A Penny for your Quotes? Assessing the impact of Biological Resource Centers on Life Sciences Research

Jeff Furman *(Boston U & NBER)*

Scott Stern *(Northwestern U & NBER)*

NCI's Biospecimen Best Practices Forum

November 5, 2007

Motivation: Knowledge accumulation is essential for economic growth

- Long-term economic growth depends on the ability to draw upon an ever-wider body of scientific & technical knowledge
  - the mere production of knowledge does not guarantee that others will be able to exploit it *(Mokyr, 2002)*
  - without mechanisms to ensure access to knowledge (at reasonable cost), researchers must literally "reinvent" the wheel *(Rosenberg; Mokyr)*
- While economic historians, institutional economists, and sociologists have emphasized the role of "institutions," the micro-foundations of knowledge accumulation are, by and large, still a "black box"
  - the "institutions of open science" depend on a subtle economic logic grounded in reputation and the ability for collective action to overcome the public goods problem of knowledge production *(Dasgupta and David)*
  - few detailed empirical studies demonstrating the impact of specific institutions/practices to the implications of this conceptual framework

Incentives for Cumulative Knowledge Production

- Establishing a knowledge hub within a technical community involves a collection action problem
  - role for public funding / cooperation among competitors
- Even if a knowledge hub is funded, the incentives to participate as a depositor may be too low without explicit rules or norms
  - social objective: maximize the impact of prior knowledge on reducing the costs to discovering new knowledge
  - as long as knowledge producers care about the impact of their knowledge (for intrinsic, career, or strategic reasons), positive deposit incentives
  - however, potential depositors trade off overall impact of knowledge with potential for rent extraction through continued control over knowledge
    - *example: lots of citations or lots of coauthorships?*

BRCs as Economic Institutions

- Economic institutions such as BRCs have the power to amplify the impact of scientific discoveries by enabling future generations to build on past discoveries
  - within the life sciences, "standing on shoulders" often requires access to specific biological materials or materials collections
    - *the precision of a given experimental design depends upon the understanding of the biological materials it employs*
- The evolution of BRCs as economic institutions seems to reflect the key collective action problems in transferring biological knowledge, via biological materials, across research generations
  - Authentication / Certification
  - Long-Term Preservation
  - Independent Access
  - Economies of Scale and Scope

BRCs as Economic Institutions

Authentication

- The fidelity of discovered knowledge cannot be guaranteed by the initial discoverer but must be able to be replicated
- Misidentification induces costly scientific errors
  - **HeLa Scandals**
  - contamination common at elite labs, as well as others
- BRCs at the forefront of ensuring biomaterials fidelity
  - nonetheless concerns persist *(Masters, 2002;PNAS, 2002)*

Long-Term Preservation

- The importance of a given piece of knowledge (and the physical materials required to exploit that knowledge) are often only recognized long after the time of initial discovery
- e.g., Brock's Unlikely Bacteria
  - 1967: Thomas Brock discovers *Thermus Aquaticus* in Yellowstone National Park geysers
  - 1983: K-Mullis conceives of PCR chain reaction, which requires extremophilie (Taq polymerase)
  - PCR becomes foundational tool for replication of DNA replication for modern molecular biology & biotechnology

BRCs as Economic Institutions

Independent Access

- Substantial gap between private and social benefits of providing independent access to data and materials
  - potential for rent extraction
  - potential to minimize discovery of errors
- BRCs support broad accessibility *(subject to scientific background)* in ways that the peer-to-peer network does not
  - IP Issues?
  - select materials?
  - democracy of science?

