Measuring Science Intensity of Industry using Linked Dataset of Science, Technology and Industry

Kenta Ikeuchi (RIETI)
Kazuyuki Motohashi (University of Tokyo)
Ryuichi Tamura (Wakkanai Hokusei Gakuen University)
Naotoshi Tsukada (NISTEP)

CEPR – RIETI Joint Workshop, 26 February 2018, NIESR, London

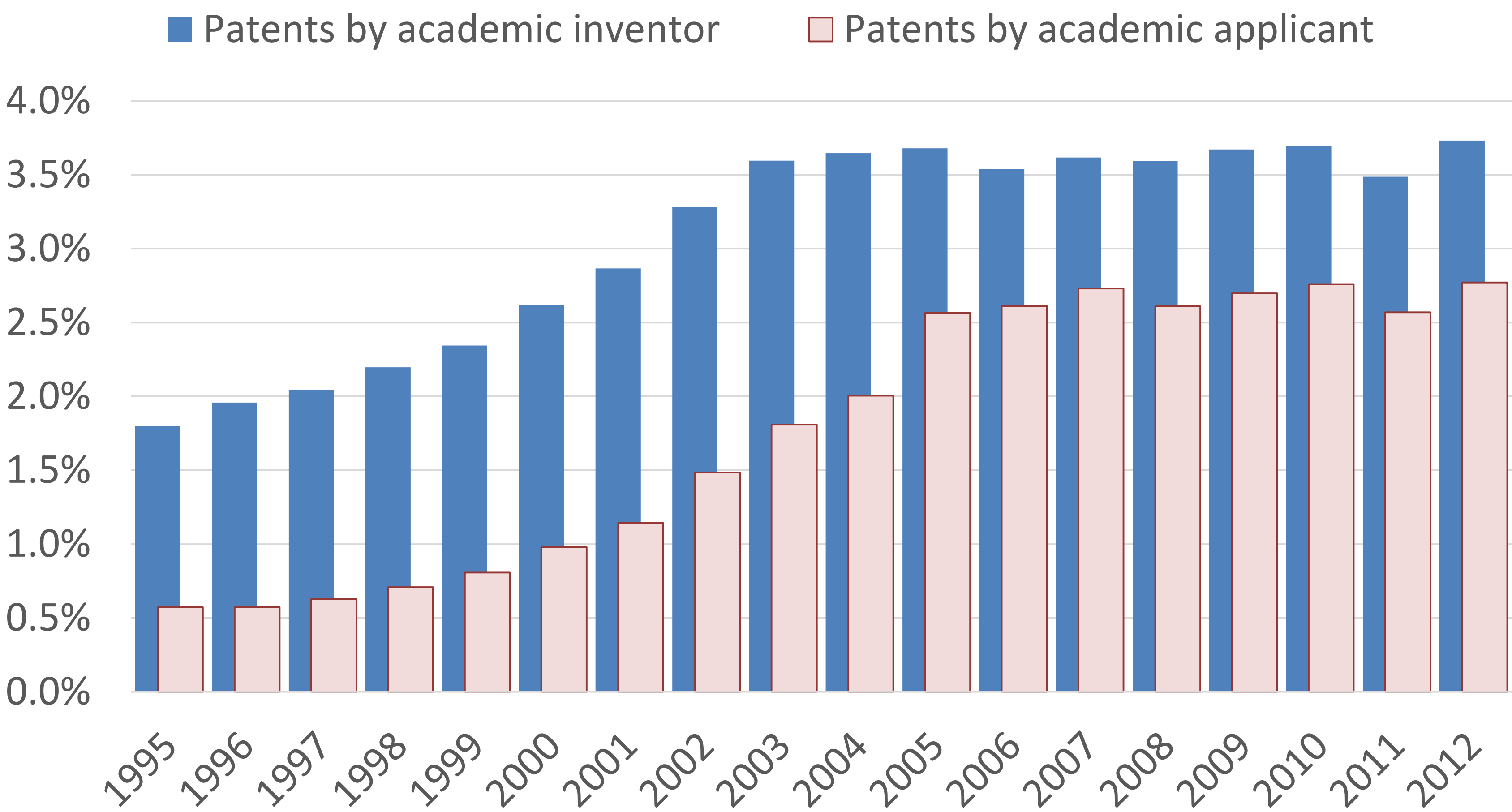
Introduction

- Scientific foundation becomes important for industrial innovation process.
  - E.g., the genome science for pharmaceutical industry
- Growing policy interests in the economic impact of pubic R&D.
  - Science sectors heavily relied on public money while severe constraint in public spending.
- Existing indicators of the contributions of science to economy.
  - Non-patent literature (NPL) citation in patents (Narin and Noma, 1985; Schmoch, 1997).
  - Patent-publication pairs: an identical content of research output/invention found in both patent and research paper.
    - Simultaneous disclosures both in a patent and a research paper (Lissoni et al., 2013) or text mining technique to measure the similarity between the contents (Magerman et al., 2015)
- We propose more comprehensive indicators of science and industry linkages, based on novel dataset of linking science, technology and industry,
  - Innovation Process Database (IPDB).

Policy Backgrounds in Japan

- S&T policy and Innovation policy converging, surge of academic patents after incorporation of national universities
- Limitation of traditional science linkage indicator, NPL citation of patents.

