Applications in academic research

Lecture 13

Louis SIRUGUE

CPES 2 - Fall 2022

Quick reminder

Standard interpretations

- When both x and y are continuous, the **general** template for the **interpretation** of $\hat{\beta}$ is:

"Everything else equal, a 1 [unit] increase in [x] is associated with an [in/de]crease of [beta] [units] in [y] on average."

- With a discrete x, the interpretation of the coefficient must be **relative to the reference category:**

"Everything else equal, belonging to the [x category] is associated with a [beta] [unit] [higher/lower] average [y] relative to the [reference category]."

- With a **binary y variable**, the coefficient must be interpreted in **percentage points:**

"Everything else equal, a 1 [unit] increase in [x] is associated with a [beta $\times$ 100] percentage point [in/de]crease in the probability that [y equals 1] on average."

Quick reminder

Interpretations with variable transformation

Standardization

- To standardize a variable is to **divide it by its SD**
  - The variation of a standardized variable should not be **interpreted** in units but **in SD**
  - For instance if x and y are continuous and x is standardized, the interpretation becomes:

*"Everything else equal, a 1 **standard deviation** increase in [x] is associated with an [in/de]crease of [beta] [units] in [y] on average."*

- If both x and y are standardized, the slope is the correlation coefficient between x and y

Log-transformation

- The log transformation allows to interpret the coefficient in percentage terms:

Interpretation of the regression coefficient

	y	**log(y)**
x	$\hat{\beta}$ is the unit increase in y due to a 1 unit increase in x	$\hat{\beta} \times 100$ is the % increase in y due to a 1 unit increase in x
log(x)	$\hat{\beta} \div 100$ is the unit increase in y due to a 1% increase in x	$\hat{\beta}$ is the % increase in y due to a 1% increase in x

Quick reminder

Regression table layout

	Birth weight	
	(1)	(2)
Household income	0.002***	0.002***
	(0.0003)	(0.0003)
Girl (ref: Boy)		-141.943***
		(34.878)
Constant	3,127.146***	3,247.126***
	(16.188)	(33.520)
Observations	1,000	963
R^2	0.047	0.063
Note:	*p<0.1; **p<0.05; ***p<0.01	

Regression tables often contain multiple regressions:

- With **one regression in each column**
- And one variable in **each row**
  - With the **point estimate**
  - And a **precision measure** below
- **General info** on each model **at the bottom**
  - Number of observations
  - $\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{n} \hat{\varepsilon_i}^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$
- A **symbology** for the **p-value** testing whether the coefficient is significantly different from 0 or not

Today: Applications in academic research

1. Causal approach (Behaghel et al., 2015)
- 1.1. Structure
- 1.2. Data
- 1.3. Analysis

2. Correlational approach (Chetty et al., 2014)
- 2.1. Empirical approach
- 2.2. National results
- 2.3. Spatial variations
- 2.4. Correlational analysis

3. Structural approach (Nerlove, 1963)
- 3.1. Motivation
- 3.2. Theoretical modeling
- 3.3. Regression expression

4. Wrap up!

Today: Applications in academic research

1. **Causal approach (Behaghel et al., 2015)**
    1.1. Structure
    1.2. Data
    1.3. Analysis

1. Causal approach (Behaghel et al., 2015)

1.1. Structure

- Research papers always start with an **abstract** that briefly **describes the study:**

Unintended Effects of Anonymous Résumés†

By Luc Behaghel, Bruno Crépon, and Thomas Le Barbanchon*

We evaluate an experimental program in which the French public employment service anonymized résumés for firms that were hiring. Firms were free to participate or not; participating firms were then randomly assigned to receive either anonymous résumés or name-bearing ones. We find that participating firms become less likely to interview and hire minority candidates when receiving anonymous résumés. We show how these unexpected results can be explained by the self-selection of firms into the program and by the fact that anonymization prevents the attenuation of negative signals when the candidate belongs to a minority. (*JEL* J15, J68, J71)

1. Causal approach (Behaghel et al., 2015)

1.1. Structure

Typical structure of an empirical research paper:

- Introduction/literature
- Data/Descriptive statistics
- Empirical framework
- Results
- (Heterogeneity)
- Robustness checks
- Conclusion

Structure of Behaghel et al. (2015) is this one:

- Introduction
- Institutional Background
- Experiment and Data Collection
  - Program and Experimental Design
  - Data Collection
- Impact of Anonymous Résumés
  - Interview Rates
  - Hiring Rates
  - Recruitment Success
  - Robustness Checks
- Mechanisms
  - Firms' Participation Decision
  - Résumé Valuation by Participating Firms
- Conclusion

1. Causal approach (Behaghel et al., 2015)

1.1. Structure

Program and Experimental Design

1. **Firm entry in the program:** Firms with more than 50 employees posting vacancies lasting at least three months at the public employment service (PES) were offered to enter the program, which consists in having a 50% chance to receive anonymized instead of standard resumes for that vacancy.
2. **Matching of resumes with vacancies:** The PES posts the vacancy on a variety of media, including a public website asking interested job seekers to apply through the PES branch. The PES agent selects resumes from these applicants and from internal databases of job seekers.
3. **Randomization and anonymization:** Resumes are randomly anonymized or not with a 50% probability and sent to the employer.
4. **Selection of resumes by the employer:** The employer selects the resumes of applicants she would like to interview and contact them (through the PES if resumes are anonymized).

1. Causal approach (Behaghel et al., 2015)

1.2. Data

Data sources

1. **Administrative data**
   - **Coverage:** All firms and all job seekers who used the public employment services in the experimental areas during (and after) the program
   - **Content:** information on the firm (size, sector), on the job position offered (occupation level, type of contract) and limited information on candidates (unless the candidate is filed as unemployed)

2. **Telephone interviews:**
   - **Coverage:** All firms entering the program, a subsample of firms that declined to participate, subsamples of applicants to vacancies posted by these two groups of firms both during and after the experiment
   - **Content:** additional characteristics of the vacancy and of the recruiter (characteristics that could be associated with a differential treatment of candidates), questions on the result of the recruitment (time to hiring and match quality)

1. Causal approach (Behaghel et al., 2015)

1.2. Data

Sample description

- **1,005 firms entered the program (608 declined):**
  - 385 firms in the control group
  - 366 firms in the treatment group
  - 254 firms not allocated because canceled or job filled too early
- **Sample of 1,268 applicants:**
  - 660 to vacancies from the control group
  - 608 to vacancies from the treatment group
  - 203 to vacancies from firms that withdrew before randomization
- **Main variables:**
  - Whether the candidates is from the minority or the majority
  - Whether the resume was anonymized
  - Whether the employer called back for an interview

- **Authors use sampling weights:**
  - Representativity of the sample
  - Non-response bias correction
  - The weight associated with an individual can be viewed as the number of individuals she represents

1. Causal approach (Behaghel et al., 2015)

1.2. Data

- Import the data

```
library(haven)
data_rct <- read_dta("data_candidates_mainsample.dta")
View(data_rct)
```

data_rct

Filter | Cols: « < 1 - 50 > »

	ID_OFFRE Identifier vacancy	ID_CANDIDAT Identifier candidate	CVA Treatment: anonymous resume	REFUSAL Recruiter refused the experiment	ENTRETIEN Interviewed	RECRUTE Hired	POIDS_SEL Sampling weights (within and out of the experiment)	PREN_MUSULMAN Muslim souding name	ZUS_CUCS Deprived neighborhood	ORIGINE_IM_1 Immigrant	ORIGINE_IM_2 Child of immigrant (father)
1	1	1759	1	NA	0	0	5.352863	0	0	0	
2	1	6177	1	NA	0	0	5.352863	0	0	0	
3	2	371	0	NA	0	0	2.676431	0	0	0	
4	3	230	0	NA	0	0	2.676431	1	0	1	
5	4	1885	0	NA	0	0	5.352863	0	0	0	
6	5	3407	1	NA	0	0	4.014647	1	1	0	
7	5	5546	1	NA	0	0	4.906791	0	0	0	
8	5	7081	1	NA	0	0	4.906791	0	0	0	
9	5	4608	1	NA	0	0	4.906791	0	0	0	
10	6	20	1	NA	0	0	28.102530	1	0	1	
11	8	5943	1	NA	0	0	2.007324	0	0	0	
12	8	6568	1	NA	0	0	2.007324	0	0	0	
13	9	4408	0	NA	0	0	21.411451	0	0	0	
14	11	6384	1	NA	0	0	21.411451	0	0	0	

1. Causal approach (Behaghel et al., 2015)

1.2. Data

- Subset the data

```
data_rct <- data_rct %>%
  filter(!is.na(CVA)) %>%                    # Keep participating firms
  select(treatment = CVA,                    # Select and rename variables
         minority = ZouI,                    # of interest
         interview = ENTRETIEN,
         weight = POIDS_SEL)

head(data_rct, 5)
```

```
## # A tibble: 5 x 4
##   treatment minority interview weight
##       <dbl>    <dbl>     <dbl>  <dbl>
## 1         1        0         0   5.35
## 2         1        1         0   5.35
## 3         0        0         0   2.68
## 4         0        1         0   2.68
## 5         0        0         0   5.35
```

→ We want to know whether anonymizing resumes helped reducing labor market discrimination toward the minority group

1. Causal approach (Behaghel et al., 2015)

1.3. Analysis

- Authors use the following **notations**
  - An indicates whether the resume is **anonymous**
  - D indicates whether the candidate is from the **minority**
  - Y indicates whether the candidate obtained an **interview**

- The **parameter of interest** then writes:

$$\delta = \underbrace{(\overline{Y}^{An=1,D=1} - \overline{Y}^{An=1,D=0})}_{\substack{\text{Difference in interview rates} \\ \text{between the majority and the minority} \\ \text{when resumes are anonymized}}} - \underbrace{(\overline{Y}^{An=0,D=1} - \overline{Y}^{An=0,D=0})}_{\substack{\text{Difference in interview rates} \\ \text{between the majority and the minority} \\ \text{when resumes are } \underline{\text{not}} \text{ anonymized}}}$$

➜ What sign do you expect for δ?

