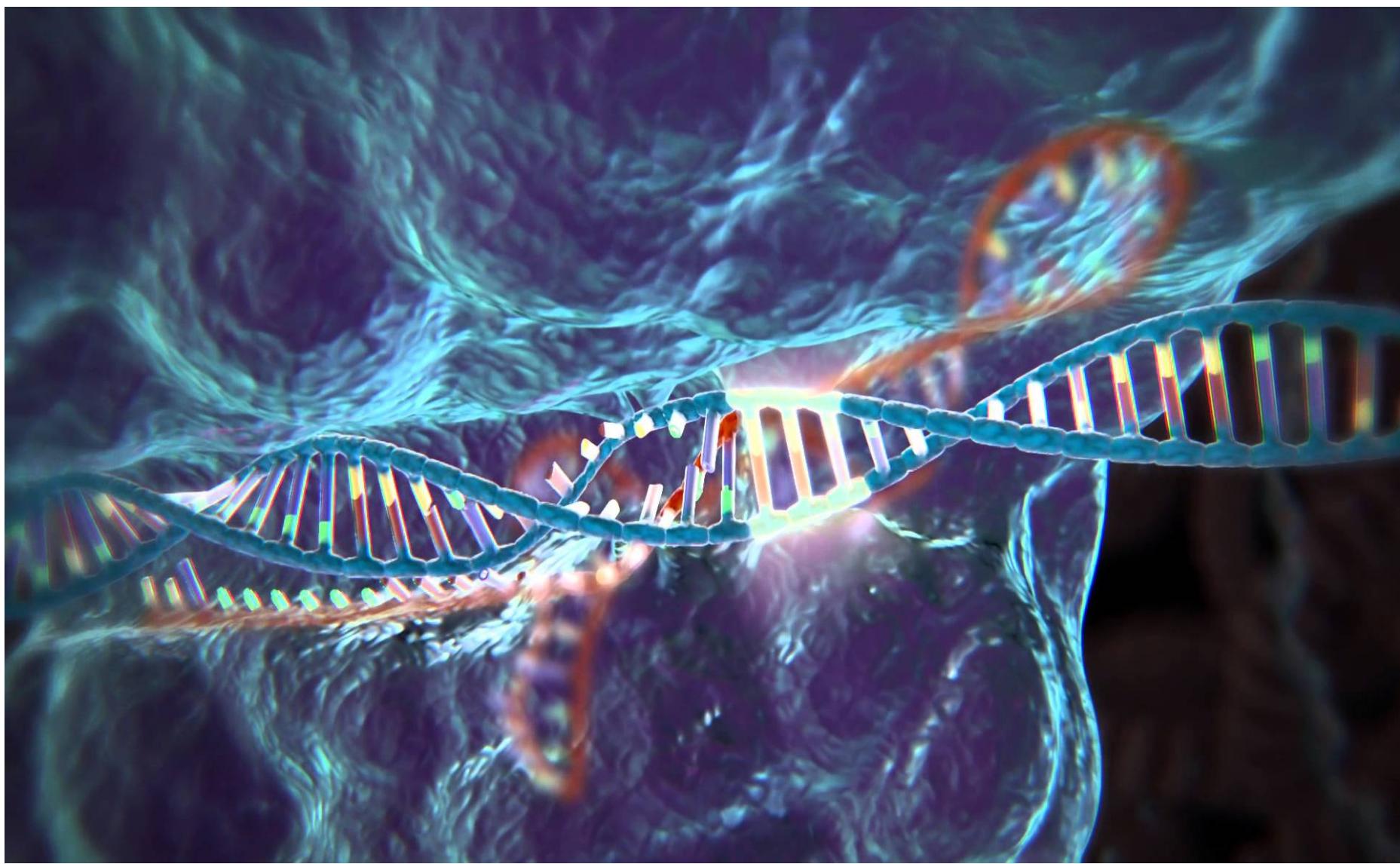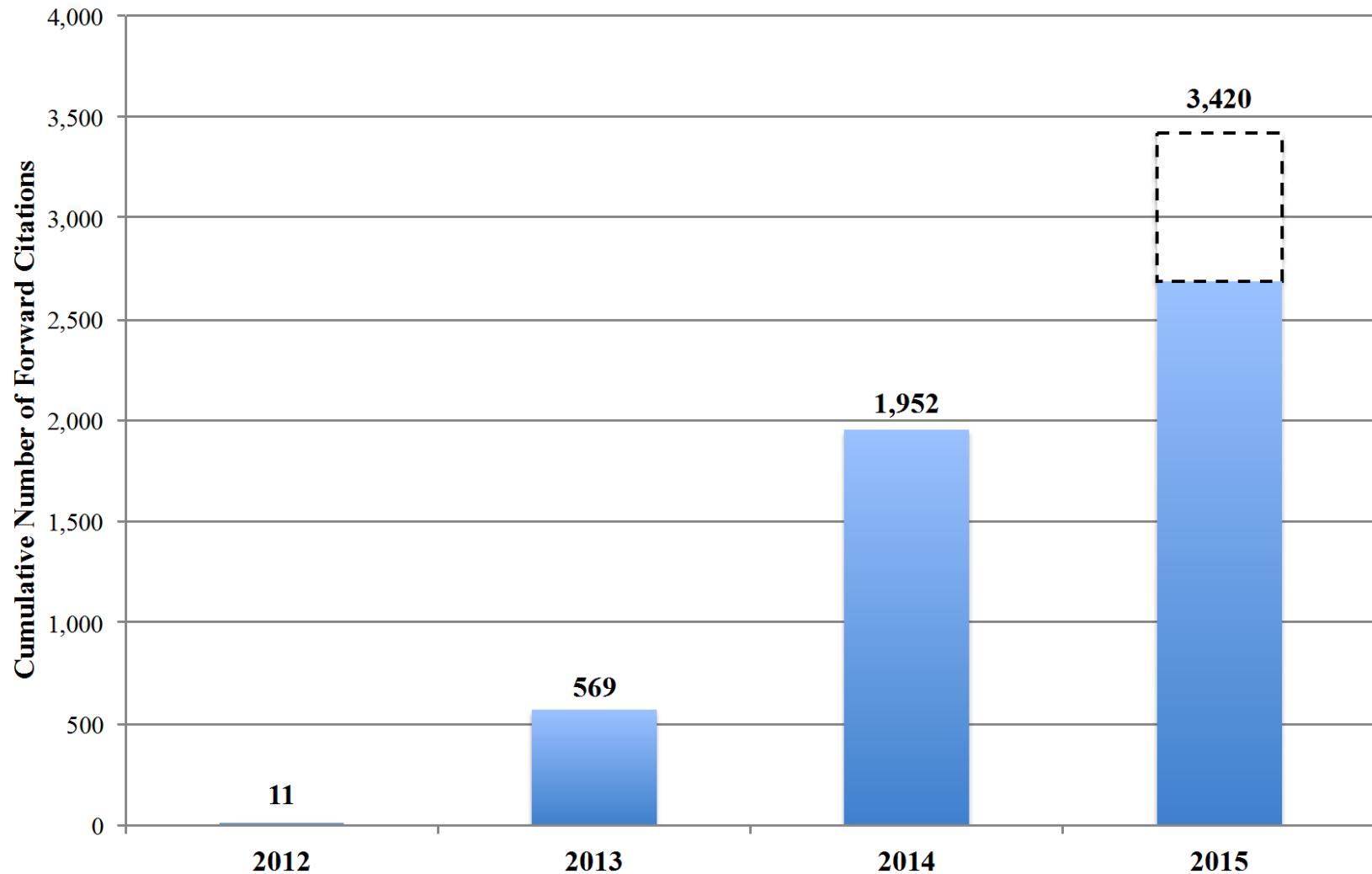# Innovation Measurement of the Future: An Agenda

Scott Stern, MIT and NBER

May 2016
NCSES/CNSTAT Workshop on
Advancing Concepts and Models of Innovative Activity and STI Indicator Systems

What is CRISPR? Drug Developers Harness The New Genome Editing Tool; But Can It Deliver Novel Drugs?