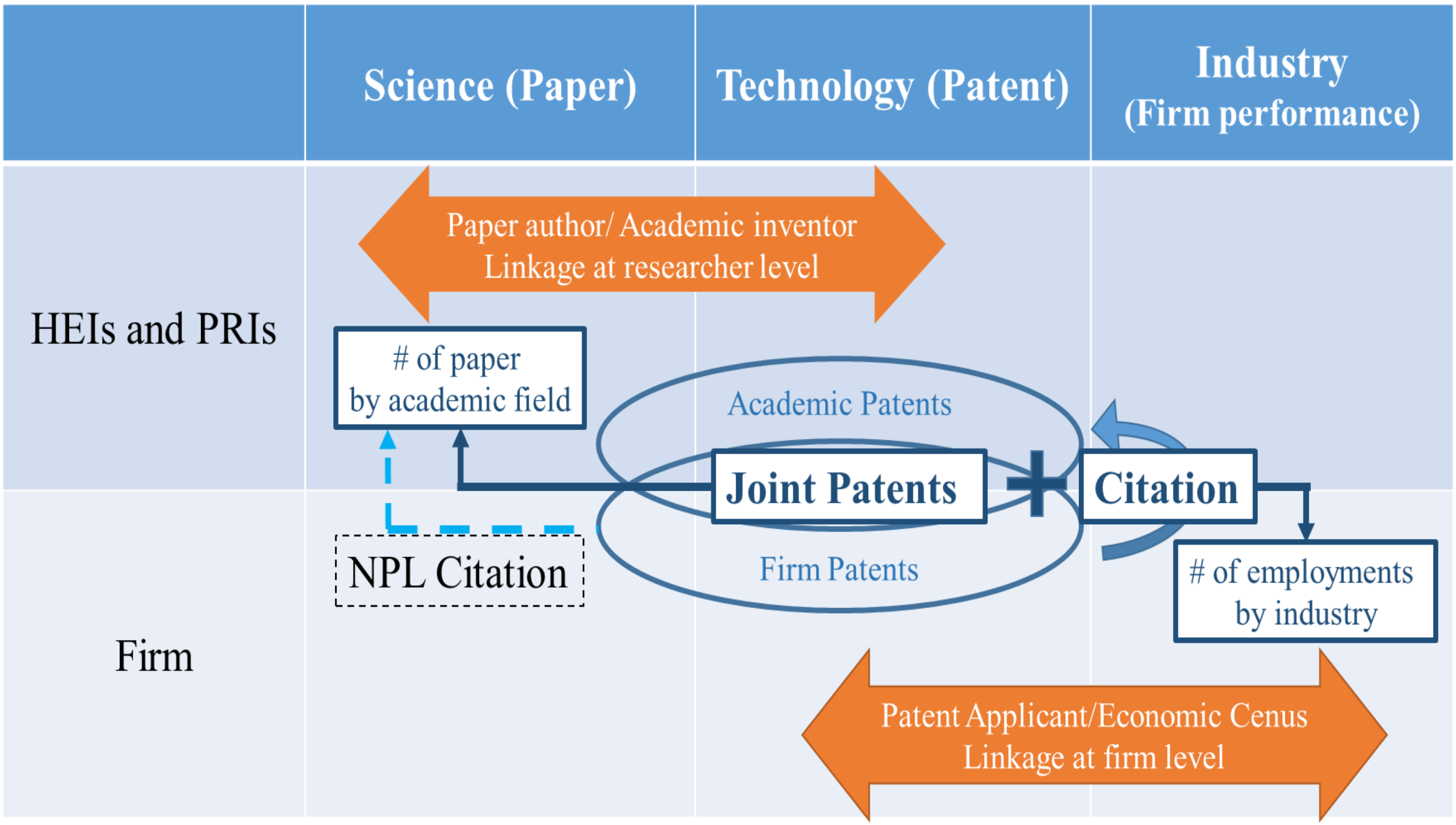
Academic patents in Japan



Policy factors

- Centralization of strategy and macro control
  - 1995: Science and Technology Basic Law : 1st Science and Technology Basic Plan (1996-2000). Now, 5th Plan (2016-2020) just started.
  - 2001: CSTP (Committee of Science Technology Policy) and changed its name to CSTI (STP-ST and Innovation Policy)
  - 2001: Ministry of education and S&T Agency merged into MEXT
- Decentralization of strategy implementation
  - 2001: Incorporation of national research laboratories (PRIs)
  - 2004: Incorporation of national universities
- Consistent Supports for U-I collaborations
  - 1998: TLO promotion law
  - 1999: Japanese Bayh-Dole Act
  - 2001: 1000 university spin-outs plan
  - 2005- : Various supports to university IP offices

