 24,
- α_{qos}: FE for the origin o by parent income decile q by birth cohort s,
- $\Delta_{odps} = \bar{y}_{pds} \bar{y}_{pos}$: difference in predicted income rank (at age 24) of permanent residents in the destination versus origin for the relevant parent income rank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

$$y_i = \underbrace{\alpha_{qosm}}_{FE}$$

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



Δ_{qosm}: (origin × parent income decile × cohort × age) FE
 b̂_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 ↑ in Δ_{odps}



Δ_{qosm}: (origin × parent income decile × cohort × age) FE
 b̂_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 ↑ in Δ_{odps}

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

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

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

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

20

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

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

$$+\sum_{m=9}^{30}b_m\mathbb{I}_{m_i=m}\Delta_{od}+\kappa\Delta_{od}+\varepsilon_{3i},$$

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

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

$$+\sum_{m=9}^{30}b_m\mathbb{I}_{m_i=m}\Delta_{od}+\kappa\Delta_{od}+\varepsilon_{3i},$$

Generalizing to various cohorts and parental income:

- -

$$y_i = \underbrace{\sum_{s=1980}^{1988} \mathbb{I}_{s_i=s}(\alpha_s^1 + \alpha_s^2 \bar{y}_{pos})}_{\text{origin effects by cohort}}$$

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

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

$$+\sum_{m=9}^{30}b_m\mathbb{I}_{m_i=m}\Delta_{od}+\kappa\Delta_{od}+\varepsilon_{3i},$$

Generalizing to various cohorts and parental income:

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

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

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

$$+\sum_{m=9}^{30}b_m\mathbb{I}_{m_i=m}\Delta_{od}+\kappa\Delta_{od}+\varepsilon_{3i},$$

Generalizing to various cohorts and parental income:

- -

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

Results: \hat{b}_m as Function of Age m

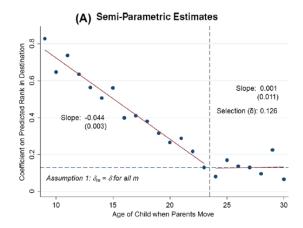
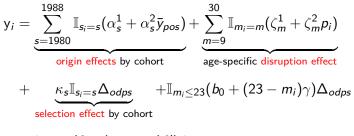


FIGURE IV

Childhood Exposure Effects on Income Ranks in Adulthood

Childhood Exposure Effect Estimates- Linear Specification



$$+ \mathsf{I}_{m_i>23}(\delta + (23 - m_i)\delta')\Delta_{odps} + \varepsilon_{3i},$$

Childhood Exposure Effect Estimates- Linear Specification

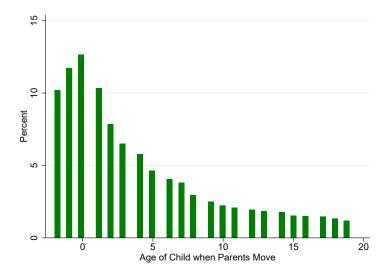
$$\begin{split} \mathbf{y}_{i} &= \sum_{\substack{s=1980\\s=1980}}^{1988} \mathbb{I}_{s_{i}=s}(\alpha_{s}^{1} + \alpha_{s}^{2}\bar{y}_{pos}) + \sum_{\substack{m=9\\m=9}}^{30} \mathbb{I}_{m_{i}=m}(\zeta_{m}^{1} + \zeta_{m}^{2}p_{i}) \\ &\xrightarrow{\text{origin effects by cohort}} \\ &+ \underbrace{\kappa_{s}\mathbb{I}_{s_{i}=s}\Delta_{odps}}_{selection effect} + \mathbb{I}_{m_{i}\leq23}(b_{0} + (23 - m_{i})\gamma)\Delta_{odps} \\ &\xrightarrow{\text{selection effect by cohort}} \\ &+ \mathbb{I}_{m_{i}>23}(\delta + (23 - m_{i})\delta')\Delta_{odps} + \varepsilon_{3i}, \end{split}$$