AstraZeneca announces four collaborations to use CRISPR technology for genome editing across its drug discovery platform

Some Characteristics of Innovation

- Cumulative

- Uncertain and Highly Skewed

- Distributed Impact and Application (beyond traditional STI Indicators)

How do these fundamental characteristics of innovation get reflected in our innovation measurement framework?

Are we able to move beyond examples?
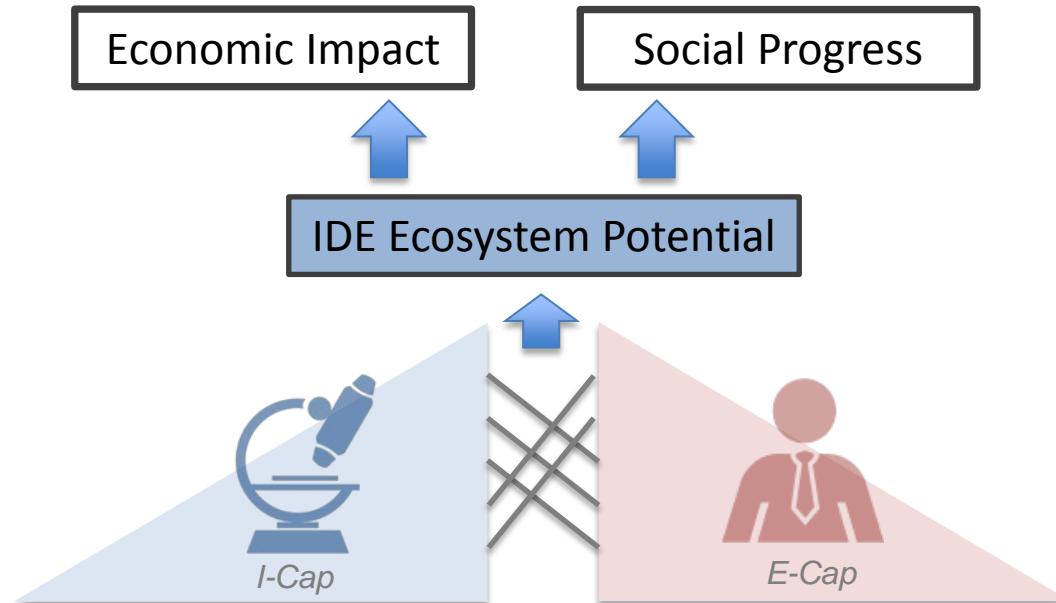
Standing on the Shoulders of Giants

- Long-term economic growth depends on the ability to draw upon an ever-wider body of scientific & technical knowledge (Rosenberg, Mokyr, Romer, Aghion & Howitt, David & Dasgupta)

- Ideas-based growth ultimately depends on increasing returns in the innovation production function

$$\dot{A} = \delta A^{\phi} L_A^{\lambda} \implies g_A = \delta A^{\phi-1} L_A^{\lambda}$$

- But, how do we know that innovation is cumulative, and how and whether the degree of cumulativeness varies by time and place?
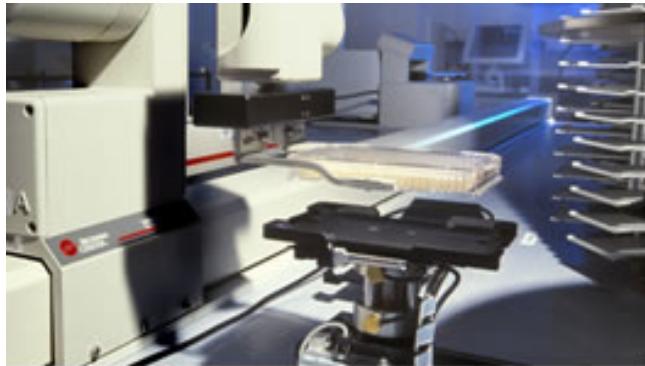
IDE Ecosystem

Economic Impact   Social Progress

IDE Ecosystem Potential

I-Cap   E-Cap

Cluster Based Comparative Advantage

Foundational Institutions

MIT REAP
REGIONAL ENTREPRENEURSHIP
ACCELERATION PROGRAM

How does we measure the cumulativeness of the innovation system?

The Impact of Biological Resource Centers (with J. Furman), *AER*
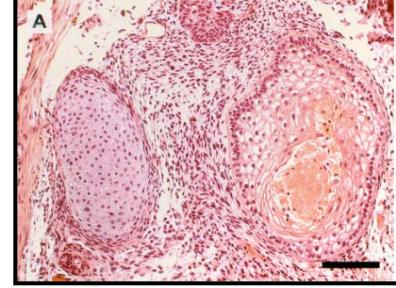
Empirical Framework:
Diffs-in-diffs analysis of citations received

Pre-period institutional setting

Post-period institutional setting

Publication

Publication

Treated

Publication

FC_{jt}

FC_{jt}

FC_{jt}

FC_{jt}

FC_{jt}

FC_{jt}

Publication

Control

FC_{jt}

FC_{jt}

FC_{jt}

FC_{jt}

FC_{jt}

Exogenous SHIFT
Measure citations before & after to estimate impact of treatment on treated "diffs-in-diffs" approach

Plot forward citations over time as a measure of scientific knowledge accumulation building on a "piece of knowledge"

How does the rate of citation of a scientific article change after the materials association with that article have been deposited in a culture collection?