Concept of indicators



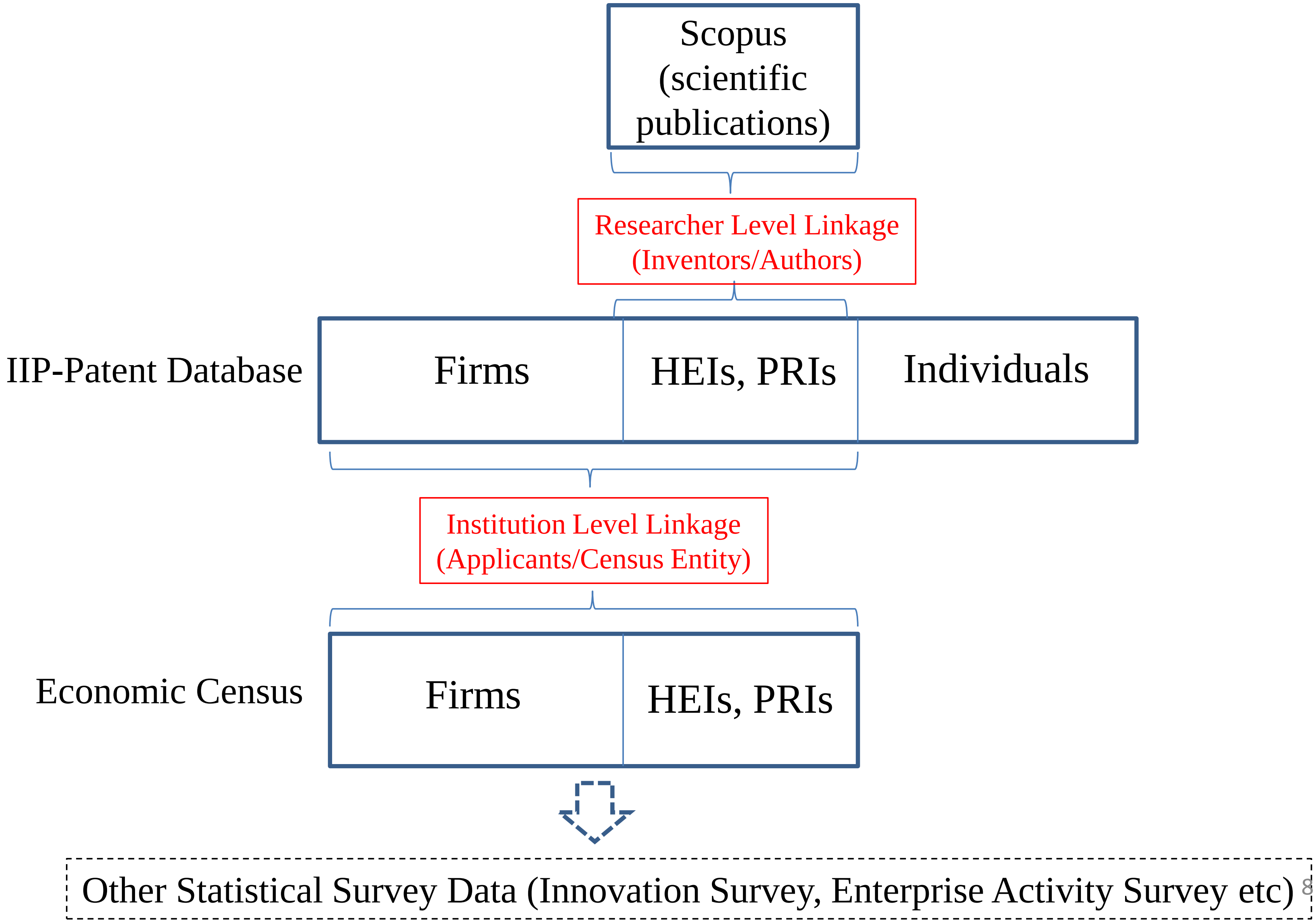
Innovation Process Database (IPDB) at a glance

- All disambiguated Japanese inventor records (1995-2013)
  - Identified Academic Inventors (HEIs and PRIs, 53K) + Industry Inventors (1.23M) in 2000-2011 JPO application patents.
- Linked with SCOPUS author data for academic inventors（26K linked),
- Linking applicants with Economic Census of Japan at firm level.

- Enable us to investigate the whole process of scientific knowledge (research papers), technological outputs (patents) and industrial performance (Econ Census).
  - Covers all technology fields and all industries.

- New channels of scientific knowledge flow from academia to industry:
  - Science knowledge embodied in academic inventors.
  - Citations from industry patent to academic patents.
  - Joint patent inventions with academia.

Structure of Innovation Process Database



Database Construction

IIP-PD

- Inventor Part (1995-2013 application)
  - ➢ Disambiguation of inventor (Li et. al, 2014)
    - using information of name, address, applicant, IPC4, co-inventor
    - Telephone directory rare name as training set
    - KAKEN info as correct reference info to decide threshold value
  - ➢ English inventor name (PATSTAT link) and inventor affiliate (by single applicant patent)

- Applicant Part (all IIP patent)
  - ➢ Standardized applicant name
  - ➢ Standardized address

NISTEP Institution ID

Scopus

- Author ID information (1999-2012 publication)
  - ➢ Author name
  - ➢ Author affiliate

Match by name and location (using establishment level info)

Economic Census

- Firm and institution ID information (2001, 2004, 2006, 2009 and 2012)
  - ➢ Name
  - ➢ Address

Patent level analysis of the effects of science linkages on performance

- Sample
  - Patents applied to JPO by Japanese corporate enterprises during 2001-2010.
- Dependent variables (quality of patent)
  - Number of non-self forward citations [Poisson model]
- Independent variables (type of science linkage of patent)
  - Citing NPL (non-patent literature) (dummy)
  - Citing academic patent (dummy)
  - Joint invention with academia (dummy)
- Control variables
  - Number of inventors, and citing patents (backward citations).
  - IPC (3 digit) and application year (dummies), IPC-specific trends

Science linkage and patent quality

Dependent variables	# of non-self forward citations				
Independent variables	[1]	[2]	[3]	[4]	[5]
Ln. # of inventors	0.156***	0.143***	0.140***	0.140***	0.141***
	[0.00117]	[0.00117]	[0.00118]	[0.00119]	[0.00119]
Ln. # of citing patents	0.258***	0.257***	0.248***	0.248***	0.247***
	[0.00117]	[0.00117]	[0.00118]	[0.00118]	[0.00118]
Citing NPL		0.241***	0.228***	0.227***	0.228***
		[0.00212]	[0.00214]	[0.00215]	[0.00215]
Citing academic patent			0.143***	0.143***	0.146***
			[0.00250]	[0.00254]	[0.00254]
Joint invention with academia				0.00468	0.0432***
				[0.00407]	[0.00438]
Joint application with academia					-0.172***
					[0.00780]
Constant	-33.01**	-30.35**	-29.41**	-29.41**	-30.02**
	[10.06]	[10.05]	[10.05]	[10.05]	[10.05]
Application year dummies	Yes	Yes	Yes	Yes	Yes
3 digit IPC dummies	Yes	Yes	Yes	Yes	Yes
3 digit IPC specific trend	Yes	Yes	Yes	Yes	Yes
N	1,823,703	1,823,703	1,823,703	1,823,703	1,823,703