Scale/Scope Economies

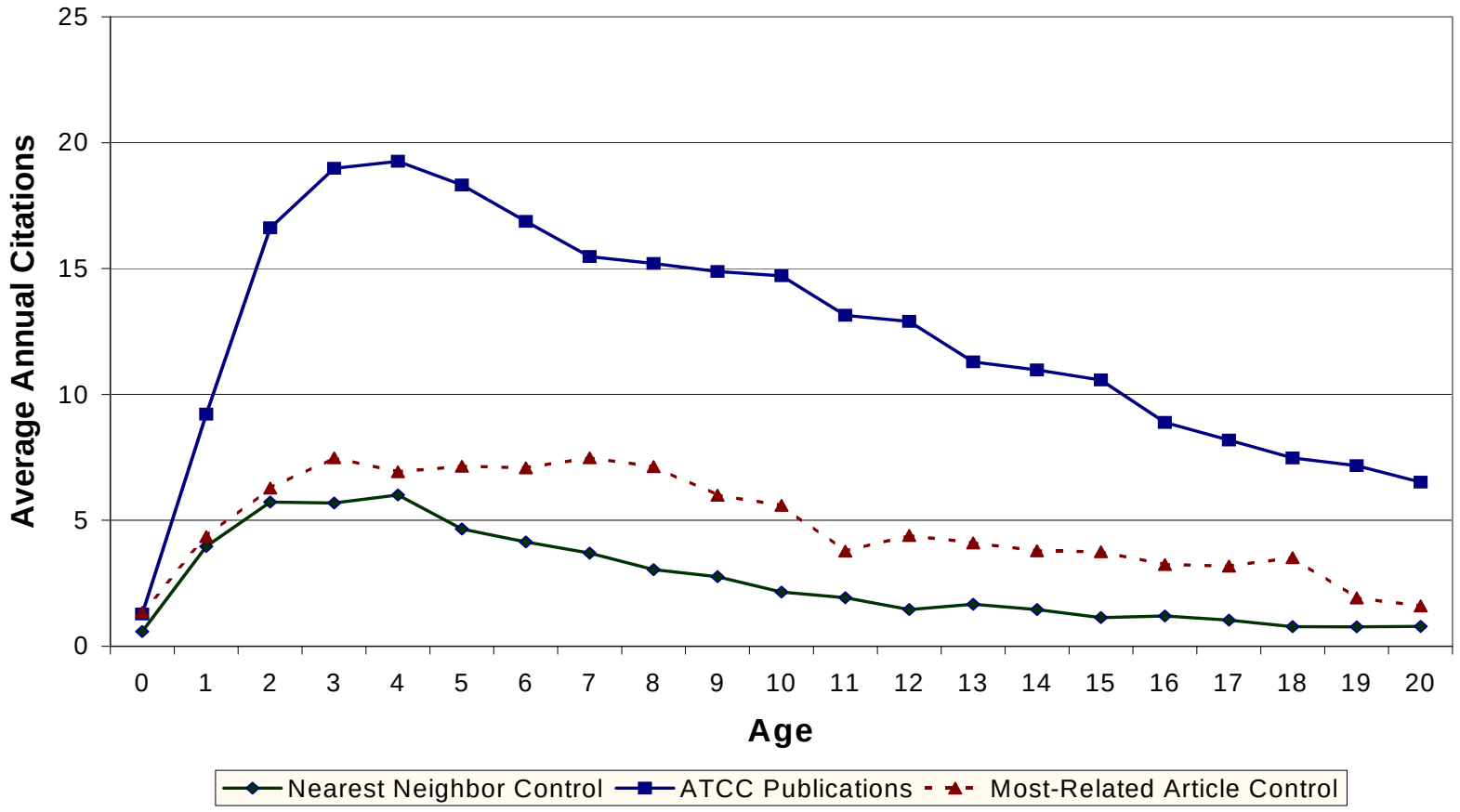
- Centralized institutions' investments in infrastructure, technology, & human capital may be cost-efficient relative to alternatives
  - substantial fixed cost component
  - learning-by-doing / specialization
  - minimizing replication of functions and collections *across* laboratories
  - establishment of a reputation as a "fair broker"
- Orphan Collections
  - even well-maintained collections are often "abandoned"

BRCs: An institutional response to the demand for public goods supporting the accumulation of useful knowledge

