Innovation Metrics:

We are working on the answers.

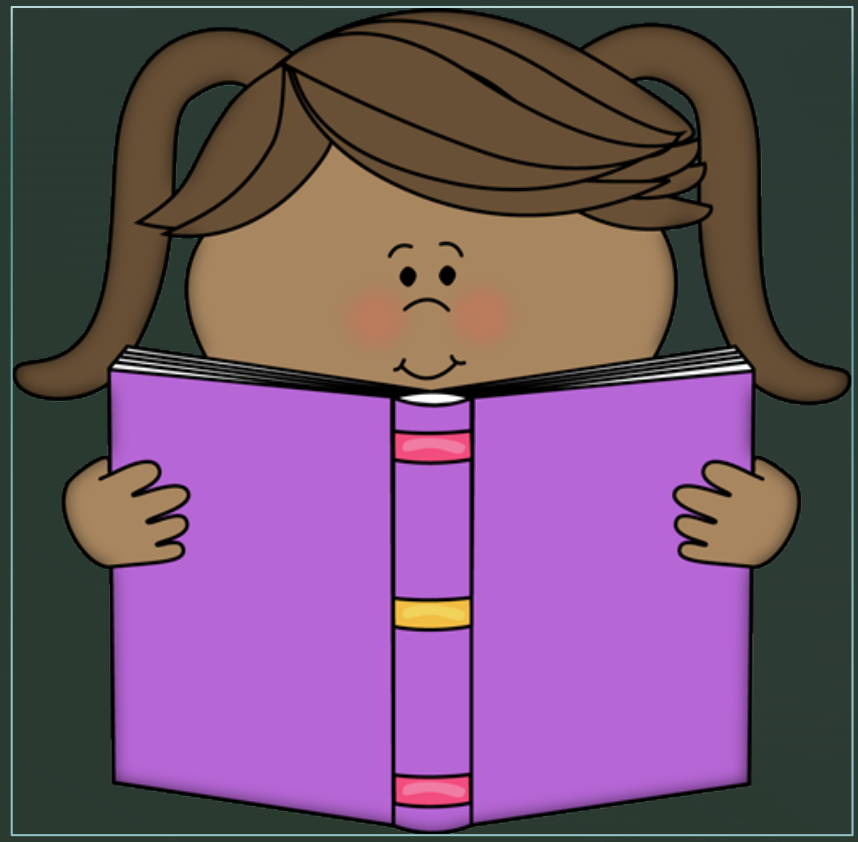
What are the interesting questions?

Adam B Jaffe
Brandeis University
MIT Sloan School
Motu Economic and Public Policy Research
Queensland University of Technology

Knowledge Creation and Innovation Workshop
Queensland University of Technology
28 October 2019

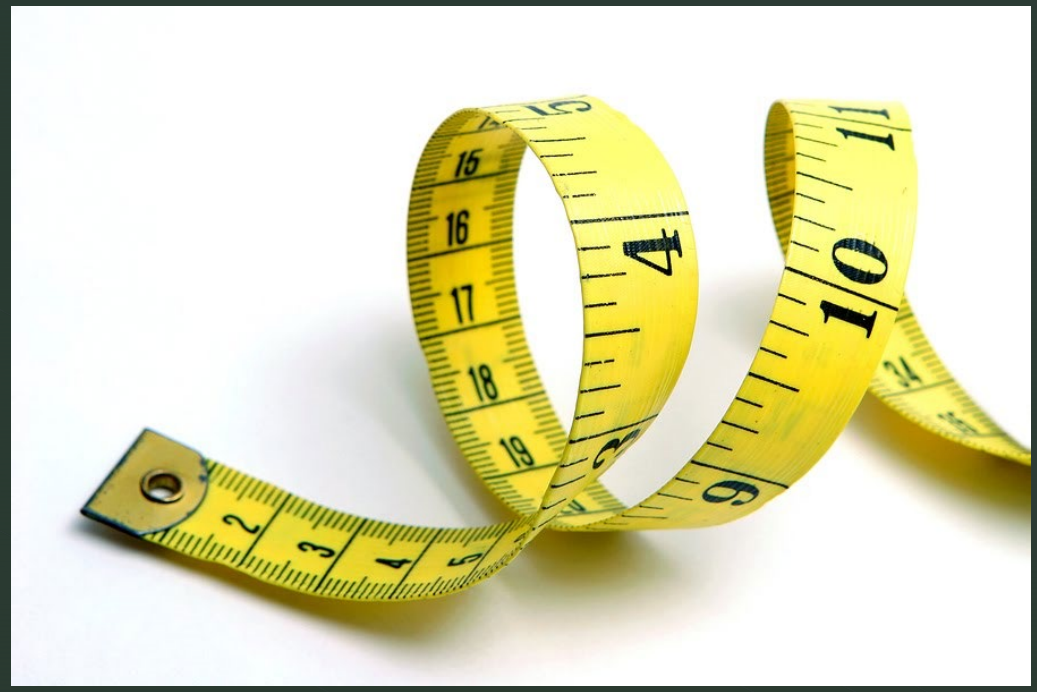
Roadmap

- What are we trying to measure and why
- Over-arching issues
- Coda: Innovation Information Initiative



