The Division of Labor Across Research Institutions: Evidence from Evolutionary Biology (1980-2005)

Ajay Agrawal, University of Toronto and NBER
John McHale, National University of Ireland, Galway
Alexander Oettl, Georgia Institute of Technology

The Changing Frontier: Rethinking Science and Innovation Policy Pre-Conference
NBER, October 26, 2012

What is Evolutionary Biology?

- Evolutionary biology is a sub-field of biology concerned with the study of the evolutionary processes that have given rise to the diversity of life on Earth

- Evolutionary biologists study the descent of species and the origin of new species

Why Evolutionary Biology?

1. Theory and applied (lab)
   - Like science, not mathematics

2. Mature subfield
   - Unlike biotechnology

3. Location not pre-determined
   - Unlike particle physics (CERN), astronomy

Our Measure of Output in Evolutionary Biology

- Journals published by the major Evolutionary Biology societies:
  - Evolution, Systematic Biology, Molecular Biology and Evolution, and Journal of Evolutionary Biology
- First, we extract all articles published in the major society journals (set **A)**
- Second, we collect all papers in any journal that have been cited at least once by an article in set **A**. This set is the corpus of influence on the field of Evolutionary Biology (set **B)**
- Third, we weight each paper in set **B** by the number of forward citations it received from the set of four society journals
- Fourth, we assign each citation-weighted article in set **B** to the institutions affiliated with the article. If an article was written by authors at three distinct institutions, then the counter for each institution is incremented

Stylized Facts

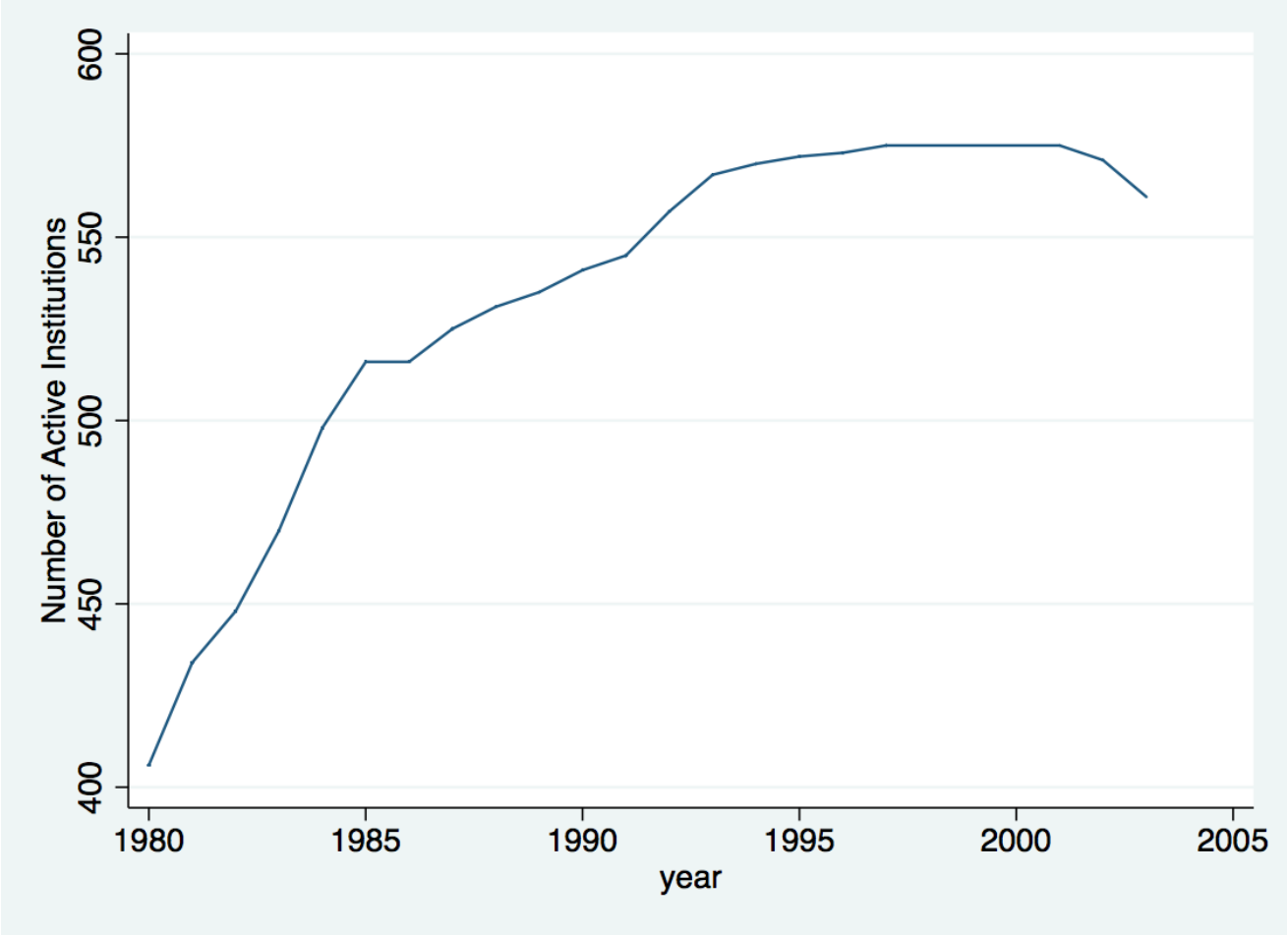
1. Increasing participation in the production of knowledge, but persistent skew in the distribution of output

2. Collaboration increasing: by volume, across rank, across geography

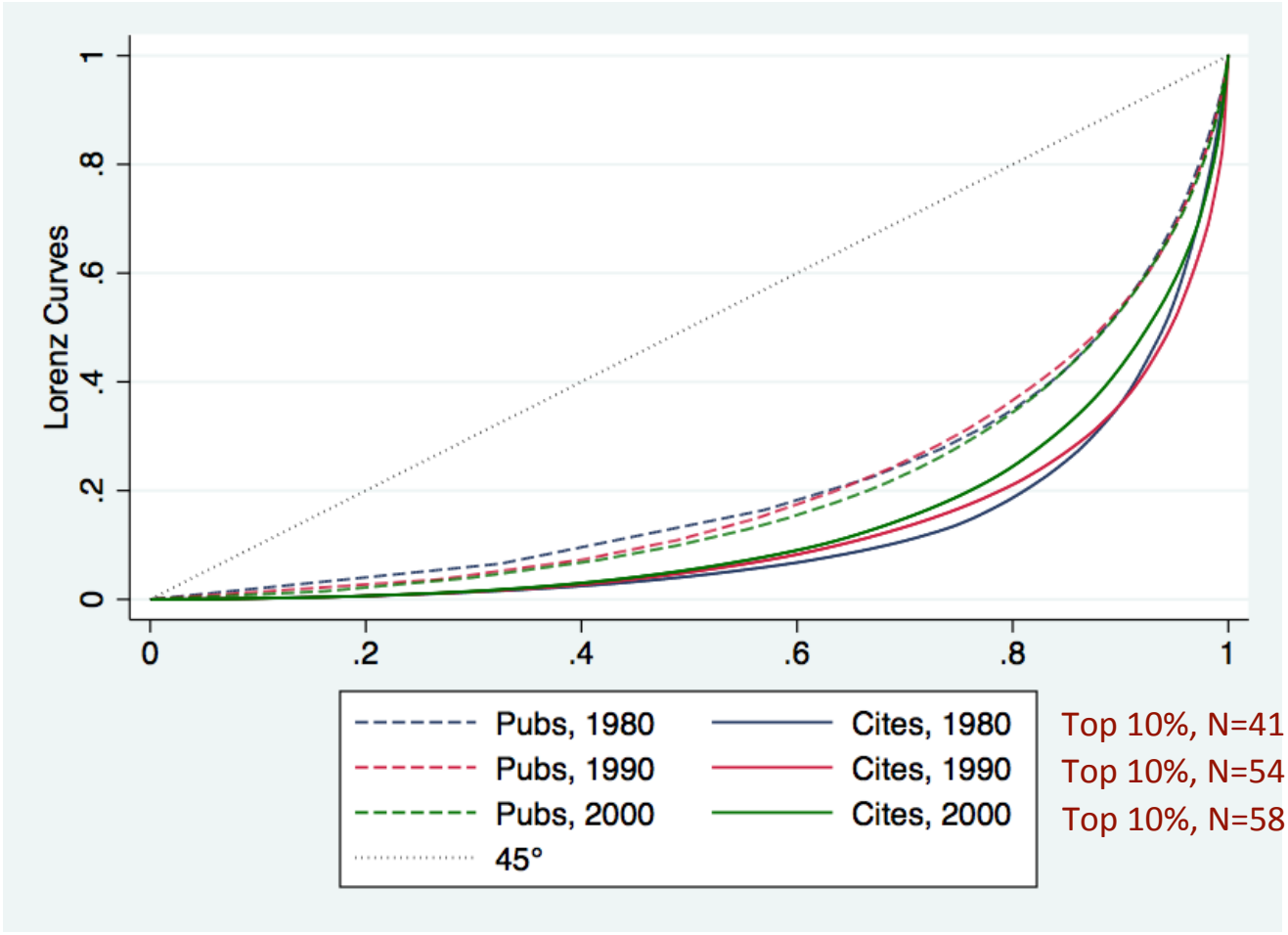
3. Leadership in knowledge production is persistent across time, although with notable exceptions
   - Theory of stars and the agglomeration of talent