1. Causal approach (Behaghel et al., 2015)

1.3. Analysis

```
means <- data_rct %>%
  group_by(treatment, minority) %>%
  summarise(means = weighted.mean(interview, weight))
```

treatment	minority	means
0	0	0.12
0	1	0.09
1	0	0.18
1	1	0.05

```
means <- means %>% group_by(treatment) %>%
  summarise(discrim = means[2] - means[1])
```

treatment	discrim
0	-0.02
1	-0.13

```
means$discrim[2] - means$discrim[1]
```

```
## [1] -0.1067092
```

The interview rate of the minority is even lower than the majority in the treatment group

Practice

08:00

1) Estimate this parameter of interest using a regression

Hint: To apply weights in a regression you can indicate the weighting variable in the `weights` *argument*

```
lm(y ~ x1 + x2 + ..., data, weights = )
```

- Reminder:

```
library(tidyverse)
library(haven)
data_rct <- read_dta("data_candidates_mainsample.dta") %>%      # read .dta data
  filter(!is.na(CVA)) %>%                                       # Keep participating firms
  rename(treatment = CVA, minority = ZouI,                      # Rename variables of interest
         interview = ENTRETIEN, weight = POIDS_SEL) %>%
  select(treatment, minority, interview, weight)                # Select variables of interest
```

$$\delta = \underbrace{(\overline{Y}^{An=1,D=1} - \overline{Y}^{An=1,D=0})}_{\text{Difference in interview rates between the majority and the minority when resumes are anonymized}} - \underbrace{(\overline{Y}^{An=0,D=1} - \overline{Y}^{An=0,D=0})}_{\text{Difference in interview rates between the majority and the minority when resumes are not anonymized}}$$

Solution

- We want to see how the effect of the minority variable varies with the treatment variable
  - In the regression framework, this is what interactions allow to capture

$$Y_i = \alpha + \beta D_i + \gamma An_i + \delta D_i \times An_i + \varepsilon_i$$

```
summary(lm(interview ~ minority + treatment + minority*treatment,
           data_rct, weights = weight))$coefficients
```

```
##                       Estimate Std. Error    t value     Pr(>|t|)
## (Intercept)         0.11638530 0.01630149  7.1395491 1.575140e-12
## minority           -0.02365790 0.02368346 -0.9989208 3.180243e-01
## treatment           0.06101349 0.02419977  2.5212424 1.181630e-02
## minority:treatment -0.10670915 0.03479712 -3.0666092 2.210982e-03
```

  - α is the interview rate for individuals in both reference groups (majority/control)
  - β is the difference in means between the minority and the majority in the control group
  - γ is the difference in means between the treatment and the control group for the majority group
  - δ is how this difference in means between the minority and the majority differ between the treatment and the control group

1. Causal approach (Behaghel et al., 2015)

1.3. Analysis

- Why the effect is negative?

TABLE 7—INTERVIEW AND HIRING RATES IN FIRMS ACCEPTING TO PARTICIPATE (*and Randomized in the Control Group*) AND IN REFUSING FIRMS

	All	Minority (D)	Majority (ND)	Gap (D–ND)
Panel A. Interview rates				
Participating (only controls) firms (c)	0.143	0.146	0.140	0.006
	(0.015)	(0.021)	(0.022)	(0.031)
Nonparticipating firms (r)	0.146	0.073	0.210	−0.137**
	(0.038)	(0.028)	(0.059)	(0.063)
Difference (c-r)	−0.004	0.073**	−0.070	0.143**
	(0.041)	(0.035)	(0.063)	(0.070)
Number of candidates	1,378	759	619	1,378
Number of vacant jobs	507	374	376	507

- Compare the **interview rates** of the **control group** to those of **non-participating firms**
  - **Non-participating firms interview way less the minority** compared to the control group
  - Only firms who interview as much from the minority as from the majority entered the program

➜ Selection bias

Overview

1. Causal approach (Behaghel et al., 2015) ✓
- 1.1. Structure
- 1.2. Data
- 1.3. Analysis

2. Correlational approach (Chetty et al., 2014)
- 2.1. Empirical approach
- 2.2. National results
- 2.3. Spatial variations
- 2.4. Correlational analysis

3. Structural approach (Nerlove, 1963)
- 3.1. Motivation
- 3.2. Theoretical modeling
- 3.3. Regression expression

4. Wrap up!

Overview

1. Causal approach (Behaghel et al., 2015) ✓
- 1.1. Structure
- 1.2. Data
- 1.3. Analysis

2. Correlational approach (Chetty et al., 2014)
- 2.1. Empirical approach
- 2.2. National results
- 2.3. Spatial variations
- 2.4. Correlational analysis

2. Correlational approach (Chetty et al., 2014)

2.1. Empirical approach

Abstract

We use administrative records on the incomes of more than 40 million children and their parents to describe three features of intergenerational mobility in the United States. First, we characterize the joint distribution of parent and child income at the national level. The conditional expectation of child income given parent income is linear in percentile ranks. On average, a 10 percentile increase in parent income is associated with a 3.4 percentile increase in a child's income. Second, intergenerational mobility varies substantially across areas within the U.S. For example, the probability that a child reaches the top quintile of the national income distribution starting from a family in the bottom quintile is 4.4% in Charlotte but 12.9% in San Jose. Third, we explore the factors correlated with upward mobility. High mobility areas have (1) less residential segregation, (2) less income inequality, (3) better primary schools, (4) greater social capital, and (5) greater family stability. While our descriptive analysis does not identify the causal mechanisms that determine upward mobility, the publicly available statistics on intergenerational mobility developed here can facilitate research on such mechanisms.

2. Correlational approach (Chetty et al., 2014)

2.1. Empirical approach

- How to characterize the **joint distribution of parent and child income?**

The intergenerational elasticity:

$$\log(y_i^c) = \alpha + \beta_{IGE} \log(y_i^p) + \varepsilon_i$$

➜ $\hat{\beta}$ **would be the expected percentage increase in child income for a 1% increase in parent income**

The rank-rank correlation:

$$\text{percentile}(y_i^c) = \alpha + \beta_{RRC} \text{percentile}(y_i^p) + \varepsilon_i$$

- **In this particular case, because the dependant and the independant variables have the same variance, the regression coefficient equals the correlation coefficient**

2. Correlational approach (Chetty et al., 2014)

2.1. Empirical approach

$$\beta = \frac{\mathrm{Cov}(x, y)}{\mathrm{Var}(x)}$$

$$= \frac{\mathrm{Cov}(x, y)}{\mathrm{SD}(x) \times \mathrm{SD}(x)} \times \frac{\mathrm{SD}(y)}{\mathrm{SD}(y)}$$

$$= \frac{\mathrm{Cov}(x, y)}{\mathrm{SD}(x) \times \mathrm{SD}(y)} \times \frac{\mathrm{SD}(y)}{\mathrm{SD}(x)}$$

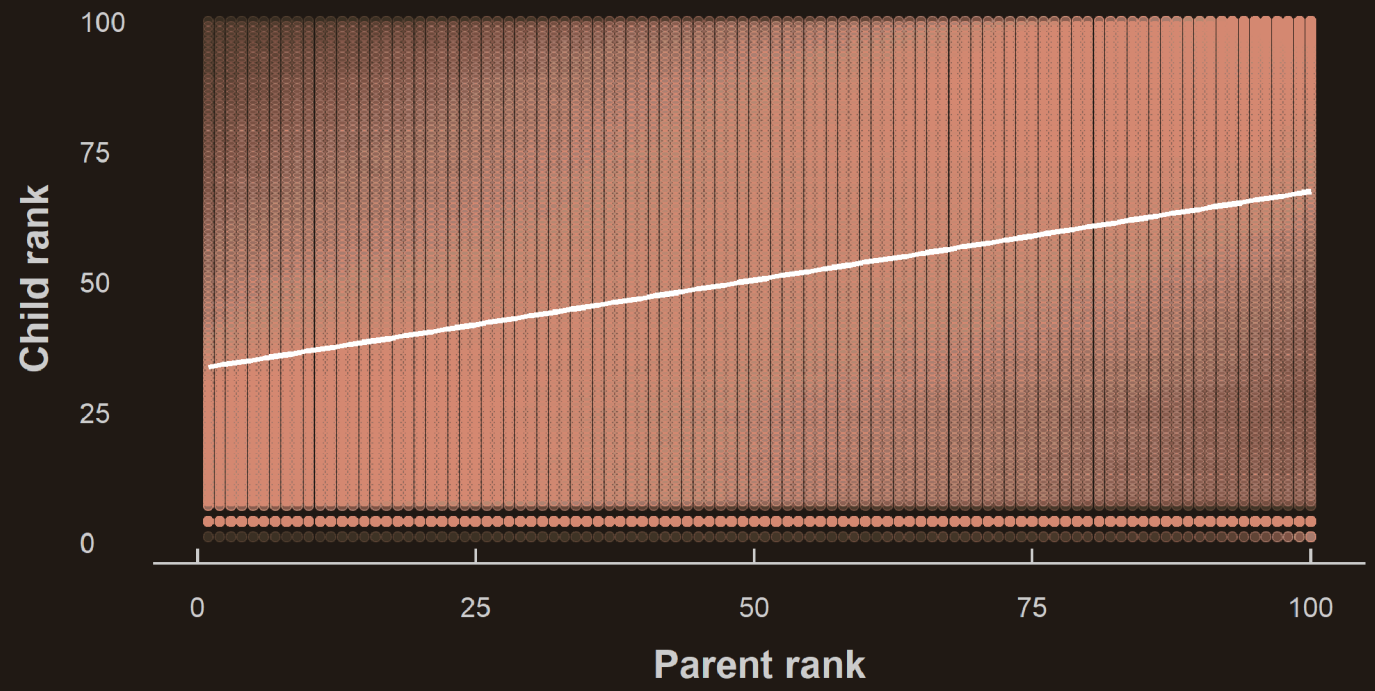
$$= \mathrm{Cor}(x, y) \times \frac{\mathrm{SD}(y)}{\mathrm{SD}(x)}$$