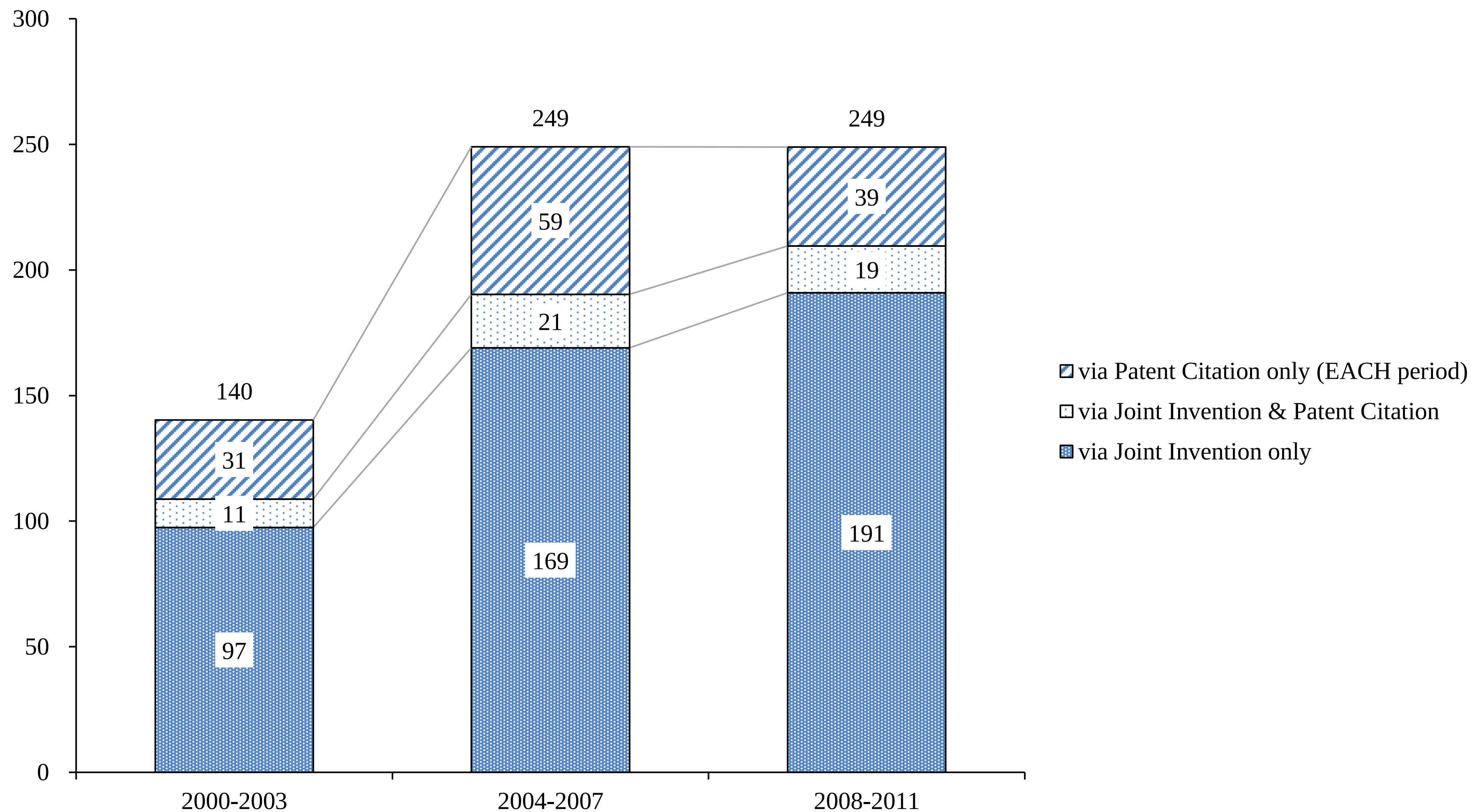
New indicators of science-industry linkage

- We propose new indicators for science industry linkage, by using the interactions of industry and academic in patenting activities:
  - Joint inventive activities (joint patent invention).
  - Firm’s patent citations to academic patents.