INCREASING PARTICIPATION, BUT PERSISTENT SKEW

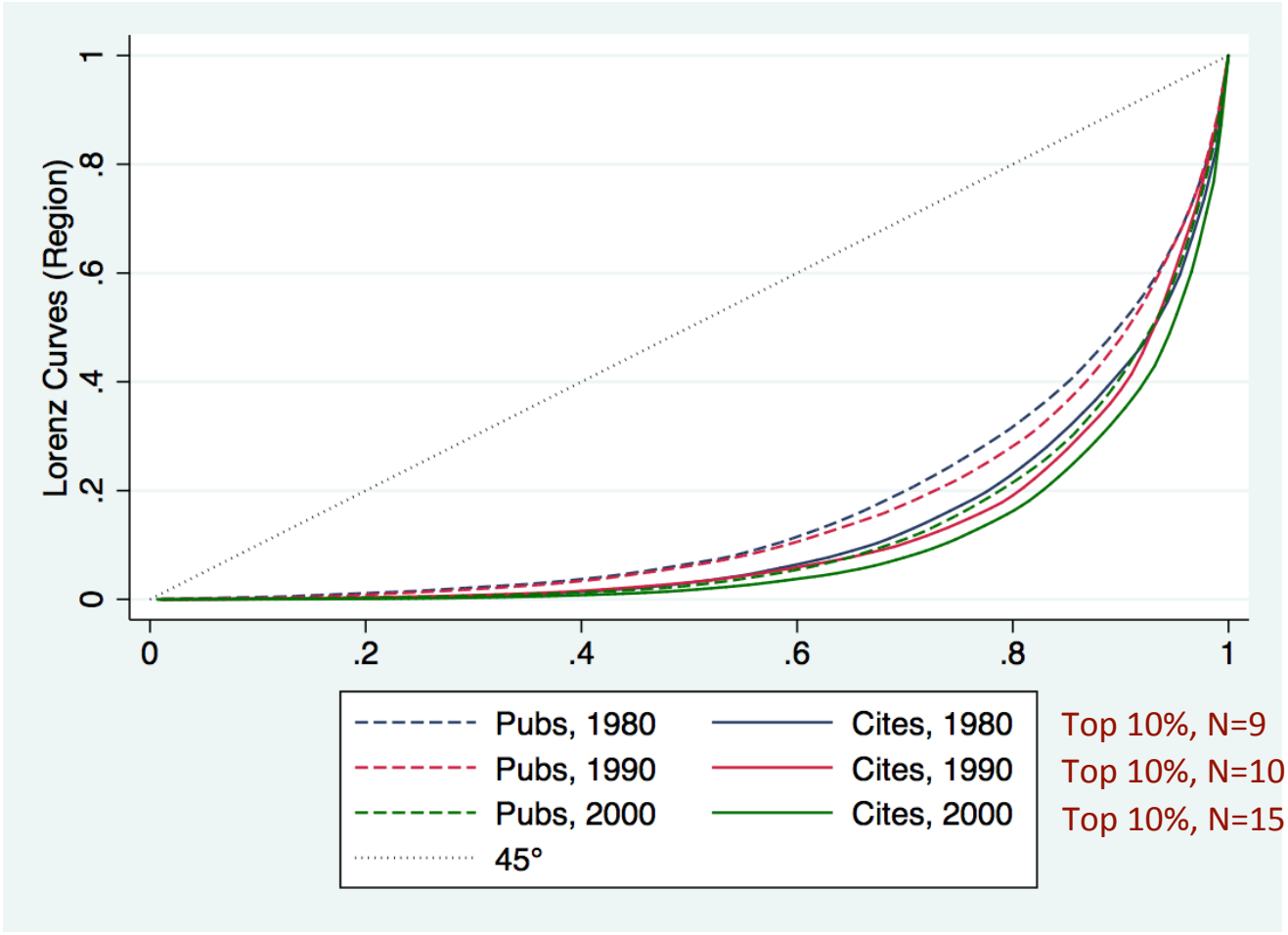
Institution Entry



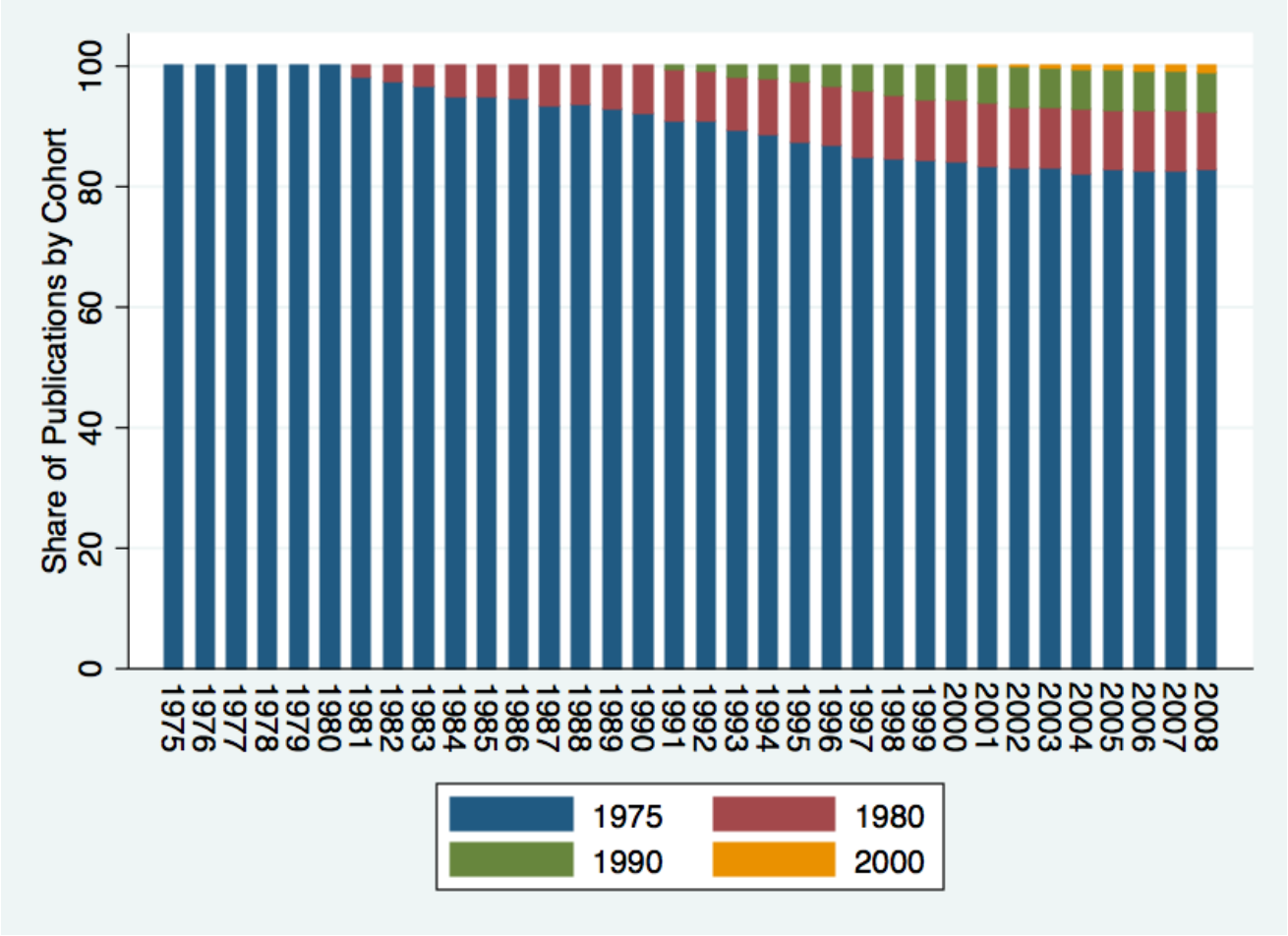
Lorenz Curves by Institution



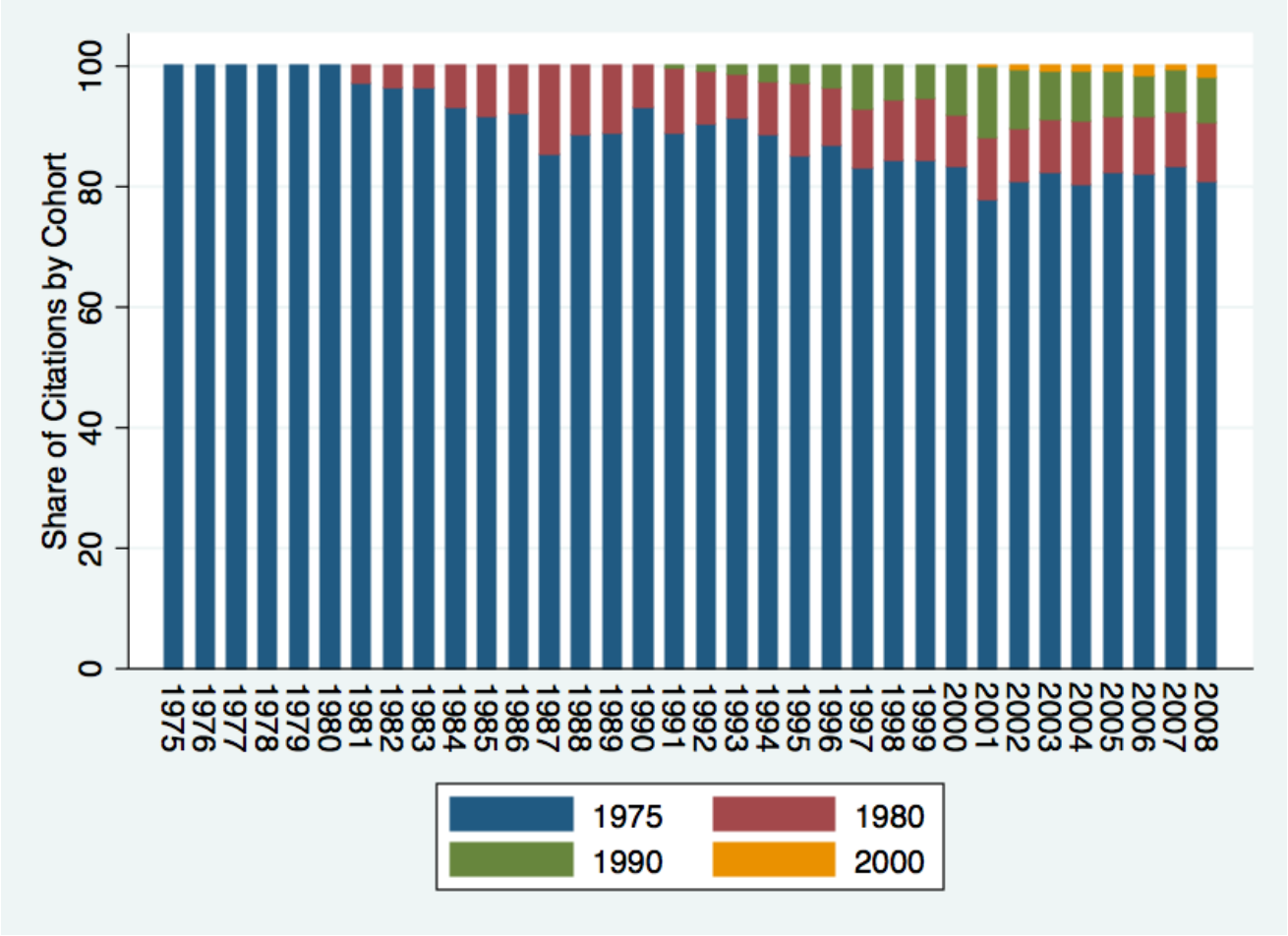
