IIR/NISTEP/RIETI Workshop

Science sources of corporate inventions in Japan: Evidence from inventor survey

Isamu Yamauchi
(Research Institute of Economy, Trade and Industry)
Sadao Nagaoka
(Research Institute of Economy, Trade and Industry/ Hitotsubashi Universiy)

This research is a part of the ongoing research on “Science sources of innovation and its economic effects” supported by JST(Japan Science and Technology Agency). We would like to thank for the database work by Yusuke Naito as well as for excellent assistance by the RAs.

March 17 2014

Purpose

- Science of Science and Innovation Policy (SciSIP)
- Inventor Survey (Additional Survey)
  - To what extent does science contribute to Japanese corporate inventions?
  - Where does important scientific knowledge for inventions come from?
  - How complete and accurate does the citation data reflect the actual knowledge flow?
- Determinants of citing the important scientific source for the conception or implementation of R&D in patent documents

Key Findings

- For about 25% of the inventions, scientific knowledge embodied in literature, equipment or research materials was essential to conceive or implement the R&D
  - Importance of the scientific knowledge as a public good
- About 70% of the scientific knowledge source is located domestically in Japan, and large part of the rest 30% are based on the scientific knowledge generated in US
  - Domestic scientific knowledge sources are more important for Japanese inventions
- Even if the inventions cite the scientific literature in the patent document, two thirds of those inventions cite only unimportant science sources; about 20% of inventions without any reference to scientific literature in the patent document have important science sources
  - Patent citation to science sources is incomplete and noisy

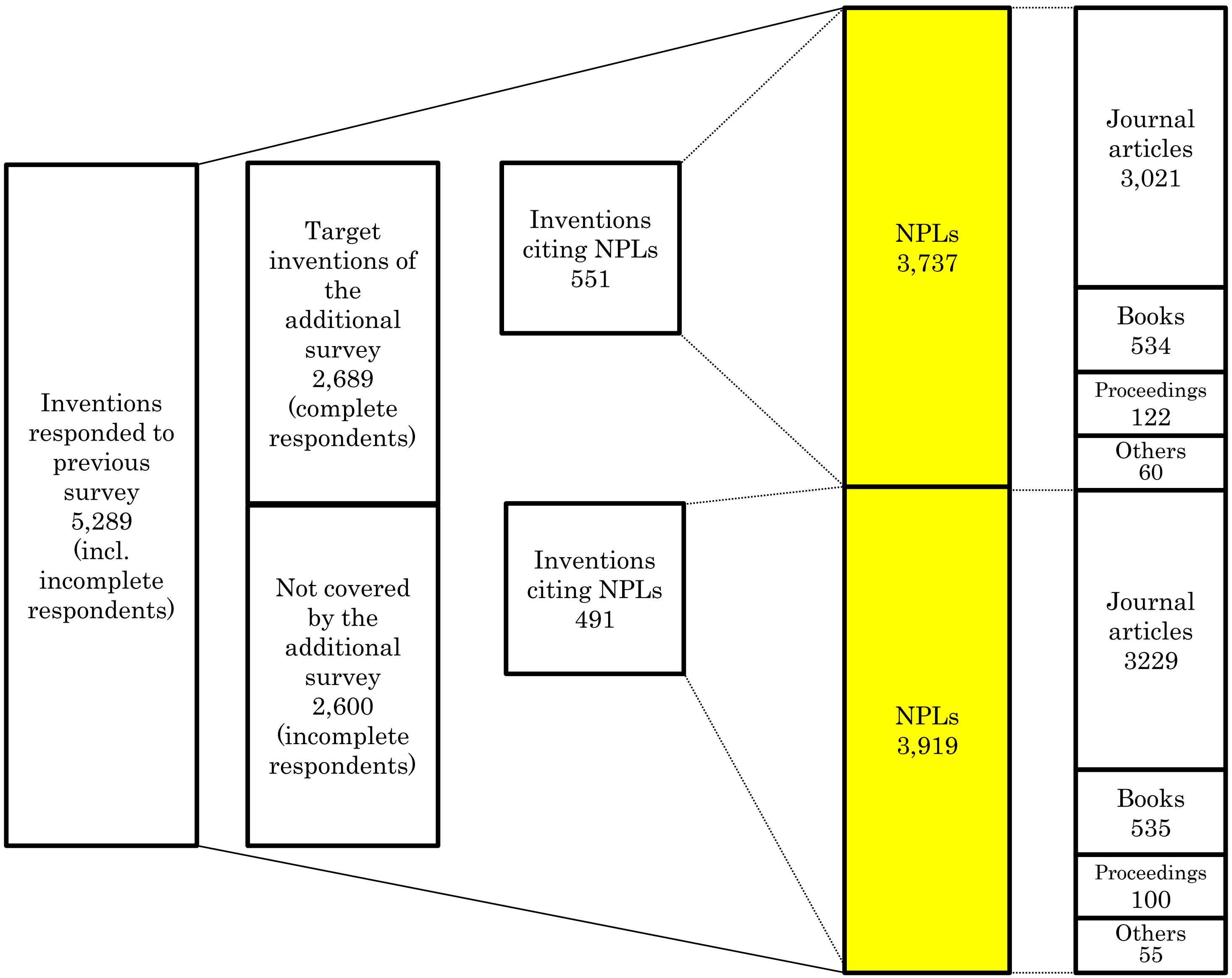
Determinants of citing the important scientific source

- Whether the inventors apply the scientific knowledge to their R&D depends on the
  - scientific absorptive capacity measured by the number of paper published
  - technology field
- Whether the inventors cite those important scientific knowledge sources in the patent document depends on the
  - value of the invention
  - firm's application strategy

Additional Survey

- **Sample**
  - 2,689 inventors who answered their email address in the previous Inventor Survey
  - 5,289 inventors responded to the previous Inventor Survey targeted the patent applications filed at both the JPO and the EPO with priority date 2003-2005
- **Survey method**
  - Web survey
- **Period**
  - September 2013 to February 2014
- **Number of responses**
  - 843（Response rate: 32.4%）

Overview of NPL dataset



Construction of NPL Dataset

- Extracted all non-patent literatures (NPLs) cited *by the inventors* in patent documents of the population (5,289 inventions) by text mining: 7,656 NPLs
- Got information on the author, affiliation, journal, publication year and title of the paper by merging the data on Web of Science and the Japanese literature database provided by Japan Science and Technology Agency (JST) to those 7,656 NPLs
- Information is manually collected for unmatched literatures
- Identified the types of literature, origin of the affiliation for almost all the NPLs cited in patent documents

Share of the inventions citing NPLs

- 1,042 out of 5289 inventions (19.7%) cite NPLs in the patent documents
  - Biotechnology, Organic chemistry, and Pharmaceuticals have high propensity to cite NPLs (about 70 % or more)

	N	Number of inventions citing NPLs	Share of the inventions citing NPLs
Biotechnology	117	97	82.9%
OrganicChem	218	157	72.0%
Pharmaceuticals/Cosmetics	197	133	67.5%
Agric&Foods	38	14	36.8%
Polymers	219	79	36.1%
…	…		
Motors	209	4	1.9%
ThermProcesses	60	1	1.7%
Transportation	292	2	0.7%
ConstrTechn	35	0	0.0%
	5289	**1041**	**19.7%**

