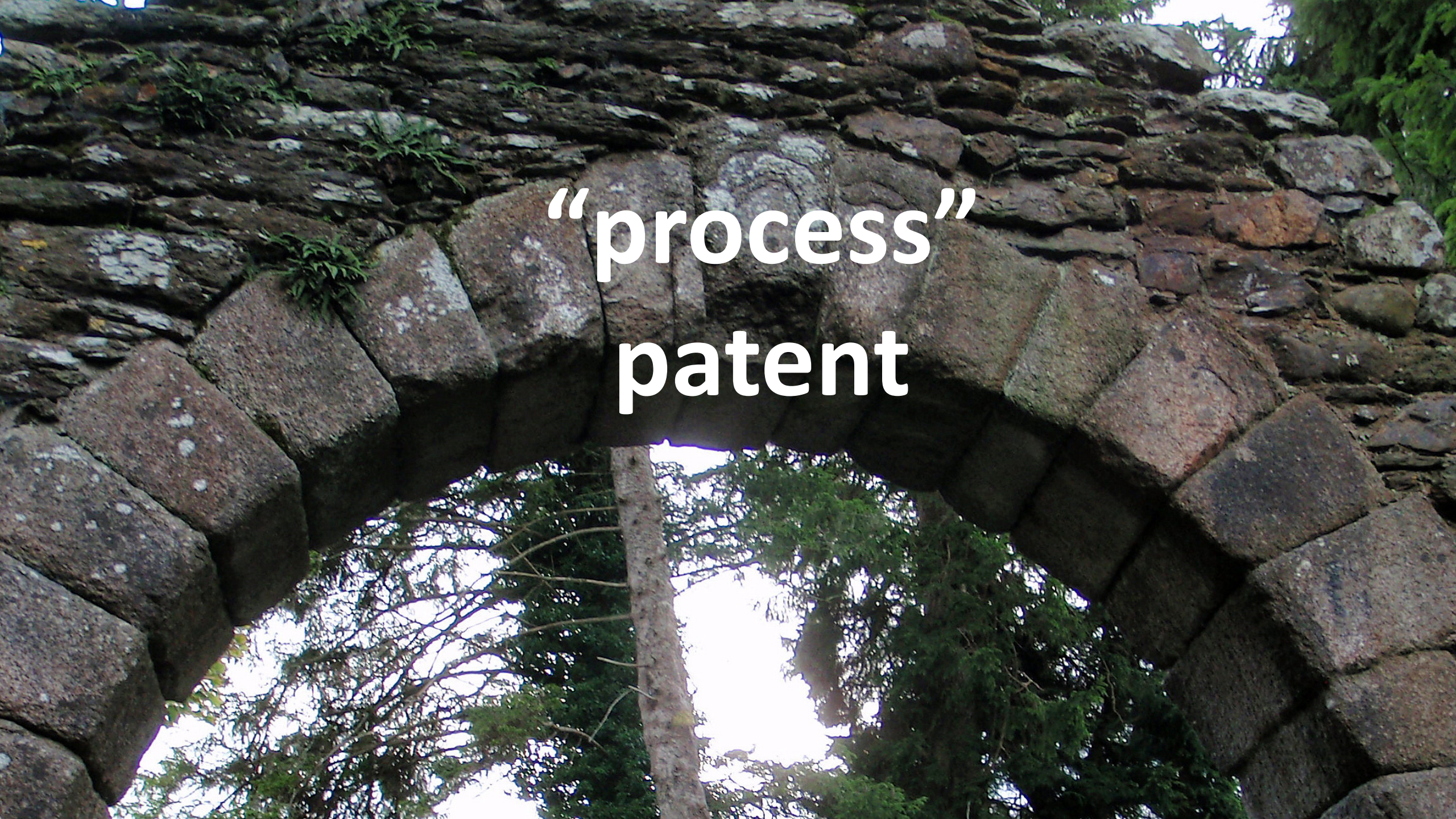
Comments on Ganglmar/Reimers, Kuhn/Thompson

Matt Marx (BU Questrom)
@marxmatt
http://relianceonscience.org

Ganglmair/Reimers: Visibility & Patenting

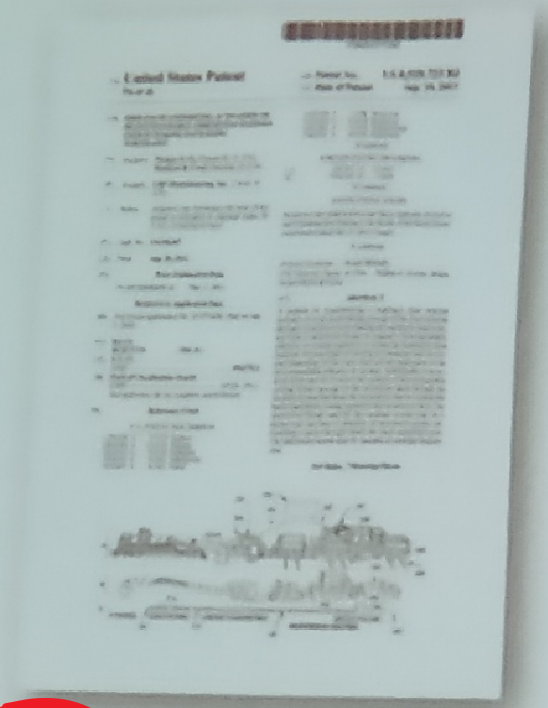
- Rare view into the world of secrecy!
- Transplants methodology of Png'17 to assess impact of UTSA
  - Staggered DD, related-law "whimsicality" IV
- Especially like assume-earlier-reform placebo test
  - Maybe randomize directionality of treatment
- Structural model I can't even hope to understand…


