Public Funding and Its Relationship to Research Outcomes

Paula Stephan
Georgia State University & NBER
UNU-MERIT/MGSoG Conference
November 2014

Research at Universities

- Often funded by government
- Rationale is public goods nature of research
- Funding relationship places governments, public agencies, and universities under pressure to show relationship between research inputs and outputs
- Part of this is of their/our own making
- Tendency to “over promise” significant spillovers that will affect the economy
- Yet lags can be quite long; “dry holes” encountered along the way and counterfactuals difficult to establish

Problems Such as These

- Lead government agencies and universities to try establish a relationship between inputs and outputs in the short run
- But this can be difficult: Lack both methodology and data to support analysis
- Focus today is to provide an example of type of analysis that can be done with new data that are becoming available from the StarMetrics Project

Acknowledgements

- Joint
  - Jacques Mairesse, Maastricht University
  - Michelle Pezzoni, University of Nice Sophia Antipolis, GREDEG and CRIOS
  - Julia Lane, American Institutes of Research
- Funding: Alfred P. Sloan Foundation
- Part of much larger STAR METRICS effort to use administrative data to understand the relationship between inputs and outputs in research

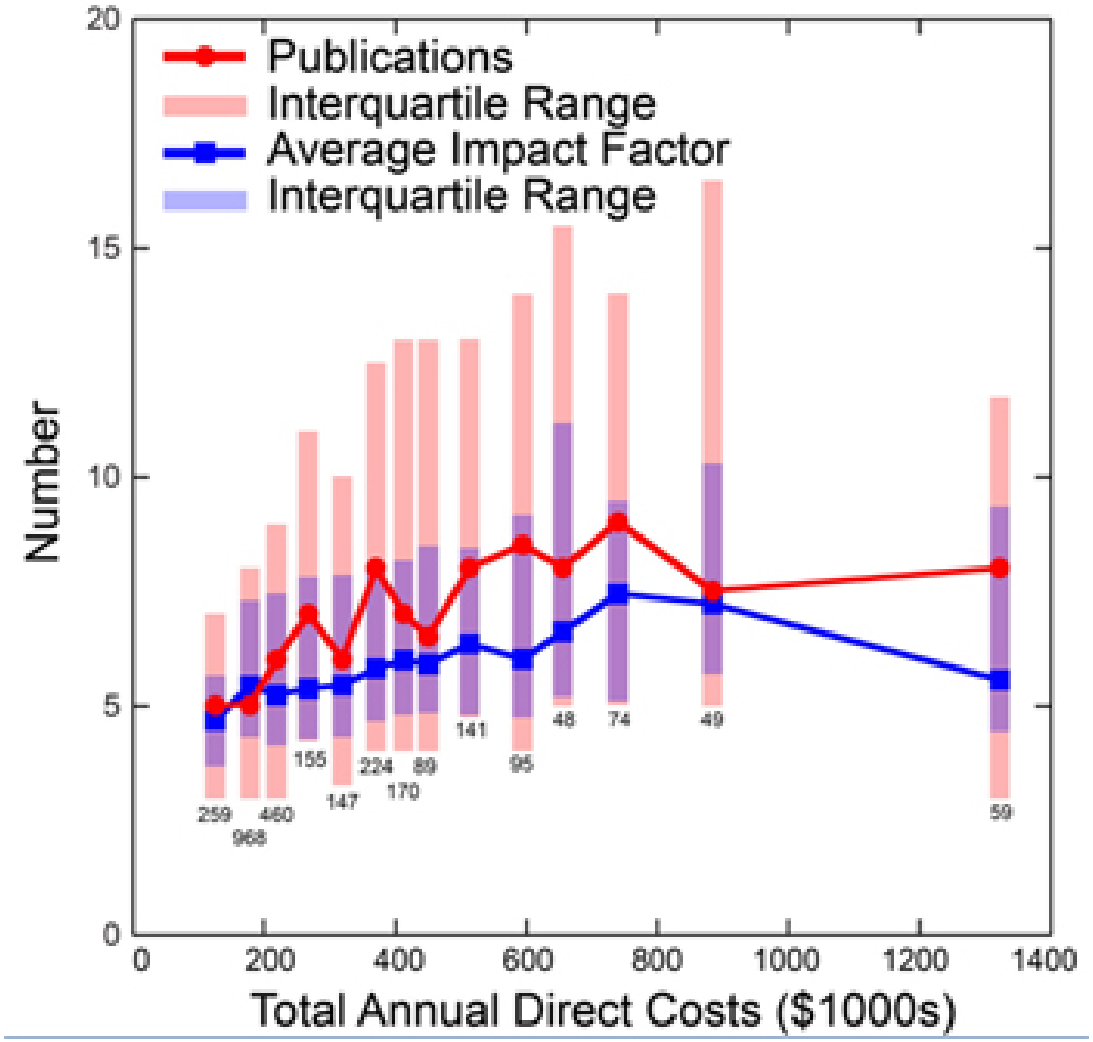
Focus: Relationship Between Research Outputs and Research Inputs