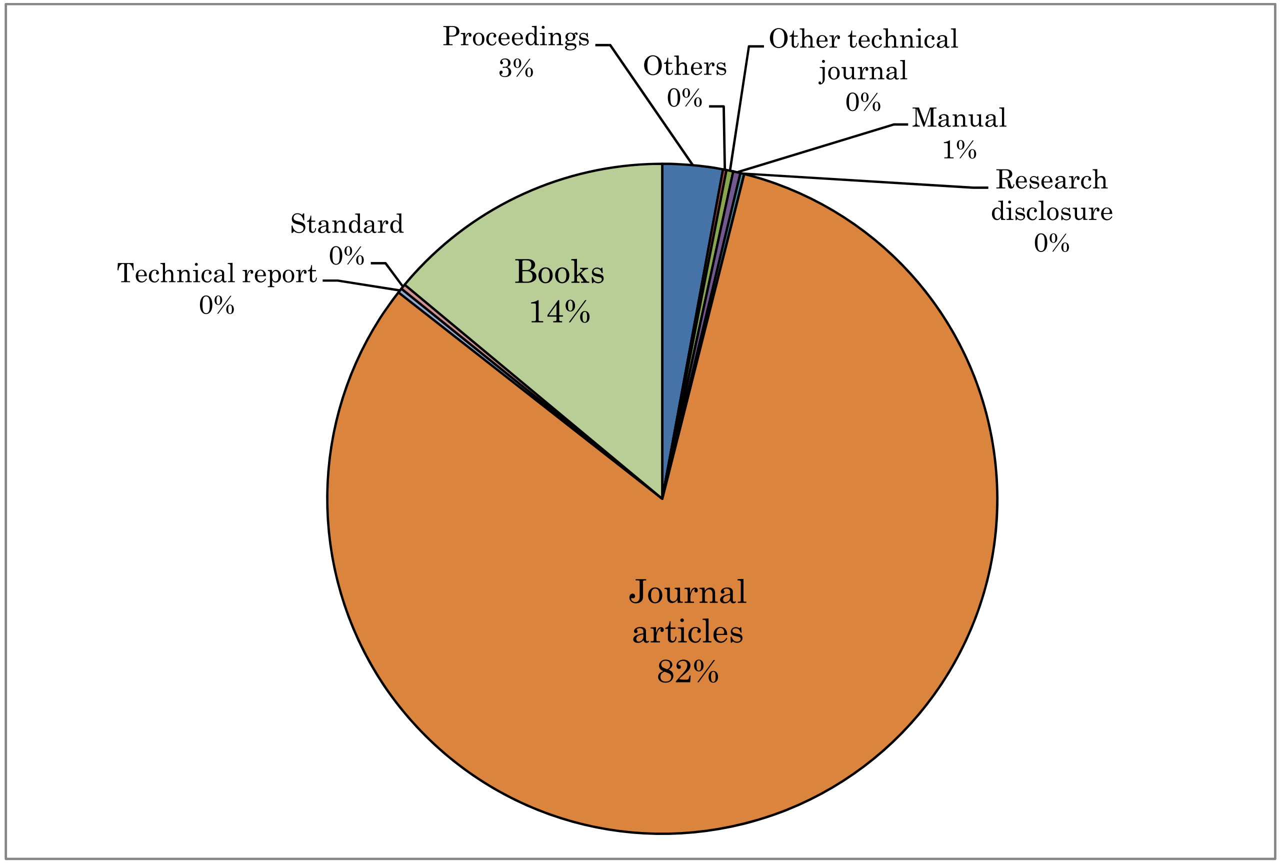
Places where NPLs are cited

- 60% of NPLs are cited at the place where the invention is described, and the rest 40% are cited where prior art is described

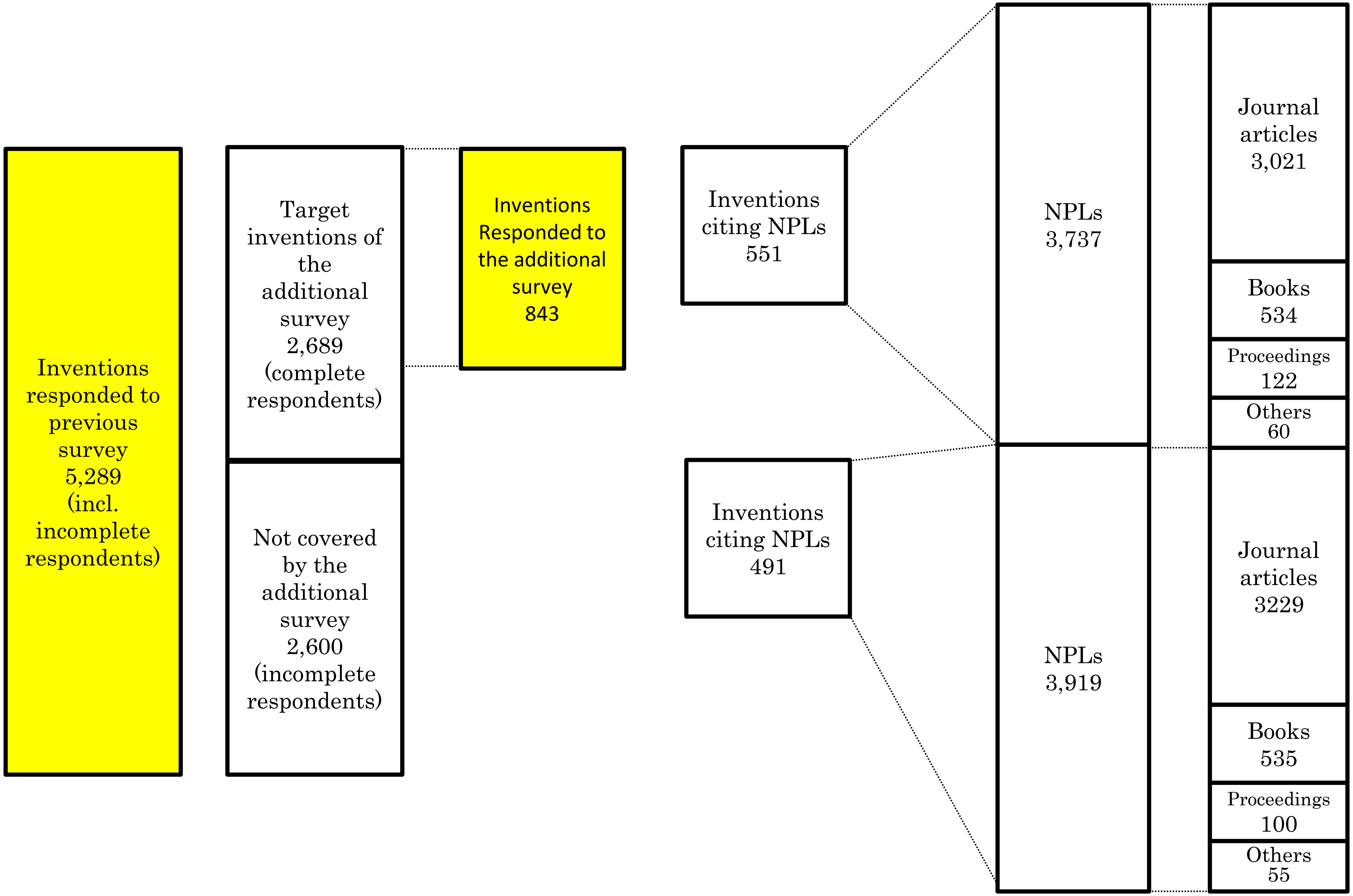
	Number of inventions	Number of cited NPLs	Number of cited NPLs per inventions	Share of NPLs cited where prior art is described	Share of NPLs cited where the invention is described
Biotechnology	117	1873	16.0	32.8%	67.2%
Pharmaceuticals/Cosmetics	197	1769	9.0	32.9%	67.1%
OrganicChem	218	1798	8.2	44.4%	55.6%
PetrolChem/materialsChem	95	162	1.7	35.2%	64.8%
…		…			
Motors	209	7	0.0	85.7%	14.3%
ThermProcesses	60	1	0.0	100.0%	0.0%
Transportation	292	2	0.0	100.0%	0.0%
ConstrTechn	35	0	0.0	-	-
Total	5289	7657	1.4	40.2%	59.8%

Types of NPLs

- 82% of 7,656 cited NPLs are journal articles
  - Journals, books and proceedings account for 99% of cited NPLs



Population and Sample



Selection bias?

EU technology class	Sample N	Sample %	Population (respondents of the previous survey) N	Population (respondents of the previous survey) %
Electr/Energy	69	8.4%	472	9.2%
Analysis/Measurement	65	7.9%	375	7.3%
Telecom	54	6.6%	342	6.6%
OrganicChem	51	6.2%	218	4.2%
IT	49	5.9%	293	5.7%
Polymers	45	5.5%	219	4.3%
Transportation	44	5.3%	292	5.7%
Matprocessing/Textiles	43	5.2%	188	3.7%
Audiovisual	37	4.5%	248	4.8%
Semiconductors	36	4.4%	214	4.2%
Pharmaceuticals/Cosmetics	36	4.4%	197	3.8%
MechElements	34	4.1%	204	4.0%
Optical	31	3.8%	246	4.8%
Motors	30	3.6%	209	4.1%
Materials	24	2.9%	159	3.1%
Handl/Printing	22	2.7%	185	3.6%
MedicalTechn	21	2.5%	203	3.9%
ConsGoods	21	2.5%	114	2.2%
SurfaceTechn	18	2.2%	113	2.2%
PetrolChem/materialsChem	17	2.1%	95	1.8%
ChemEngineering	17	2.1%	106	2.1%
Biotechnology	15	1.8%	117	2.3%
Environment	14	1.7%	50	1.0%
MachineTools	14	1.7%	120	2.3%
…	…		…	
Total	824		5145	

Selection bias?