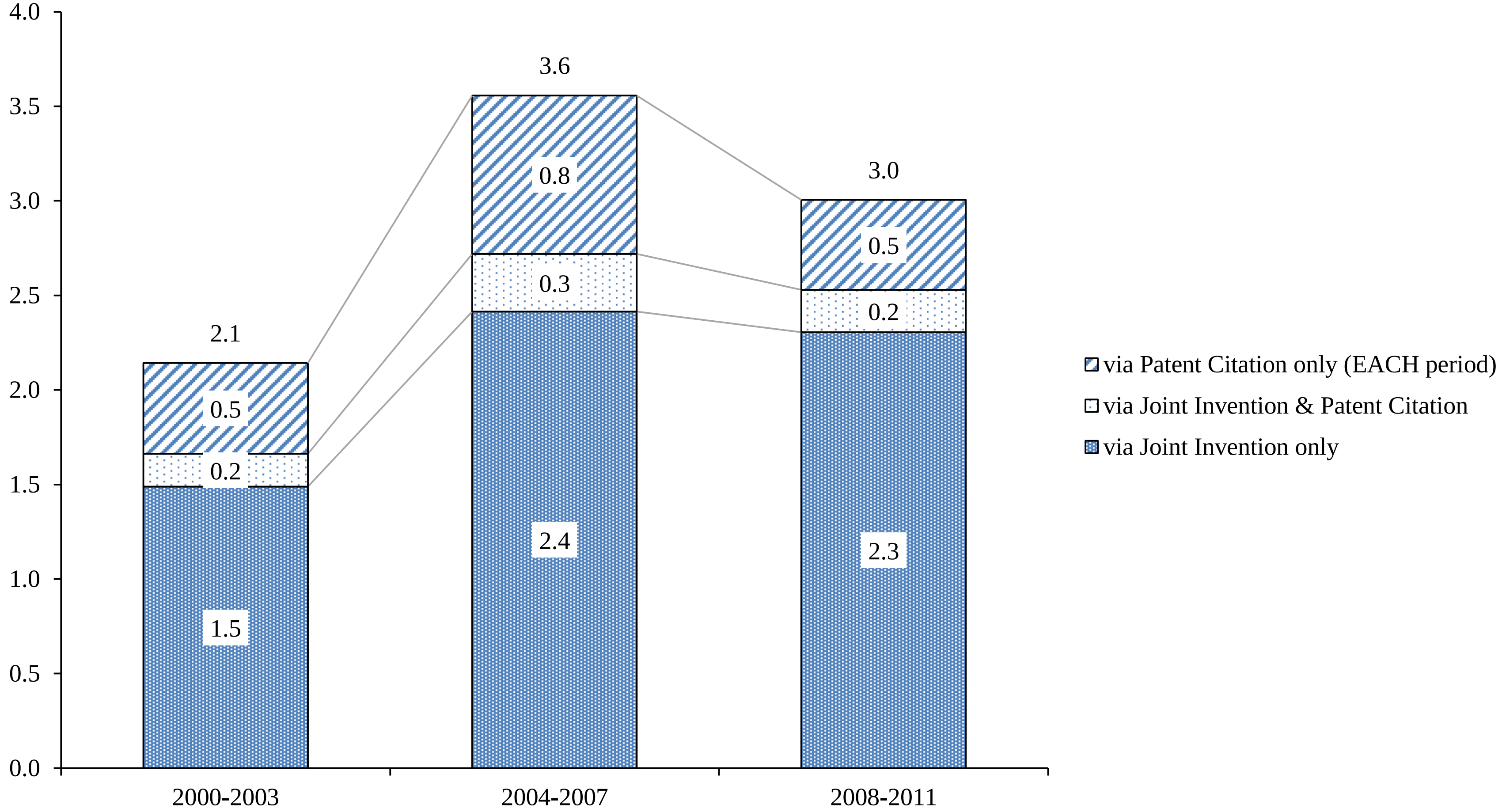
- Our indicators capture the mechanism of involvement of scientific knowledge in industrial innovation via patenting.
  - Science intensity (SI) of a firm = # of academic papers utilized by the firm per inventor/employee.
  - Utilization rate of science knowledge (URSK) = Share of papers utilized by firms in the total number of academic papers.

- Key assumption:
  - “New” scientific knowledge produced by academic inventors during a certain period (measured by the number of “new” scientific papers) is utilized by firms through joint research (patent invention) with and/or citing patents invented by the academic inventors during the same period (4 years).

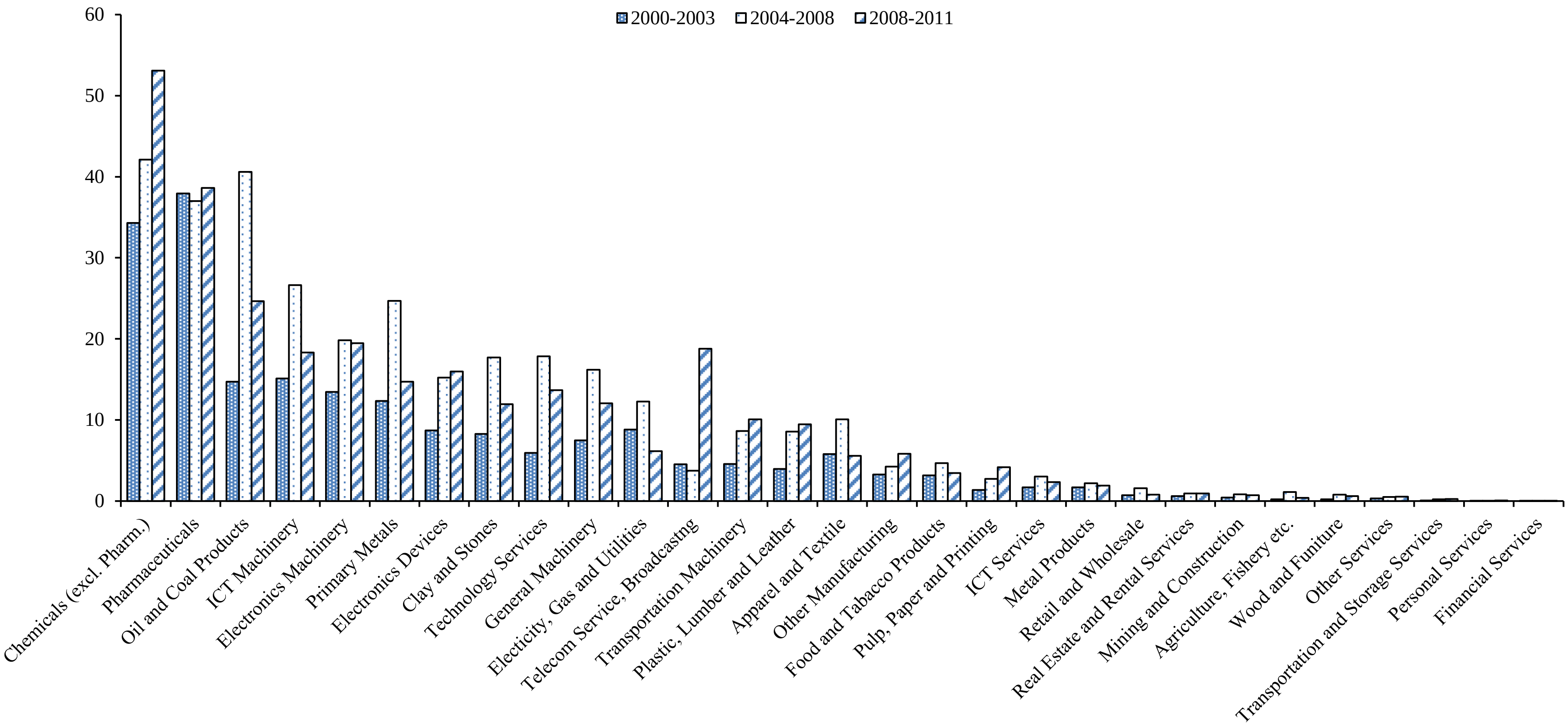
SIINV: # of Papers/Inventors (at firm)