Impact of Deposit Grows Over Time and Does Not Exist Prior to Deposit


Furman and Stern, 2011
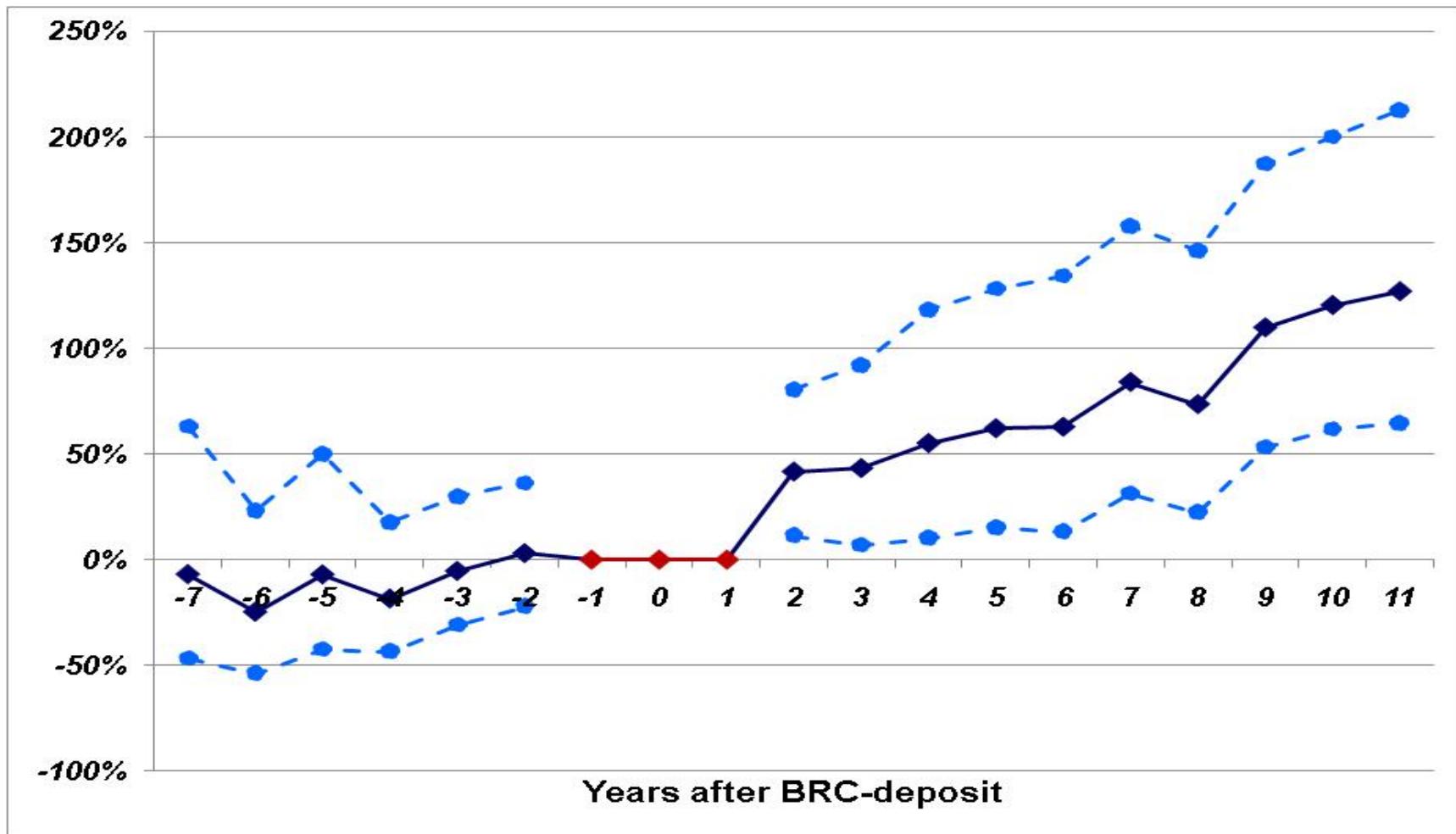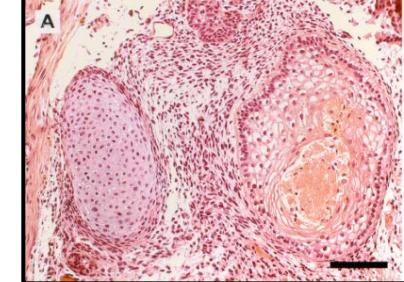
But, how can we aggregate our study of cumulativeness (and the role of institutions and policy in shaping cumulativeness and outcomes) beyond the level of individual "pieces" of knowledge?

MAR DE SALA, vel de
Bachu, olim Caspium.

NATOLIAE PARS

AR
MENIA

ANADOLE
PEGIAN
GENECH
Curdi

DIARBECH
Mesopotamia

Caldaran
campus

SERVAN olim
MEDIA

DIARGVMENT
olim Hyrcania

GILAN

TARABAT

ISTIGIAS

BVCCARA

MESANi
DARAN

CASM-IN

Cyprus

SARCIi olim
Assyria

A S I AE

MARIS MEDI:
TERRANEI PARS

SYRIA

ARAC
PARTHIA

CAPACOP
SERNERI

CANDA
HAR
RASAN

ARDEN

CVSISTAN olim
Susiana

PERSIA
olim
IXD

Sigistam

FLAPRAN

FARSI

CIRCA

MOGORES

Chesimur

AE
GYP
TVS

ARABIA
DESERTA

ORMVE

DEL

INDIA

INDIA
EXTRA GANGEM

Sinus
Persicus

GVZARATE

SANGA

INTRA

GANGEM

BENGALA

Ayam·an

Cam
ba

INDOSTAN

TROPICVS CANCRI

ARABIA
FELIX

MAFFA

ALBINALI

MARE

DELL

ORIXA

GOLFO
DE
BENGALA

IRMIN

AMANSIRIFDIN

ARABICVM

VA

HERIT

ZIBIT

ET

SINVS

INDICVM

GANGE
TICVS

AFR
DAFRILA

BARNAGASSO

AE

LACCA

Andemaon

DANGALI

DOBAS

DOARA

I.Caremeubar

I.Rasa

Ide Nicubar

ANGO
TE

MAGADOXO

PARS

Gomespola

Maldiuæ
Insulæ

Su
matra

Henricus F.ab Langren
Sculpsit.

ÆQVINOCTIALIS

# Intellectual Property Rights and Innovation: Evidence from the Human Genome (Heidi Williams, MIT)

- During the final years of the HGP, competition between HGP and Celera, with *temporary* licensing rights for Celera sequences occuring prior to HGP coverage
  - Only lasted 2 years at most
- Williams examines whether follow-on research on individual genes in the post-HGP era were impacted by Celera IPR claims
- Results suggest an ~30% reduction in subsequent publications, phenotype-genotype linkages, and diagnostic tests for genes first sequenced by Celera
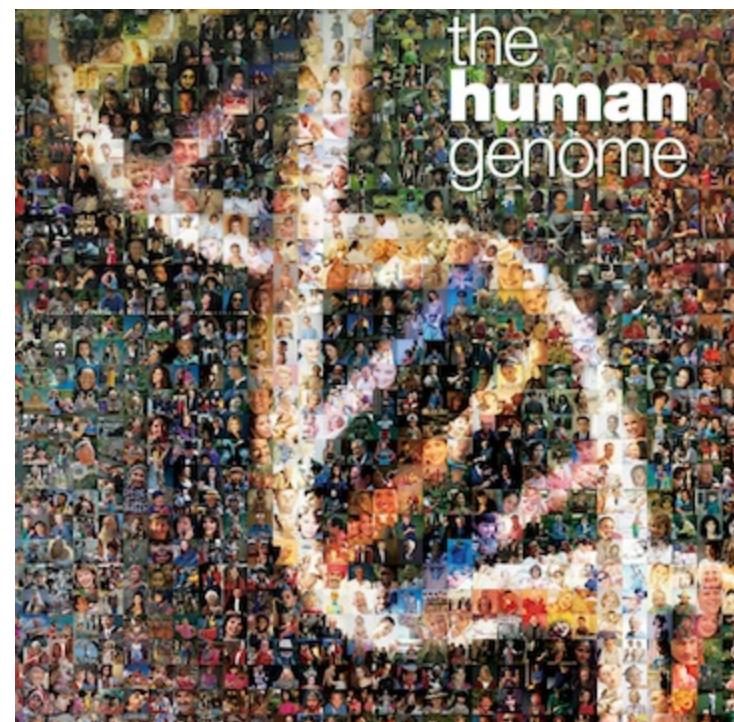
Some Characteristics of Innovation