- Establishing effective institutions *(funding, leadership, etc.)* is subject to a public goods problem
- Even if a research-enhancing institution is funded, the incentives to participate as a depositor may be too low without explicit rules or norms
- The growth in importance of BRCs as a key intermediary reflects systematic efforts over time to overcome these collective action problems

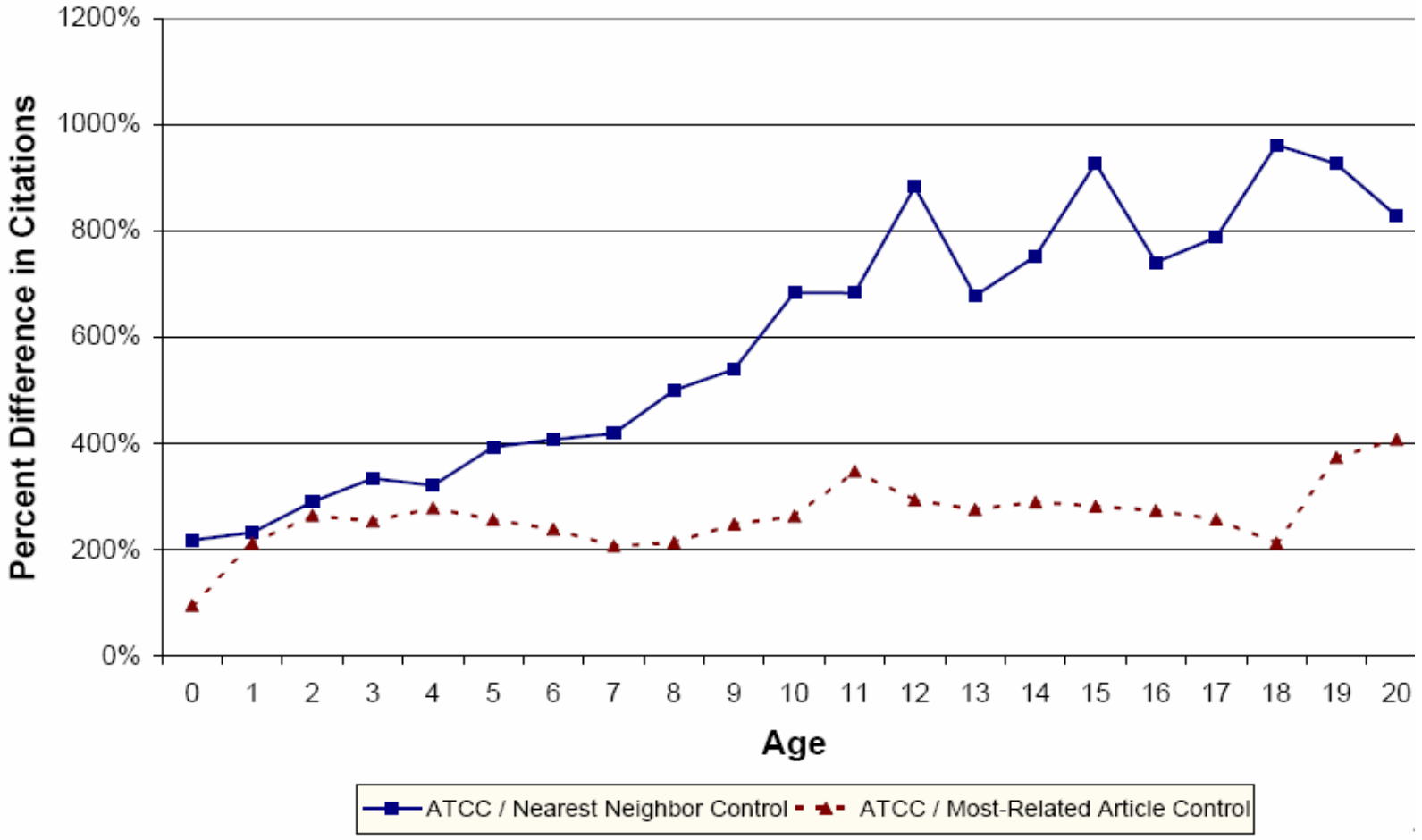
✪ ***But, do BRCs actually enhance the diffusion of scientific knowledge? How?***

In the cross-section, BRC-referenced articles have a higher rate of citation….



