Research Opportunities Using RDC-Accessible Data

Richard K. Mansfield
University of Colorado-Boulder

October 16, 2016

Overview

1. Overview of LEHD
2. Description of Current Research
3. Other Avenues for Future Research
4. Disclosure Avoidance Review

Overview of the LEHD

- Longitudinal Employer-Household Dynamics (LEHD) database:
    - Job-level sampling frame covering a near universe of worker-establishment matches from the U.S.
    - Based on quarterly UI records
    - Records contain worker IDs, firm IDs and quarterly job earnings
    - Merged with firm and estab. info. (e.g. geo. location, industry, firm age)
    - Merged with worker info. (age, race, gender, country of origin, some residential location)
    - First cluster of states enter in 1990, nearly all states present by 2004
    - Potential to merge with survey data (CPS, ACS, SIPP)

Project #1: Projecting the Incidence of Local Labor Demand Shocks

- Have U.S. local labor markets become more geographically integrated since the internet arrived?
- If so, for which types of workers and which types of industries?
- What are the implications of the structure of local labor market networks for the incidence of local labor demand shocks?
- How does the distribution of employment and wages across geographic areas and across worker skill types change when:
  - A Toyota plant moves from Ohio to Georgia?
  - A large Google office moves from California to Colorado?
  - A set of tech start-up subsidies is provided in Boise?
  - A hurricane wipes out all businesses located in a ten mile stretch along the Florida coast?
  - A stimulus project creates 300 construction jobs in Casper?

Project #1: Projecting the Incidence of Local Labor Demand Shocks

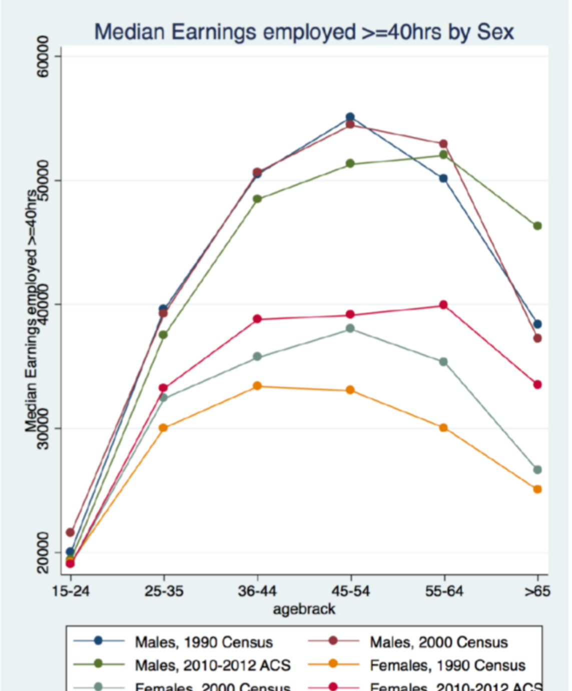
- Since observe all job-to-job transitions between t and $t+1$, and can pinpoint estab. locations:
    - $\Rightarrow$ Assign each worker-firm match at time t to types (rows) based on observable worker, firm, job characteristics.
    - Same for time $t+1$ job matches (columns) ...
    - $\Rightarrow$ Can populate a transition matrix capturing the transition from set of time t to time $t+1$ job matches in quarter or year

Project #1: Projecting the Incidence of Local Labor Demand Shocks

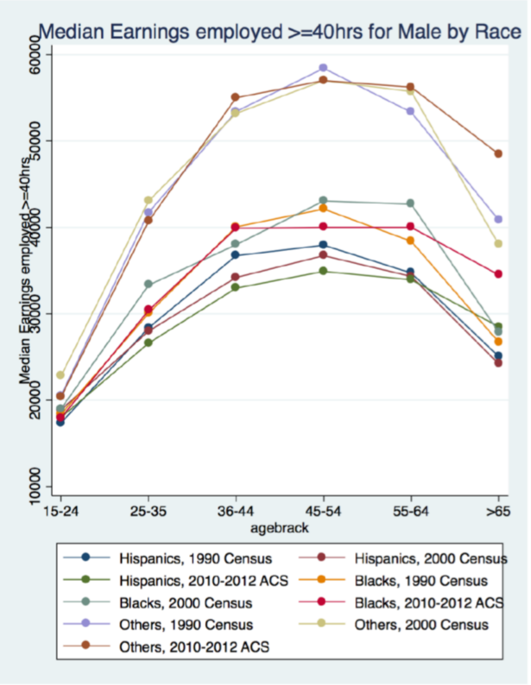
- Advances in the econometrics of estimating two-sided matching models $\Rightarrow$
- Distill underlying propensity for certain transition types to take place above and beyond what could be predicted based on beginning- and end-of-year job composition.
- Use inferences about job-matching technology to construct simulations of how labor market will respond to labor demand shocks of the types suggested above.
- Akin to a big game of musical chairs!
- How does the composition of the labor demand shock (and the composition of the local labor supply) affect the incidence of the shock?