- Cumulative

- Uncertain and Highly Skewed

- Distributed Impact and Application (beyond traditional STI Indicators)

Underlying Structure of Uncertain and Highly Skewed S&T and E&I Outcomes
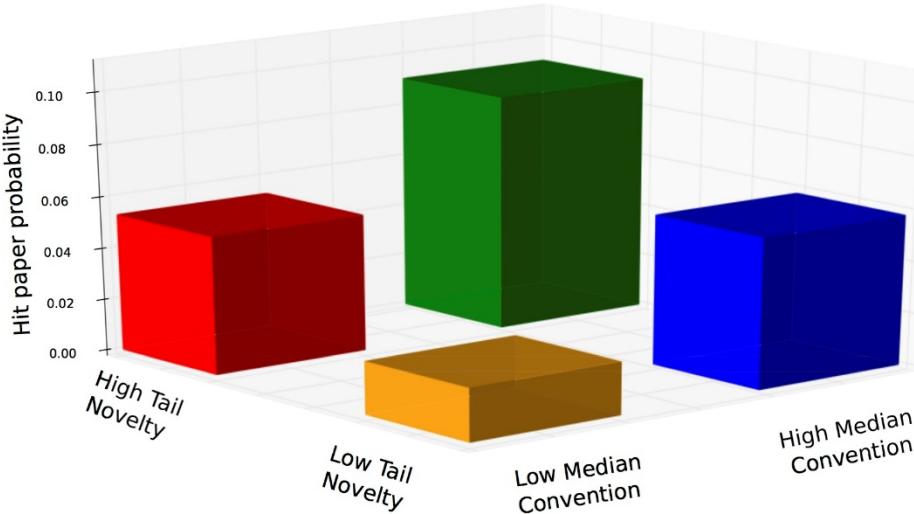

Structure to the Skew of Entrepreneurial Quality (Fazio, Guzman, et al, 2013)

Structure to the Skew of Cumulative Knowledge (Jones, Uzzi, 2013)

10-Fold Test of Predictive Quality of Model*
Top 1% includes 51% of growth outcomes (range: [49%, 53%])
Top 5% includes 69% of growth outcomes (range: [65%, 72%])
Top 10% includes 75% of growth outcomes (range: [70%, 79%])