- Distribution of affiliations

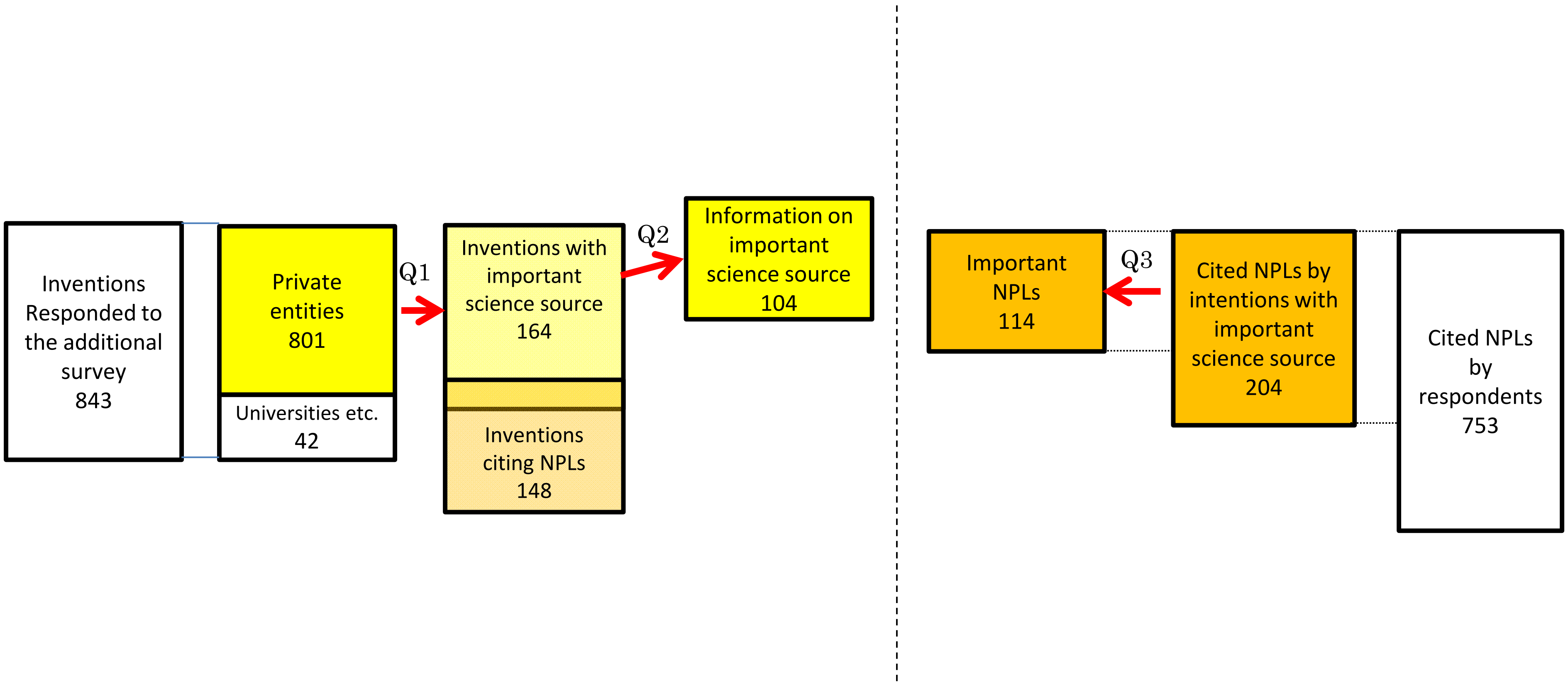
	Sample		Population (respondents of the previous survey)	
	N	%	N	%
private enterprise	797	94.7%	4,158	94.3%
public research institute	12	1.4%	64	1.5%
university/educational institute	26	3.1%	140	3.2%
other governmental organization	0	0.0%	3	0.1%
hospital/nonprofit organization	4	0.5%	23	0.5%
Other	3	0.4%	23	0.5%
Total	842		4,411	

- Number of the inventions and the papers of the inventors

	Sample				Population			
	N	mean	median	SD	N	mean	median	SD
number of inventions	823	57.9	30	76.0	3779	56.2	30	92.7
number of papers	825	12.2	2	47.5	3833	10.9	1	54.9
Ln(number of invention	823	3.44	3.43	1.18	3779	3.36	3.43	1.19
Ln(number of papers)	825	1.19	1.10	1.31	3833	1.11	0.69	1.30

→Distribution of technology classes and affiliations is quite similar, though the number of papers is a little bit larger for the sample

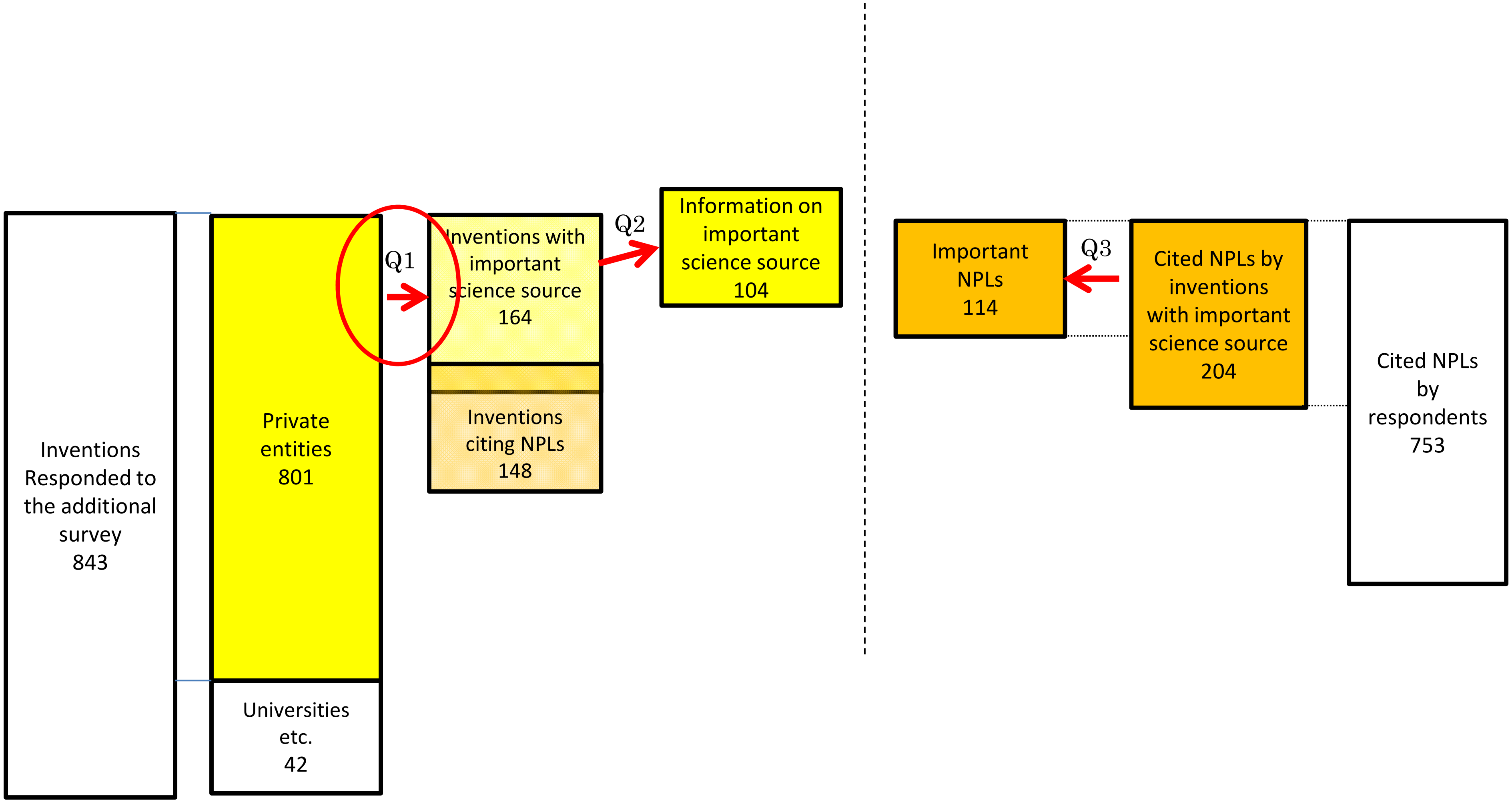
Survey strategy



Survey strategy

1. First, before showing the literature list cited in the patent document, we ask whether there is any science source that is essential to conceive or implement the R&D
2. Then, to identify such important science sources, we ask the detailed information such as the name of the researcher who developed the science source, his affiliation, journal name, title of the paper etc.
3. After that, we show the literature list when we ask whether the literature cited in the patent document is essential/important to conceive or implement the R&D