 $\hat{\gamma} pprox 0.04$

Neighborhood Exposure Effects in Denmark

Distribution of Child's Age when Family Moves

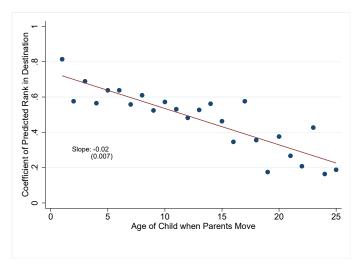
Figure: DISTRIBUTION OF THE CHILD'S AGE WHEN PARENTS MOVE



Neighborhood Exposure Effects

Childhood Exposure Effects on Inc. Ranks

Figure: Childhood Exposure Effects on Income Ranks in Adulthood



Dependent Variable: Child's Income Rank in Adulthood (Age 30)									
Specification:							Family FE		
	Pooled	$\mathrm{Age} <= 23$	Age < 18	No cohort controls	Family Income	Child nbhd FE	Baseline	No cohort controls	Time- varying controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
US: Exposure Effect (γ)	0.040	0.040	0.037	0.036	0.041	0.031	0.044	0.031	0.043
	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)	(0.002)	(0.008)	(0.005)	(0.008)
Denmark: Exposure Effect (γ)	0.023	0.023	0.019	0.016	0.016	0.021	0.020	0.017	0.023
	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.013)	(0.009)	(0.015))
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

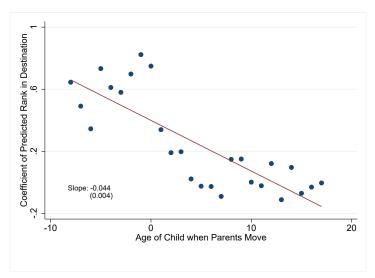
- Examine the credibility of the estimation strategies for identifying long-run neighborhood effects
- The extent to which nbhd exposure estimates are driven by the sorting of heterogeneous families across nbhd with different amenities rather than by causal impacts of nbhd on children
- Data on birth 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

$$\begin{split} bl_i &= \sum_{s=1997}^{2005} \kappa_s I(s_i = s) (\alpha_s^1 + \alpha_s^2 \bar{b} l_{pos}) + \sum_{m=1}^{20} I(m_i = m) (\zeta_m^1 + \zeta_m^2 p_i) \\ &+ \sum_{m=1}^{20} \beta_m I(m_i = m) \Delta_{odps}^{bl} + \sum_{s=1997}^{2004} \kappa_s^d I(s_i = s) \Delta_{odps}^{bl} + \epsilon_{3i}, \end{split}$$

where bl_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 $\Delta_{odps}^{bl} = \bar{b}l_{pds} - \bar{b}l_{pos}$ 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

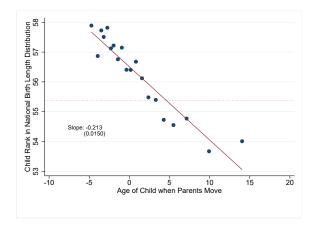


$$\begin{split} bl_i &= \sum_{s=1997}^{2005} \kappa_s I(s_i = s) (\alpha_s^1 + \alpha_s^2 \bar{b} l_{pos}) + \sum_{m=1}^{20} I(m_i = m) (\zeta_m^1 + \zeta_m^2 p_i) \\ &+ \sum_{s=1997}^{2004} \kappa_s^d I(s_i = s) \Delta_{odps}^{bl} + I(m_i \ge 0) (b_0 + m_i \gamma) \Delta_{odps}^{bl} \\ &+ I(m_i < 0) (\delta_0 + m_i \delta') \Delta_{odps}^{bl} + \epsilon_{3i}, \end{split}$$