- Question of interest to
  - Policy makers at national level (such as OST) who would like to know degree to which more funding leads to more research
  - Funding agencies such as ERC, NIH, Wellcome Trust, would like to know degree to which research (often measured by publications) can be attributed to funds foundation invests in researchers
- Recently addressed by Jeremy Berg of NIH

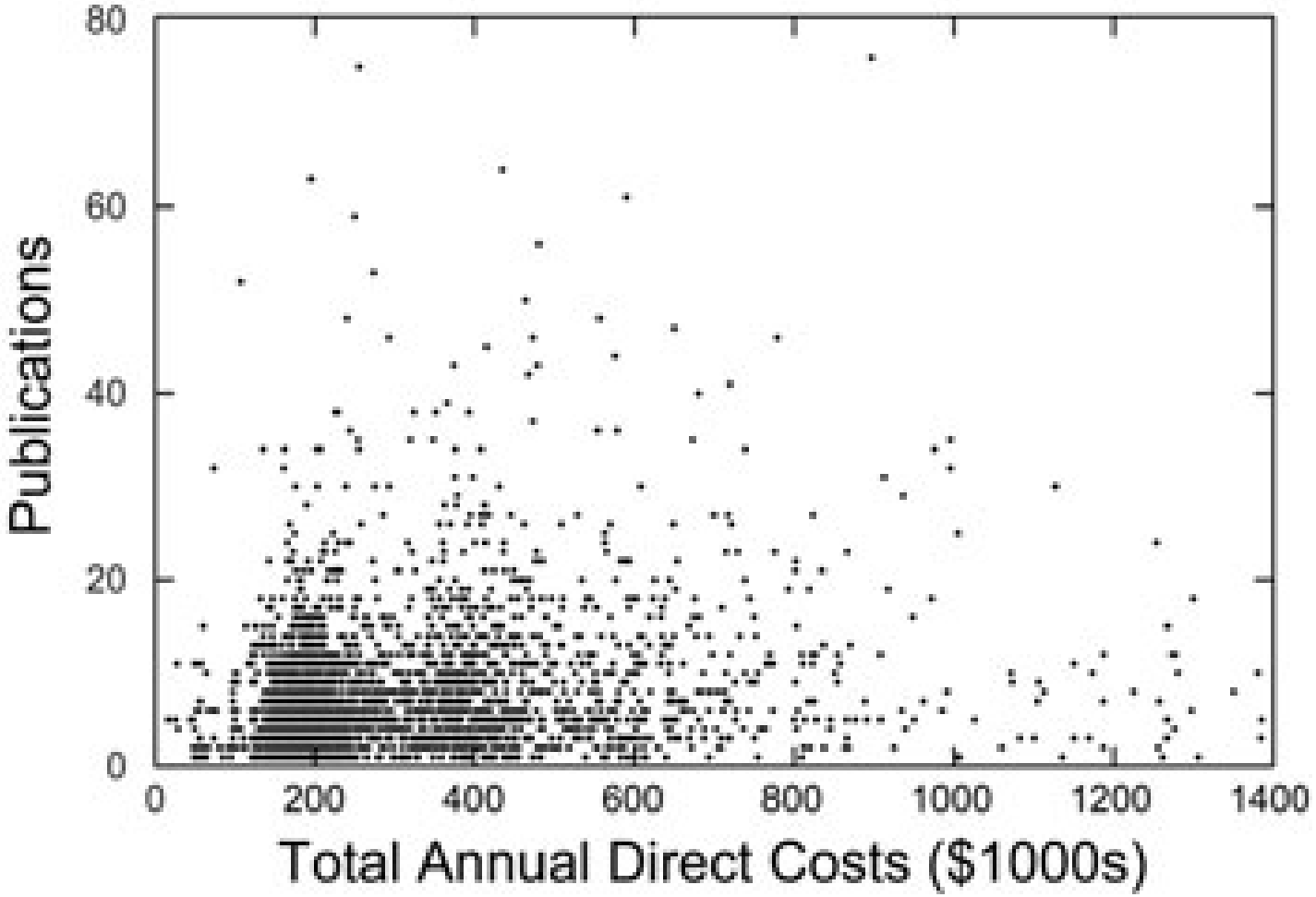
Jeremy Berg

- 2010 postings by Jeremy Berg, Director at the time of NIGMS at NIH, took a step at answering the question for NIH
- Related amount of funding NIGMS investigators received in fiscal 2006 to number of articles published during period 2007-2010

“Buckets of Funding”



Investigator Level



Berg Approach Highlights Two Issues

- Attribution: exactly which articles should be attributed to what funding stream?
- Does way in which funding is bundled matter? What we refer to in our work as the neutrality assumption.

Attribution

- Approach used by Berg attributes all articles published in next three years to all NIH funding received in 2006
  - Some of these articles undoubtedly related to funding received prior to 2006; some to funding after 2006
  - Some related to funding from other agencies
- Both types of attribution issues lead to an overstatement of relationship funding in one year and publication

Neutrality

- Neutrality assumes that manner in which funding is bundled has no effect on productivity of lab
- Two dimensions of the assumption
  - It assumes that it makes no difference whether PI has four grants a year that sum to $1 million or one grant that sums to $1 million.
  - It assumes that relative size of grants does not affect output—makes no difference in terms of output whether PI has four grants that sum to $1 million with one grant being very small or four of approximately equal size

Fixed Costs Can Matter

- Certain grant-related activities, such as the preparation of proposals and submission of progress reports, are required regardless of grant size
- Means that a portion of time required for grant administration relates to number of grants

Relative Size of Grants Can Also Matter

- Amount of time required for grant administration may be associated with size of grant
- To extent that there are economies of scale, time required for grant administration may decrease as funding portfolio becomes increasingly concentrated in a small number of grants

Non-trivial Questions

