Marx, Strumsky, and Fleming: Mobility, skills and the Michigan non-compete experiment

Discussion: Bronwyn H. Hall

UC Berkeley and U of Maastricht

Overview

- Interesting and carefully-done paper, deserving of the prize
- Main result: non-compete agreements reduce the probability that a knowledge worker will move from one firm to another
- Topic is important, and the subsequent paper makes it clear that it matters, in the sense of causing brain drain to states that do not enforce non-competes
- Comments
  - K-sharing equilibria example
  - Using patents to track inventors

Knowledge-sharing equilibria

- Paper cites Rabaut (2006) – employers become wary of non-competes – they are both a "hiring shield" and a "hiring sword"
- That is, all may benefit from not having enforceable non-competes, but each individual firm will be tempted to use one if available
- Gambardella and Hall (2006), Research Policy – model of knowledge-sharing equilibria, showing that cooperative equilibrium is not stable in the absence of some kind of enforcement

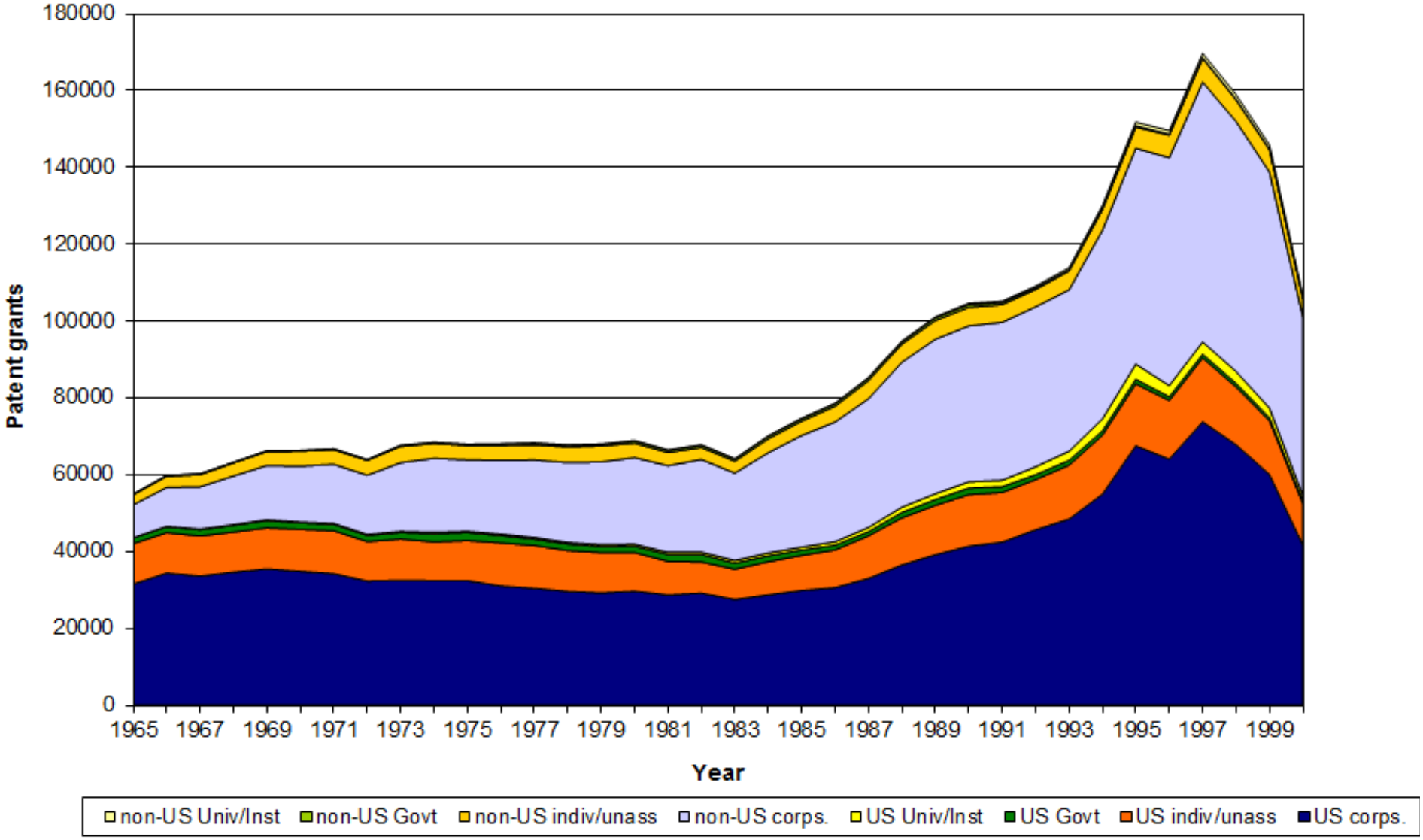
Gambardella-Hall (2006)