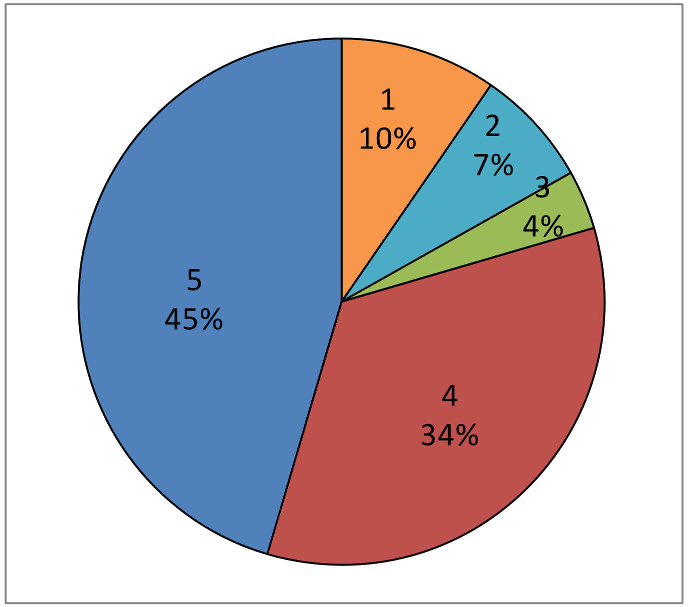
Survey strategy



Essentiality and importance of scientific literature（N=801）

- For about 20% inventions, scientific literature was important knowledge source for the conception and the implementation of the R&D
- 9.6% of inventions could not have been conceived if it were not for the scientific literature

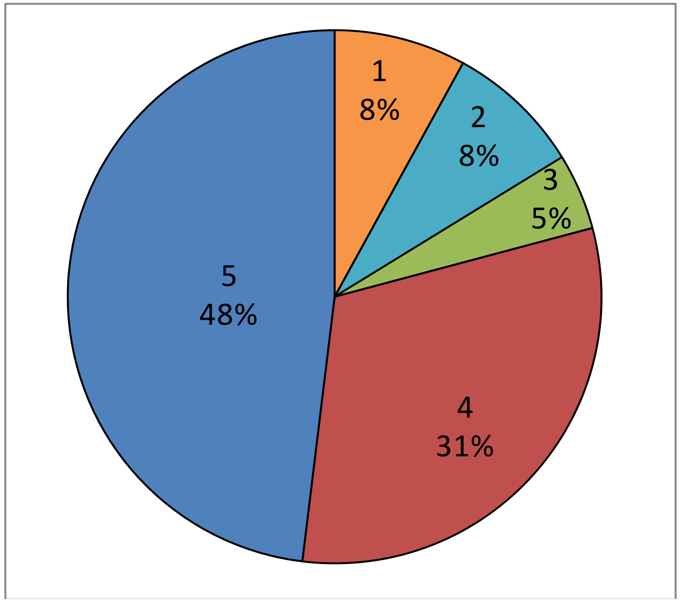
	If it were not for the scientific literature published in the last 15 years,	frequency	share
1	the idea for the invention could not have been conceived	77	9.6%
2	implementation of R&D would have been extremely difficult	58	7.2%
3	getting a result would have been significantly delayed	29	3.6%
4	getting a result would have been slightly delayed	273	34.1%
5	it would have had no impact	364	45.4%



Essentiality and importance of research equipment and materials

- For about 20% of inventions, scientific research equipment and materials had important impact on the conception or implementation of the R&D

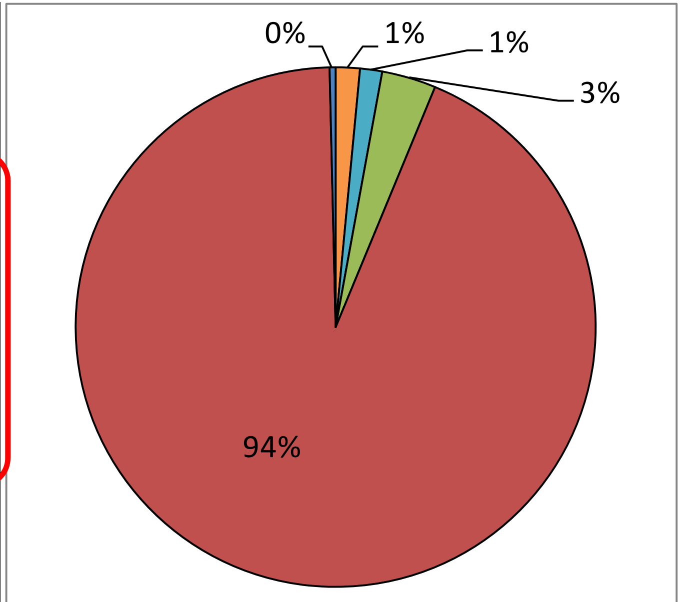
If it were not for the equipment or research materials those embodies scientific knowledge in the last 15 years,		frequency	share
1	the idea for the invention could not have been conceived	64	7.8%
2	implementation of R&D would have been extremely difficult	66	8.5%
3	getting a result would have been significantly delayed	37	4.7%
4	getting a result would have been slightly delayed	249	30.6%
5	it would have had no impact	385	48.4%



Essentiality and importance of collaboration with the university

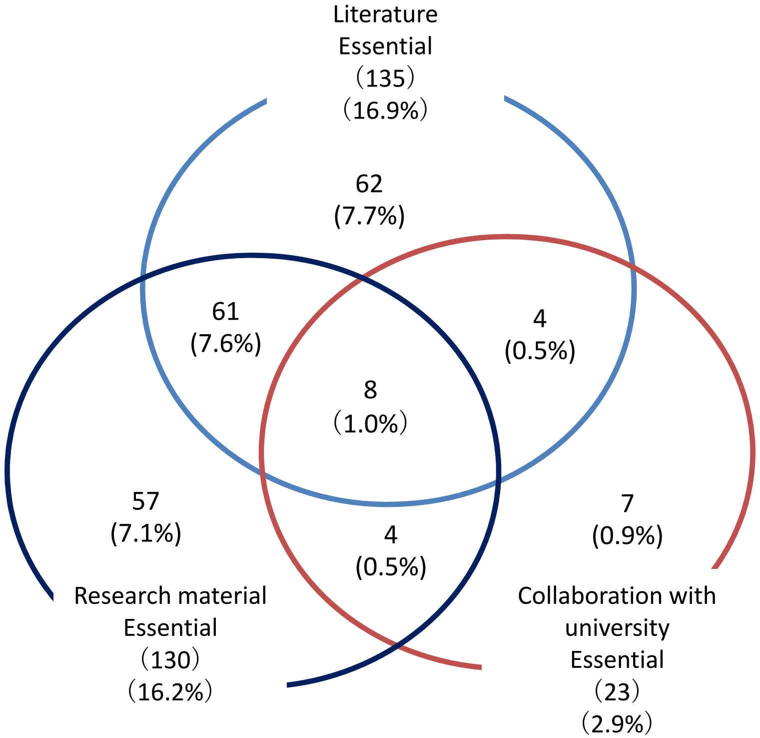
- For about 6% inventions, collaboration with the university had important impact on conception or implementation of the R&D
- For about 3% inventions, collaboration with the university was essential for the R&D

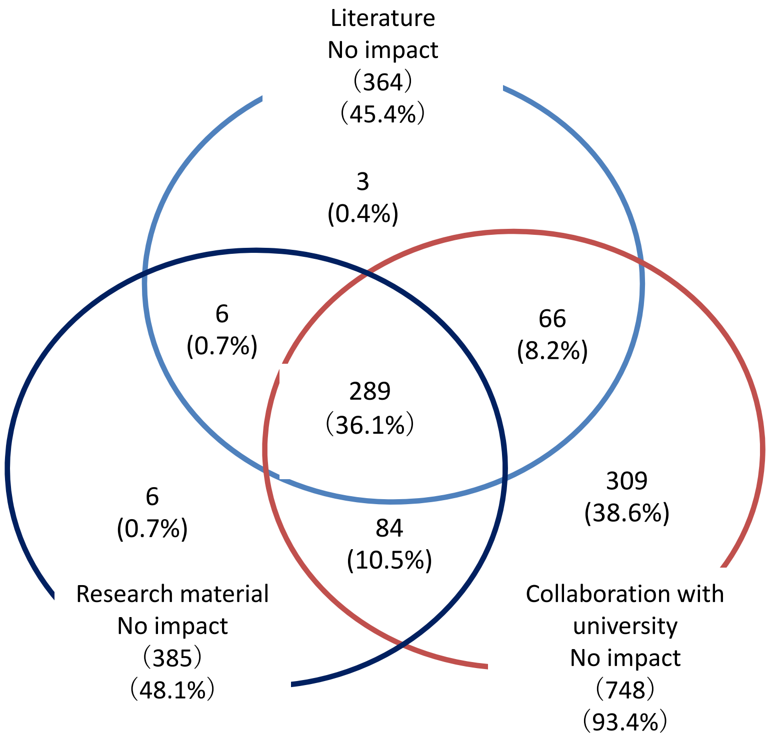
	If it were not for the collaboration with the universities or research institute,	frequency	share
1	the R&D project would not have exist	12	1.5%
2	the implementation of the R&D would have been given up	11	1.4%
3	the size of the R&D would have been significantly reduced or the R&D project would have been significantly delayed	27	3.4%
4	it would have had no impact	748	93.4%
5	Others	3	0.4%