- Time associated with grant administration is non-trivial
- 2007 report of the Faculty Standing Committee of Federal Demonstration Partnership found faculty working on federally supported research spent 42 percent of their research time on pre-and post-grant award administration activities—not on active research.

Current Research

- Sets out to address attribution and neutrality issues
- In order to do so must have panel data that allow us to model annual research output of PI at grant level and then aggregate up to the PI level
- Data at this level of detail are becoming available through the StarMetrics project

StarMetrics Data

- Goal is to have data for all research universities in US
- Currently have data for 11 universities
- Data come from administrative records
- Initially for Federal projects; eventually data will include non-Federal projects
- University with most complete data is Caltech; results I present today are for Caltech

StarMetrics Data

- Records are by month
- For each grant we can determine at monthly level
  - number of people working on a grant
  - amount of time allocated to project by each individual and amount of money charged off the grant
  - occupational status of each (faculty, postdoc, staff scientist, etc.)
- Also have information on such things as purchase of equipment and supplies associated with project, subcontracts

CALTECH

- Size
  - Approximately 300 faculty
  - 980 undergraduates
  - 1250 graduate students
  - 650 (approximate) postdocs
- Focus
  - Science and engineering
- Top rated

Data Matched to

- Patents (NBER data)
- Publications—quantity, Impact Factor of Journal; citations (not yet collected)
- Placement of students and postdocs (Census)
- Startup of new firms (Census)

Attribution

Several Ways to Attribute Output to Grants

- "Everything Counts" approach of NIGMS
  - We reject as too "gross"
- Acknowledgement in footnotes
  - Problems with overstatement in some cases and not-reported in others
  - Only applies to articles; not to other forms of output
- Topic modeling
  - Likely to result in over attribution—key words can be shared by a number of grants
- Weighting by some factor

Weighting

- Here attribute outcomes to inputs from a specific grant by weighting the outcome by relative share of all funds coming from the grant—both within grants and across grants
- Assume articles begin to appear within a year of starting the research and continue to appear for up to three years