- Increasing (as a %) in the time since publication
- Robust to publication age, calendar, & article characteristic controls

In the cross-section, BRC-referenced articles have a higher rate of citation….



- Increasing (as a %) in the time since publication
- Robust to publication age, calendar, & article characteristic controls

Diffs-in-Diffs: Substantial Selection & Marginal Effects (Baseline Specification)

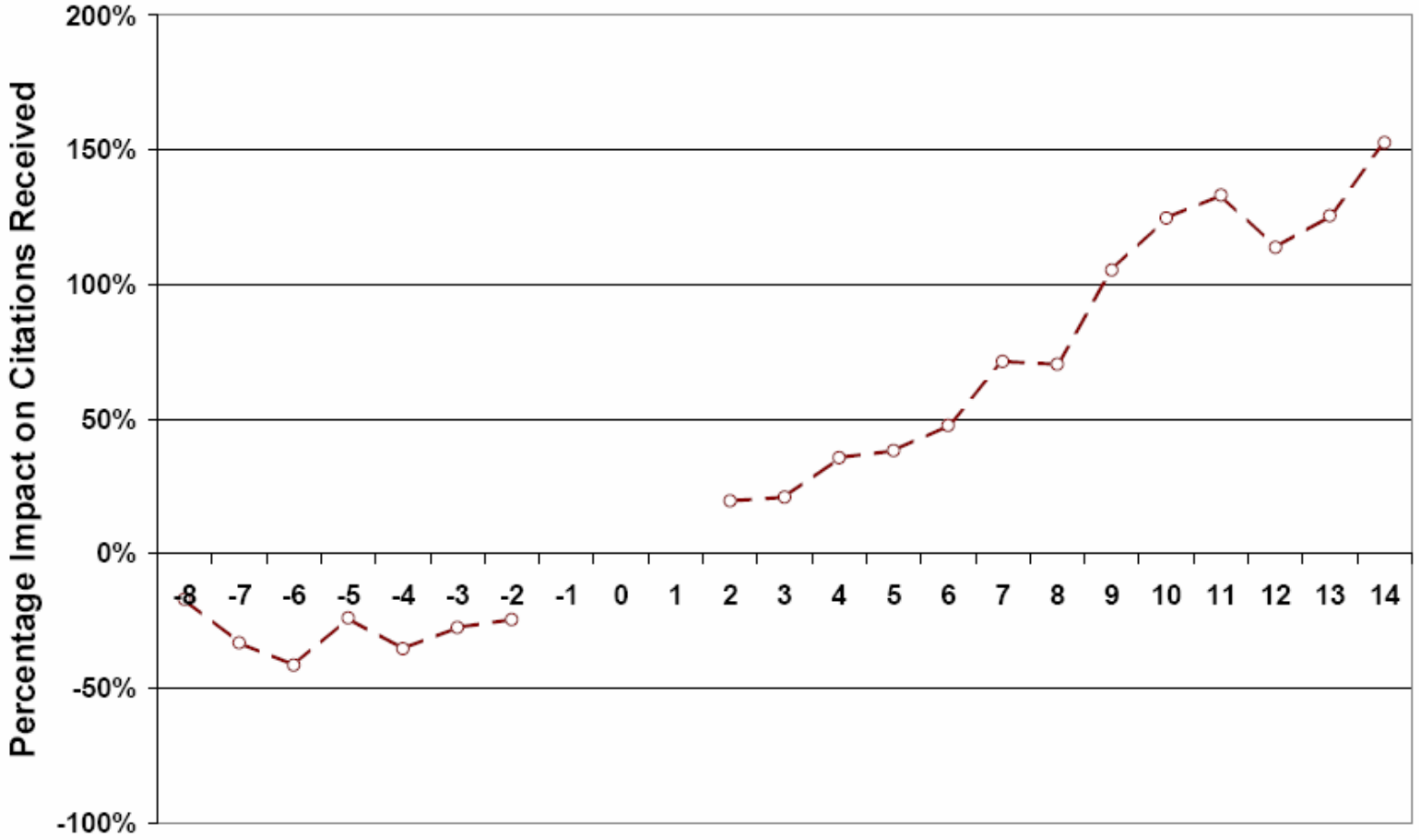
Negative Binomial Models	Forward Citations
	(4-3) Selection vs. Marginal
BRC-Article (*Selection*)	**2.08** **(0.31)**
BRC-Article,Post-Deposit *(Marginal)*	**1.91** **(0.43)**
Article Family FE	**X**
Article FE	
Age FE	**X**
Calendar Year FE	**X**