*10-Fold analysis of model separates the model into 10 random samples and then uses each of those sample as a test sample. We report the average value as well as minimum and maximum (range) of such.
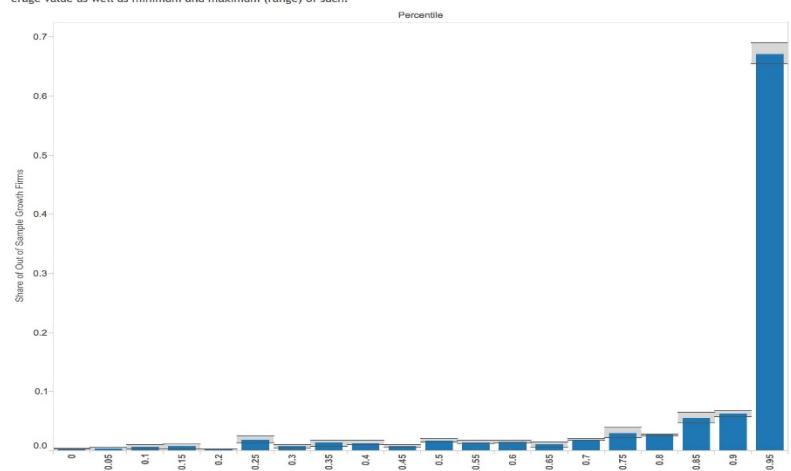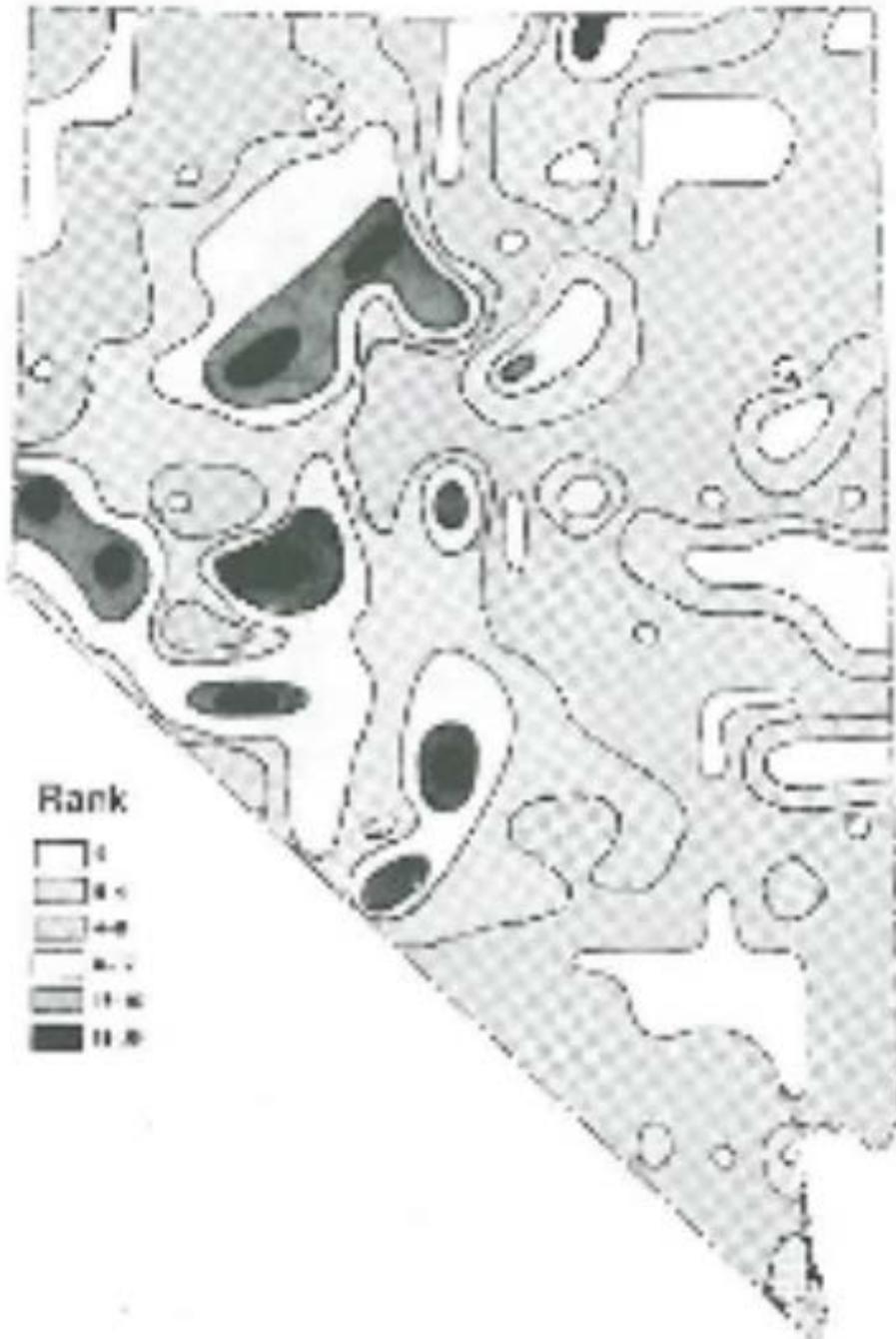
Maps and Innovation: Evidence from the Gold Exploration Industry, Abhishek Nagaraj, MIT Sloan, 2016
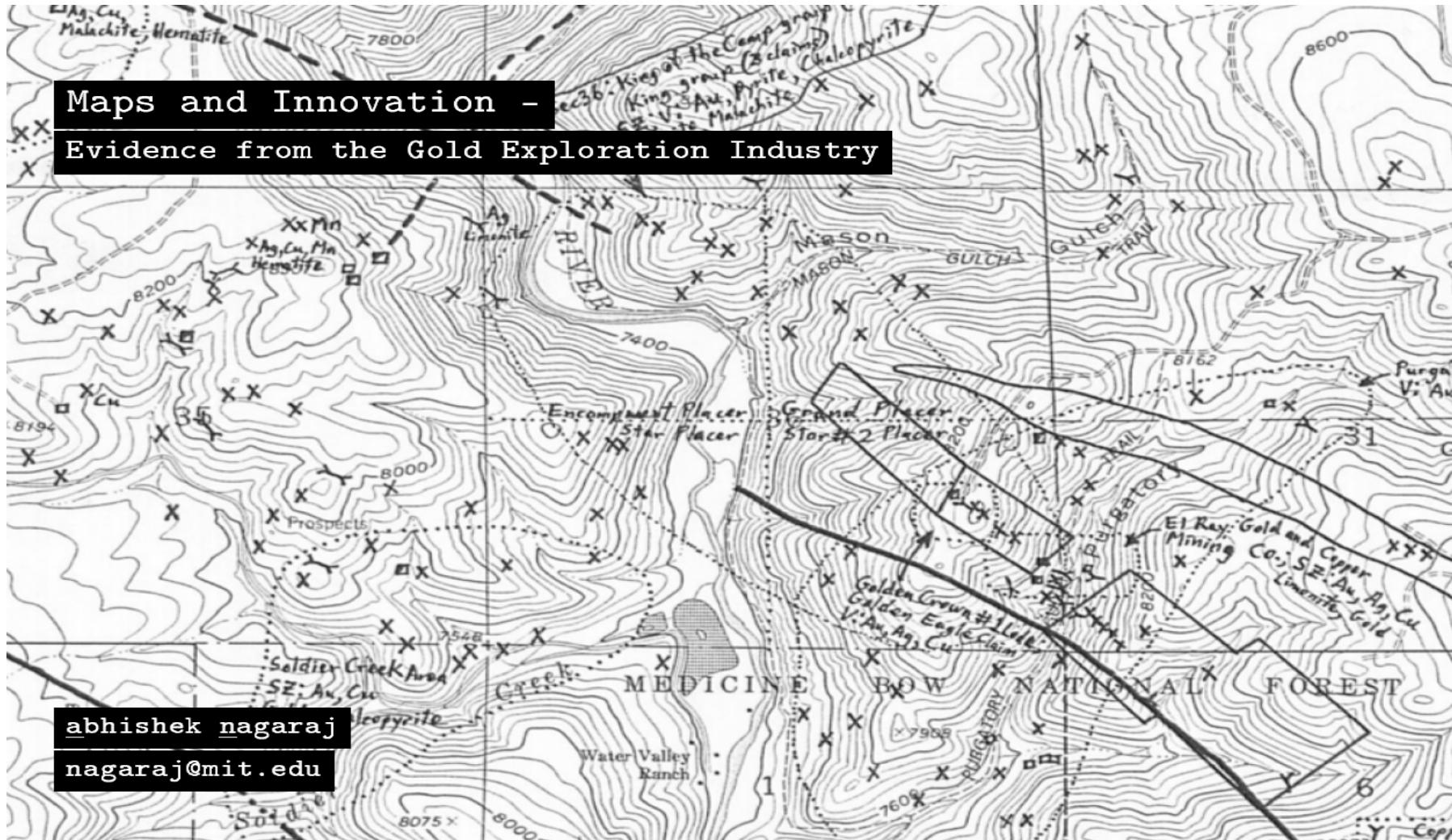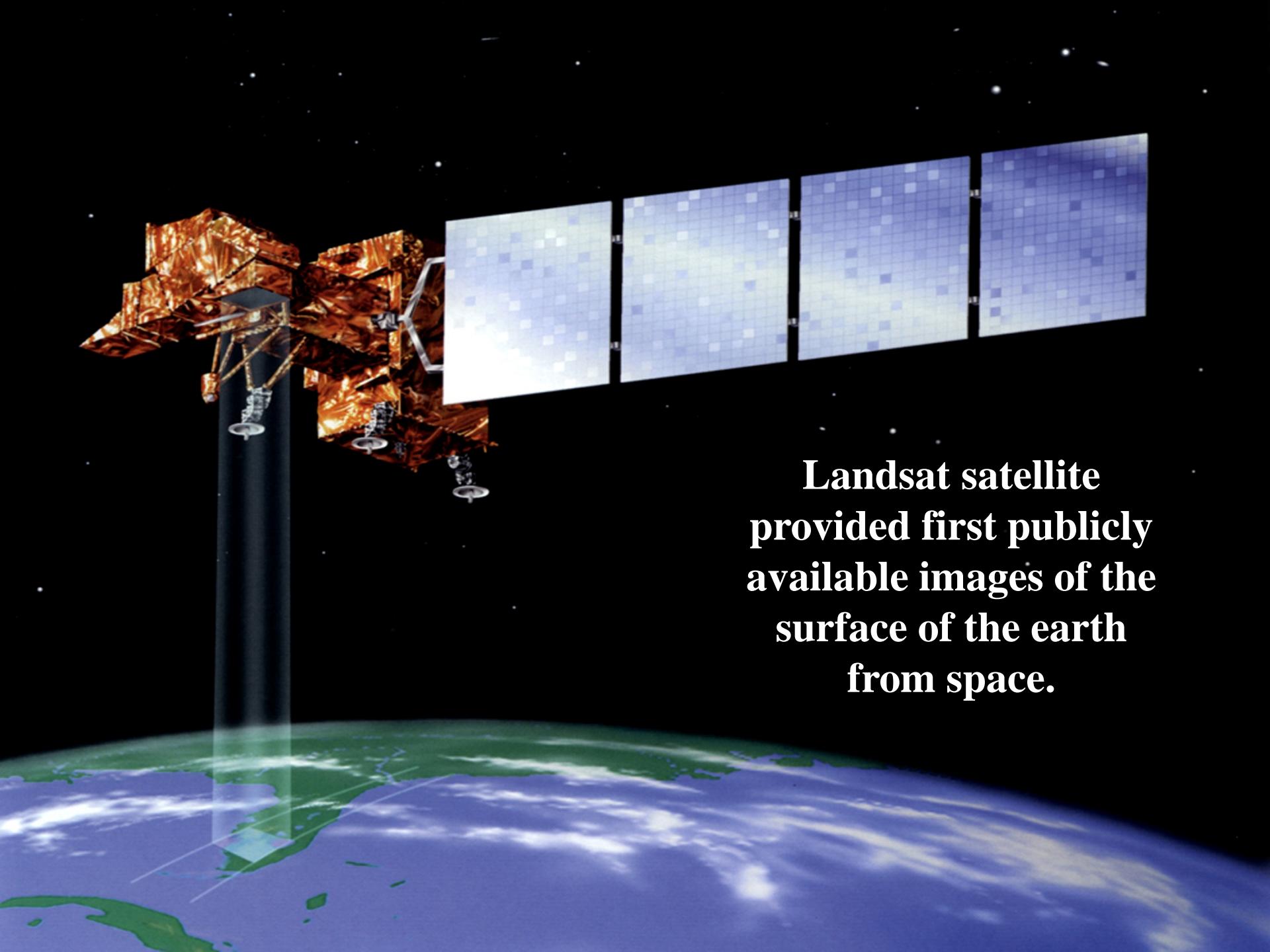
Landsat satellite provided first publicly available images of the surface of the earth from space.