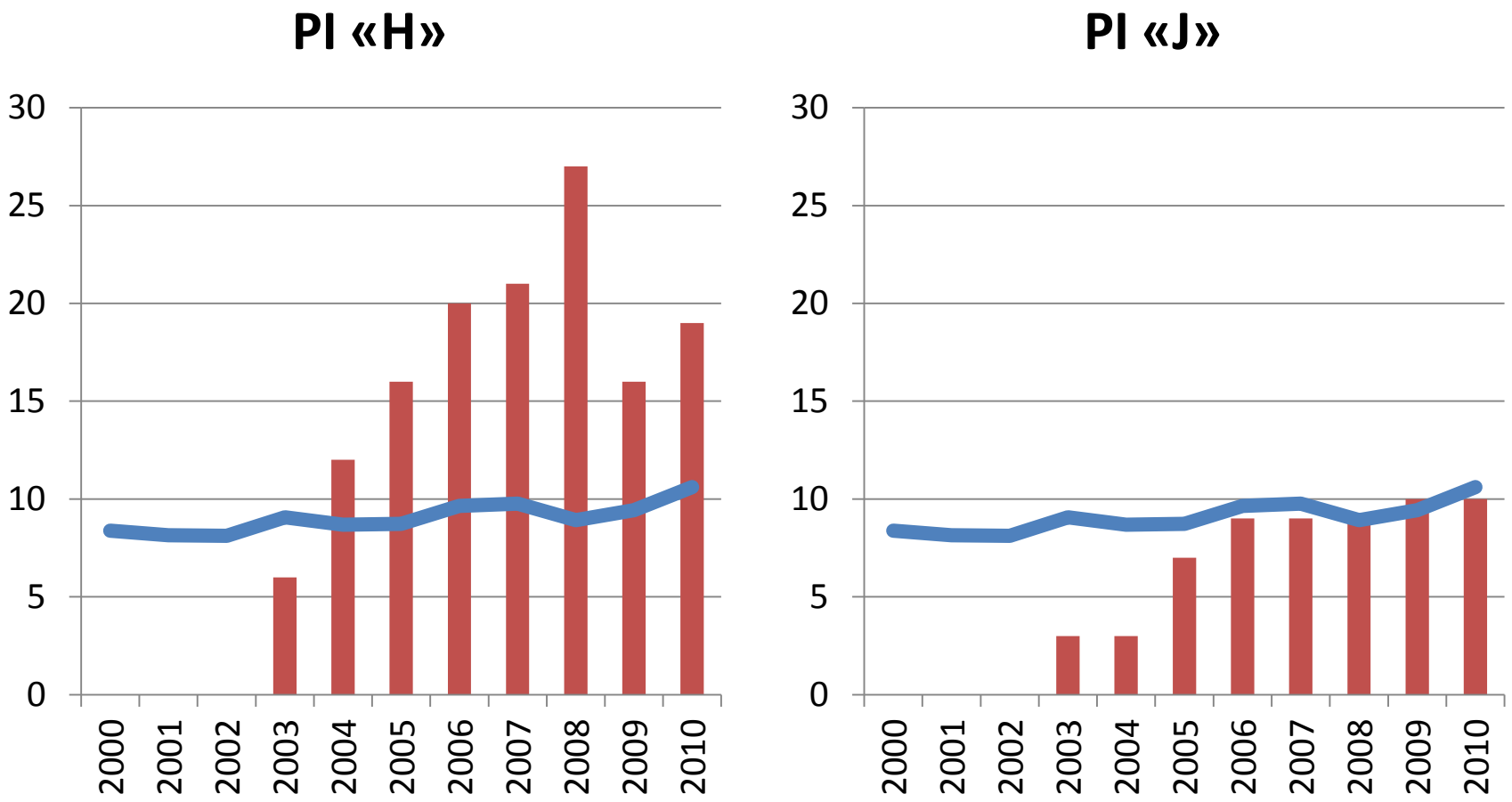
Unbalanced Panel

- Estimate at the PI level where observation is PI-flow of funds in a specific year from all grants (2258 observations)
- At the PI-grant-year level of observations: 7232 PI-grant-year observations
- Examine findings to see what they say about attribution
- Aggregate from PI-grant-year level to see what results say about way in which funds are bundled