Dependent Variable: Child's Birth Length Rank									
		· ·					Family FE		
Specification:	Pooled	Age >= 0	$\mathrm{Age} < 22$	No cohort controls	Family Level	Child nbhd FE	Baseline	No cohort controls	Time- varying controls
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
US: Exposure Effect (γ)	0.040	0.040	0.037	0.036	0.041	0.031	0.044	0.031	0.043
	(0.002)	(0.002)	(0.005)	(0.002)	(0.002)	(0.002)	(0.008)	(0.005)	(0.008)
Denmark: Placebo Effect (γ)	0.044	0.045	0.031	0.044	-	0.043	0.028	0.033	0.029
	(0.006)	(0.006)	(0.006)	(0.006)	-	(0.006)	(0.014)	(0.014)	(0.014)
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:

$$y_i = \alpha + \beta_m \Delta_{odps} + \epsilon_i, \quad (4)$$

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:

$$y_i = \alpha + \beta_m \Delta_{odps} + \beta_e edu_i^p + u_i, \quad (5)$$

Then,

$$Plim \ \hat{\beta}_{m} = \beta_{m} + \beta_{e} \frac{cov(edu_{i}^{p}, \Delta_{pds})}{var(\Delta_{pds})}$$
$$= \beta_{m} + \beta_{e} \delta_{m}$$

Plim
$$\hat{\gamma}_m = (\beta_m - \beta_{m+1}) + \beta_e(\delta_m - \delta_{m+1})$$

Figure: Intensity of Sorting b/w Parent's Education and Quality of Move

(a) Parental Education (b) Selection Bias by Age 74 8-Mother's Years of Schooling 12.5 13 13.5 Selection Bias 05 0 Slope: -0.037 (0.0013) Slope: -0.005 (0.000005) 4 25 10 15 20 5 10 15 20 Age of Child when Parents Move Age of Child when Parents Move

25

To evaluate the size of the bias, $\beta_e(\delta_m - \delta_{m+1})$:

- **1** Using equ (5), obtain some estimates for β_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 $(\delta_m \delta_{m+1})$: $(\delta_m \delta_{m+1}) \approx 0.005$

Life Cycle Heterogeneity in the Neighborhood Sorting Process

Selection and Age of Child at Move:

(A) Parental Characteristics

Figure: Age of Child at Move and Parental Edu. by Ownership Status

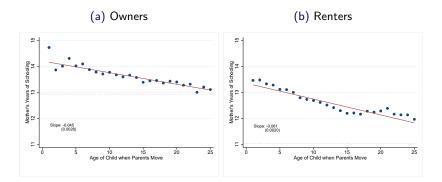


Figure: Parental Income Rank and Age of Child when Parents Move

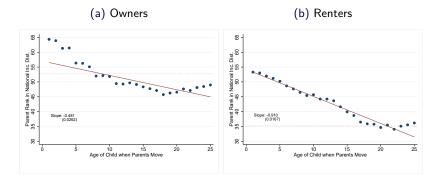
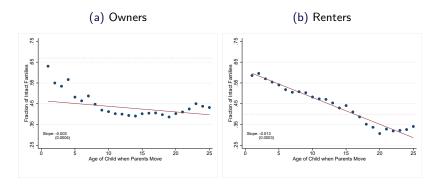


Figure: Fraction of Intact Families and Age of Child when Parents Move



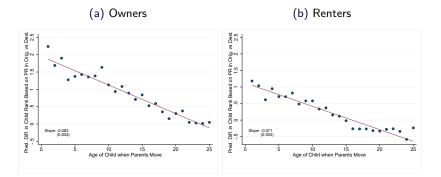
Selection and Age of Child at Move:

(B) Quality of Moves

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

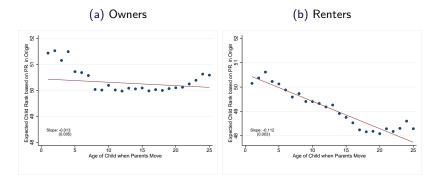
The Quality of Moves and Age of Child at Move