“process”
patent

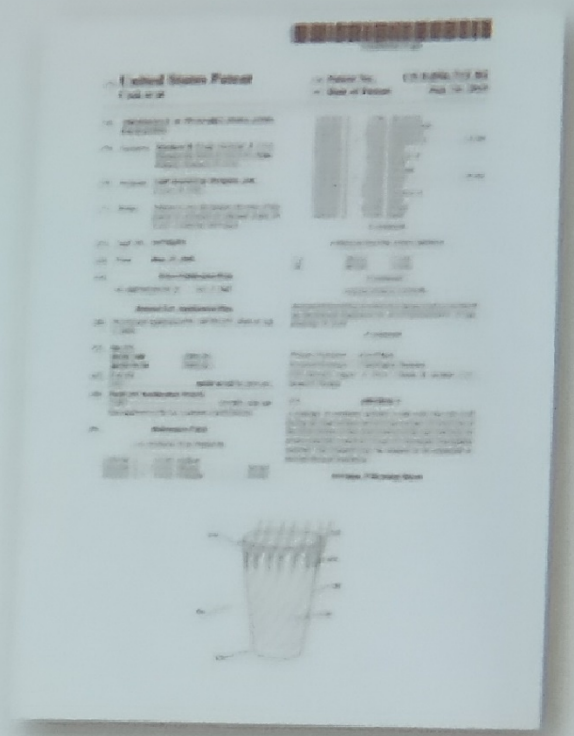
3.3 Indicators for Process and Product Patents

To categorize process and product patents, we use data constructed by Ganglmair et al. (2019), who employ text-analytical methods to identify invention types of independent claims in a given patent.[22] Claims can be of three distinct types: (1) process (or method) claims describe the sequence of steps which together complete a task such as making an article; (2) product-by-process claims define a product through the process employed in the making of a product; and (3) product claims describe an invention in the form of a physical apparatus, a system, or a device.

We aggregate the claim-level information in Ganglmair et al. (2019) to obtain an indicator for the invention type at the patent level. Specifically, we classify a patent as a *process patent* if at least one of its independent claims is either a process claim or a product-by-process claim, and as a *product patent* otherwise.[23]



Process of expediting activation of heat-expandable adhesives/coatings used in making packaging substrates



Thermally activatable insulating packaging

Table 1: Summary Statistics

	N	Mean	Median	SD	Min	Max
Process patent	1461240	0.472	0	0.499	0	1
Number of process claims	1473486	0.863	0	1.402	0	60
Number of product claims	1473486	1.903	2	1.884	0	104
Indep. claims	1463686	2.873	2	2.283	1	116
Length of first claim (words)	1463682	168.969	148	106.535	1	7078
Length of description (chars.)	1473876	26031.370	15628	39648.204	4	3608036
Generality	1114531	0.639	0.719	0.244	0	1
Originality	1295568	0.626	0.694	0.244	0	1
4th year renewal	1379555	0.825	1	0.380	0	1
Observations	1473878					

Quick & dirty variable reconstruction

- Processed full-text of 97,498,661 patent claims, 1976-2019
- Marked process patents if any claims contained the words
  - “process”
  - “method”
- 47.9% of 6,910,405 patents classified as process patent

- Would like to see validation of measure
  - false positives, false negatives, ideally via 3rd-party test set

11
10
Getty Images

Kuhn & Thompson, Patent Races

- Two main contributions
  1. Winner of a patent race performs more follow-on research
  2. 10-11% of patents came from "twin" *applications*

- Identify "twin" patent applications via *novelty* rejections
  - Similar applications unobservable for up to 18 months
  - Papers (Bikard'18,19; Bikard/Marx'19; Bikard/et.al'19; Marx/Hsu'19)
  - Patent (Ganguli/et.al'19): 1,200 decisions on interference proceedings

- Heroic efforts to digitize PAIR proceedings!