Roadmap

- **What are we trying to measure and why**
- Over-arching issues
- Coda: Innovation Information Initiative



Sources/inputs to innovation

- innovation builds on internal and external inputs
  - firm's previous inventions
  - spillovers from other private agents
  - public investments in science and technology
  - tools and instruments
- We use backward citations as indicators of where these inputs come from: geography, organizational boundaries, patents versus scientific articles
- Citations have multiple origins and multiple functions; examiner versus applicant; EPO categories; in text versus front page

Interactions within the innovation system—one interaction at a time

- spillovers and knowledge flows
  - ‘spillovers’ are of particular interest because they are the source of dynamic increasing returns to invention/innovation.
  - knowledge flows are an important source of spillovers, but not all knowledge flows are spillovers nor vice versa
- most patents are associated with teams of inventors: important for the innovation process, and interesting in and of itself

Interactions within the innovation system—network approach

- A *system* is a set of actors that interact in such a way that the performance of the set has attributes or outputs not obviously predictable by aggregation of the attributes or outputs of the individual actors
- need to identify connections (‘edges’ in networkspeak)
- need to characterize the ‘system’ as distinct from its elements

(Meaningful) rate of innovation

- To characterize the innovativeness of a firm, region, etc., need some way to ‘add up’ innovations
- lots of words
  - *value*: private or public
  - *breadth* or *scope*: how much of technology space is covered
  - *novelty*: how big a leap from what came before
  - *impact*: how much is the course of technology evolution changed
  - …
- not just semantics: patent quality is fundamentally a multidimensional concept

(Meaningful) rate of innovation

- ‘value’ has natural units; other concepts less obvious how to quantify even in principle
  - ordinal ranking is not sufficient if the goal is aggregation across portfolio
- metrics may create false concreteness
  - not just a matter of measurement error (e.g. plus/minus epsilon); may have entirely wrong curvature

Policy evaluation/recommendation

- Measure ‘output’ of public programs
  - productivity of patent examiners
  - output of public research support programmes
- Characterize the ‘but for’ world
- Identify possible policy levers (see ‘inputs’, impacts, ‘innovation system’)

Prediction of innovative success

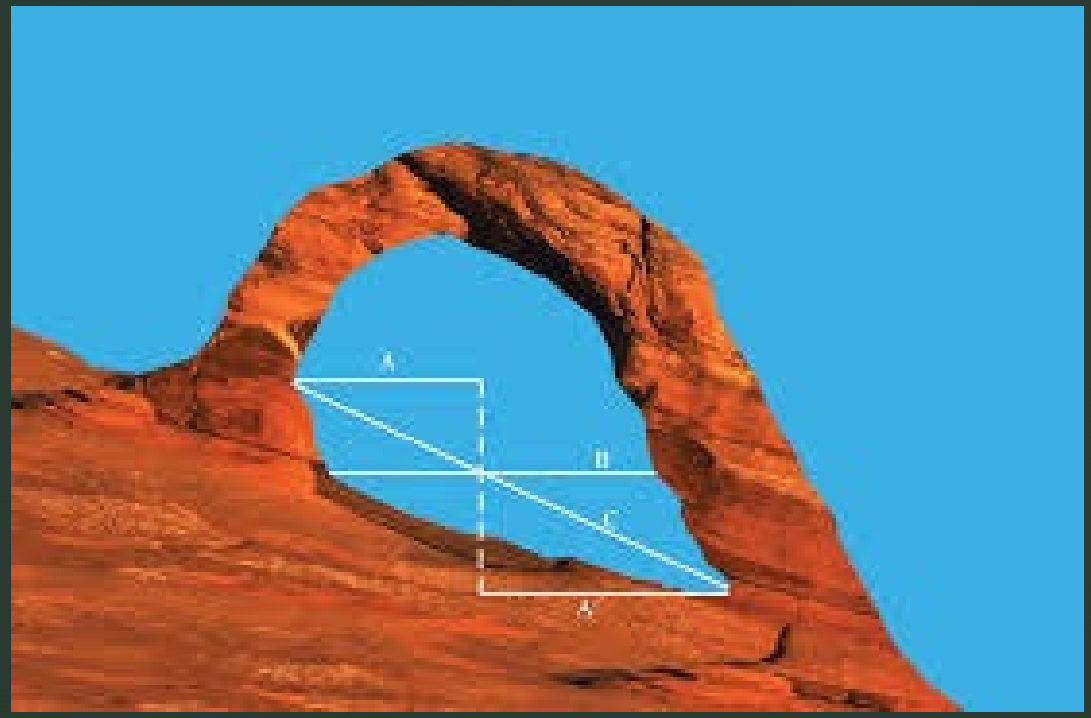
- Most inventions, new products, new firms (etc.) fail.
- High return to being able to predict (even poorly) the likelihood of success for a given new introduction.
- Are there observable attributes at or close to t=0 that are systematically related to subsequent success
- ‘Pure’ prediction (e.g. neural nets) or causal/structural