- $\mathrm{SD}(\log(y_i^c)) \lessgtr \mathrm{SD}(\log(y_i^p))$
  - The standard deviation of log income can be viewed as a measure of inequality
  - The IGE is sensitive to relative inequality across generations
- $\mathrm{SD}(\mathrm{percentile}(y_i^c)) = \mathrm{SD}(\mathrm{percentile}(y_i^p))$
  - The RRC is *not* sensitive to relative inequality across generations
  - And the regression coefficient indeed equals the correlation coefficient

2. Correlational approach (Chetty et al., 2014)

2.1. Empirical approach

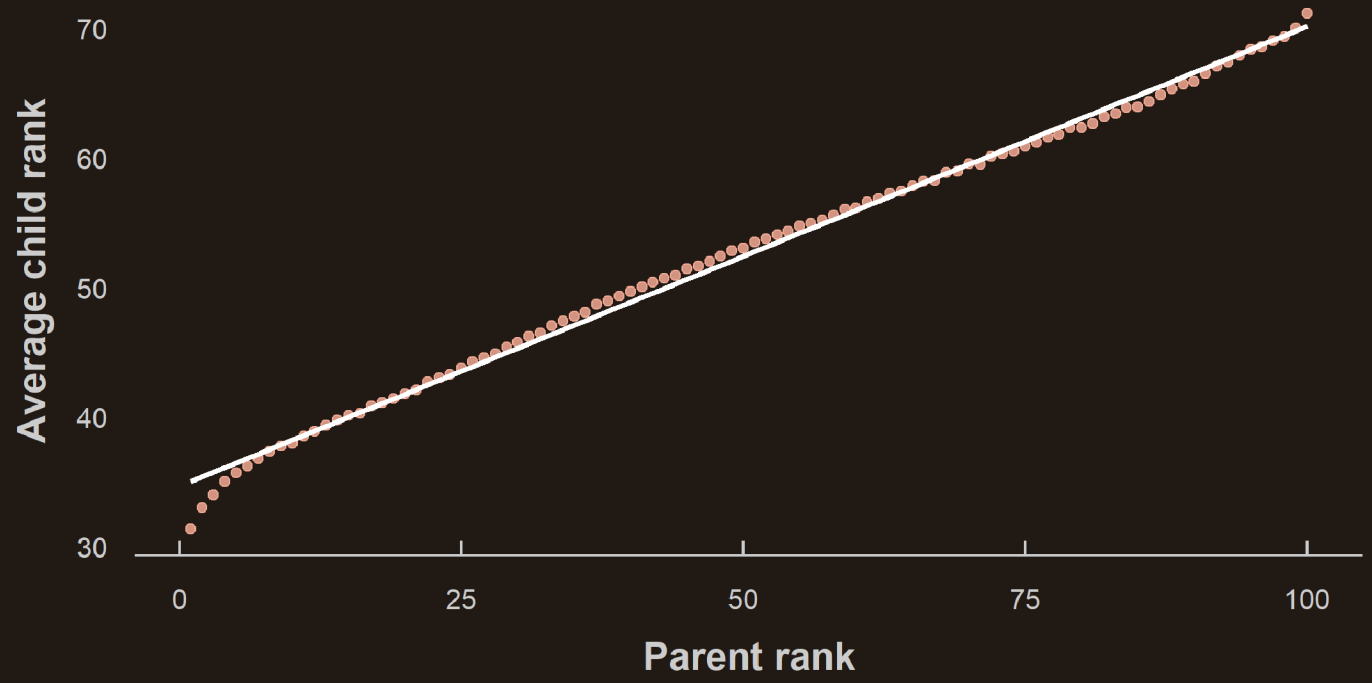
- Here is what the fit of the relationship between parents and child income ranks looks like
  - We can't see much, even at 1% opacity
  - We can't even tell whether or not a linear fit is appropriate



2. Correlational approach (Chetty et al., 2014)

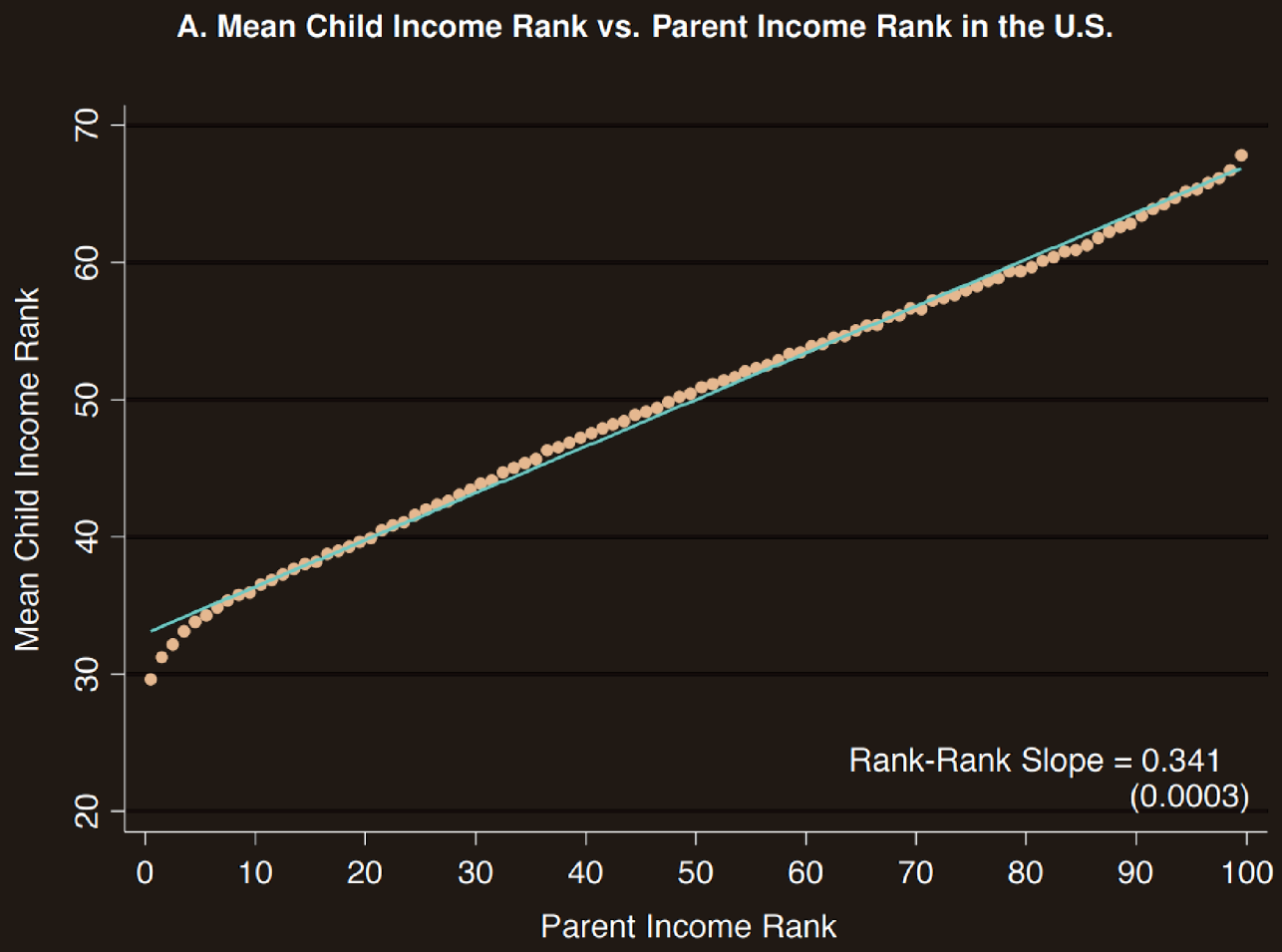
2.1. Empirical approach

- So authors compute the average child rank for each parent percentile group
  - The resulting visual representation is much clearer
  - And it allows to see whether or not a linear specification is appropriate