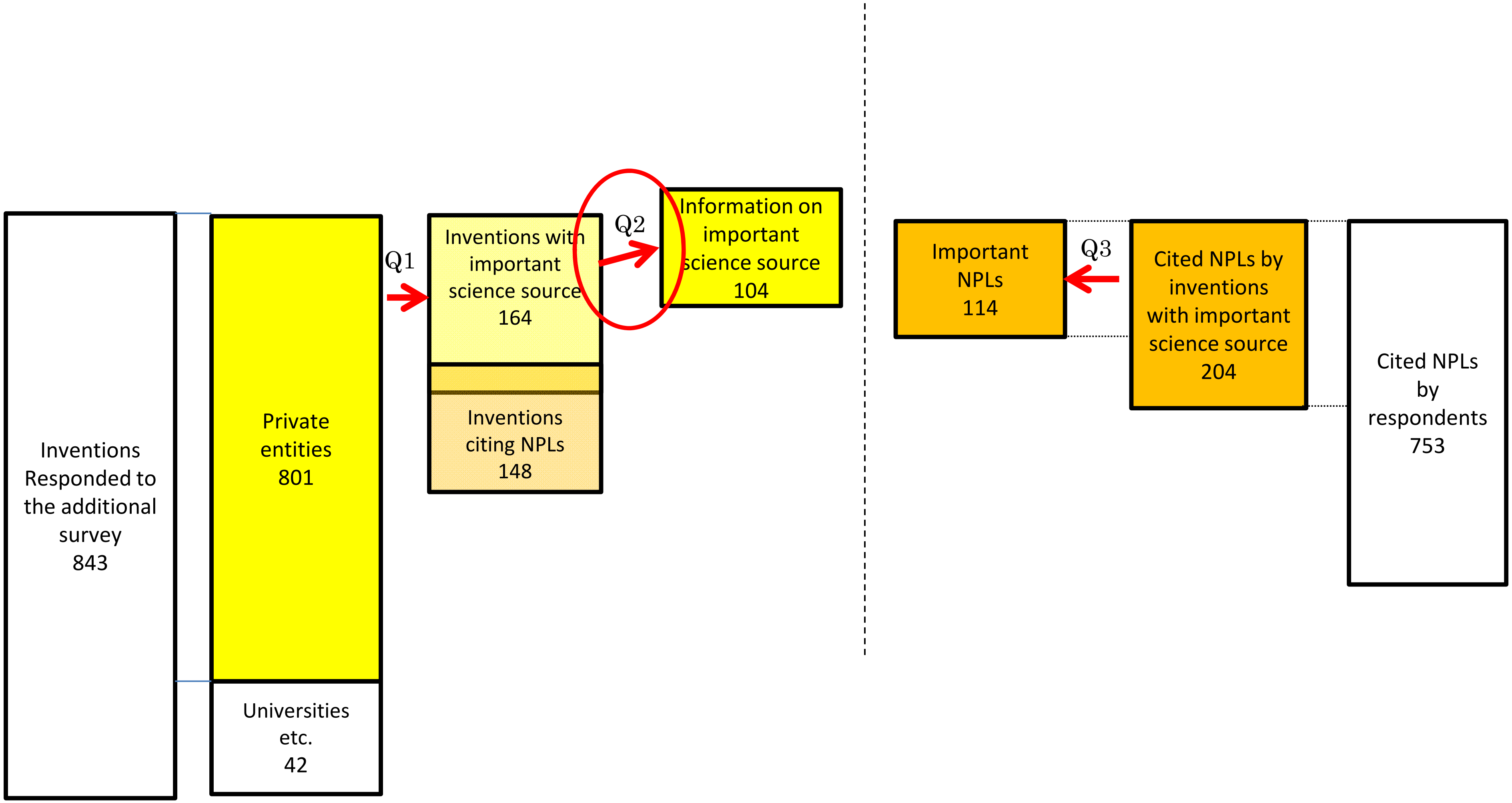
Essential science sources（N=801）

- Recent scientific advances were essential for a quarter of the inventions
  - For about 25% of inventions, recent scientific advances, embodied in literature or equipment or research material, were essential to conceive or implement the R&D
  - For about 3% of inventions, direct collaboration with universities was essential for the invention
- Recent scientific advances had no direct impact on the invention
  - 36.1% → for the rest 65.9% inventions, scientific knowledge had some impact