Caltech Data

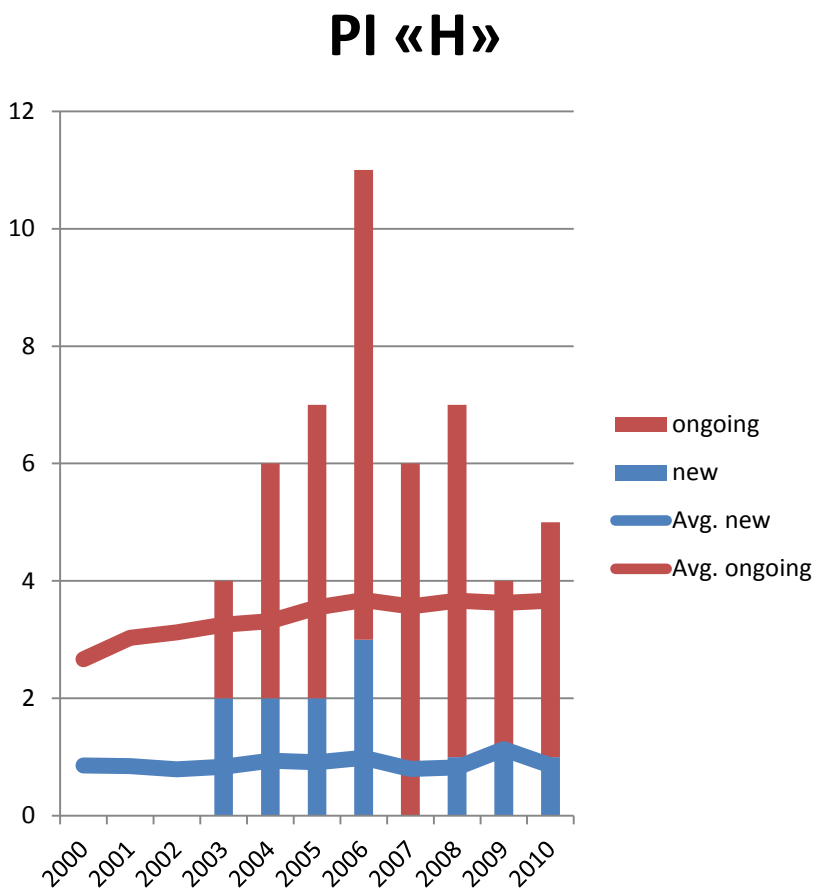
- Federal grants from 2000-2010
- PIs who hold faculty appointment at anytime during 11 year period: 279
- Maximum number of years observe a PI is 11; average number of years we observe a PI is 7.7
- Average duration of a grant is 4.1 years
- Annual flow of funds at grant level (including indirect) is $147,000
- Average flow of funds from all grants is $496,000