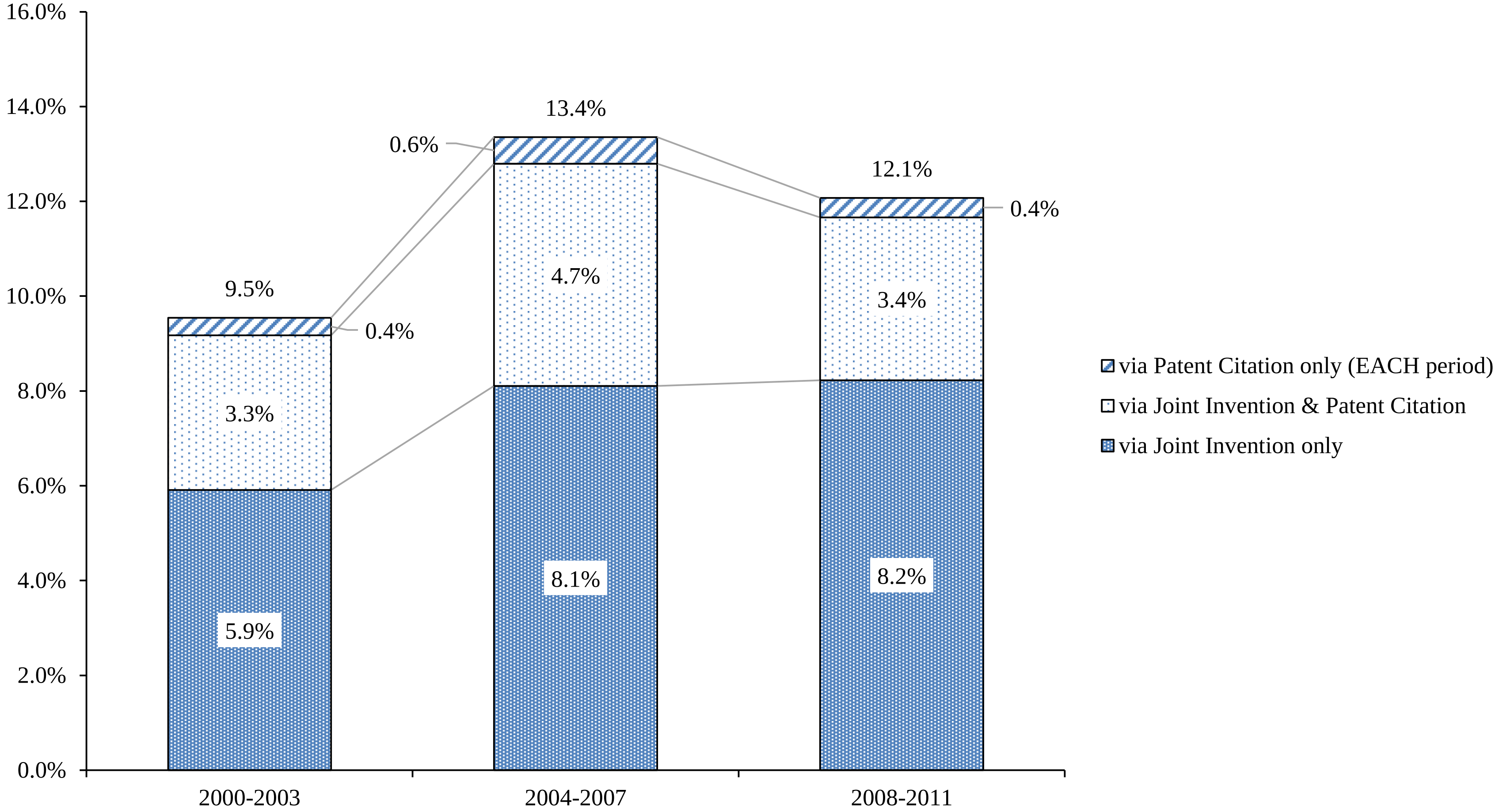
SIEMP: # of Papers/Employees (at firm)