Lorenz Curves by Region



Cohort Publishing Contribution

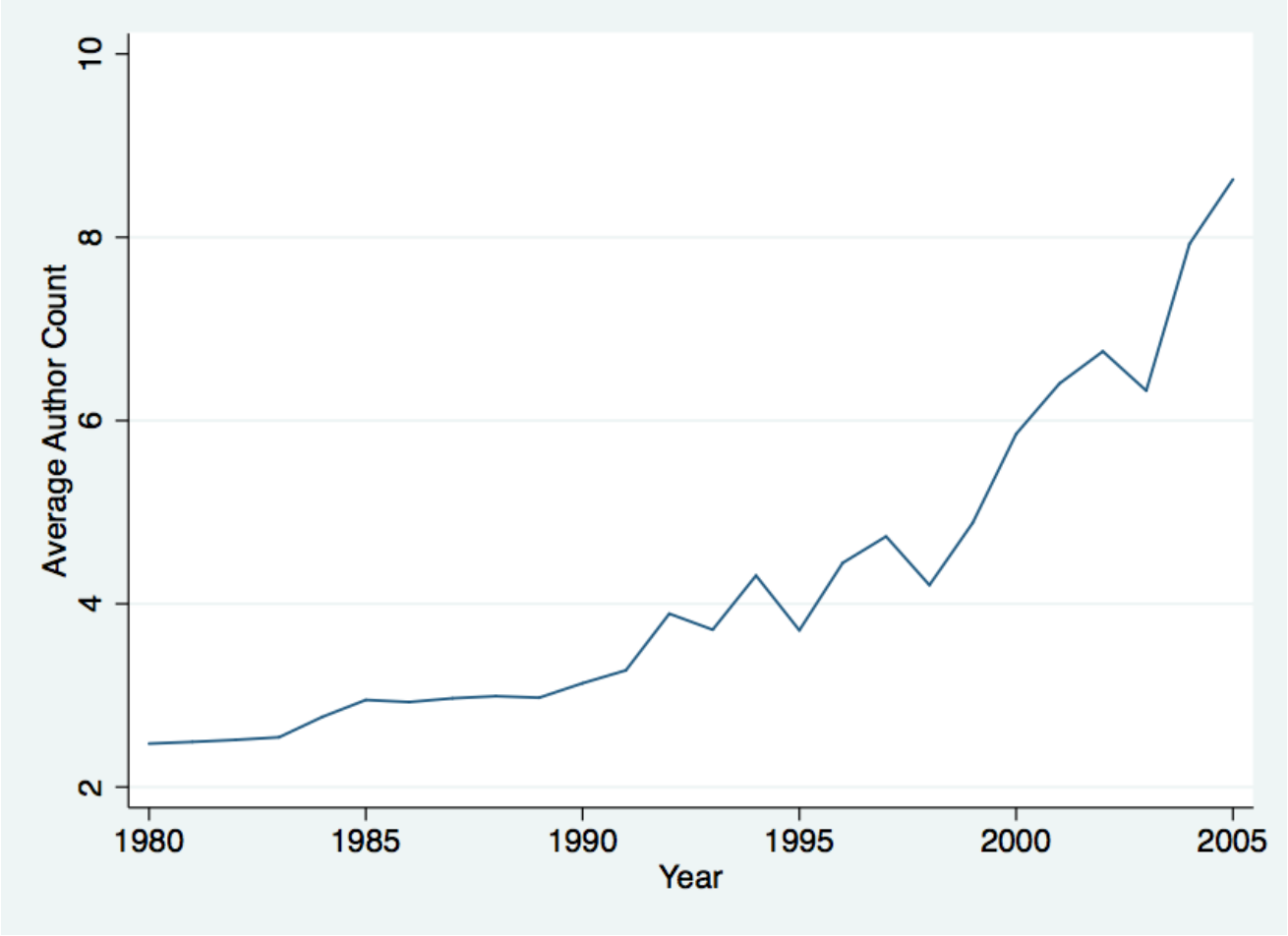


Cohort Citation Contribution

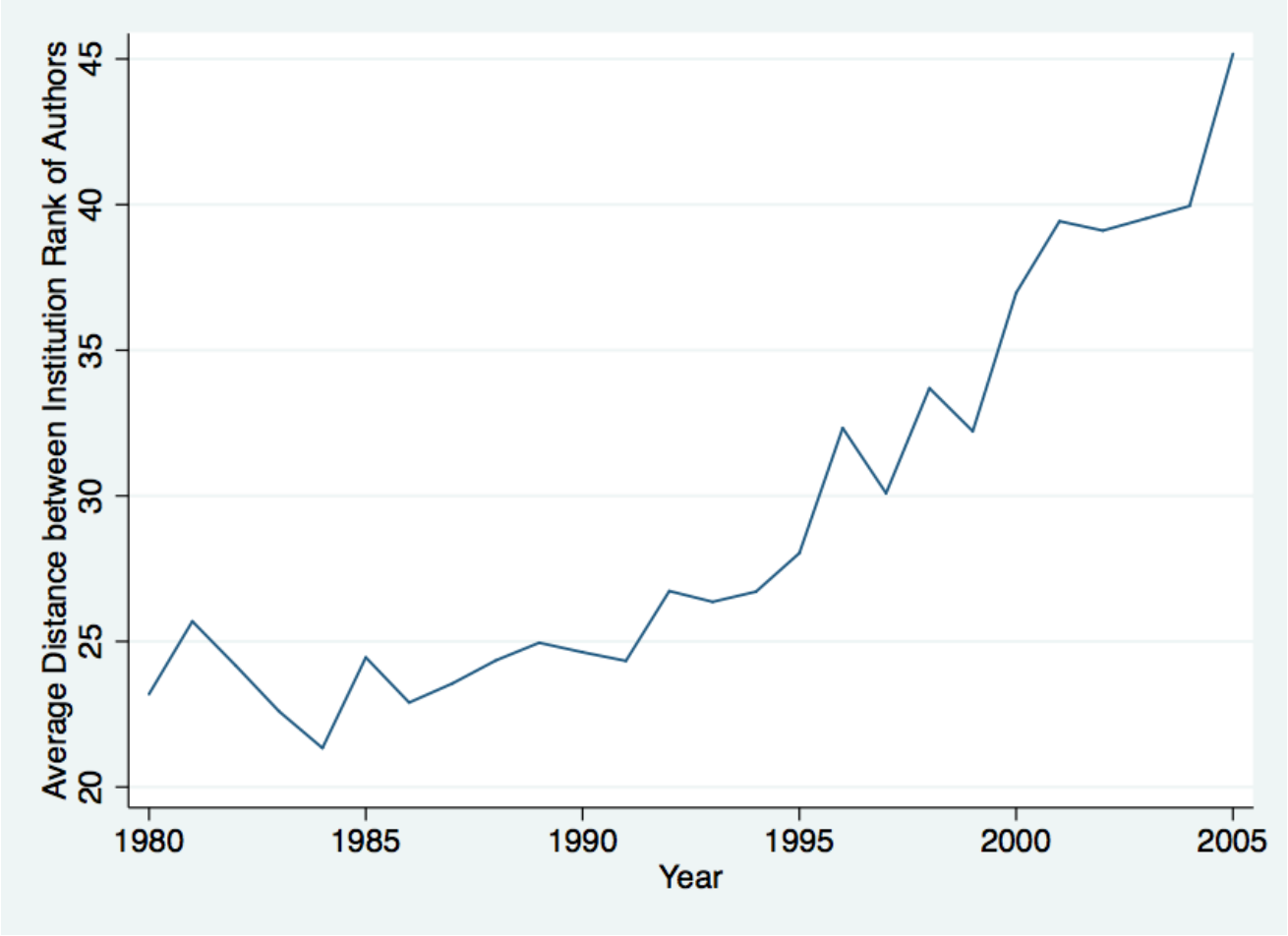


COLLABORATION

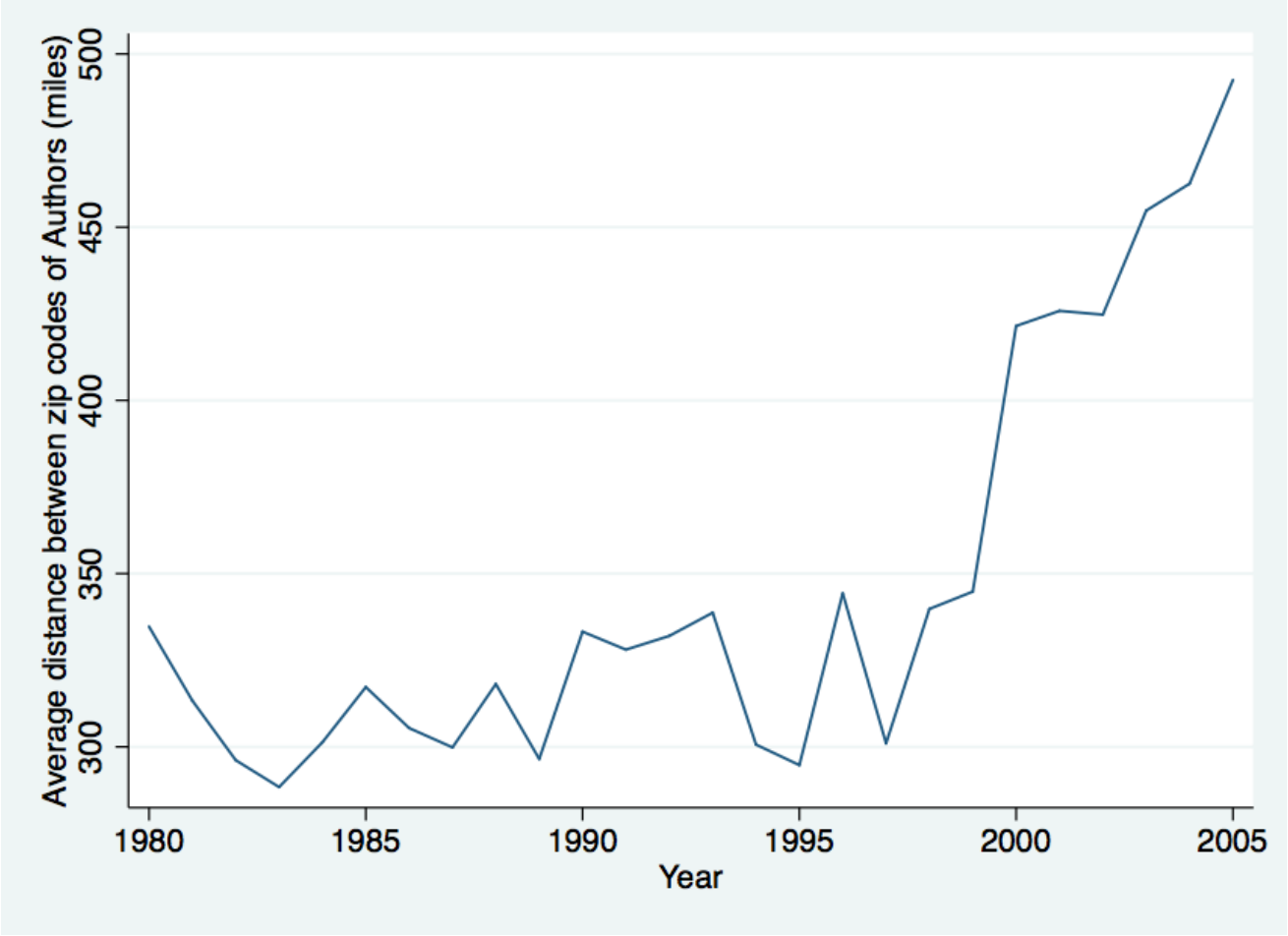