Impacts of innovation

- Should be defined relative to objectives
  - For firms: profits, growth
  - For scientists and inventors: fame and fortune
  - For society: income and wealth; consumer surplus; health; satisfaction; environment
  - For government: tax dollars; re-election
- Again, some of these have ‘natural’ units, others do not
- Patents are clearly closer to the lamp-post than they are to the most important underlying concepts

Roadmap

- What are we trying to measure and why
- **Over-arching issues**
- Coda: Innovation Information Initiative



Measurement issues versus conceptual issues

- Start with a concept, not an indicator.
- Think about both Type 1 and Type 2 error.
- Think about how agents' behavior affects metrics, and about how their incentives affect their behavior. (Understand the rules governing the data generation process.)
- Think about what level of aggregation is important to you, and how aggregation affects the error and bias
- Think about Hawthorne effect, particular for metrics that might be incorporated into policy or procedure

Ex ante versus ex post

- Metrics reflect ‘state’ of the system and information about that system at the time they are created.
- Ex ante metrics reflect innate characteristics, plus expectations about how those characteristics will likely affect future developments.
- Ex post metrics reflect innate characteristics, developments to date, and expectations about the future.

Agents and spillovers versus networks

- Different ways of asking related questions:
  - How do firm characteristics X, and ‘neighbors’ characteristics Z affect firm performance?
  - How does the performance of a firm that has a certain position in a network differ from other firms within that network?
  - How does the overall performance of a network of firms that has certain network properties differ from another network with different network properties?
- Dual role of citations: indicators of impact and also identifiers of network connections

Dynamics and endogeneity

- If we take spillovers seriously, then neither the arrival of patents nor the arrival of citations are independent events
- Preferential attachment (success appears to breed success) is a pervasive feature of publication/certification/credit systems such as patents and scholarly publications
- Conceptually, observed patterns could be generated by:
  - intrinsic unobserved heterogeneity
  - dynamic increasing returns (initial success makes you truly better)
  - fad/fashion and self-reinforcing popularity

Semantic analysis

- A patent is really a bag of words. On some level, analysis of patents has to be semantic analysis
- Characterization of an invention: breadth or scope; complexity
- Characterization of the patent document: transparency or obfuscation of underlying invention art
- Relationships among patents:
  - technological similarity (pairwise)
  - uniqueness (lack of similarity to contemporaries)
  - novelty (lack of similarity to predecessors)
  - impact (similarity to successors)

Validation

- What does it mean to ‘validate’ a metric?
- Consistent with empirical expectations *given maintained hypotheses*
- Consistent with other metrics
- Consistent with expert opinion (really just another metric)
- Capable of out-of-sample ‘prediction’
- Capable of prediction across time
- Again face issue of concepts that are fundamentally hard to quantify

Coda: Innovation Information Initiative

- **New Sloan-Foundation-funded collaborative focused on creating and sharing innovation metrics based on patent data**
- **initial institutional participants: Yarn Labs at MIT, Lens.org, NBER, BU and EPFL**

Coda: Innovation Information Initiative

- **NBER will host:**
  - **Technical Working Group will bring together anyone interested working on such metrics to discuss collaboration, best practices, validation, inter-operability, and priorities for new work**
  - **Annual research conference at Summer Institute**

Coda: Innovation Information Initiative

- First Technical Working Group meeting is Dec 6-7, 2019 in Cambridge, USA
- Register your interest in attending at:

  http://conference.nber.org/confer/reply/IIIf19.html

- Register your interest in making some kind of presentation at this meeting at:

  http://papers.nber.org/confsubmit/backend/cfapp?id=IIIf19