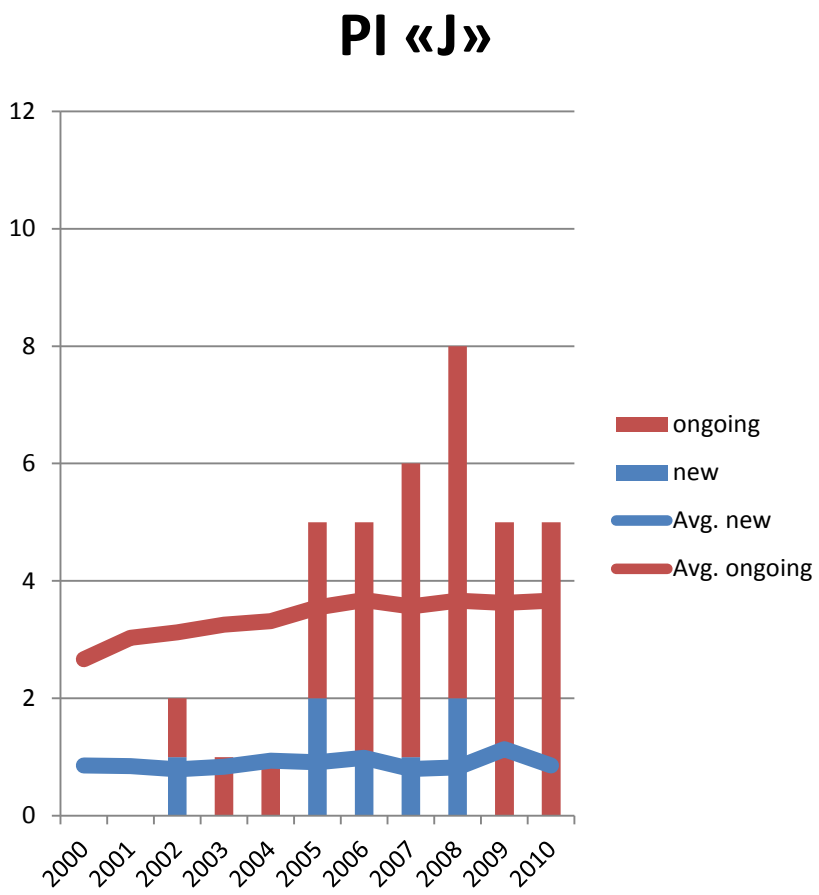
Number of Publications



blue solid line is the average per PI

Number of ongoing and new awards





Estimate a log linear model: log of publications is dependent variable

- An underlying two dimensional unbalanced panel at the PI, year level (2258)
- An underlying three dimensional unbalanced panel grant *g*, PI *i* and year *at Grant level (7236)*
  - This model can distinguish between flow of funds from the focal grant and spillovers from other grants active at the time

Regressions at the PI Level

Variables	Berg Approach Log (Publication count)	Weighted approach Log (Publication count)
log(Flow of funds)	0.33***	0.33***
year dummies	yes	yes
department dummies	yes	yes
Constant	1.94***	1.17**
Observations	2138	2,138
R-squared	0.174	0.225

Regressions at the PI Level

Variables	Berg Approach Log (Publication count)	Within weighted Log (Publication count)
log(Flow of funds)	0.33***	0.33***
year dummies	yes	yes
department dummies	yes	yes
Constant	1.94***	1.17**
Observations	2138	2,138
R-squared	0.174	0.225

Measures publications associated with flow of funds in year t to be all pubs in years t+1, t+2, t+3

Accounts for attribution by weighting publications by percent that funds received by PI in year t represents of all funds that could contribute to output in next 3 years

Effects of Attribution

- Elasticity does not change once one controls for attribution by weighting by flow of funds
- Constant does, indicating that predicted level of output resulting from given level of funds drops considerably once attribution has been taken into consideration

Regressions at the Grant level for weighted number of publications 2000-2010

VARIABLES	Quantity	Quantity
log(flow of funds_grant level)	0.70***	0.86***
log(flow of funds_spillovers)	-0.30***	-0.16***
Dummy one grant	-0.049	0.16***
Year (centered)	yes	yes
Field (centered)	yes	yes
PI Fixed effects		yes
Constant	-0.21***	0.023
Observations	7,232	7,232
R-squared	0.413	0.792

Interpretation

- Coefficient on focal grant is significant and implies a 10 percent increase in flow of funds from focal grant leads to 7.0 percent increase in publications associated with grant in years t+1, t+2, t+3
- Coefficient on spillover variable is negative and significant: 10% increase in funds available from other grants is associated with a 3% decrease in output related to focal grant
- Intuition underlying negative effect is that other funds "compete" with funds from focal grant and although these funds are expended on resources there are sufficient fixed inputs in lab to create negative spillover

Relevant Policy Question

- At national/international level, relevant policy question is what happens to output when all funds increase by 10%
- Here we find answer to be 4% increase—consistent with what we found at the PI level
- Conceptually congruent with fact that PI's essentially are running firms. In this entrepreneurial capacity, PI's arguably care less about the productivity related to one grant and more about the productivity of their firm.