Collaboration:
Average Number of Authors



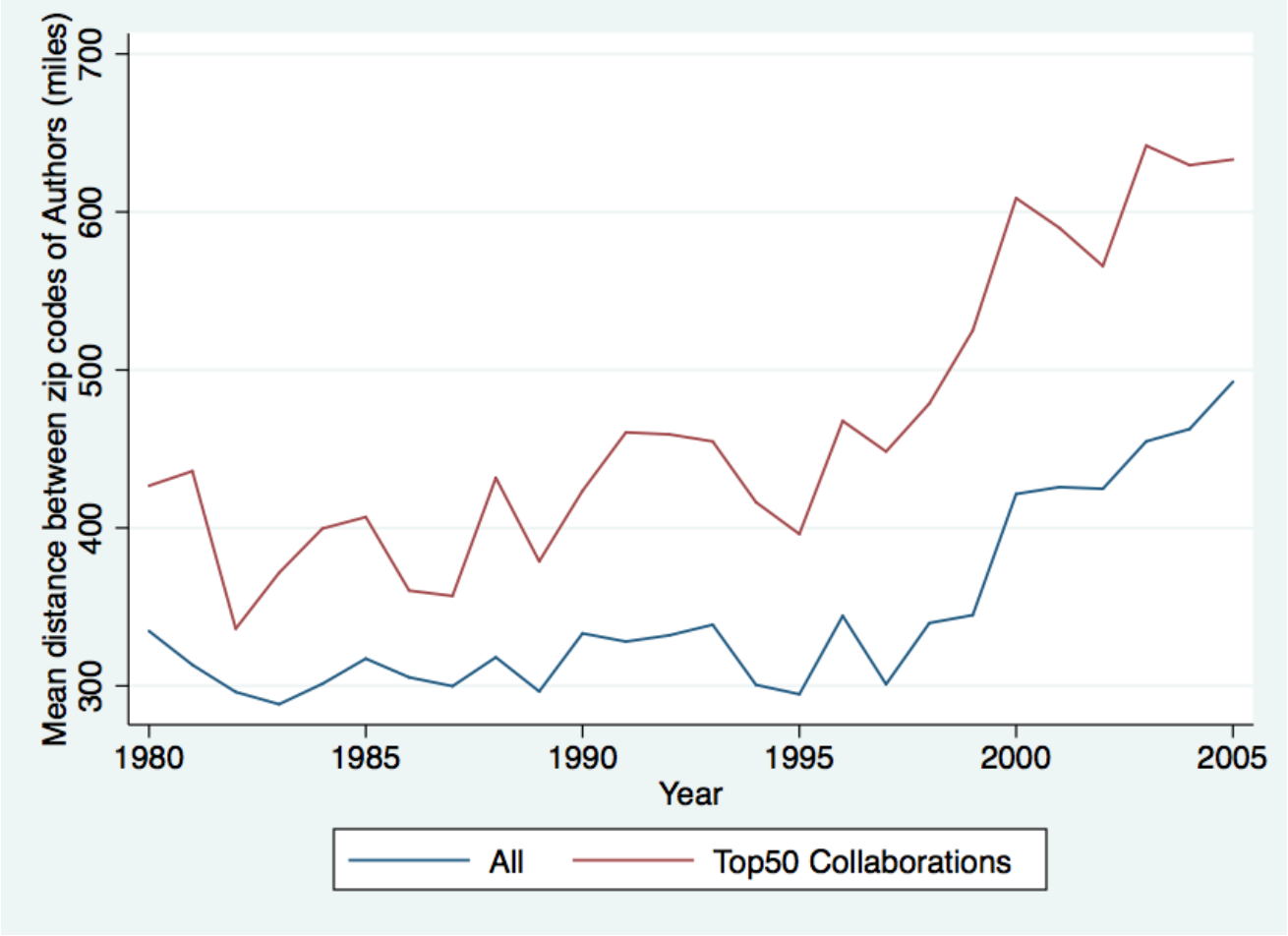
Collaboration:
Average Distance Between Rank



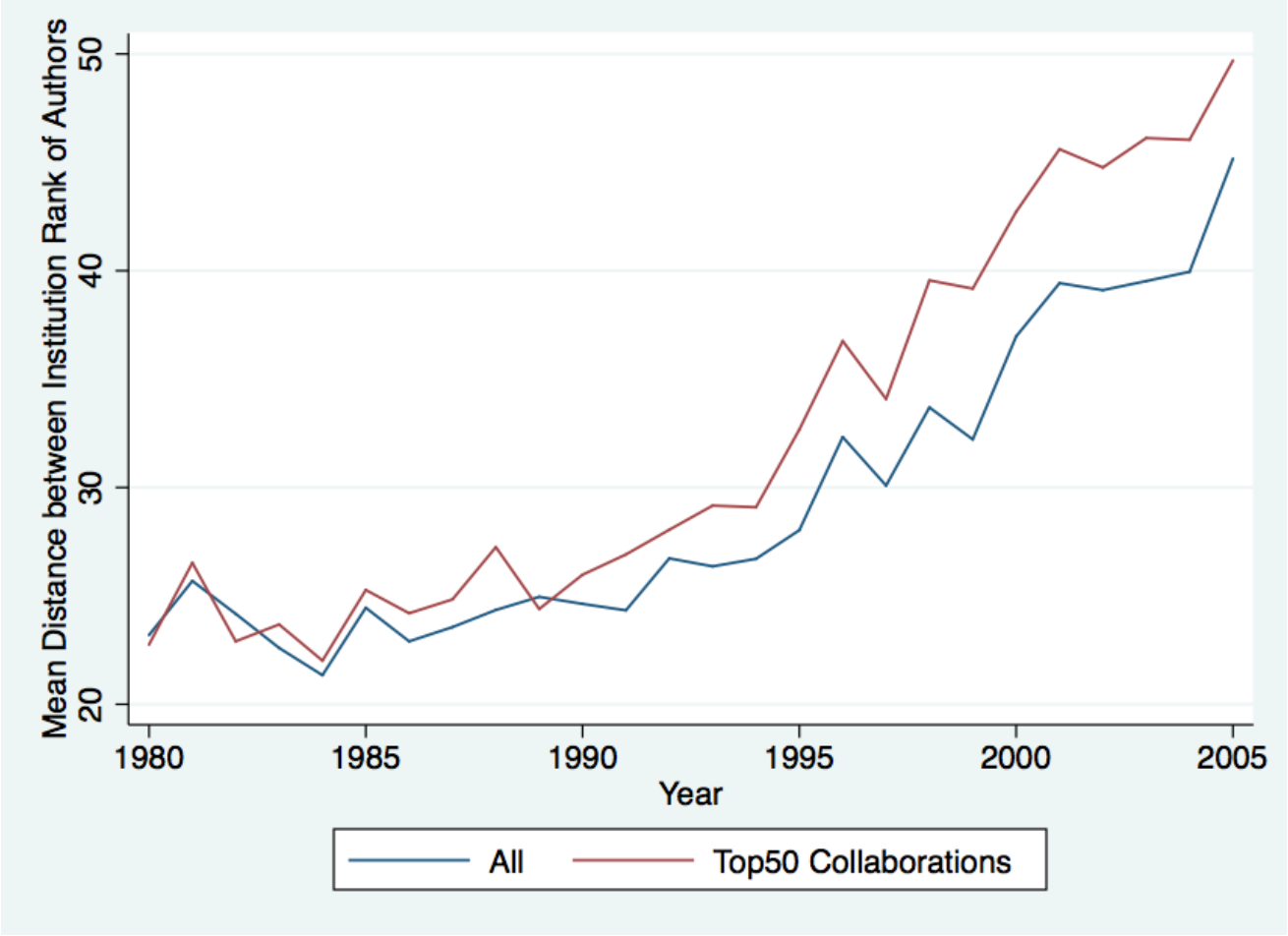
Collaboration:
Average Distance Between Coauthors