It is possible to use the information from these images to develop a "heatmap" for potential new gold discoveries (and other resources)
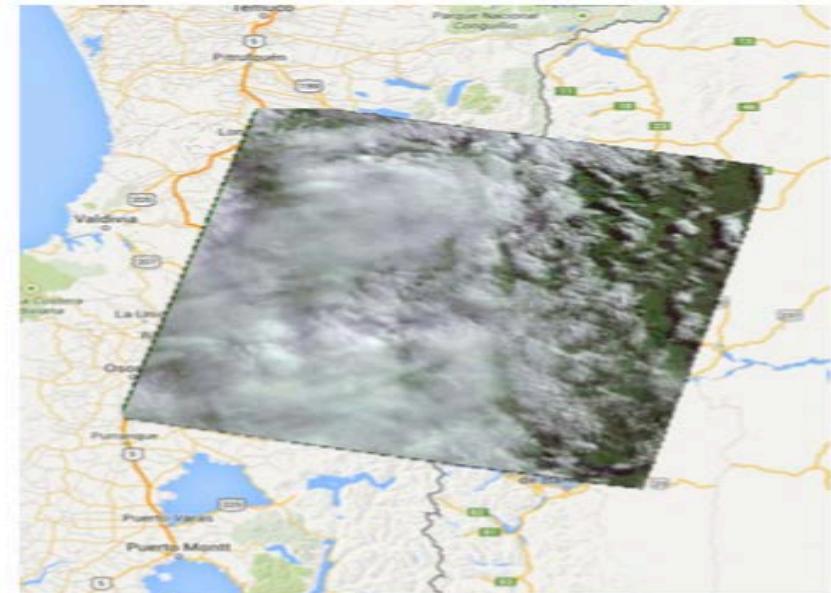
Rank

How did the availability of "open" maps impact discovery and entrepreneurship in the gold industry?

- Nagaraj takes advantage of the fact that the timing of a "clear" image from the Landsat program had a large random element
  - Wide variation in the date at which photos were taken
  - Clear images depended on "no cloud cover" images

(1) Block 25177, Chile
Cloud free imagery available
Amax Gold Discovery reported in 1980

(2) Block 24988, Chile
No cloud free imagery available by 1983
No discovery reported till date

How did the availability of "open" maps impact discovery and entrepreneurship in the gold industry?

- Nagaraj takes advantage of the fact that the timing of a "clear" image from the Landsat program had a large random element
  - Wide variation in the date at which photos were taken
  - Clear images depended on "no cloud cover" images

- Nagaraj then compares the rate of gold discovery and mining from locations with access to "open" public images with those where a public image is not yet available

- Explores not only the impact on gold discovery and investment, but on whether these new opportunities are taken advantage of by entrepreneurs (speculators) or established firms (vertically integrated mining firms)
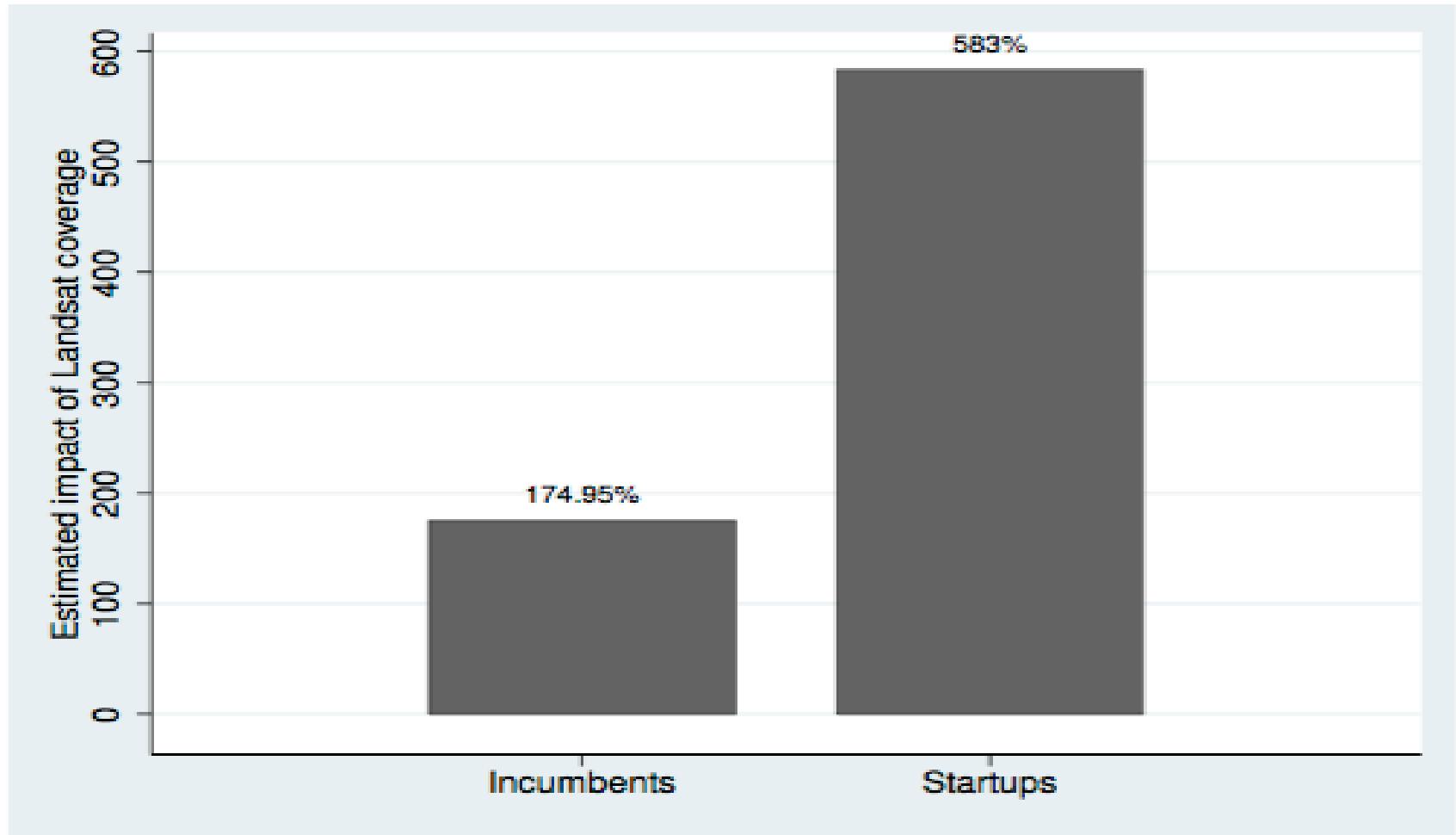
There is a large and persistent difference in the rate of gold discovery depending on the availability of an open-access map

