Comments on "Why are some regions more innovative than others? The role of firm size diversity" by Agrawal, Cockburn, Galasso and Oettl

Saul Lach
The Hebrew University

May 2012

Main hypothesis

- Large firms (labs) produce ideas but do not develop/commercialize all of them because not all ideas “fit”.
- Spin-outs are formed to commercialize them.
- When more spin-outs are created, fewer ideas abandoned $\rightarrow$ innovation productivity is higher.
- Assumption: spin-out formation cheaper when there are more spin-outs.
- $\Longrightarrow$ more spin-outs and innovations in regions with large and many small labs (given same number of inventors).

Comments on model

- In fact, optimal diversity is when **exactly one** large firm and many small firms ($N > \underline{N}$) coexist, because of assumed scale economies in ideas production.
  - this cannot always be true: optimal number of large labs depends on shape of scale economies and # of scientists in MSA.
  - Generalization in appendix? need "size normalization".
- Costs of spin-out formation decrease with number of spin-outs. How realistic is this assumption? (post 1960s) evidence?
  - a large part of costs is scientific manpower whose supply may be highly inelastic $\Longrightarrow$ spin-out formation costs may increase (if not just redistribution of scientists).
  - costs of spin-out formation can be focus of policy tools.
- Predictions (proposition 2) are for social planner but data are generated by individual, profit-maximizing firms/entrepreneurs...what predictions would a market equilibrium deliver?

Errors in measurement

- Not all inventions are patented $\Longrightarrow$ number of inventions and of inventors biased downward.
- How do these measurement errors affect diversity measure and estimates of its effect on innovations/spin-outs?
  - Inventions invented by multiple inventors and/or of higher quality may be more likely to be patented.

Fixed effects?

- Long sample period: most unobserved effects not so fixed. FE will be missing a lot of unobservables.
  - Use supplementary data on infrastructure and labor market conditions which vary over time and MSAs
- IV estimation is therefore useful.
  - income taxes may affect self-employment but less so measure of diversity given its discrete nature (strength of the instrument).
  - IV results much larger than OLS. What is driving this?
    - omitted variables negatively correlated with diversity measure?
    - reverse causality: more innovation decreases diversity?. Patent thickets would imply more patents and more difficulties in setting up spin-outs (less diversity) ... but you are measuring patent citations not just counts.

Model specification

- A diverse MSA has **at least** one large lab. x
- Finding a positive effect of such diversity is interesting but not clear how useful it is for policy purposes.
- An estimated 17% increase in citations when MSA increases its number of small labs from 139 to 140? Isn't this too large an effect?
- Problem is with coarseness of diversity measure.
- Estimate effect of different configurations of large & small labs allowing for this effect to vary with MSA size. . . more demanding of the data and perhaps unfeasible.
  - Use sets of dummies for zero large labs, for exactly one large lab, for 2 or more large labs, etc.. . . and for small labs.
    - similar to Table A2 but with finer large lab dummies classifications and interaction terms.

Additional comments on empirics

- Tables 8 & 9 show a lower effect of diversity on innovations in MSA's with enforceable non-compete agreements and in MSAs where large labs conduct less focused research.
- Rationale for the effect of non-compete agreements and "technological focus" on **innovations** is that they affect the **formation of spin-outs**.
  - Estimate effect of non-compete agreements and "technological focus" on spin-outs **directly**.
  - Similar to Tables 8 & 9 using number of spin-outs as dependent variable.
- If all spin-outs are small labs doesn't this generate a positive relationship between number of spin-outs and diversity measure?

Arbitrary decisions

- Many arbitrary decisions due to reliance on data originating from patent document.
  - but checks for robustness of some decisions.
  - are there external data on number of labs/spin-outs?
- Threshold number of labs ($\underline{N}$) depends on profits of inventions (π) and on cost of spin-out formation ($k(\cdot)$).
  - π and $k(\cdot)$ likely to vary across technologies, time, and location $\rightarrow$ changing definition of "large" number of small labs.
  - How would results be affected by allowing for different thresholds?