Project #2: Decomposing and Projecting Race- and Gender- Earnings Gaps Over the Life Cycle

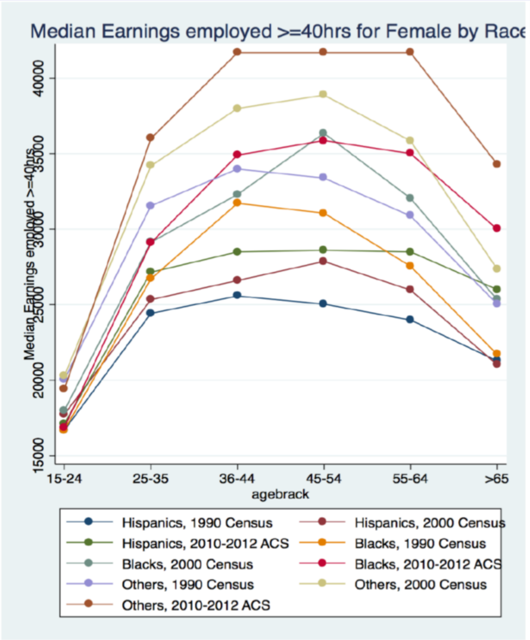
- Earnings Distributions by Race and Gender tend to diverge from labor market entry ($\sim$ age 22) through mid career ($\sim$ age 55)
- Exact patterns differ by race-gender combination and by cohort.
- How much of this divergence can be attributed to:
  1. Differential sorting to early career job types with different career paths (e.g. dead-end jobs!)?
  2. Differential rates of promotion within-firms conditional on job type?
  3. Differential rates of firm-to-firm job transitions conditional on job type?
- Appropriate policy response to close gaps depends critically on the source of the divergence.



Median Earnings employed >=40hrs by Sex
Median Earnings employed >=40hrs
20000
30000
40000
50000
60000
15-24
25-35
36-44
45-54
55-64
>65
agebrack
Males, 1990 Census
Males, 2000 Census
Males, 2010-2012 ACS
Females, 1990 Census
Females, 2000 Census
Females, 2010-2012 ACS



Median Earnings employed >=40hrs for Male by Race
Median Earnings employed >=40hrs
10000
20000
30000
40000
50000
60000
15-24
25-35
36-44
45-54
55-64
>65
agebrack
Hispanics, 1990 Census
Hispanics, 2000 Census
Hispanics, 2010-2012 ACS
Blacks, 1990 Census
Blacks, 2000 Census
Blacks, 2010-2012 ACS
Others, 1990 Census
Others, 2000 Census
Others, 2010-2012 ACS



Median Earnings employed >=40hrs for Female by Race
Median Earnings employed >=40hrs
15000
20000
25000
30000
35000
40000
15-24
25-35
36-44
45-54
55-64
>65
agebrack
Hispanics, 1990 Census
Hispanics, 2000 Census
Hispanics, 2010-2012 ACS
Blacks, 1990 Census
Blacks, 2000 Census
Blacks, 2010-2012 ACS
Others, 1990 Census
Others, 2000 Census
Others, 2010-2012 ACS

Project #2: Decomposing and Projecting Race- and Gender- Earnings Gaps Over the Life Cycle

- Same two-sided matching framework as first project.
- Again, construct transition matrix with rows and columns being job match types at year t and $t + 1$
- This time, origin job types (existing positions) defined by (age, race, gender, earnings category, firm size, occupation, education)
- Destination job types (positions to be filled) defined by (earnings category, firm size, occupation)
- From transition matrix, can distill relative propensity for each position type to hire or promote hispanic women with a college education relative to non-hispanic white men with a high school education.

Project #2: Decomposing and Projecting Race- and Gender- Earnings Gaps Over the Life Cycle

- Given estimated willingness of destination job types to substitute across different worker types . . .
- Can estimate how earnings gaps at each age would converge if:
  1. education distribution were equalized across race-gender combos for incoming cohort
  2. initial sorting to job types were equalized across race-gender combos
  3. propensity to promote workers within firms were equalized across r-g combos
  4. propensity to hire workers of each type from other firms were equalized across r-g combos.

Project #2: Decomposing and Projecting Race- and Gender- Earnings Gaps Over the Life Cycle

- From surveys (Dec. Census, CPS, ACS), can project distribution of race-gender for next 20 cohorts to enter labor market.
- Also can project the flow of retirees over next 20 years.
- And BLS projects changes in labor demand by occupation.
- $\Rightarrow$ Use model to predict how earnings distributions of age-race-gender-educ. combos will evolve over next 20 years!
- Which groups will experience substantial earnings growth?
- Which should we be particularly worried about?
- How would the projections change if we made a major educational investment to keep young men from dropping out?
- How would projected earnings change if automation dramatically shrank the number of jobs in the transportation industry . . . or the retail industry?

Other Potential Avenues of Research Featuring the LEHD

- Which jobs are dead-end jobs?
- What happens to firms when workers retire? What happens to the labor market for young and mid-career workers when many workers retire (or don't!)?
- Recent decreases in labor market and firm volatility: are the two related? If so, which causes which?
- Which occupations and industries are the most deleterious to long-run health?

Disclosure Avoidance Review

- Design study to guarantee minimal risk that worker or firm confidentiality will be breached.
- $\Rightarrow$ hard to analyze particular small geographic area, large firm competition.
- Have a plan for what you would like to eventually disclose at the outset of the project.
- Construct primary sample in SAS or STATA, since Census-provided DAR code is only in these languages.
- Minimize the number of samples you use, and the number of times you request disclosures.
- Don't hardcode any numbers into programs, since these must also pass through DAR.