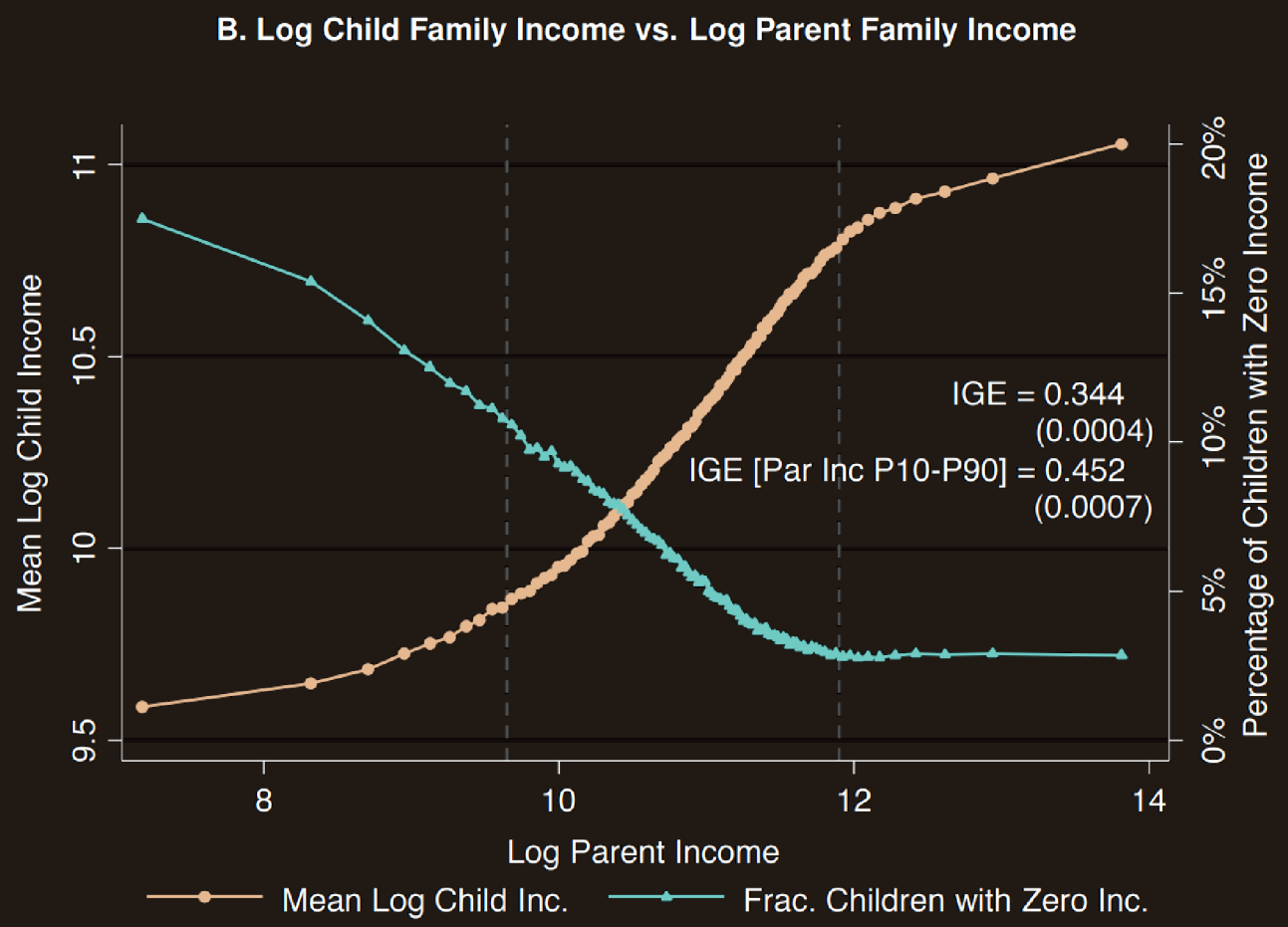
2. Correlational approach (Chetty et al., 2014)

2.2. National results



2. Correlational approach (Chetty et al., 2014)

2.2. National results



- Authors do the same for the IGE
- For each parent income percentile:
  - x: Mean parent log income
  - y: Mean child log income
- The relationship is non-linear

2. Correlational approach (Chetty et al., 2014)

2.3. Spatial variations

- Then authors estimate the rank-rank regression separately for each commuting zone

$$\text{percentile}(y_i^c) = \alpha + \beta_{RRC}\text{percentile}(y_i^p) + \varepsilon_i$$

- From these local estimations they derive **two statistics:**

Relative mobility: $\hat{\beta}_{RRC}$

- The slope of the rank-rank relationship
  - **Expected rank increase** for a children had their parents been ranked 1 percentile higher
  - The estimated increase indicates where the children would locate in **relative terms**

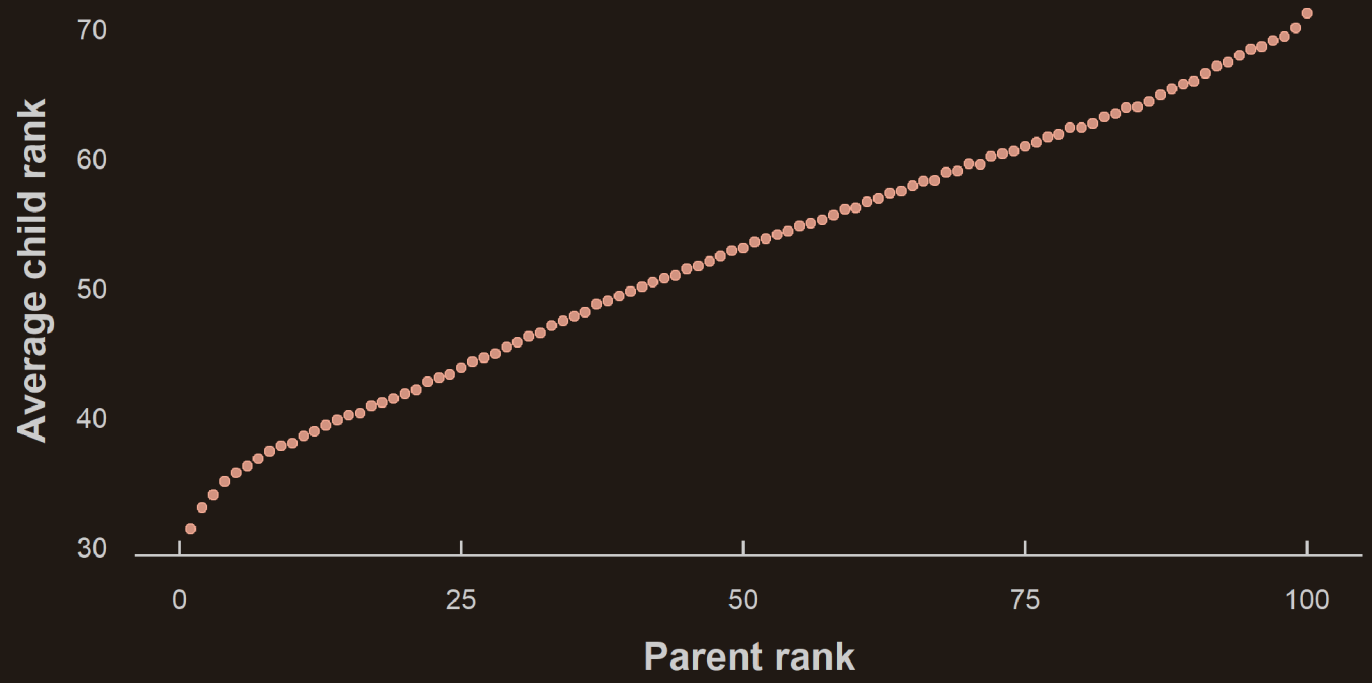
Absolute mobility: $\widehat{\alpha} + 25 \times \hat{\beta}_{RRC}$

- The fitted value at $x = 25$
  - **Expected percentile rank** for children whose parents locate at the 25^{th} percentile
  - The estimated percentile indicates where the children would locate in **absolute terms**

2. Correlational approach (Chetty et al., 2014)

2.3. Spatial variations

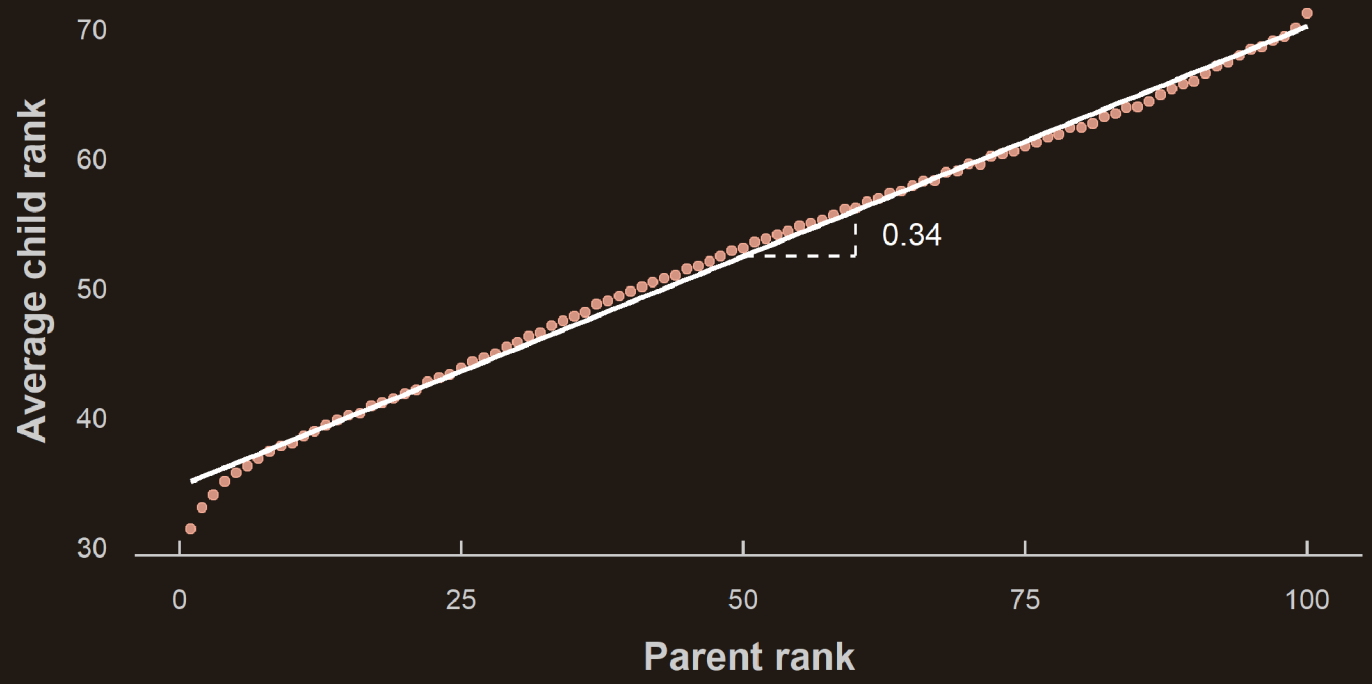
- Here is an illustration on the national-level relationship:
  -
  -