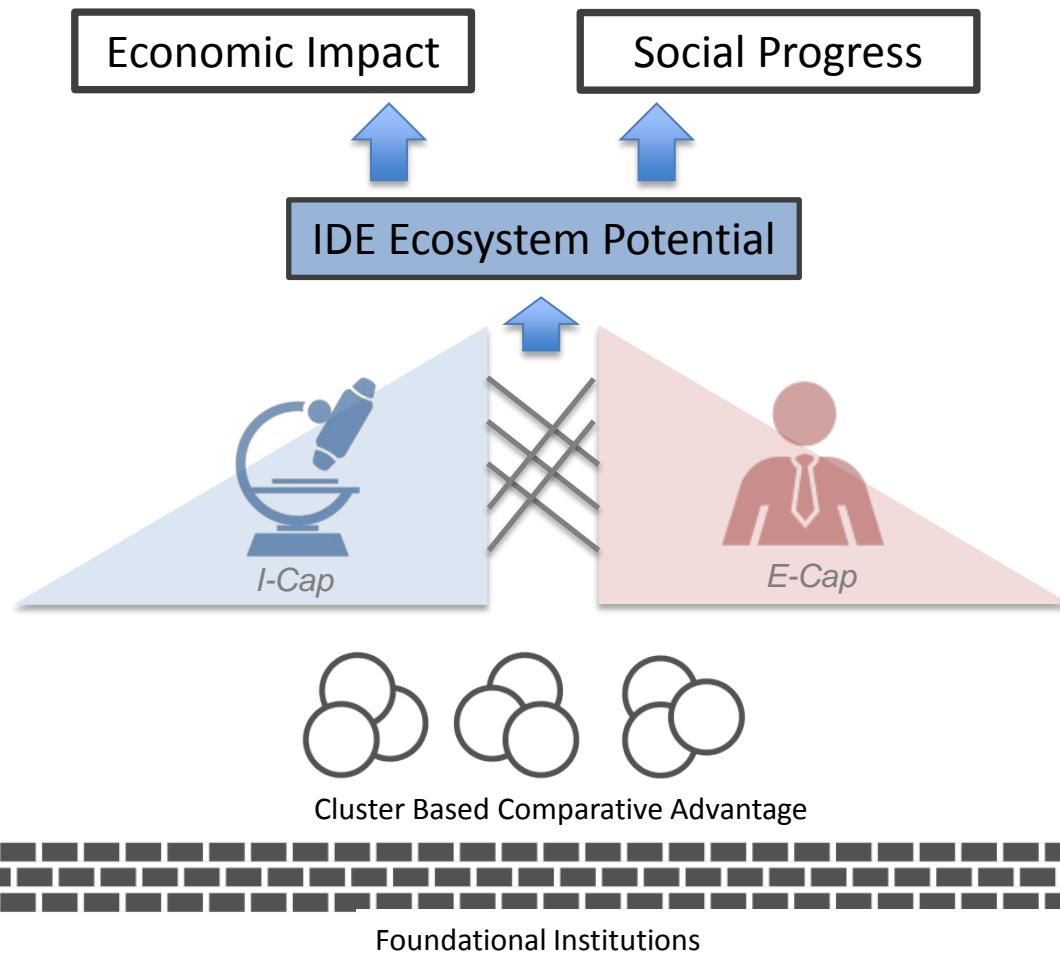Years Since Recieving Low Cloud Image
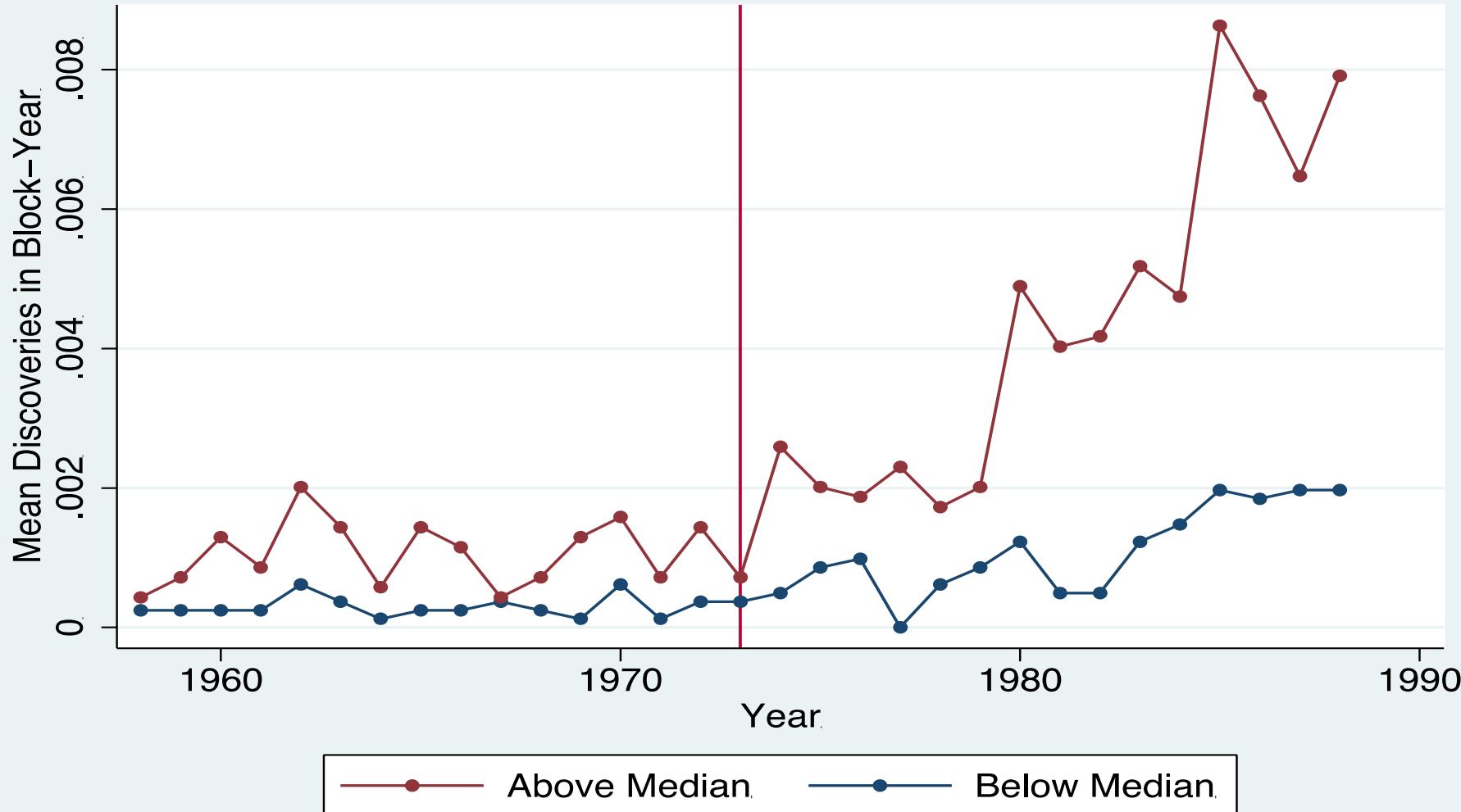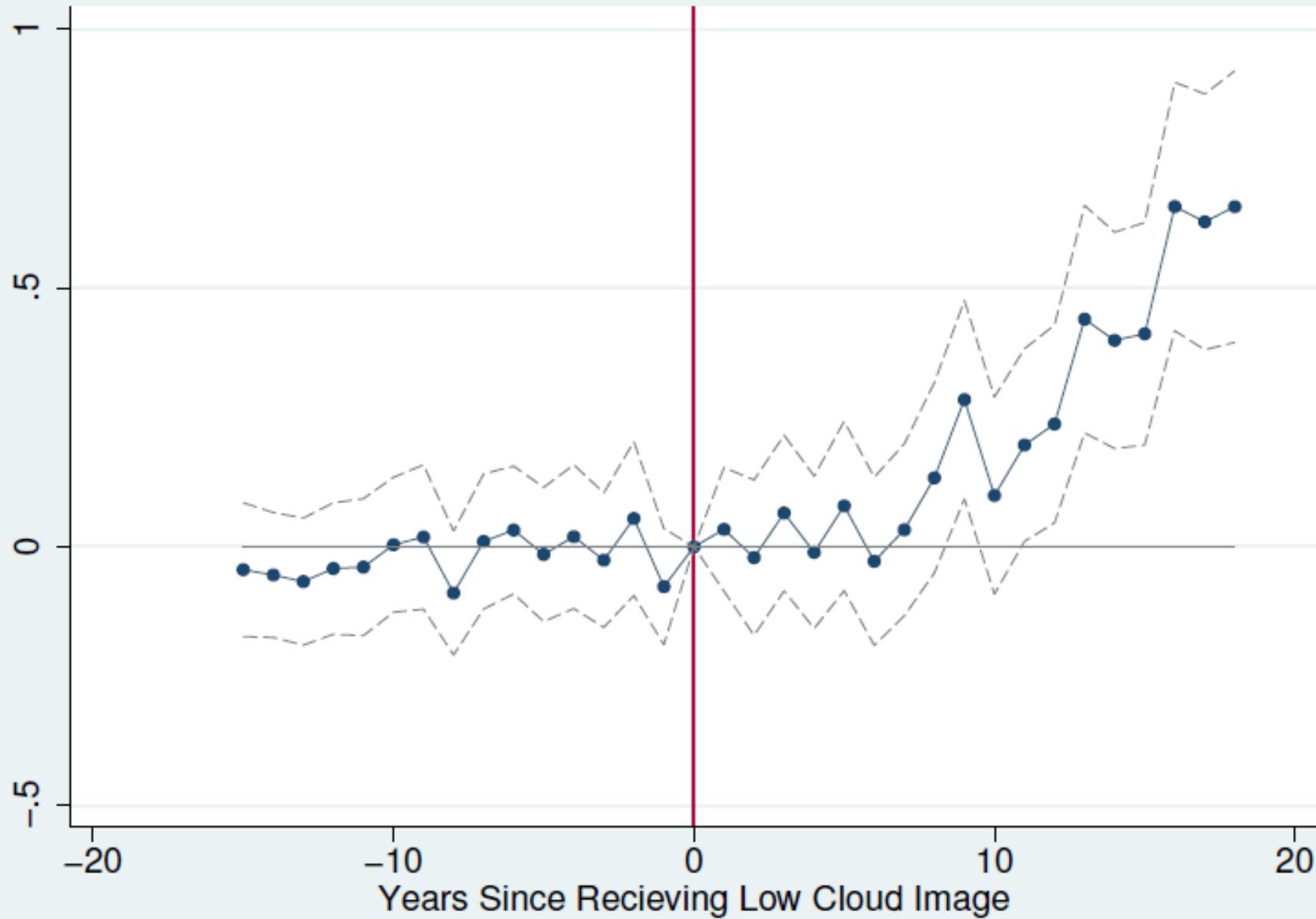
Entrepreneurs are Far More Likely to Take Advantage of Open Access Maps than Established Firms