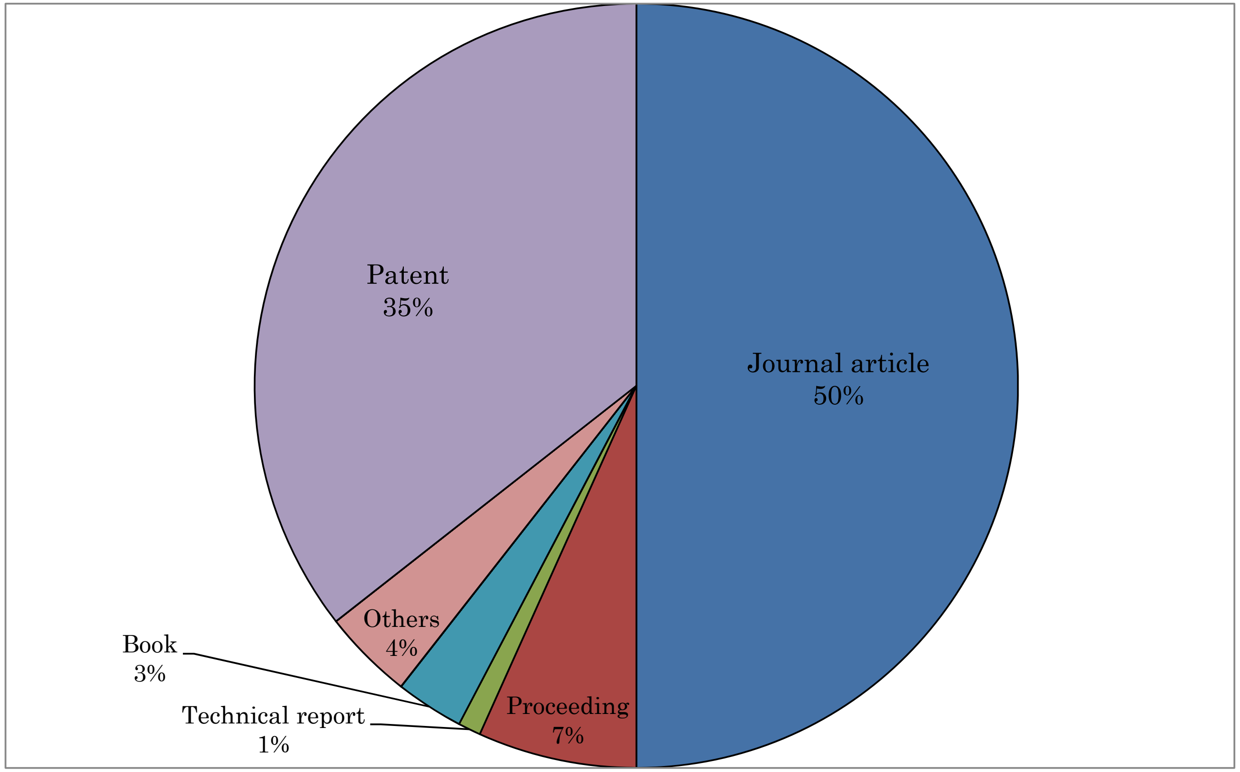


Survey strategy



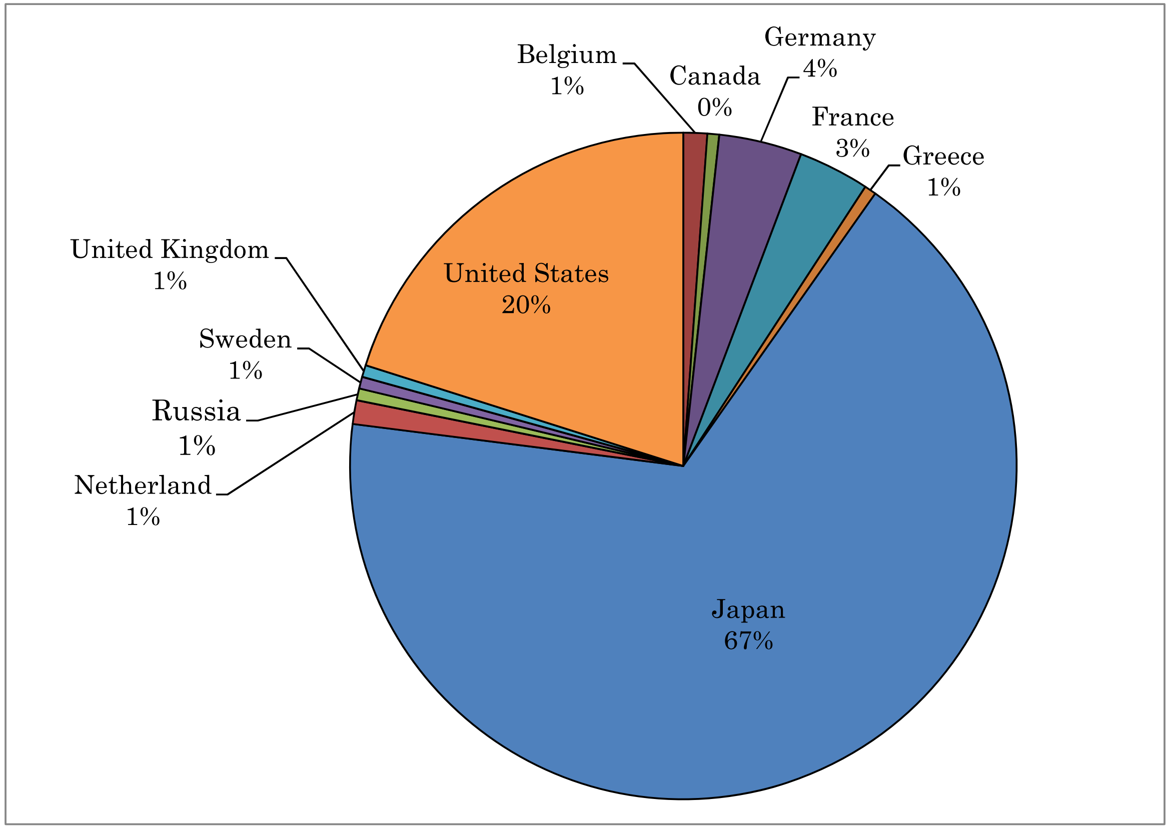
Types of important science sources embodied in literature (N=104)

- Important scientific literature to conceive or implement the R&D
  - articles: 52 (50%), patent: 37 (35%), proceedings: 7 (7%)
- Papers, books and proceedings account for 92.5% if exclude patent



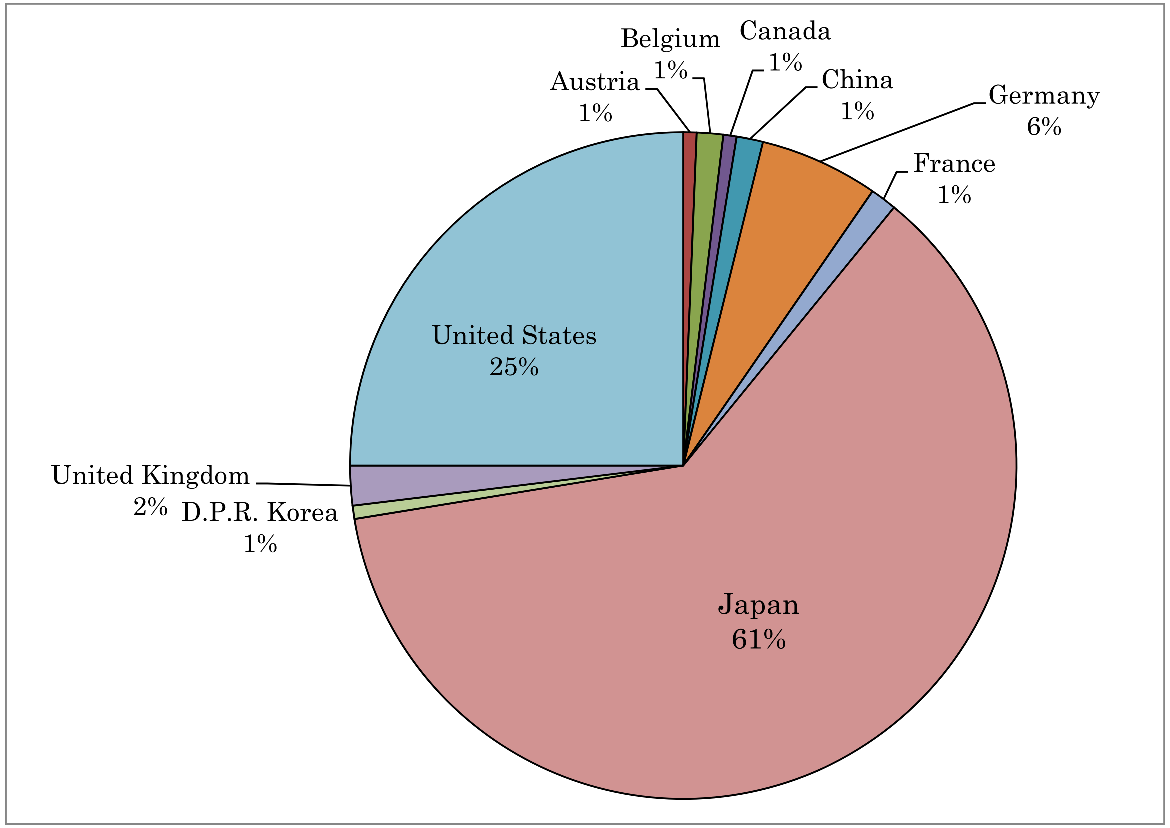
Location of the important science sources embodied in literature

- 67% of the important scientific knowledge sources are developed by the researchers located in Japan
- 20% is generated by US researchers

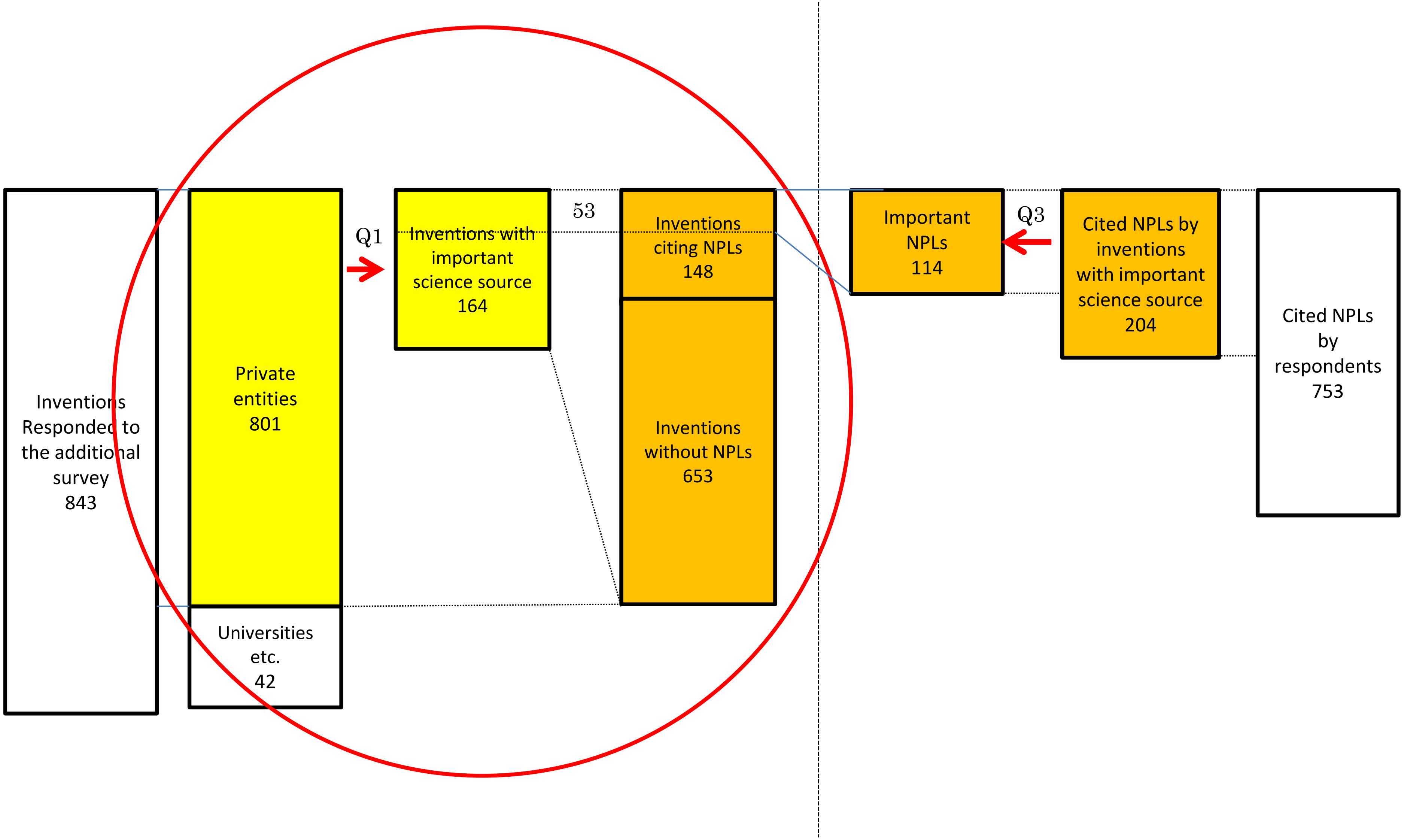


Location of the important science sources embodied in equipment or research materials