2. Correlational approach (Chetty et al., 2014)

2.3. Spatial variations

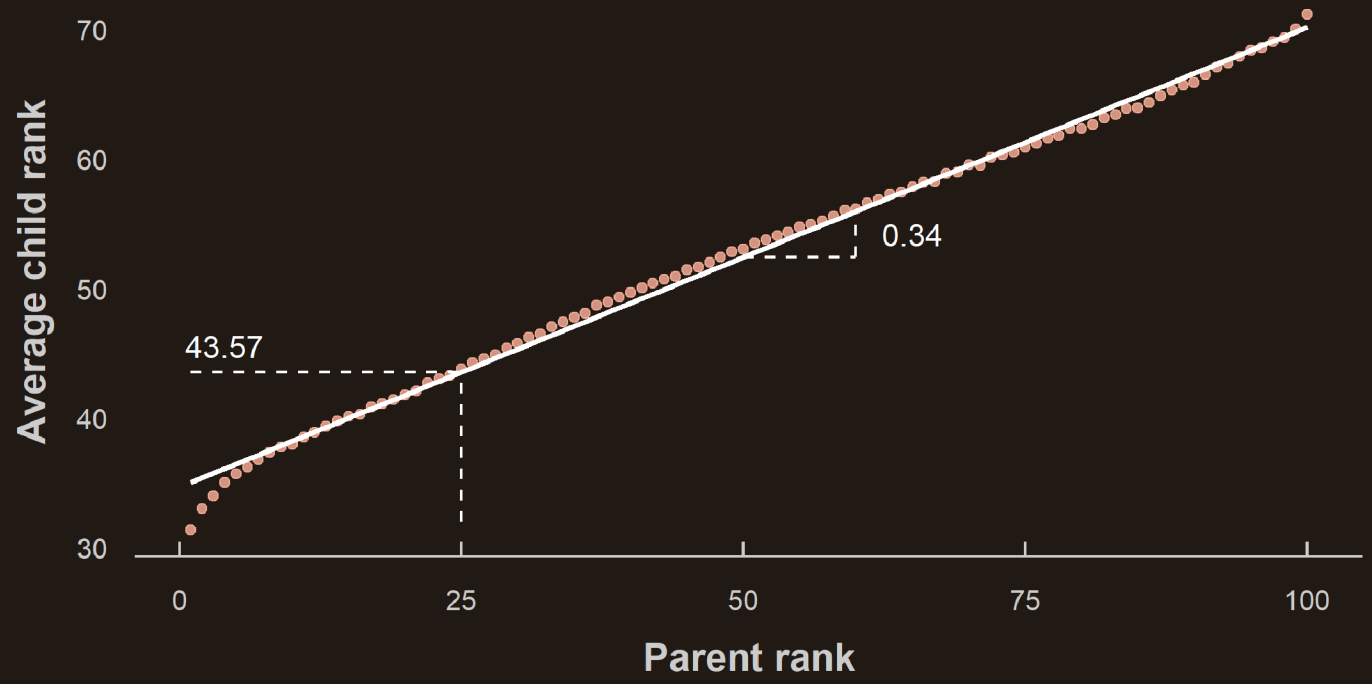
- Here is an illustration on the national-level relationship:
  - The **relative mobility** is the slope - the rank-rank correlation
  -



2. Correlational approach (Chetty et al., 2014)

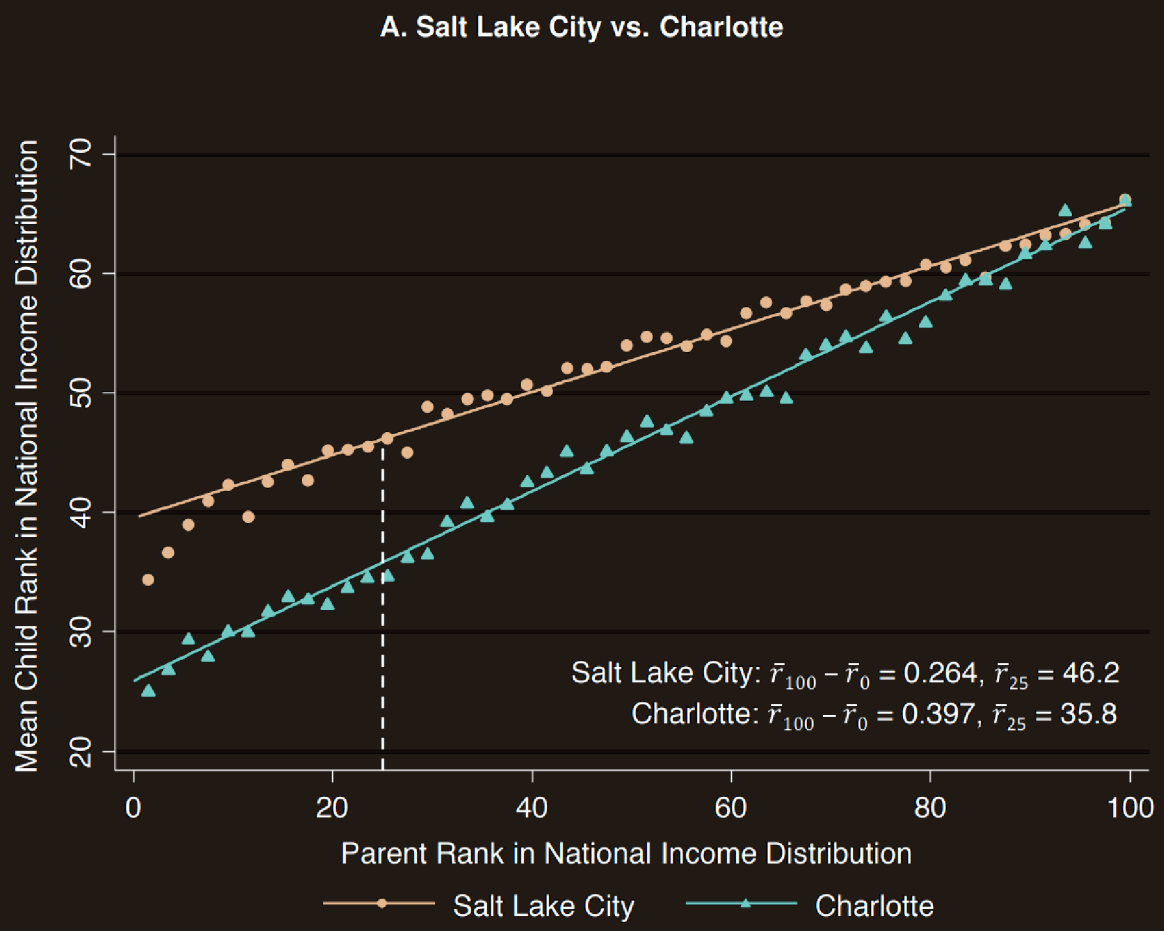
2.3. Spatial variations

- Here is an illustration on the national-level relationship:
  - The **relative mobility** is the slope - the rank-rank correlation
  - The **absolute mobility** is the fitted value for x = 25