Collaboration with Top50 Institutions: Geographic Distance

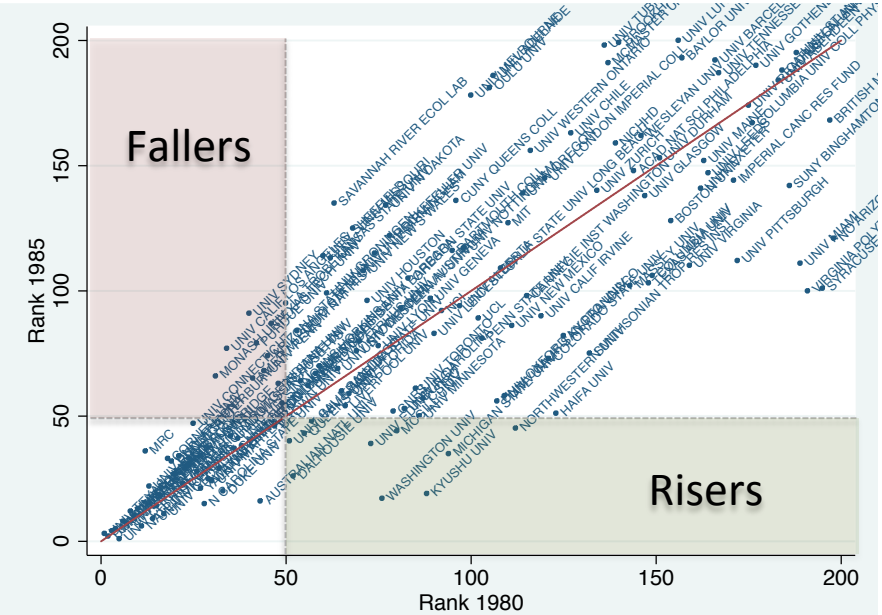


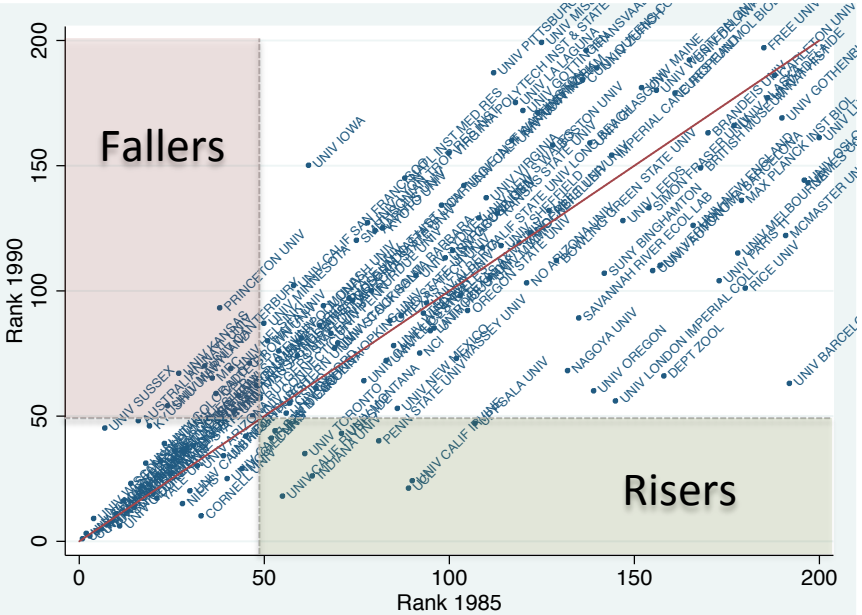
Collaboration with Top50 Institutions: Rank Distance

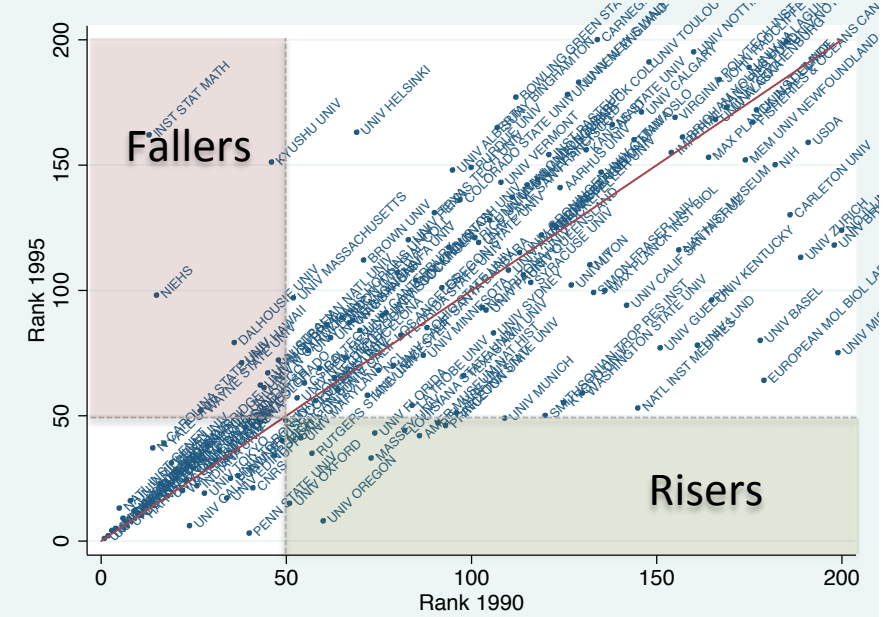


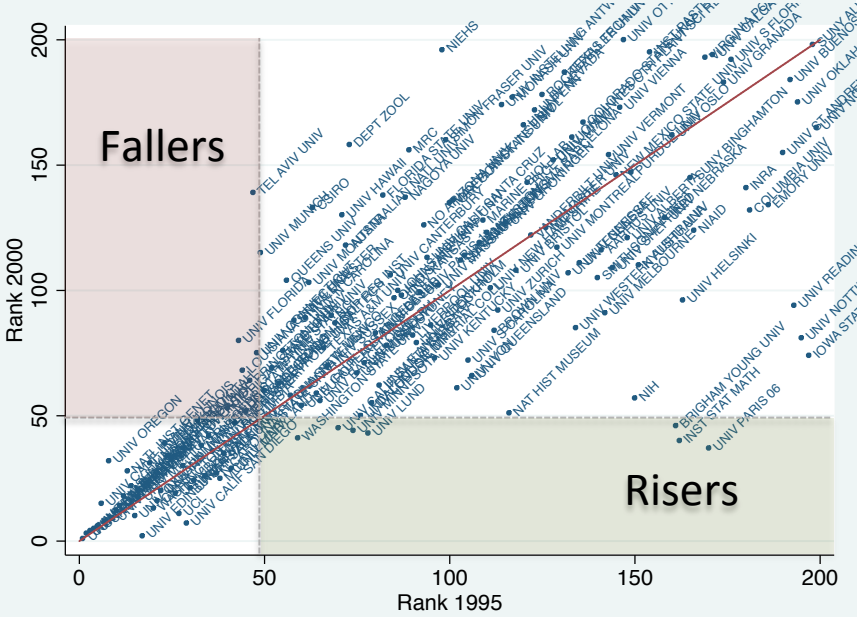
PERSISTENCE AT THE TOP, BUT EXCEPTIONS

Institution Rank (5Yr Increments)