Figure: The Quality of Moves by Ownership Status



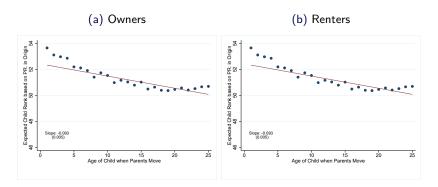
The Quality of Origin and Age of Child at Move

Figure: The Quality of Moves by Ownership Status



The Quality of Destination and Age of Child at Move

Figure: The Quality of Moves by Ownership Status



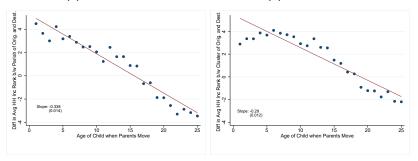
(B.2) NBHD Avg Inc Rank at Orig. vs Dest.

NBHD Income Rank and Age of Child at Move

Figure: Change in NBHD Inc Rank and Age of Child

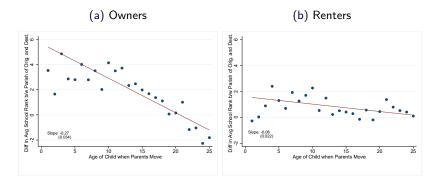
(a) Parish Level

(b) Cluster Level



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

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



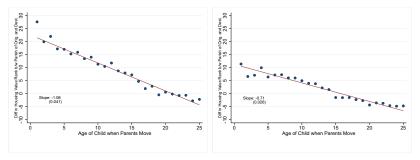
(B.4) Average Neighborhood House Price Rank

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

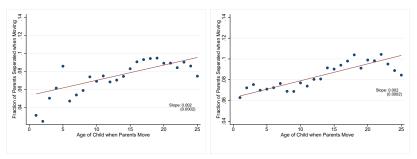


(C) Timing of Moves and Lifecycle Shocks

Figure: Age of Child at Move & Frac. of Parents Separated when Moving



(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 → overestimating NBHD impacts on children
- The placebo tests clearly showcase the methodological problems of the popular studies in the literature.

Thanks!

Appendix

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

		0.1.1		N. ()
	Mean	Std. dev.	Median	Num. of obs.
Variable	(1)	(2)	(3)	(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				
Child individual income at 30	24,880	10,007	24,846	258,295
Child family income at 30	41,732	19,911	42,257	257,744
Child cohabiting at 30	0.65	0.48	1.00	258,592
Child years of schooling by 30	14.50	2.55	14.50	251,296
Child individual property value at 30	69,105	92,740	47,726	255,337
Parent family income	43,586	13,549	41,948	252,652
Parent property value	94,273	77,781	86,069	251,903
Nuclear (intact) Family	0.39	0.49	0.00	234,262
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
reacted (numer) running	0.10	0.00	0.00	110/1/2

Table 1: Summary Statistics for Municipality Permanent Residents and Movers

Figure: Number of Moves by Education Level

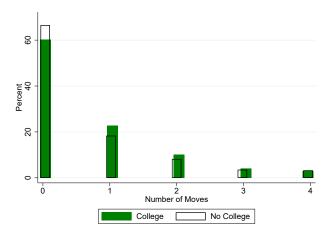
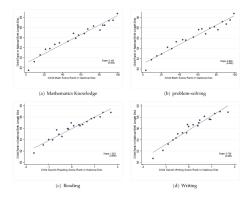


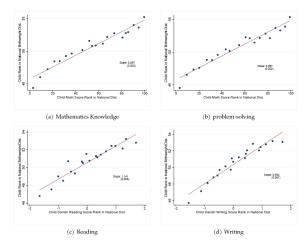


Figure: Birth Length Rank and the Age of the Child at the Time of the Move



Birth Weight and Academic Achievement

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



Academic Achievement and Income Ranks

Figure: Test Scores and Adulthood Income Rank