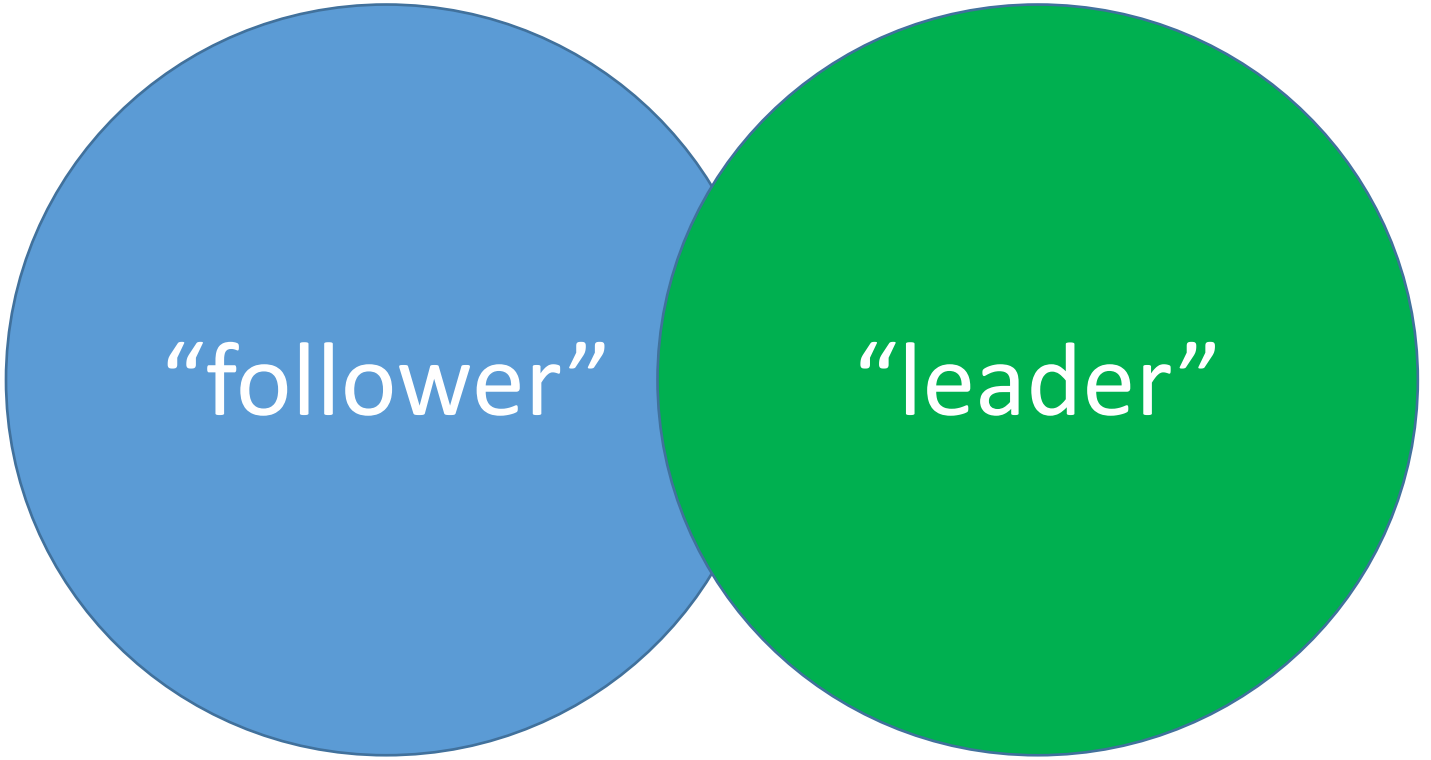
“twin”
patent

Defining patent “twins”

- 10-11% is a lot of twins!
- Two conditions
  1. 18-month temporal overlap
  2. “sufficient technological overlap”
     - # or % of claims rejected to qualify as twin?

GoPro

Uh-Oh



Defining patent “twins”

- 10-11% is a lot of twins!
- Two conditions
  1. 18-month temporal overlap
  2. “sufficient technological overlap”
- # or % of claims rejected to qualify as twin?
- Paper shows greater TF*IDF similarity among twins than in class
- Remedies
  - Show us distribution of claims overlap (conf. score?)
  - Hand-check for false positives esp. with low claim overlap
  - Use patent-interference decisions to test for false negatives

Another interpretation of the core result…

- Rosenberg: Why do firms do research with their own money?
- Popular answer: absorptive capacity (Cohen/Levinthal, 39k cites)
  - Publishing is “ticket to the game” of scientific community (Chai/et.al’14)
  - Scientists reluctant to join “closed” firms (Stern’04; Hicks’95)
- Absorptive-capacity argument is *extramural*
- Are there *intramural* reasons to do R&D?
  - Theory, cases, surveys, correlations…little if any causal evidence
  - Key challenge: firms decide to protect inventions with IPR

→ Your results suggest that R&D has inherent value

Open dataset! http://relianceonscience.org

- Others have linked patents to Web of Science; can't share ☹
- We found 16.7MM linkages to Microsoft Academic Graph
  - USPTO front-page citations since 1947
  - Applicant/examiner indicator
  - Confidence score (99% accuracy with ~~90%~~ 93% coverage)
- Metadata for all papers (+ JIF, locations)
- Coming soon: body-text citations since 1926