Fixed Effects

- Find elasticity to increase once we control for fixed effects
- Suggests that underlying relationship between output and funds gets closer to constant returns to scale after we control for inputs specific to the PI such as ability, and persistence

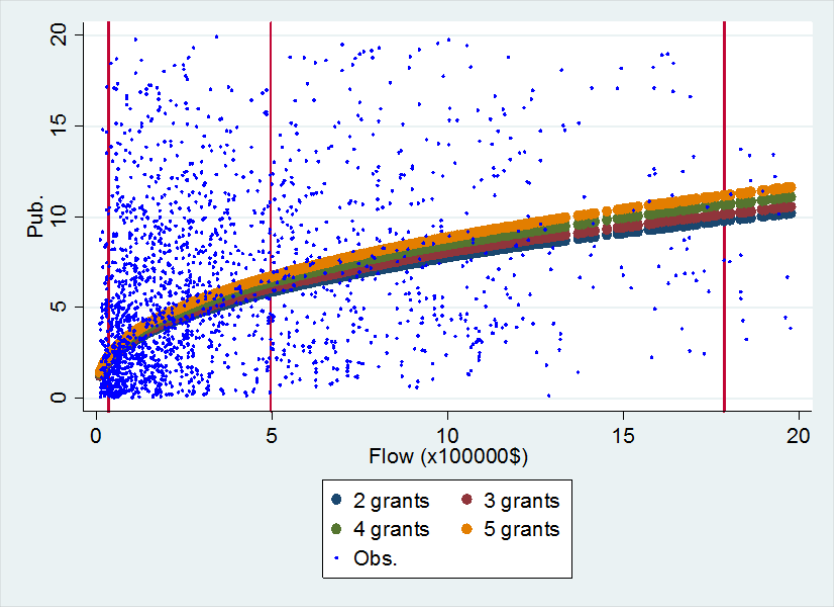
Neutrality Assumption: Does It Matter How PI's Funding Portfolio is Bundled?

Method of Analysis

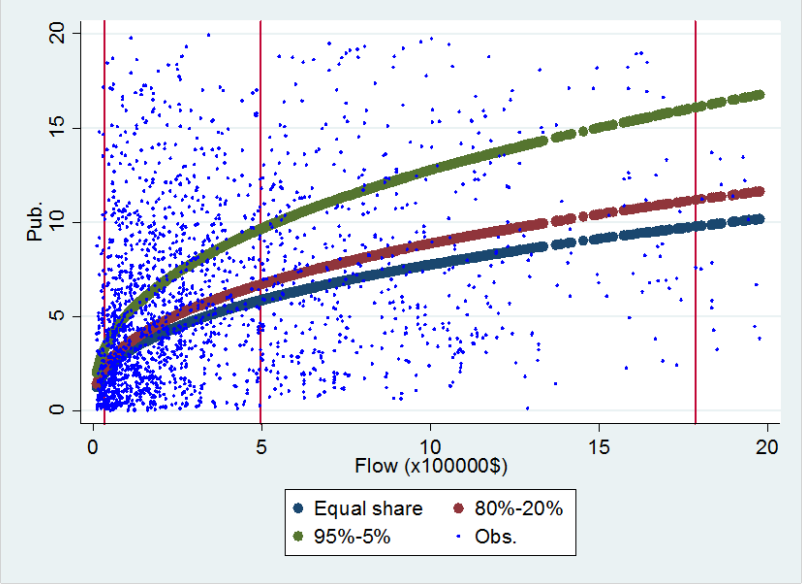
- Use estimates to simulate isoquants showing relationship of output to funding under circumstances where
  - Shares are equal, number of grants varies
  - Shares are not equal; number of grants constant

Comparison

Equal Share: Number Varies



Two grants: different shares



Summary of Simulation Results

- Find
  - Output is almost invariant to number of grants of equal size, holding funding constant
  - But holding funding constant, quantity increases as funding becomes more concentrated

Conclusion

Summary

- Two contributions
  - Accounting for attribution matters in terms of predicting quantity of output related to a specific flow of funds; but it makes no difference in measuring elasticity
  - Simulations suggest that neutrality assumption matters
    - More concentrated are a PI's grants, more productive is PI's lab as measured by the quantity of publications
    - A possible reason for the finding is economies of scale in grant administration.

Neutrality Finding

- Congruent with proposed MIRA pilot initiative announced by NIGMS in July 014 that would support an investigator's research through a single grant rather than through separate project grants
- Also congruent with National Cancer Institute's newly established Outstanding Investigator Awards that will replace a PI's project-based grants with 7 years of funding of up to $600,000 a year in direct costs

Caveats

- Results are for a highly selective university and for federal grants and are preliminary
- At some point hope to be able to incorporate funding from private foundations in the analysis
- Must wait to see how generalizable our results are when extended to other settings
- We invite others to use our methodology and do so

Questions/Comments

- pstephan@gsu.edu