But how do we aggregate this to the level of a regional or national innovation system?

Economic Impact

Social Progress

IDE Ecosystem Potential

I-Cap

E-Cap

Cluster Based Comparative Advantage

Foundational Institutions

Lessons for Policymakers

- Innovation statistics and metrics are increasingly being used to evaluate and track innovation systems (at multiple levels of granularity).

  - Need to develop a meaningful connection (often between existing data sources) that allows cohesive assessment of the role of different elements of the innovation system (inputs) and different outputs of the innovation system over time

  - Recognizing Cumulativeness, Uncertain and Highly Skewed, and Distributed Impact

Lessons for Policymakers

- Innovation statistics and metrics are increasingly being used to evaluate and track innovation systems (at multiple levels of granularity).

  - Need to develop a meaningful connection (often between existing data sources) that allows cohesive assessment of the role of different elements of the innovation system (inputs) and different outputs of the innovation system over time

  - Recognizing Cumulativeness, Uncertain and Highly Skewed, and Distributed Impact

- Particularly important for areas of innovation beyond traditional "tech to market" applications of science, particularly as it relates to emerging uses of digital knowledge, maps, and non-"science" knowledge systems

How Likely Is It that the Innovation System Metrics we Developed in the Wake of Vannevar Bush's The Endless frontier Happen to be "Just Right" for the 21st Century?

Thank you!

sstern@mit.edu