Of Mice and Academics: Examining the Effect of Openness on Innovation
Fiona Murray Philippe Aghion Mathias Dewatripont Julian Kolev Scott Stern

Discussion: Bronwyn H. Hall
University of California at Berkeley and University of Maastricht
Concern that upstream IPR may be counterproductive for research progress
- E.g., EPFL – materials transfer agreements (Aebischer)

David – conflict between norms of open science and IP
- OS: rewards are reputational, etc., encourage citations
- IP: rewards are due to right to exclude, which reduces citation activity
“Optimal” incentives for cumulative innovation

- Give first innovator IP rights
- After costs are sunk, take them away
- That’s what happened here – not an experiment that can be repeated very often
- But OS incentives (with public funding) deliver the first innovation regardless
Research question

- How does openness affect innovation?
  - Well-known tradeoff between incentives for first and second generation researchers
  - How does this operate in the case of academic biotechnology research?

- Two parts to paper:
  - Use a simple model to derive predictions
  - Test them using a large panel of sci papers and D in D methodology
    - confirmation rather than rejection
Model predictions

- Lowering cost of access to research inputs is expected to
  - Increase quantity of follow-on research
    - At a point in time
    - Over time
  - Increase diversity of follow-on research
    - More researchers
    - Different types of research
  - Increase basic research relative to applied research
- Did we need the model to make these predictions? (I am not convinced)
Empirical evidence

- Well-executed and very compelling
  - Relates annual citations received by papers that cite or do not cite mice which have been made open access in 1998.
  - Breaks it down:
    - Cites from new v prior researchers
    - Cites from new v prior institutions
    - Cites with new v prior keywords
    - Cites in new v prior journals
  - Fairly large impacts, all in the right direction
Sample plot

Citation impact - illustrative example - 1994 papers

- Overall
- Window
- Age (common)
- Trend (common)
Comments and suggestions

- Paper needs more explanation of exactly what was estimated and why
- Discussion of any effects due to avoidance of “visibility”
- Separate trends for the two groups – plot?
- To what extent are there false “new” cites due to spelling errors?
  - probably does not affect the D in D
- Identification problem for age, year, fixed paper effects (next slides)
The identification problem

- Want to measure citations as a function of age of the article, publication date (or fixed effect), and time period (current year)
- Well-known that the identity $\text{age} = \text{year (period)} - \text{year of birth (pub. date)}$ implies all 3 cannot be identified in a linear model
- Less well-known that identification can be achieved in a dummy variable model by dropping a small number of variables
  - Berndt and Griliches (*J of Econometrics 1991*)
  - Hall, Mairesse, Turner (*EINT 2007*)
Models

saturated: $p_{it} = a_{ct} + \varepsilon_{it}$

three-way: $p_{it} = \alpha_{c} + \beta_{t} + \gamma_{a} + \varepsilon_{it}$

two-way: $p_{it} = \alpha_{c} + \beta_{t} + \varepsilon_{it}$

and so forth....

where $i = 1,\ldots,N$ papers

t $= 1,\ldots,T$ years

c $= 1,\ldots,C$ pub date (or fixed effect)
a $= t-c$ (age)
## Saturated model

<table>
<thead>
<tr>
<th>Pub. Date: Year ↓</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>$a_{10,1}$</td>
<td>$a_{10,2}$</td>
<td>$a_{10,3}$</td>
</tr>
<tr>
<td>11</td>
<td>$a_{11,1}$</td>
<td>$a_{11,2}$</td>
<td>$a_{11,3}$</td>
</tr>
<tr>
<td>12</td>
<td>$a_{12,1}$</td>
<td>$a_{12,2}$</td>
<td>$a_{12,3}$</td>
</tr>
<tr>
<td>13</td>
<td>$a_{13,1}$</td>
<td>$a_{13,2}$</td>
<td>$a_{13,3}$</td>
</tr>
<tr>
<td>14</td>
<td>$a_{14,1}$</td>
<td>$a_{14,2}$</td>
<td>$a_{14,3}$</td>
</tr>
</tbody>
</table>
Identification

- Oneway – all dummies are identified (but no intercept)
- Twoway – drop one dummy
- Threeway – drop two dummies
- Threeway where $a = t-c$:
  - Drop one additional dummy! (Berndt and Griliches 1991)
- How robust are the results to the choice of dummy to drop?
Suggestion for further work

Belenzon finds positive feedback effects to firm \textit{j} from:

- \textit{pat (firm j) \rightarrow pat (firm i) \rightarrow pat (firm j)}
- In this context, how are second generation cites by original researcher affected?
- Does he/she benefit more from reverse spillovers?
Wider applicability?

- Publicly funded science
  - Downstream sources of revenue for funding unlikely or remote or highly risky
  - Benefits of diversity high, incentive effects not greatly harmed (since they are mostly reputational)

- Private R&D?
  - IBM’s 500 patents