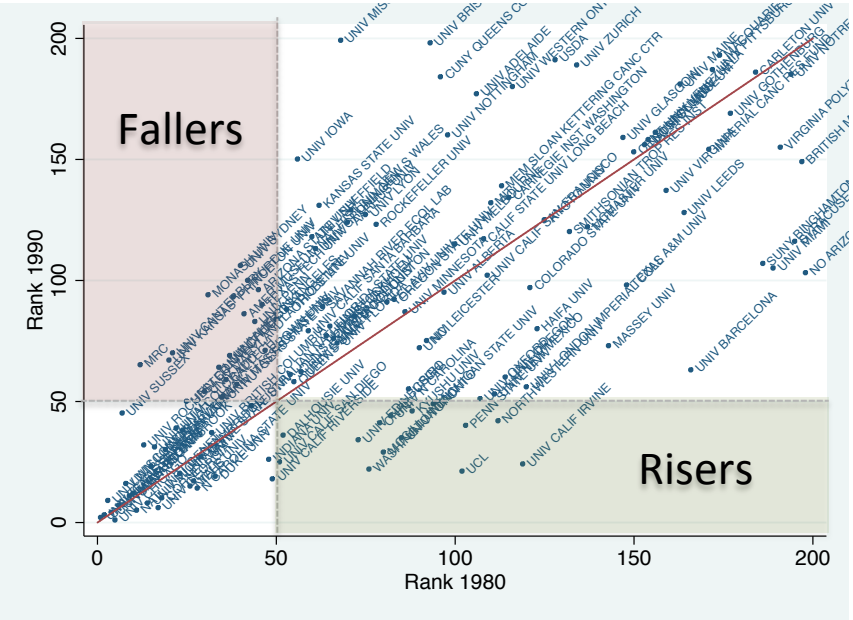




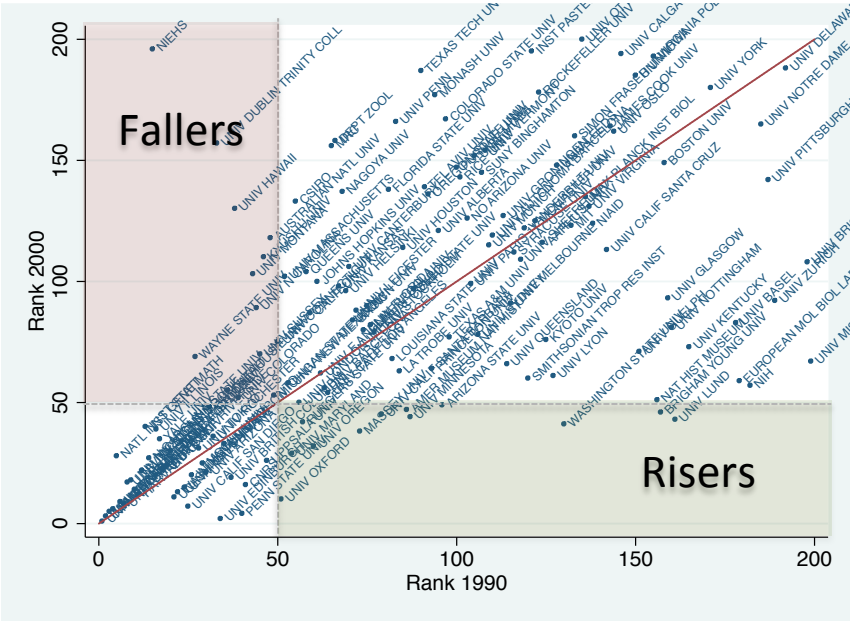


Institution Rank Change (10-Yr Increments)