2. Correlational approach (Chetty et al., 2014)

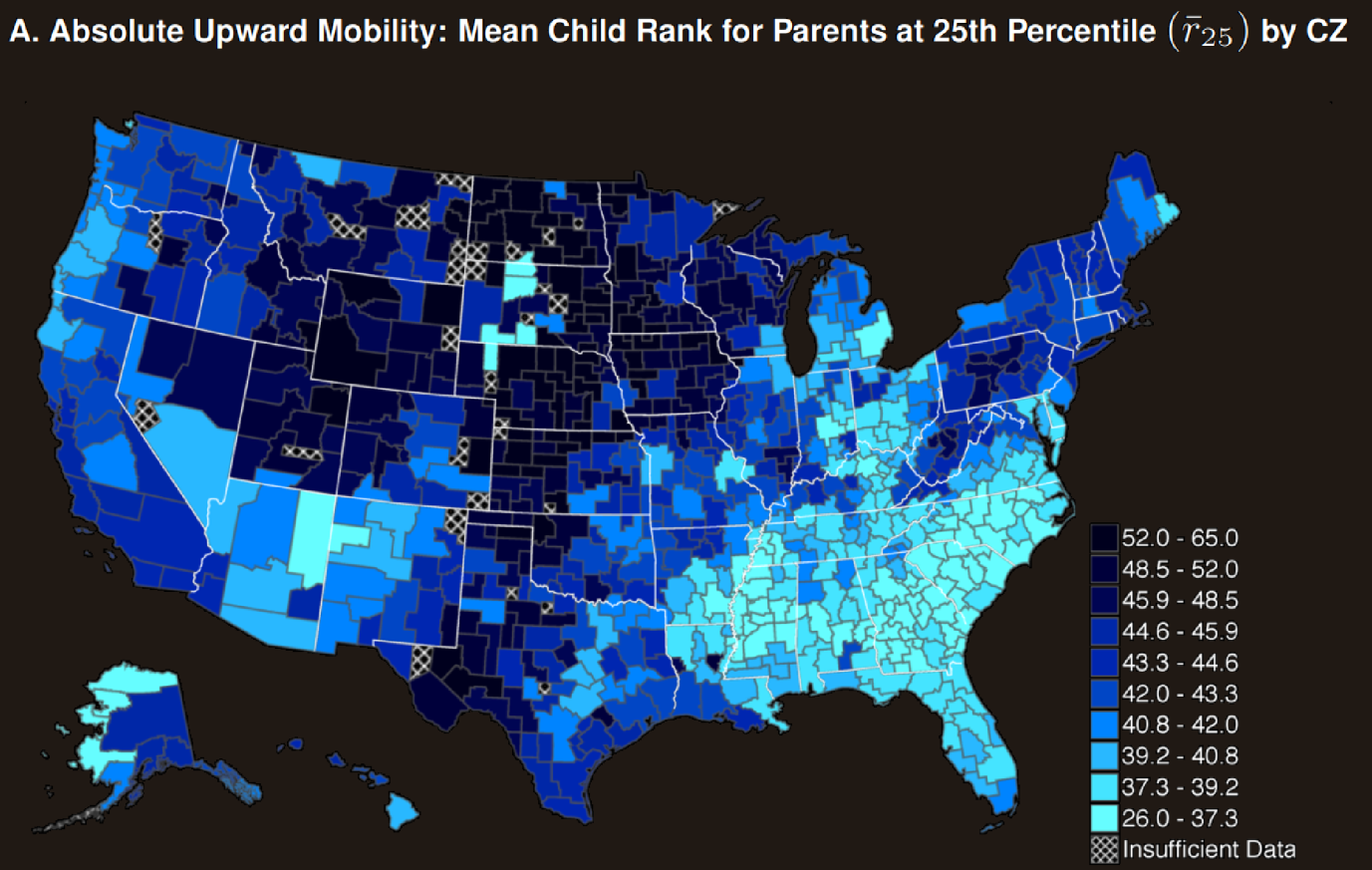
2.3. Spatial variations



- Authors compute intergenerational persistence in each commuting zone separately
- And plot the results on a map

2. Correlational approach (Chetty et al., 2014)

2.3. Spatial variations



2. Correlational approach (Chetty et al., 2014)

2.4. Correlational analysis

- Then they investigate whether local **characteristics of commuting zones** are related to mobility
- But regressing directly upward mobility on different characteristics would give:
  - **Lower coefficients** for variables with **bigger metrics** (test scores)
  - **Higher coefficients** for variables with **smaller metrics** (fraction of single mothers)

- So authors **standardize** their **variables** for the comparability of their estimates

$$\beta = \frac{\mathrm{Cov}(\frac{x}{\mathrm{SD}(x)}, \frac{y}{\mathrm{SD}(y)})}{\mathrm{Var}(\frac{x}{\mathrm{SD}(x)})}$$

- To simplify this equation, you need to know that:
  - $\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$
  - $\mathrm{Cov}(aX, bY) = ab\mathrm{Cov}(X, Y)$

$$\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$$

2. Correlational approach (Chetty et al., 2014)

2.4. Correlational analysis

$$\beta = \frac{\text{Cov}(\frac{x}{\text{SD}(x)}, \frac{y}{\text{SD}(y)})}{\text{Var}(\frac{x}{\text{SD}(x)})}$$

$$= \frac{\frac{1}{\text{SD}(x)\text{SD}(y)}\text{Cov}(x,y)}{\frac{1}{\text{SD}(x)^2}\text{Var}(x)}$$

$$= \frac{\text{Cov}(x,y)}{\text{SD}(x)\text{SD}(y)} \times \frac{\text{SD}(x)^2}{\text{Var}(x)}$$

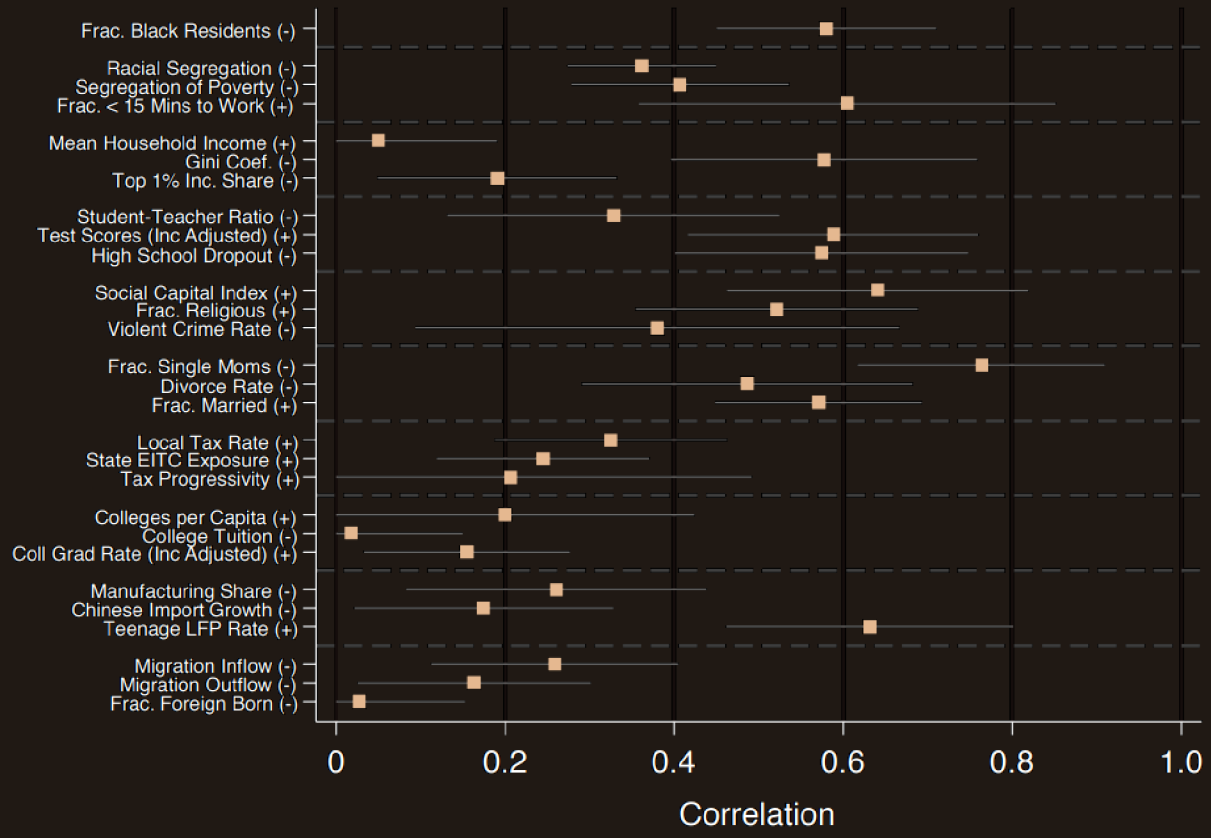
$$= \text{Corr}(x,y)$$

➜ ***Standardizing variables allows to obtain a correlation coefficient from a regression***

2. Correlational approach (Chetty et al., 2014)

2.4. Correlational analysis

FIGURE VIII: Correlates of Spatial Variation in Upward Mobility



Note that these coefficients combine:

- A neighborhood effect
- A selection effect

Overview

1. Causal approach (Behaghel et al., 2015) ✓
- 1.1. Structure
- 1.2. Data
- 1.3. Analysis

2. Correlational approach (Chetty et al., 2014) ✓
- 2.1. Empirical approach
- 2.2. National results
- 2.3. Spatial variations
- 2.4. Correlational analysis

3. Structural approach (Nerlove, 1963)
- 3.1. Motivation
- 3.2. Theoretical modeling
- 3.3. Regression expression

4. Wrap up!

Overview

1. Causal approach (Behaghel et al., 2015) ✓
- 1.1. Structure
- 1.2. Data
- 1.3. Analysis

2. Correlational approach (Chetty et al., 2014) ✓
- 2.1. Empirical approach
- 2.2. National results
- 2.3. Spatial variations
- 2.4. Correlational analysis

3. Structural approach (Nerlove, 1963)
- 3.1. Motivation
- 3.2. Theoretical modeling
- 3.3. Regression expression

3. Structural approach (Nerlove, 1963)

3.1. Motivation

- The **structural approach** refers to the following methodology:
  1. **Theoretical modeling** of the phenomenon of interest
  2. **Expressing the model** parameters **as** the coefficients of **a regression**
  3. Run the corresponding regressions on data to **estimate the parameters** of the model

- Structural papers are **more and more complex** on the theoretical side
  - The current standards in this literature are beyond the scope of this course
  - So we are going to explore a quite **old study** for this section

- Nerlove (1963) studies the **returns to scale in the electricity supply industry**
  - What is the output elasticity of each input?
  - Are the returns to scale **positive or negative?**

3. Structural approach (Nerlove, 1963)

3.1. Motivation

- Nerlove (1963) assume the following **production function:**

$$Y = AL^{\lambda}K^{\kappa}F^{\varphi}u$$

- And the following **cost function:**

$$C = p_L L + p_K K + p_F F$$

With:

Inputs	Output elasticities	Prices	Other parameters
L : Labor input	λ : OE of labor	p_L : Wage rate	A : Total factor productivity
K : Capital input	κ : OE of capital	p_K : Price of capital	
F : Fuel input	φ : OE of fuel	p_F : Price of fuel	u : Efficiency residual

3. Structural approach (Nerlove, 1963)

3.1. Motivation

- **Theoretically** we could **estimate** the output elasticities **directly from the production function**
  - The trick is to **put everything in log** such that the exponents become the parameters of the equation
  - We call this transformation a **log-linearization**

$$\begin{aligned}\log(Y) &= \log\left(AL^{\lambda}K^{\kappa}F^{\varphi}u\right) \\ &= \log(A) + \log\left(L^{\lambda}\right) + \log\left(K^{\kappa}\right) + \log\left(F^{\varphi}\right) + \log(u) \\ &= \underbrace{\log(A)}_{\text{Constant}} + \lambda\log(L) + \kappa\log(K) + \varphi\log(F) + \underbrace{\log(u)}_{\text{Residuals}}\end{aligned}$$