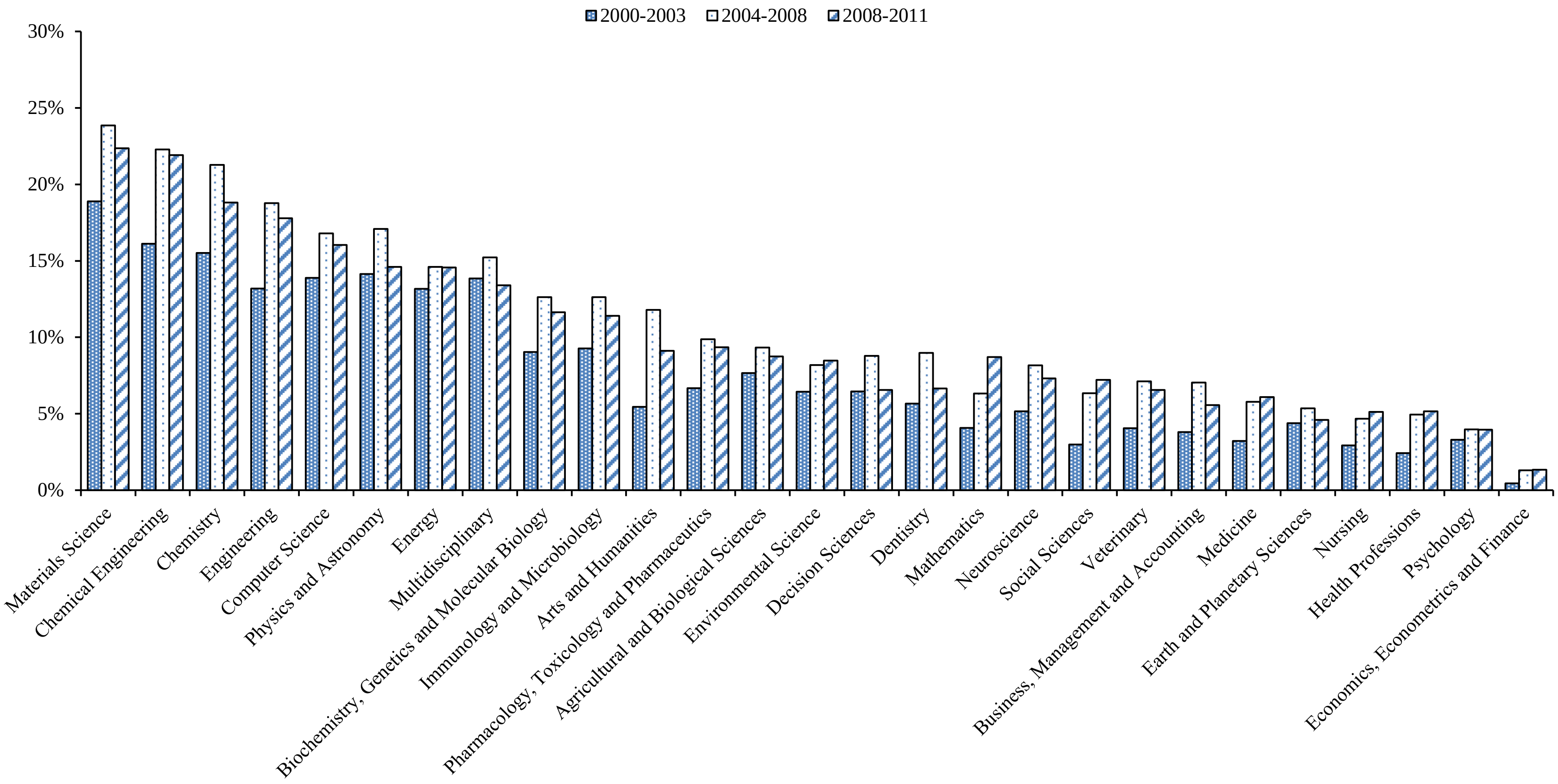
SIEMP: # of Papers/Employees (at firm)



URSK: # of Used Papers/All Papers (at academia)

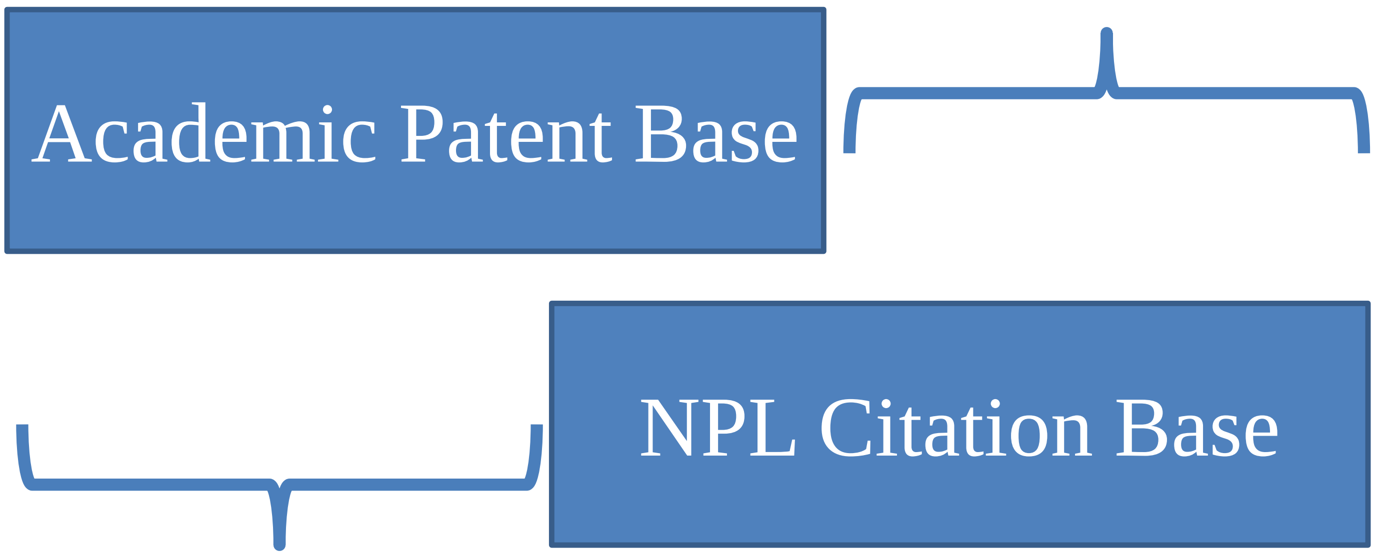


URSK: # of Used Papers/All Papers (at academia)