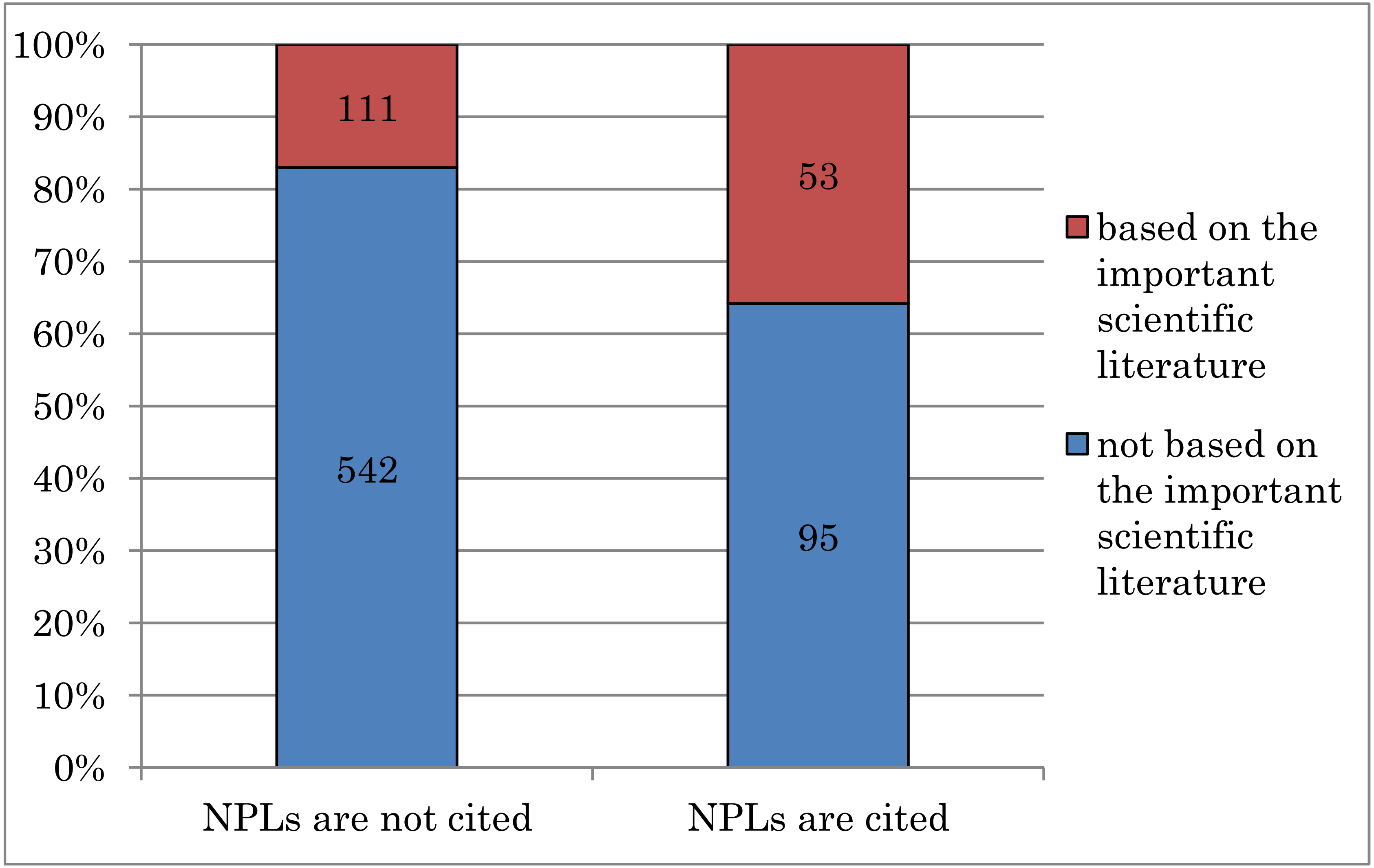
- Most of the important science sources embodied in equipment or research materials is developed by domestic firms; and most foreign science source is located in US
  - Domestic scientific knowledge sources are more important for Japanese inventions



Knowledge flow



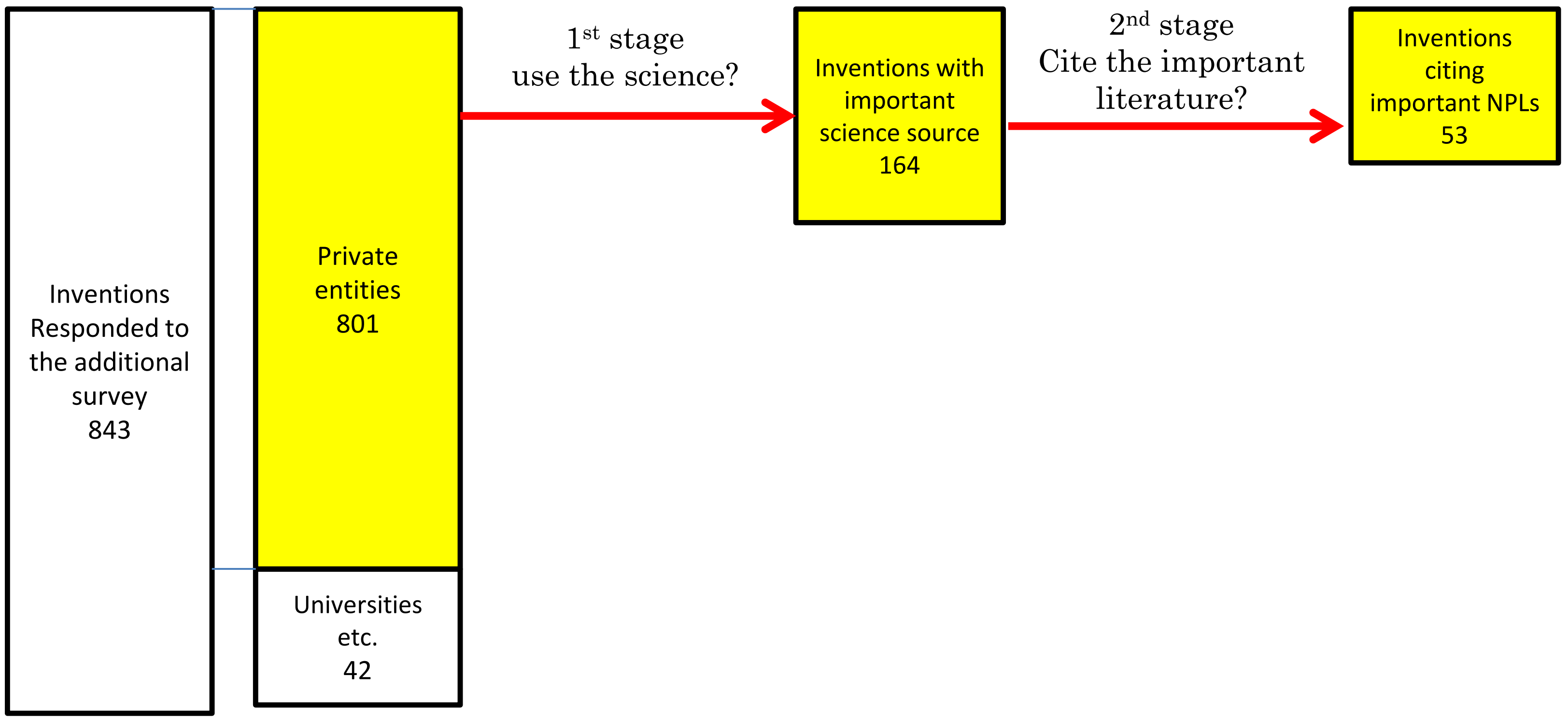
Citation and Knowledge flow for inventions



Citation and Knowledge flow

- Even if the inventions cite the scientific literature, 64% of those inventions do not have any important scientific literature to conceive and implement the R&D
  - Much of the scientific literature cited in the patent documents are not the scientific knowledge source but the prior art to explain the invention and its patentability
- Even if the inventions do not cite the scientific literature, 17% of those inventions have an important scientific literature for the R&D
  - There are many cases that the important science sources are not cited in the patent documents

Determinants of citing important scientific literature



Determinants of citing important scientific literature

- Who uses the science for the invention?
  - Scientific absorptive capacity
    - Science activity (writing papers)
    - Education (Ph.D)
  - Technology field
- Under what conditions do inventors cite such important science sources in the patent documents?
  - Value of the invention (inventive step)
  - Firm size
  - Background of inventor
    - Science activity (writing papers)
    - Education (Ph.D)
  - Technology field

Variables

■First stage

- Dependent variable
  - Importance of science for the conception or implementation of R&D (Dummy variable that takes 1 if important)
- Independent variables
  - Scientific absorptive capacity (number of papers / Ph.D. holder dummy)
  - Technology field