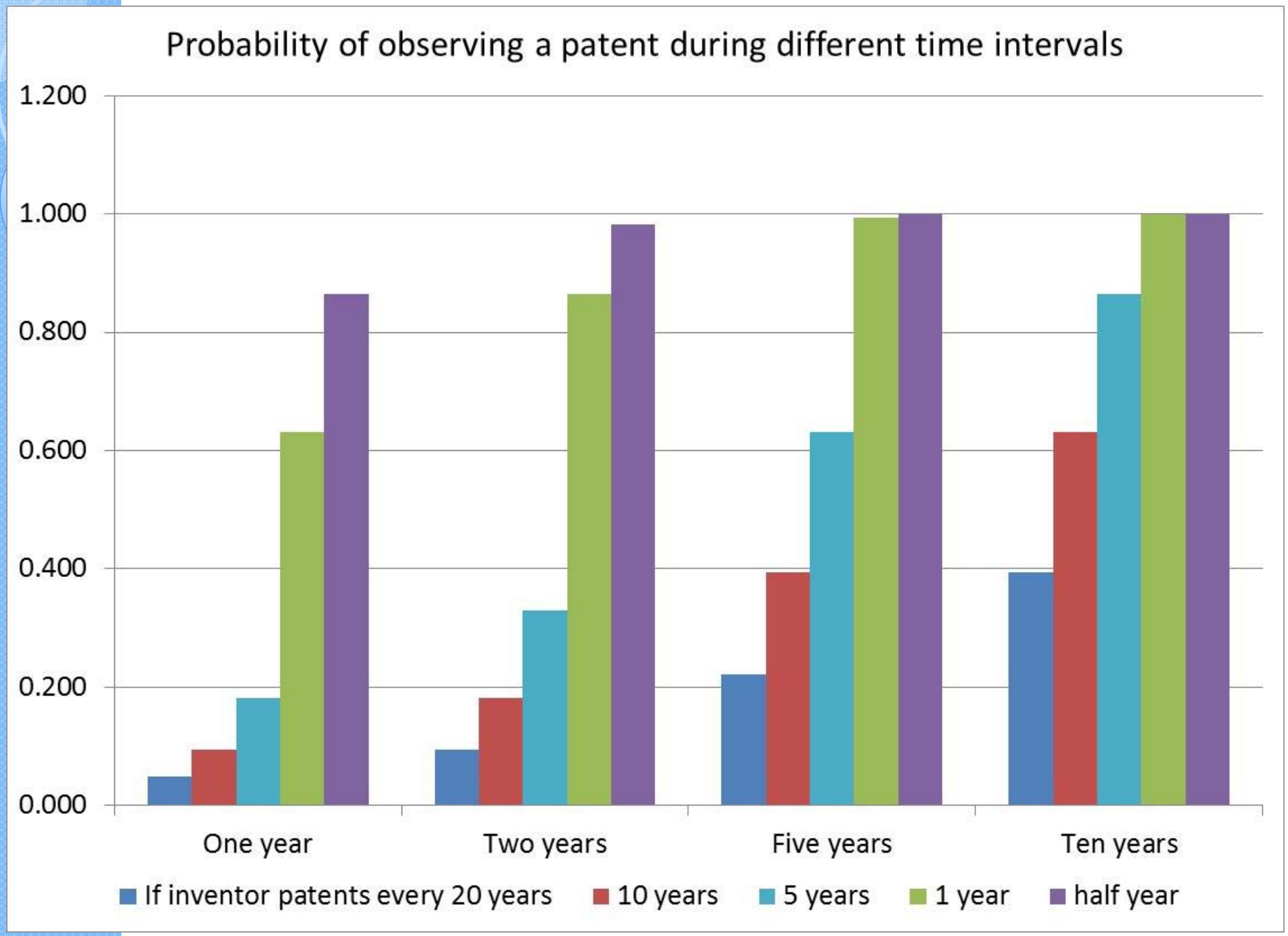
- A model of the interaction between public domain and privatized provision of research outputs (e.g., software, knowledege, etc.)
- Find an equilibrium where there is a fraction of researchers in a field whose discrete contribution to the public good exceeds their private profits
  - They therefore participate in PD
- However, as public domain grows and returns to privatizing increase, incentive for a public domain researcher to switch to private provision increase
  - Coordination required to preserve the public domain
- Same idea applies here – firms may be better off if all agree not to use non-competes – legislatively enforced