- Regressing log output on the log inputs directly gives the **elasticities**
  - But Nerlove (1963) does not have access to data on firms' inputs
  - Still, he has **data on the price** of each input
  - His solution is to derive an expression that allows to **estimate the elasticities from the price**

3. Structural approach (Nerlove, 1963)

3.2. Theoretical modeling

- To **simplify** algebra, we're going to consider **capital and labor only**
  - But the principle remains the same
  - We need to **solve the model** by minimizing the cost constrained by the production function

$$\begin{cases} \min & C = p_L L + p_K K \\ \text{s.t.} & Y = AL^\lambda K^\kappa u \end{cases} \iff \min \ \mathcal{L} = p_L L + p_K K + \mu(Y - AL^\lambda K^\kappa u)$$

- Equate partial derivatives to 0

$$\frac{\partial \mathcal{L}}{\partial L} = 0 \Leftrightarrow p_L = \mu A \lambda L^{\lambda-1} K^\kappa u$$

$$\frac{\partial \mathcal{L}}{\partial K} = 0 \Leftrightarrow p_K = \mu A \kappa L^\lambda K^{\kappa-1} u$$

$$\frac{\partial \mathcal{L}}{\partial \mu} = 0 \Leftrightarrow Y = AL^\lambda K^\kappa u$$

- Same with K and L as functions of each other

$$\frac{p_L}{p_K} = \frac{\mu A \lambda L^{\lambda-1} K^\kappa u}{\mu A \kappa L^\lambda K^{\kappa-1} u} = \frac{\lambda L^{\lambda-1} K^\kappa}{\kappa L^\lambda K^{\kappa-1}} = \frac{\lambda K}{\kappa L}$$

$$L = K \frac{p_K}{p_L} \frac{\lambda}{\kappa} \qquad K = L \frac{p_L}{p_K} \frac{\kappa}{\lambda}$$

3. Structural approach (Nerlove, 1963)

3.2. Theoretical modeling

- Express Y as a function of L only and solve for L

$$Y = AL^{\lambda}\left(L\frac{p_L}{p_K}\frac{\kappa}{\lambda}\right)^{\kappa}u$$

$$L^{\lambda+\kappa} = \frac{Y}{Au}\frac{1}{\left(\frac{p_L}{p_K}\frac{\kappa}{\lambda}\right)^{\kappa}}$$

$$L = \left[\frac{Y}{Au}\frac{1}{\left(\frac{p_L}{p_K}\frac{\kappa}{\lambda}\right)^{\kappa}}\right]^{\frac{1}{\lambda+\kappa}} = \left(\frac{Y}{Au}\right)^{\frac{1}{\lambda+\kappa}}\left(\frac{p_K}{p_L}\frac{\lambda}{\kappa}\right)^{\frac{\kappa}{\lambda+\kappa}}$$

- Same with K

$$Y = A\left(K\frac{p_K}{p_L}\frac{\lambda}{\kappa}\right)^{\lambda}K^{\kappa}u$$

$$K^{\lambda+\kappa} = \frac{Y}{Au}\frac{1}{\left(\frac{p_K}{p_L}\frac{\lambda}{\kappa}\right)^{\lambda}}$$

$$K = \left[\frac{Y}{Au}\frac{1}{\left(\frac{p_K}{p_L}\frac{\lambda}{\kappa}\right)^{\lambda}}\right]^{\frac{1}{\lambda+\kappa}} = \left(\frac{Y}{Au}\right)^{\frac{1}{\lambda+\kappa}}\left(\frac{p_L}{p_K}\frac{\kappa}{\lambda}\right)^{\frac{\lambda}{\lambda+\kappa}}$$

3. Structural approach (Nerlove, 1963)

3.2. Theoretical modeling

- Inject K and L back in the cost function and factorize

$$C = p_L \left(\frac{Y}{Au}\right)^{\frac{1}{\lambda+\kappa}} \left(\frac{p_K}{p_L}\frac{\lambda}{\kappa}\right)^{\frac{\kappa}{\lambda+\kappa}} + p_K \left(\frac{Y}{Au}\right)^{\frac{1}{\lambda+\kappa}} \left(\frac{p_L}{p_K}\frac{\kappa}{\lambda}\right)^{\frac{\lambda}{\lambda+\kappa}}$$

$$C = \left(\frac{Y}{Au}\right)^{\frac{1}{\lambda+\kappa}} \left[p_L \left(\frac{p_K}{p_L}\frac{\lambda}{\kappa}\right)^{\frac{\kappa}{\lambda+\kappa}} + p_K \left(\frac{p_L}{p_K}\frac{\kappa}{\lambda}\right)^{\frac{\lambda}{\lambda+\kappa}}\right]$$

$$C = \left(\frac{Y}{Au}\right)^{\frac{1}{\lambda+\kappa}} \left[p_L^{1-\frac{\kappa}{\lambda+\kappa}} p_K^{\frac{\kappa}{\lambda+\kappa}} \left(\frac{\lambda}{\kappa}\right)^{\frac{\kappa}{\lambda+\kappa}} + p_K^{1-\frac{\lambda}{\lambda+\kappa}} p_L^{\frac{\lambda}{\lambda+\kappa}} \left(\frac{\kappa}{\lambda}\right)^{\frac{\lambda}{\lambda+\kappa}}\right]$$

$$C = \left(\frac{Y}{Au}\right)^{\frac{1}{\lambda+\kappa}} \left[p_L^{\frac{\lambda}{\lambda+\kappa}} p_K^{\frac{\kappa}{\lambda+\kappa}} \left(\frac{\lambda}{\kappa}\right)^{\frac{\kappa}{\lambda+\kappa}} + p_K^{\frac{\kappa}{\lambda+\kappa}} p_L^{\frac{\lambda}{\lambda+\kappa}} \left(\frac{\kappa}{\lambda}\right)^{\frac{\lambda}{\lambda+\kappa}}\right]$$

3. Structural approach (Nerlove, 1963)

3.2. Theoretical modeling

$$C = \left(\frac{Y}{Au}\right)^{\frac{1}{\lambda+\kappa}} \left[p_L^{\frac{\lambda}{\lambda+\kappa}} p_K^{\frac{\kappa}{\lambda+\kappa}} \left(\frac{\lambda}{\kappa}\right)^{\frac{\kappa}{\lambda+\kappa}} + p_K^{\frac{\kappa}{\lambda+\kappa}} p_L^{\frac{\lambda}{\lambda+\kappa}} \left(\frac{\kappa}{\lambda}\right)^{\frac{\lambda}{\lambda+\kappa}}\right]$$

$$C = \left(\frac{Y}{Au}\right)^{\frac{1}{\lambda+\kappa}} p_L^{\frac{\lambda}{\lambda+\kappa}} p_K^{\frac{\kappa}{\lambda+\kappa}} \left[\left(\frac{\lambda}{\kappa}\right)^{\frac{\kappa}{\lambda+\kappa}} + \left(\frac{\kappa}{\lambda}\right)^{\frac{\lambda}{\lambda+\kappa}}\right]$$

- Isolate what's constant, each variable, and the residual term:

$$C = \underbrace{\left[\frac{\left(\frac{\lambda}{\kappa}\right)^{\kappa} + \left(\frac{\kappa}{\lambda}\right)^{\lambda}}{A}\right]^{\frac{1}{\lambda+\kappa}}}_{\text{Constant}} \times \underbrace{Y^{\frac{1}{\lambda+\kappa}}}_{\text{Output}} \times \underbrace{p_L^{\frac{\lambda}{\lambda+\kappa}}}_{\text{Wage}} \times \underbrace{p_K^{\frac{\kappa}{\lambda+\kappa}}}_{\text{Price of Capital}} \times \underbrace{u^{\frac{1}{\lambda+\kappa}}}_{\text{Residual term}}$$

3. Structural approach (Nerlove, 1963)

3.3. Regression expression

- At this stage we can **log-linearize the equation:**

$$C = \underbrace{\left[\frac{\left(\frac{\lambda}{\kappa}\right)^{\kappa} + \left(\frac{\kappa}{\lambda}\right)^{\lambda}}{A}\right]^{\frac{1}{\lambda+\kappa}}}_{\text{Constant}} \times \underbrace{Y^{\frac{1}{\lambda+\kappa}}}_{\text{Output}} \times \underbrace{p_L^{\frac{\lambda}{\lambda+\kappa}}}_{\text{Wage}} \times \underbrace{p_K^{\frac{\kappa}{\lambda+\kappa}}}_{\substack{\text{Price of} \\ \text{Capital}}} \times \underbrace{u^{\frac{1}{\lambda+\kappa}}}_{\substack{\text{Residual} \\ \text{term}}}$$

$$\log(C) = \log\left(\left[\frac{\left(\frac{\lambda}{\kappa}\right)^{\kappa} + \left(\frac{\kappa}{\lambda}\right)^{\lambda}}{A}\right]^{\frac{1}{\lambda+\kappa}}\right) + \log\left(Y^{\frac{1}{\lambda+\kappa}}\right) + \log\left(p_L^{\frac{\lambda}{\lambda+\kappa}}\right) + \log\left(p_K^{\frac{\kappa}{\lambda+\kappa}}\right) + \log\left(u^{\frac{1}{\lambda+\kappa}}\right)$$

$$\log(C) = \underbrace{\log\left(\left[\frac{\left(\frac{\lambda}{\kappa}\right)^{\kappa} + \left(\frac{\kappa}{\lambda}\right)^{\lambda}}{A}\right]^{\frac{1}{\lambda+\kappa}}\right)}_{\alpha} + \underbrace{\frac{1}{\lambda+\kappa}}_{\beta}\log(Y) + \underbrace{\frac{\lambda}{\lambda+\kappa}}_{\gamma}\log(p_L) + \underbrace{\frac{\kappa}{\lambda+\kappa}}_{\delta}\log(p_K) + \underbrace{\log\left(u^{\frac{1}{\lambda+\kappa}}\right)}_{\varepsilon}$$