108% More Than Controls

91% Boost After Deposit

** Cond FE Neg. Bin. Models, coefficients as IRRs; bootstrapped SEs*

Impact of Deposit Grows Over Time and Does Not Exist Prior to Deposit



- This suggests that deposit is, indeed, exogenous and that diffs-in-diffs approach usefully identifies marginal (post-deposit) effects
- *FE NB model*

Rate-of-return analysis: Should the marginal $ go to another experiment or ensuring that funded experiments are accessible to the next generation?

- Biological Research Social Planner's Objective: In each period, maximize the growth in the stock of knowledge available for future periods
- Compare how BRC accession expenditures compare to traditional research expenditures in creating a pool of knowledge for future researchers
- Counterfactual: Compare the "cost per citation" (i.e., the productivity of the citation production function)
- Combining estimates from a variety of sources, the results suggest a 3x – 10x higher rate of return to investments in authentication and access, relative to simply funding another experiment

Calculating the Cost Per Citation

- Traditional university-based biological research: Adams & Griliches (1996) estimate the *cost per citation* for a large sample of university biology departments in the U.S. during the 1980s
  - for each university department, they compute
    - Total research expenditures
    - Total paper production
    - 5-year citation rate
  - estimated cost per citation = $2,400 - $4,200

- BRC Accession Costs: Calculate the cost per deposit and the expected boost to citations per deposit (over 5 year window)
  - OECD (2001) estimates "high" cost per BRC-deposit ~$10,000
  - Citation boost ranges from 8.1 – 36.8 citations, depending on reference group

BRC Cost-Effectiveness Calculation

Calculation	Baseline Citation Cost	BRC Accession Cost		BRC Citation Boost		BRC Citation Cost	BRC Cost-Effective-ness Index*
BRC-Linked Article Citation Boost	$2,400	$10,000	÷	36.8	=	$271	10.63
"Top Ten" Uni. Citation Boost	$2,400	$10,000	÷	13.9	=	$719	4.01
Random Uni. Citation Boost	$2,400	$10,000	÷	8.1	=	$1238	2.81

Implications

- *Knowledge Enhancing Institutions*, such as BRCs, are closely associated with the process of cumulative knowledge production within specific scientific or technological fields
- They can influence future knowledge accumulation in 2 ways:
  - the ***Selection Effect:*** Institutions can attract researchers and knowledge production with intrinsically high cumulative value
  - the ***Marginal Effect:*** The impact of knowledge is amplified by its association with a knowledge enhancing institution
  - the magnitude of that impact can *grow* with time
- BRCs evidence all of these characteristics
  - further, relative to traditional grant mechanisms, BRCs seem effective at maximizing the knowledge pool available for future research
- *Though cumulative knowledge production is central to economic growth, the extent of knowledge growth depends on the effectiveness of often "invisible" institutions, whose features can be influenced by public policy*