Patents as a measure

- More than half of inventions/innovations not patented (see *Hall et al. 2013* on UK data, *Fontana et al. 2013* on R&D 100, many others)
- Very skewed value distribution, suggesting skewed input cost also
- Differences pre and post-1985?
- For many inventors, a relatively rare event
  - For some, we observe no patents
  - For others, we observe very few
- Moves identified via patents will therefore be undercounted
- Next 2 slides:
  - Aggregate patent grants by date of application
  - The probability of seeing at least one patent for inventors with different frequencies of discovery

Aggregate US Patent Grants by Year of Application 1965-2000




Probability of observing a patent during different time intervals
1.200
1.000
0.800
0.600
0.400
0.200
0.000
One year
Two years
Five years
Ten years
If inventor patents every 20 years
10 years
5 years
1 year
half year

Simulating the patenting probability

- Model developed for discussion of *Breschi et al* (in forthcoming WIPO volume edited by Fisk and Miguelez)
  - Probability inventor *i* applies for a patent is Posson with parameter λ_i
  - Inventors are heterogeneous, so λ_i is drawn from some suitable distribution (log-normal or Pareto)
  - Probability that inventor moves is a small number, assumed to be the same for all inventors (0.09 pre-1985 and 0.06 post-1985)
  - Inventors are observed over the 1963-2006 period (abstracting from entry and exit)
  - Latham et al. (2006) - US inventors average 0.05 patents per year; inventors here average 0.09 patents per year (=3.79/43)
    - Calibrate distribution of λ_i using distribution of inventors in Latham, but use the mean patenting rate from this paper
  - Next slide: a table that compares a simple model of actual moves with those that are observed, in an attempt to assess possible bias

Bias from non-observed data

	Logit coefficient for change in prob(move) after 1985					
	using +/-22 years		using +/-11 years		using +/-5 years	
	Actual (all data)	Observed data	Actual (all data)	Observed data	Actual (all data)	Observed data
D(post-85)	-0.45 (0.01)	-0.68 (0.02)	-0.44 (0.02)	-0.47 (0.03)	-0.44 (0.03)	-0.41 (0.04)
Observations	372,896		182,112		78,048	
Moves	28,212	12,206	13,706	6,542	5,950	2,950

Sample is all inventors with at least one patent.

Conclusions:

1) Fewer than one half of actual moves are observed
2) Estimated effect is biased away from zero only for the longest sample, possibly because there are edge effects
3) Results in Marx et al. are probably not much affected by this problem, because they are based on comparisons (DOD)