3. Structural approach (Nerlove, 1963)

3.3. Regression expression

- Finally we end up with this **regression model:**
  - Where **coefficients** are **composite** objects of the **parameters** of the structural model

$$\log(C) = \alpha + \beta \log(Y) + \gamma \log(p_L) + \delta \log(p_K) + \varepsilon$$

- But note that to **test for CRS**, we don't even need to derive κ and λ explicitely

$$\gamma = \frac{\lambda}{\lambda + \kappa} \quad ; \quad \delta = \frac{\kappa}{\lambda + \kappa}$$

- Indeed, the **null hypothesis** for constant returns to scales writes

$$H_0 : \lambda + \kappa = 1 \quad \Leftrightarrow \quad \frac{\lambda + \kappa}{\lambda + \kappa} = \frac{1}{1} \quad \Leftrightarrow \quad \frac{\lambda}{\lambda + \kappa} + \frac{\kappa}{\lambda + \kappa} = 1 \quad \Leftrightarrow \quad \gamma + \delta = 1$$

Practice

08:00

1) Import the dataset from Nerlove (1963)

```
library(haven)
nerlove <- read_dta("nerlove63.dta")
str(nerlove, give.attr = F)
```

```
## tibble [145 x 5] (S3: tbl_df/tbl/data.frame)
##  $ totcost: num [1:145] 0.082 0.661 0.99 0.315 0.197 ...
##  $ output : num [1:145] 2 3 4 4 5 9 11 13 13 22 ...
##  $ plabor : num [1:145] 2.09 2.05 2.05 1.83 2.12 ...
##  $ pfuel  : num [1:145] 17.9 35.1 35.1 32.2 28.6 ...
##  $ pkap   : num [1:145] 183 174 171 166 233 195 206 150 155 188 ...
```

2) Estimate the parameters of this regression:

$$\log(C) = \alpha + \beta \log(Y) + \gamma \log(p_L) + \delta \log(p_K) + \varepsilon$$

3) Use `linearHypothesis()` from the `car` package to test for CRS

Solution

Estimate the parameters of this regression:

$$\log(C) = \alpha + \beta \log(Y) + \gamma \log(p_L) + \delta \log(p_K) + \varepsilon$$

```
summary(lm(log(totcost) ~ log(output) + log(plabor) + log(pkap), nerlove))$coefficients
```

```
##                Estimate Std. Error    t value     Pr(>|t|)
## (Intercept) -3.82905582 1.87712121 -2.0398554 4.323183e-02
## log(output)  0.70706725 0.01819229 38.8663200 3.187712e-77
## log(plabor)  0.89957689 0.28572006  3.1484555 2.003726e-03
## log(pkap)    0.06079561 0.35250457  0.1724676 8.633172e-01
```

Use `linearHypothesis()` from the `car` package to test for CRS

```
library(car)
linearHypothesis(lm(log(totcost) ~ log(output) + log(plabor) + log(pkap), nerlove),
                 "log(plabor) + log(pkap) = 1")
```

Solution

```
linearHypothesis(lm(log(totcost) ~ log(output) + log(plabor) + log(pkap), nerlove),
                 "log(plabor) + log(pkap) = 1")
```

```
## Linear hypothesis test
##
## Hypothesis:
## log(plabor)  + log(pkap) = 1
##
## Model 1: restricted model
## Model 2: log(totcost) ~ log(output) + log(plabor) + log(pkap)
##
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1    142 24.333
## 2    141 24.332  1 0.0011159 0.0065  0.936
```

- The p-value is equal to 93.6%
  - We cannot reject the hypothesis of constant returns to scale
  - $\hat{\gamma} + \hat{\delta} = .96$ is not sufficiently far from 1 to reject that $\gamma + \delta = 1$

Overview

1. Causal approach (Behaghel et al., 2015) ✓
- 1.1. Structure
- 1.2. Data
- 1.3. Analysis

2. Correlational approach (Chetty et al., 2014) ✓
- 2.1. Empirical approach
- 2.2. National results
- 2.3. Spatial variations
- 2.4. Correlational analysis

3. Structural approach (Nerlove, 1963) ✓
- 3.1. Motivation
- 3.2. Theoretical modeling
- 3.3. Regression expression

4. Wrap up!

4. Wrap up!

Causal approach (Behaghel et al., 2015)

- Applicants resumes randomly anonymized or not before being sent to employers

$$Y_i = \alpha + \beta D_i + \gamma An_i + \delta D_i \times An_i + \varepsilon_i$$

- $\hat{\delta}$ captures how the difference in interview rates between the minority and the majority differs between the treated and the control employers

```
summary(lm(interview ~ minority + treatment + minority*treatment,
           data_rct, weights = weight))$coefficients
```

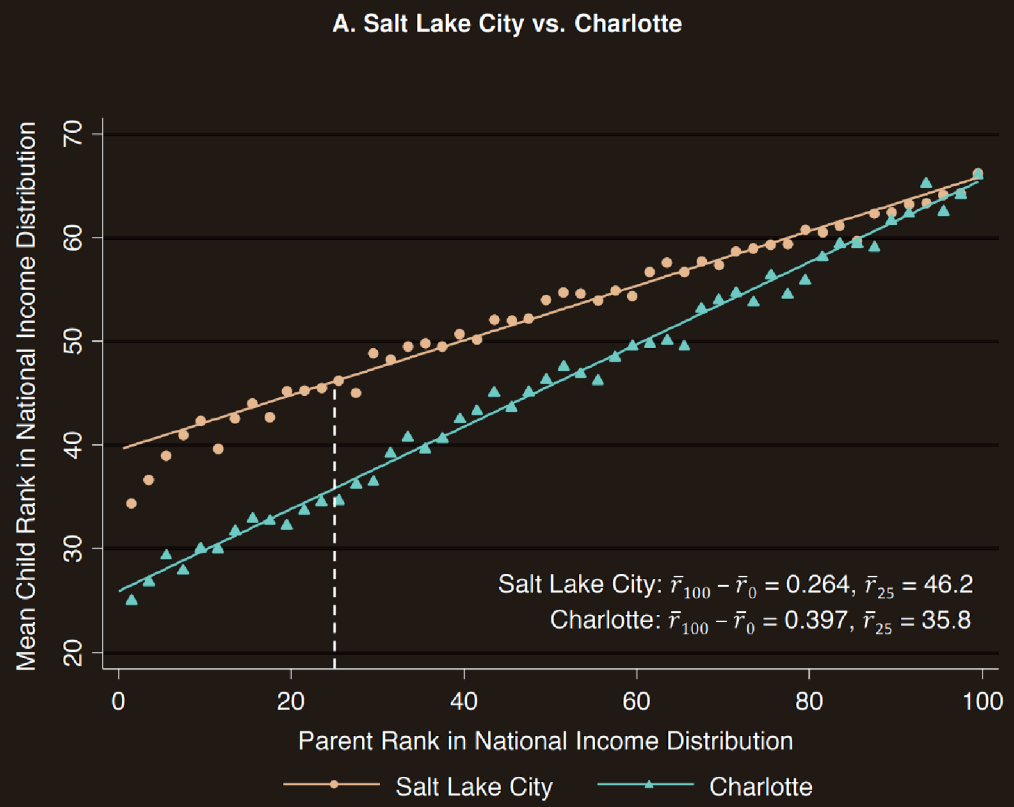
```
##                      Estimate Std. Error    t value     Pr(>|t|)
## (Intercept)         0.11638530 0.01630149  7.1395491 1.575140e-12
## minority           -0.02365790 0.02368346 -0.9989208 3.180243e-01
## treatment           0.06101349 0.02419977  2.5212424 1.181630e-02
## minority:treatment -0.10670915 0.03479712 -3.0666092 2.210982e-03
```

➜ Self-selection issue: discriminatory employers did not enter the program

4. Wrap up!

Correlational approach (Chetty et al., 2014)

$$\text{percentile}(y_i^c) = \alpha + \beta_{RRC}\text{percentile}(y_i^p) + \varepsilon_i$$



Relative mobility: $\widehat{\beta_{RRC}}$

Absolute mobility: $\widehat{\alpha} + 25 \times \widehat{\beta_{RRC}}$

- Strong persitence in the United-States
- Large variations across commuting zones
- Intergenerational mobility correlated with characteristics of childhood environment

4. Wrap up!

Structural approach (Nerlove, 1963)

- **Theoretical modeling**

$$\begin{cases} \min & C = p_L L + p_K K \\ \text{s.t.} & Y = A L^{\lambda} K^{\kappa} u \end{cases} \iff \min \ \mathcal{L} = p_L L + p_K K + \mu(Y - A L^{\lambda} K^{\kappa} u)$$

- **Regression expression**

$$\log(C) = \underbrace{\log\left(\left[\frac{\left(\frac{\lambda}{\kappa}\right)^{\kappa} + \left(\frac{\kappa}{\lambda}\right)^{\lambda}}{A}\right]^{\frac{1}{\lambda+\kappa}}\right)}_{\alpha} + \underbrace{\frac{1}{\lambda+\kappa}}_{\beta}\log(Y) + \underbrace{\frac{\lambda}{\lambda+\kappa}}_{\gamma}\log(p_L) + \underbrace{\frac{\kappa}{\lambda+\kappa}}_{\delta}\log(p_K) + \underbrace{\log\left(u^{\frac{1}{\lambda+\kappa}}\right)}_{\varepsilon}$$

- **Estimation**

$$\log(C) = \alpha + \beta\log(Y) + \gamma\log(p_L) + \delta\log(p_K) + \varepsilon \quad \Rightarrow \quad H_0 : \gamma + \delta = 1$$