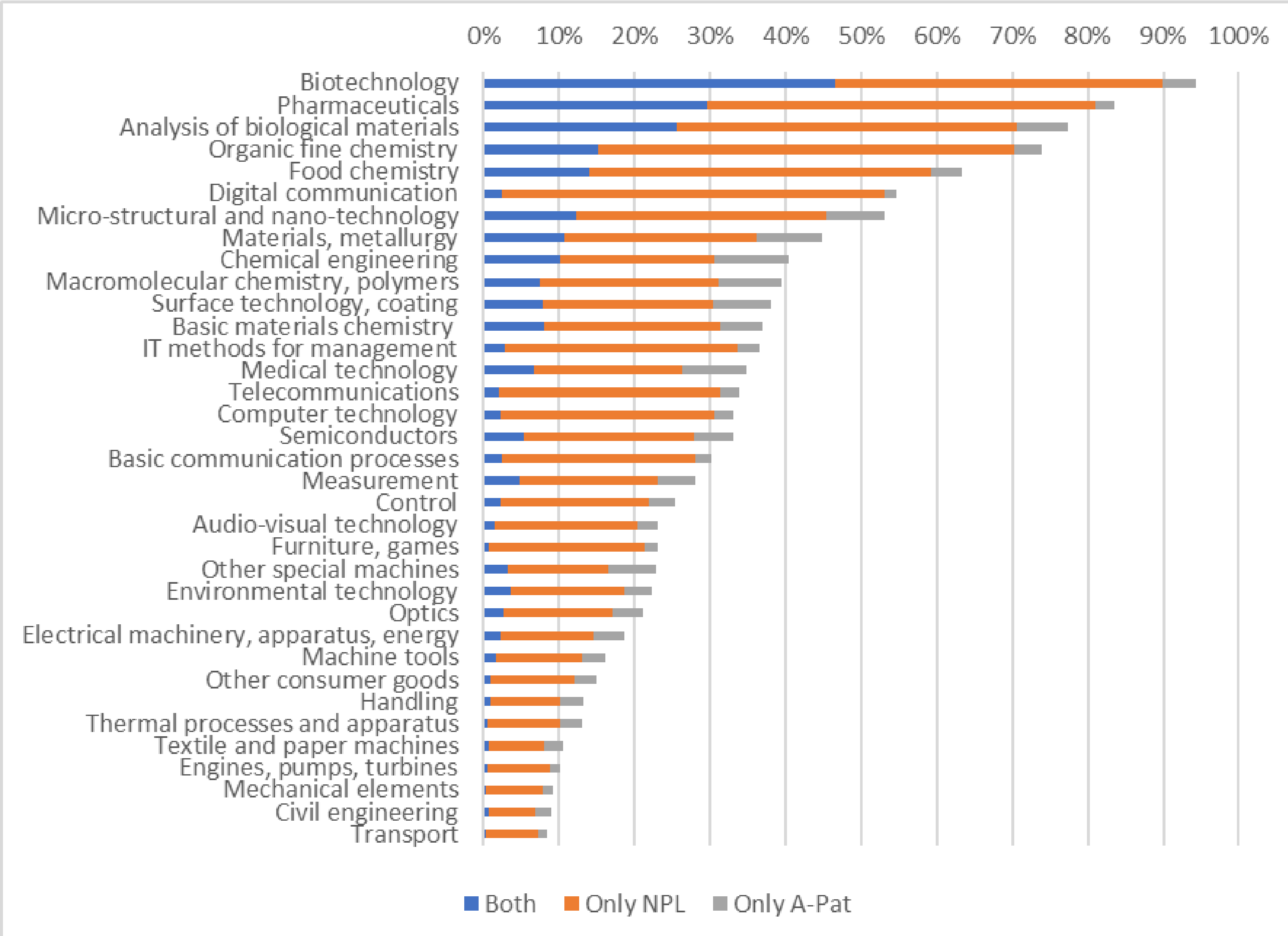
Comparison with NPL Indicator

SI not measured by our indicator and NPL not related to academic outputs (such as industry documents)



SI not measured by NPL (Academia inclining toward industry)

Share of patents with science linkage



Discussion and Conclusion

- We construct S-T-I Linkage datasets for whole population of Japanese research community and industries.
- We develop new indicators development to reflect science intensity and compared to NPL citation.

- Policy actions (UI promotion, institutional reforms) lead to “scientification” of not science based industries (primary metals, clay-stone, general machinery, telecom service) as well (Gini-coefficient by industry decreases).
- Also leads to increases in industry utilization rate of scientific outputs in all academic discipline (Gini-coefficient by academic field decreases).
- New indicator can measure SI linkage not measured by NPL based on (technologies not directly linked with science).

On going works

- Alternative machine learning methodologies for disambiguation of inventors.
- Scopus firm authors matching (based on NISTEP institution dictionary for private firms)
- NPL information incorporation for more comprehensive picture of science linkage.
- International extensions.

Thank you for your attention!

Contact: Kenta Ikeuchi

ks.ikuc@gmail.com