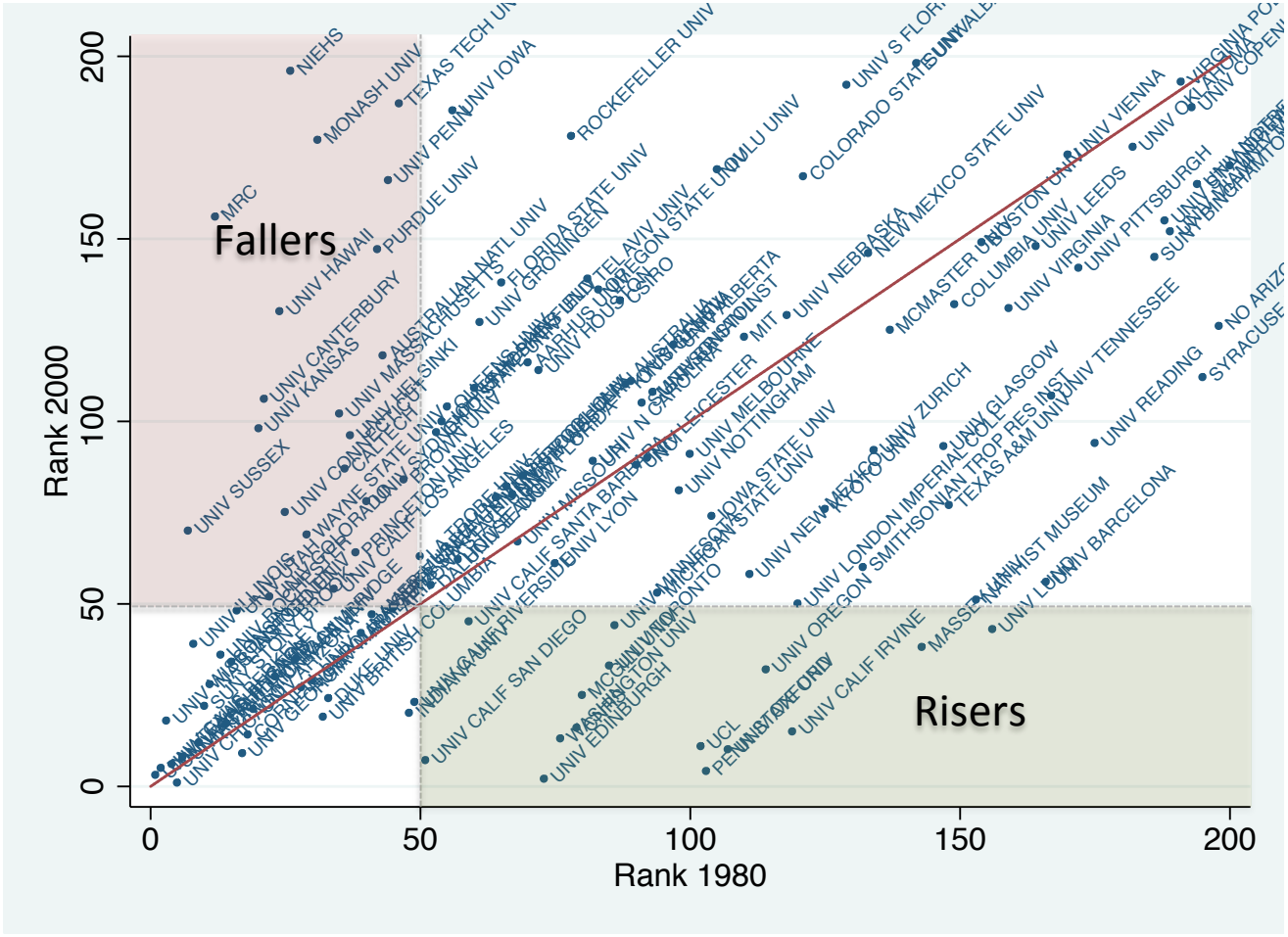
1980-1990



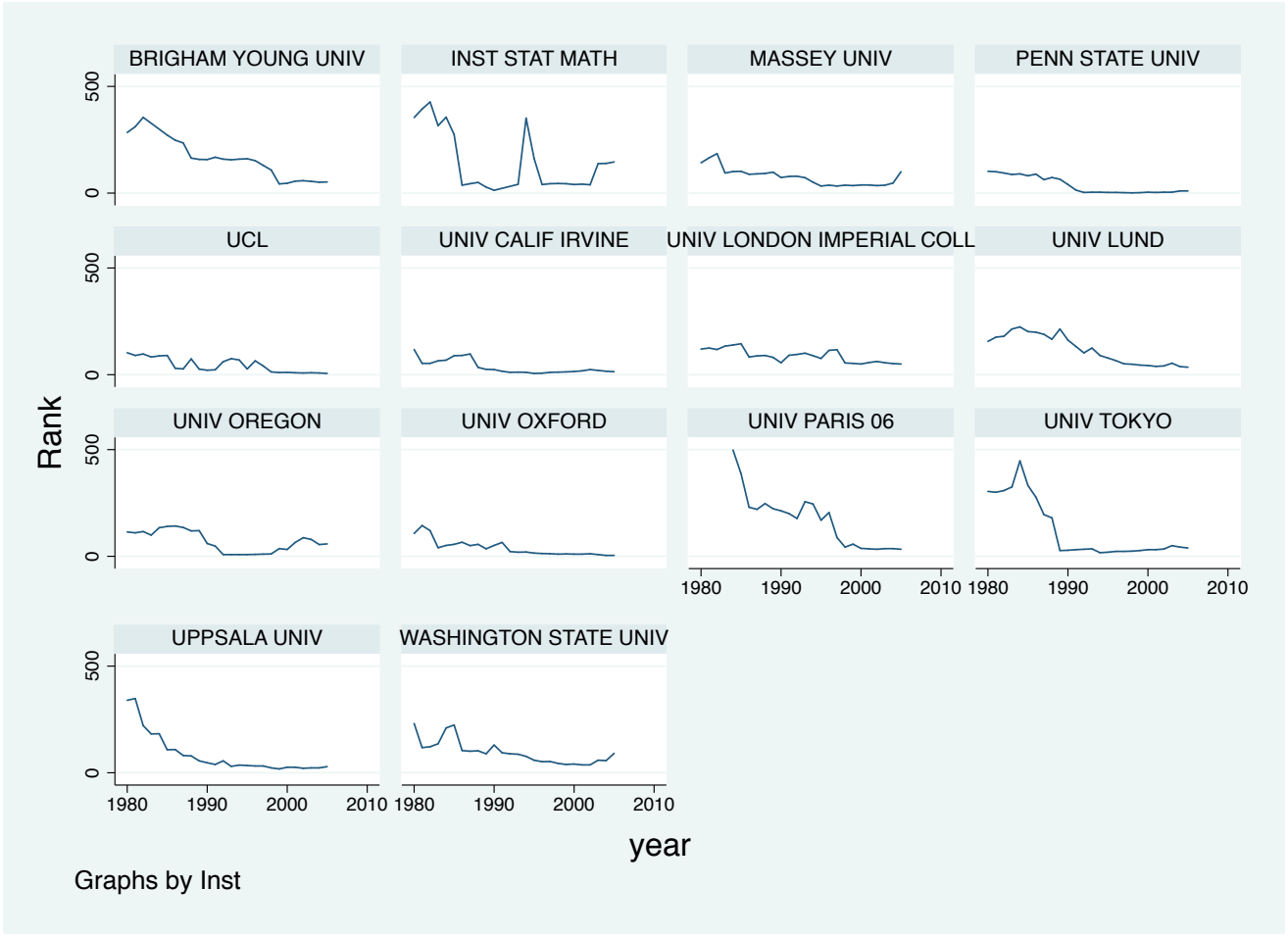
1990-2000



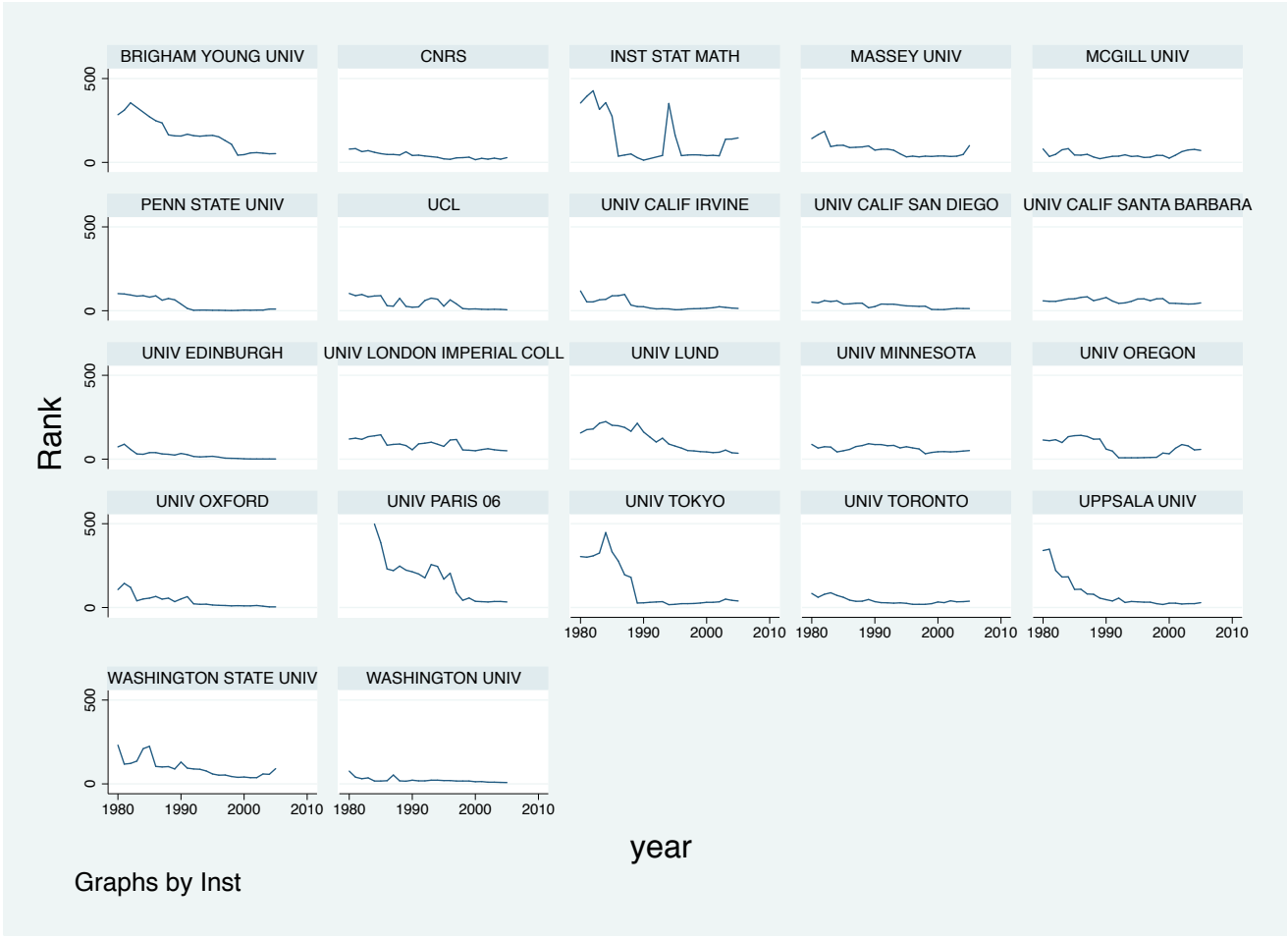
Institutions Rank Change: 1980-2000



Rank Change for Institutions not in Top 100 in 1980 but in Top 50 by 2000



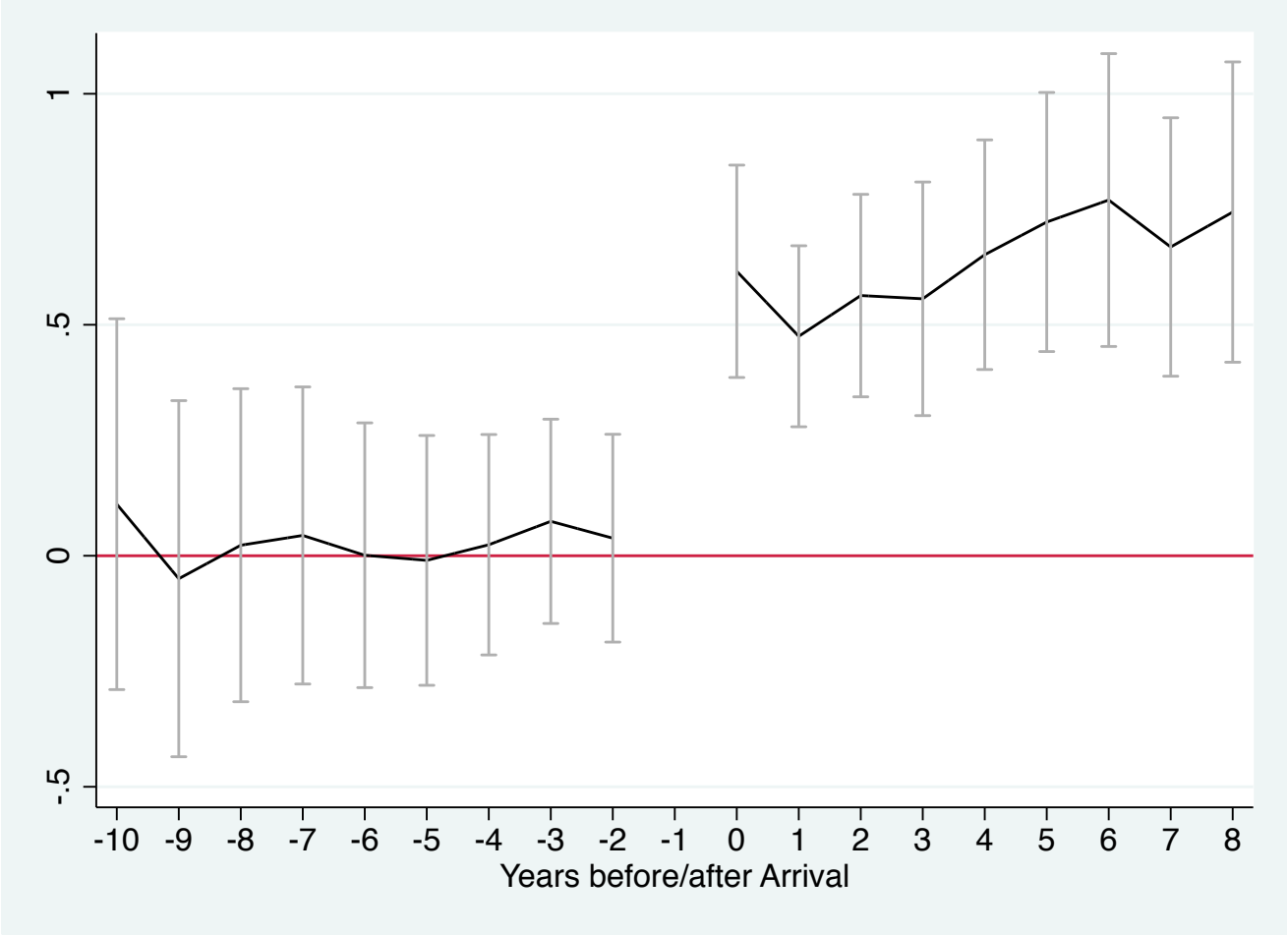
Rank Change for Institutions not in Top 50 in 1980 but in Top 50 by 2000



Stars and Institution-Level Productivity

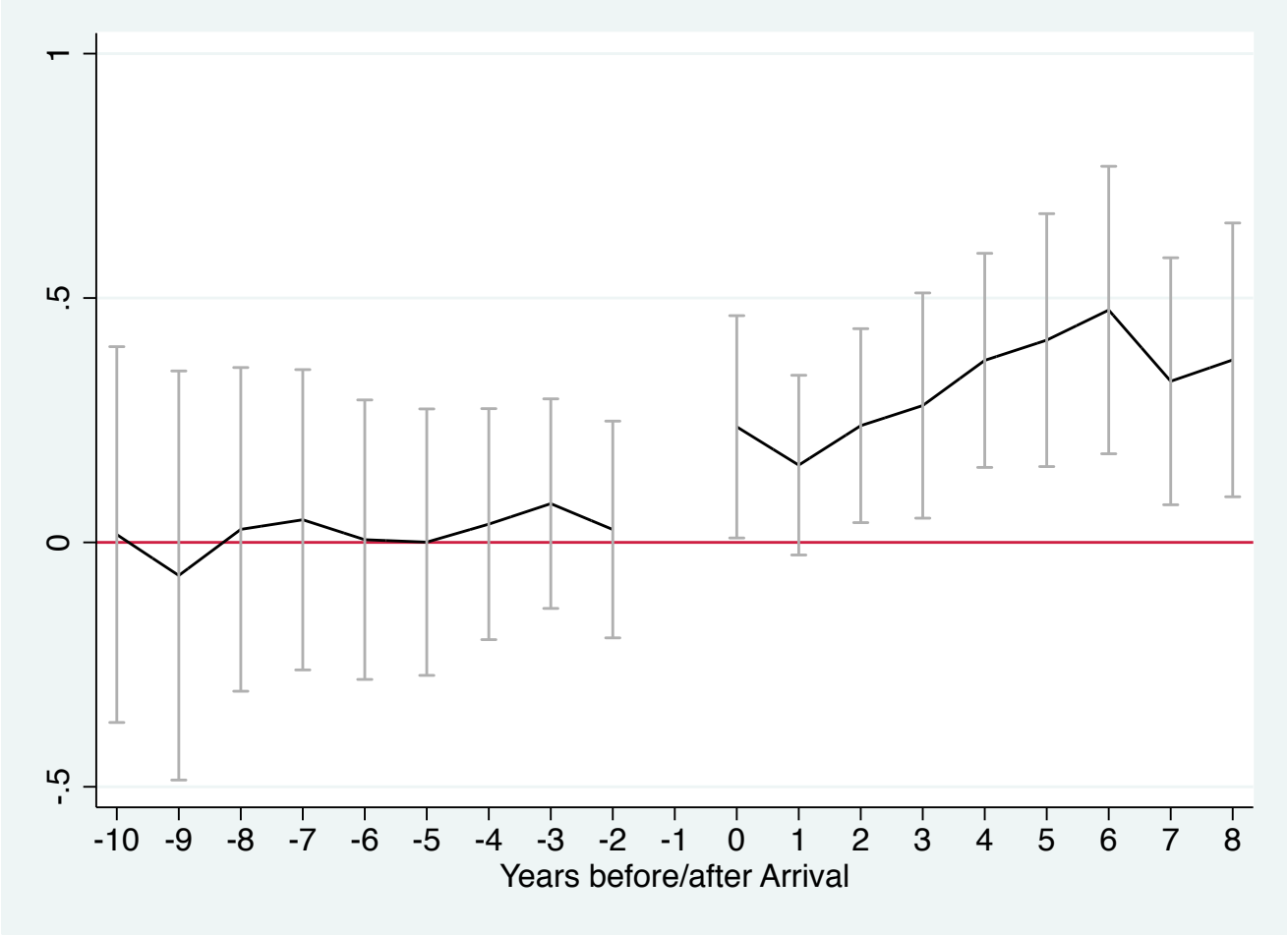
- Estimating Equation
- $Y_{it} = \exp(StarArrival_{t-1} + Scientists_{it} + d_t + r_i)$
- Where
  - Y: citation-weighted publication count for institution *i* in year *t*
  - StarArrival is set to one the year institution i receives a star for the first time and stays on thereafter
  - Scientists is the number of scientists at institution i in year t.
  - d is a series of year fixed effects
  - r is a series of institution fixed effects

Institution-Level Productivity Before and After Star Arrival

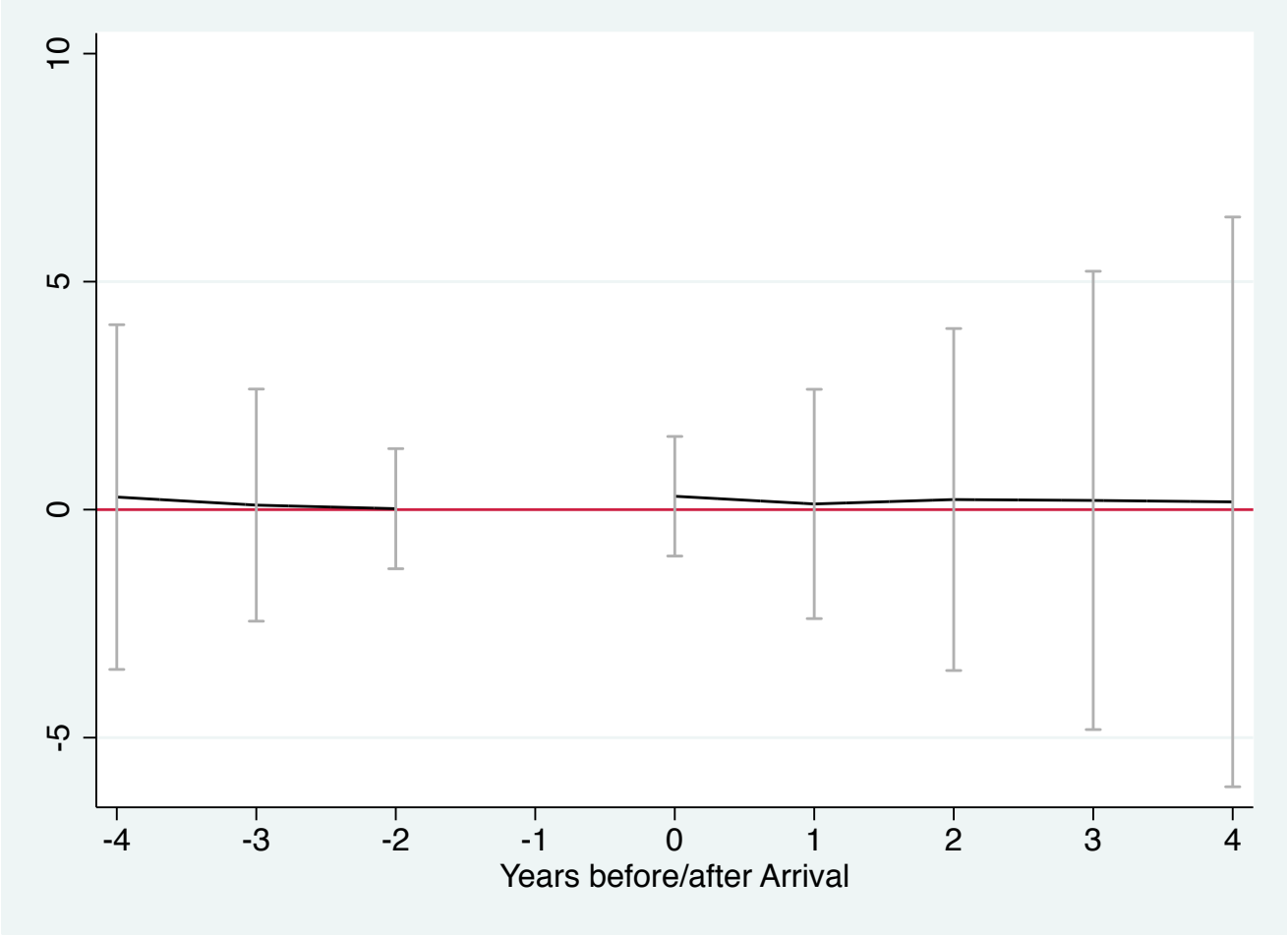


Baseline is year before star arrival

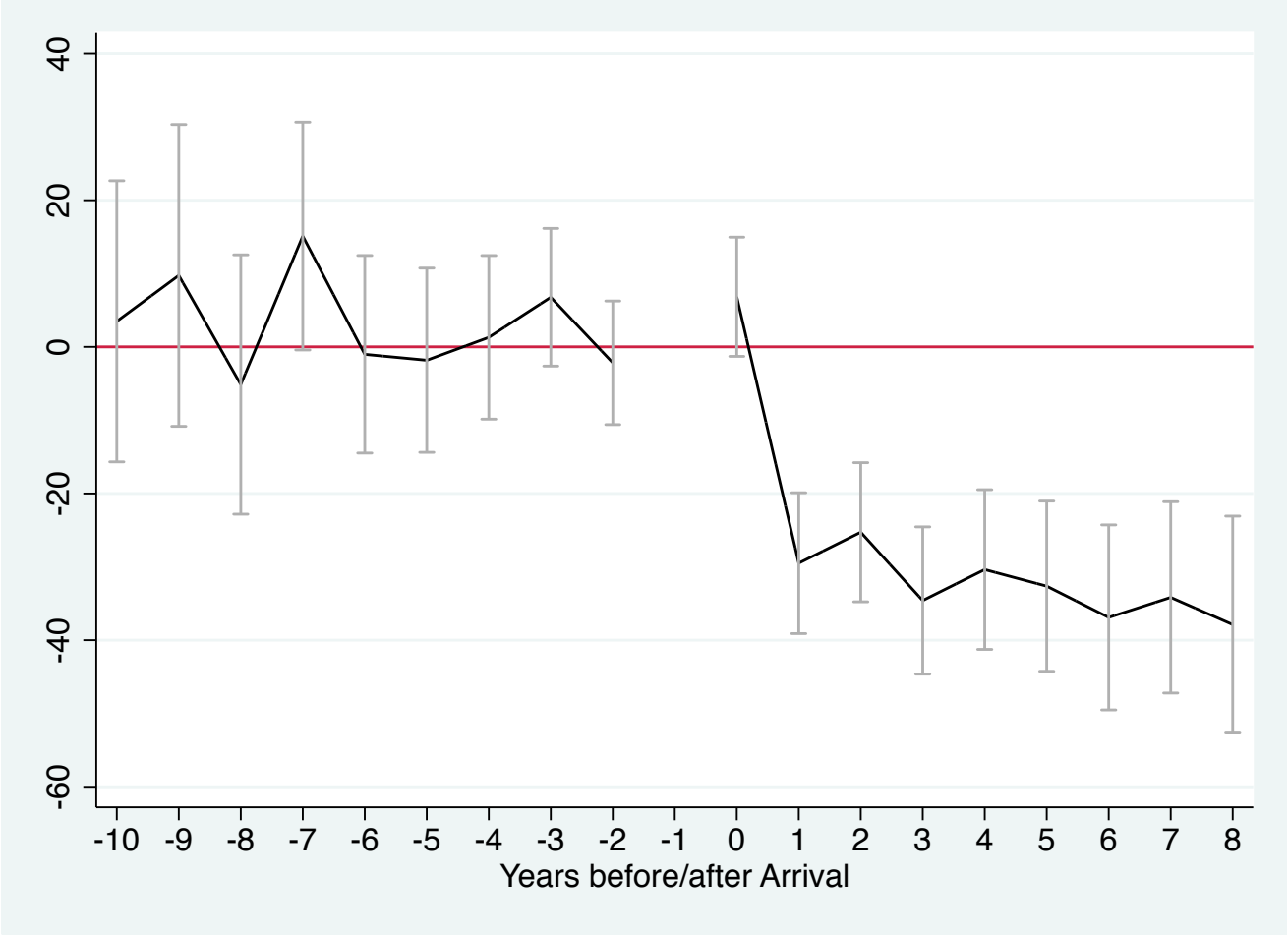
Institution-Level Productivity with Star Data Removed



No Effect on Incumbents



Source Institution Rank for New Hires Before and After Star Arrival



THANK YOU