■Second stage

- Dependent variable
  - Citing the important scientific literature in patent documents (Dummy variable)
- Independent variables
  - Inventive step (dummy variable taking 1 if extremely high or high)
  - Firm size (number of employees)
  - Scientific absorptive capacity (number of papers / Ph.D. holder
  - Technology field
  - Number of citing literature

Heckman

	first stage			second stage		
	Using the scientific literature			Citing the important literature		
	(1)	(2)	(3)	(1)	(2)	(3)
ln(num. of papers)	0.053***		0.048***	-0.069	-0.022	
	(3.20)		(3.31)	(-0.20)	(-1.00)	
Ph.D. holder	-0.032	0.048		0.065		0.042
	(-0.66)	(1.04)		(0.38)		(0.73)
num. of inventors	-0.044*	-0.045*	-0.044*	-0.008	-0.005	-0.066
	(-1.65)	(-1.71)	(-1.67)	(-0.03)	(-0.07)	(-1.45)
ln(num. of NPLs)				0.513***	0.513***	0.514***
				(17.63)	(17.77)	(17.86)
high inventive step				0.115***	0.115***	0.115***
				(2.99)	(3.00)	(2.99)
employees_1000 or more				0.056	0.056	0.054
				(0.75)	(0.75)	(0.73)
employees_5000 or more				0.151**	0.152**	0.150**
				(2.04)	(2.05)	(2.05)
Constant				0.395	0.324	-0.237
				(0.12)	(0.59)	(-0.99)
technology fields	yes	yes	yes	yes	yes	yes
Observations	732	745	733	732	745	733

coefficients in the first stage are marginal effects
z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Results

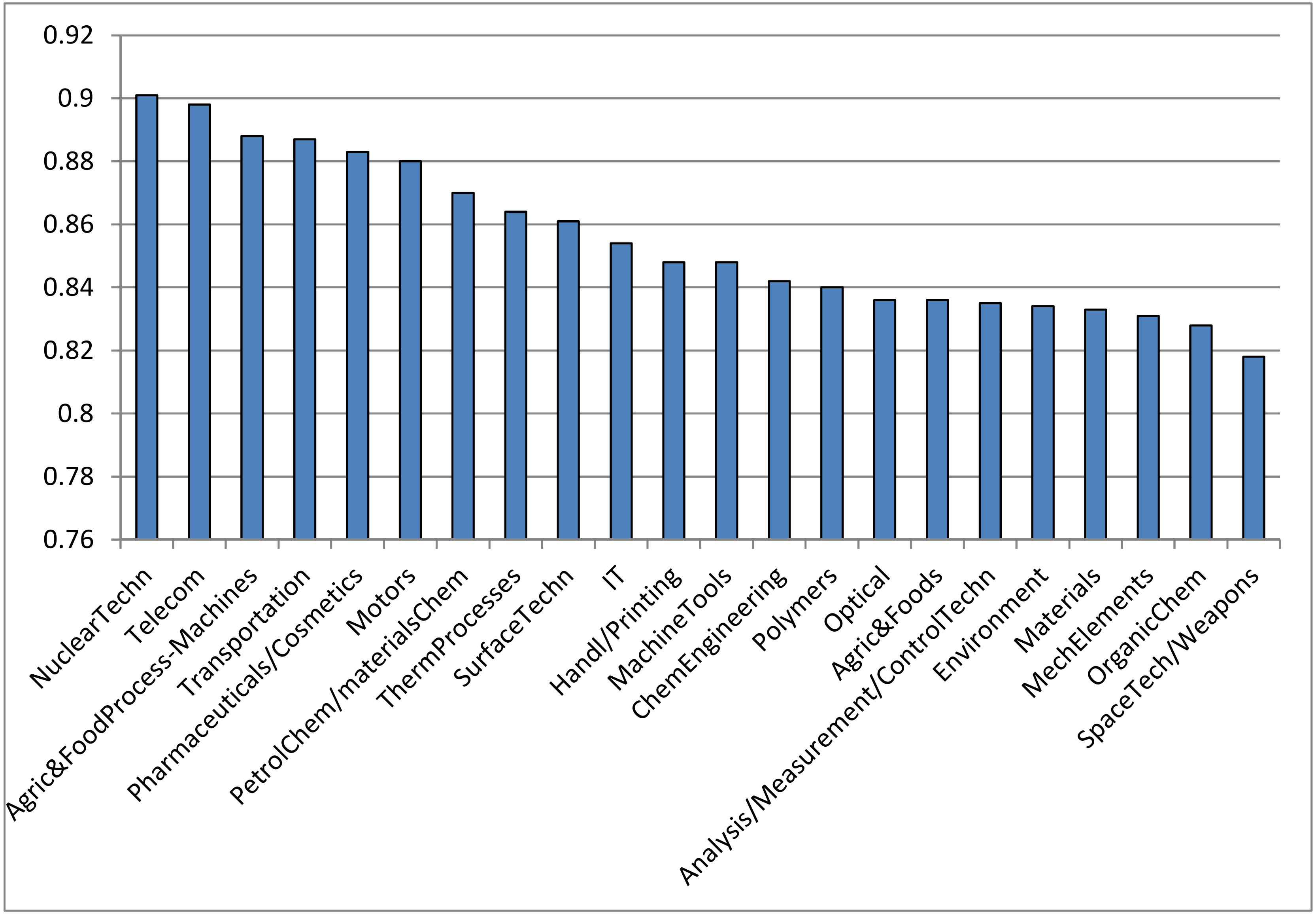
- In the first stage decision, the coefficients of the number of papers are positive and significant
  - Scientific absorptive capacity of inventors matters for the application of the scientific knowledge to the corporate inventions

- In the second stage, we cannot see such significant effect of the scientific absorptive capacity
- The inventions with high inventive step are likely to cite the important literature in patent document
- Large firm tends to describe the important literature in patent document
  - Firm's application strategy is more important determinants for citing the important literature

Heckman (technology field)

	first stage			second stage		
	Using the scientific literature			Citing the important literature		
	(1)	(2)	(3)	(1)	(2)	(3)
Audiovisual	0.809	0.809***	0.809	-0.085	-0.021	0.150
Telecom	0.898***	0.900***	0.899***	0.005	-0.013	-0.052
IT	0.854***	0.853***	0.854***	0.168	0.146	0.140
Optical	0.836***	0.834***	0.836***	-0.053	-0.076	0.085
Analysis/Measurement/Co	0.835***	0.833***	0.837***	0.144	0.131	-0.026
NuclearTechn	0.901***	0.901***	0.901***	0.248	0.218*	0.199*
OrganicChem	0.828***	0.826***	0.828***	-0.012	-0.045	-0.017
Polymers	0.840***	0.839***	0.840***	0.230	0.219	0.007
Pharmaceuticals/Cosmetic	0.883***	0.884***	0.883***	-0.026	-0.034	0.128
Agric&Foods	0.836***	0.835***	0.837***	0.302	0.284	0.203
PetrolChem/materialsChe	0.870***	0.869***	0.871***	0.112	0.104	0.105
SurfaceTechn	0.861***	0.861***	0.861***	-0.027	-0.027	-0.014
Materials	0.833***	0.832***	0.834***	0.084	0.053	0.101
ChemEngineering	0.842***	0.846***	0.843***	0.029	0.017	-0.034
Handl/Printing	0.848***	0.849***	0.848***	0.058	0.032	0.011
Agric&FoodProcess-Machi	0.888***	0.888***	0.888***	-0.024	-0.041	0.096
Environment	0.834***	0.833***	0.834***	0.108	0.081	0.078
MachineTools	0.848***	0.848***	0.848***	0.095	0.064	0.196
Motors	0.880***	0.881***	0.881***	0.180	0.167	0.255**
ThermProcesses	0.864***	0.865***	0.864***	-0.026	-0.031	0.115
MechElements	0.831***	0.830***	0.832***	0.039	0.016	0.131
Transportation	0.887***	0.889***	0.887***	-0.043	-0.063	-0.027
SpaceTech/Weapons	0.818***	0.816***	0.819***	-0.132	-0.135	-0.007

Heckman (technology field)



Results (technology field)

- Difference in the technology field has large effect on the decision on applying the scientific knowledge to the R&D
- Especially, in the technology fields such as Nuclear Technology, Telecommunication, Agriculture, Transportation, and Pharmaceuticals the scientific knowledge plays an important role for R&D

- However, in the second stage, the technology field dummies do not have any effect

Conclusion (1)

- Science significantly contribute to the R&D activities of Japanese private company
  - For about 25% of the inventions, recent scientific advance was essential to conceive or implement the R&D
- Most of the Japanese R&D activities are based on the local science sources
  - For about 70% Japanese inventions, science sources are located in Japan
- Citation data is incomplete and noisy
  - Even if the inventions cite the scientific literature in the patent document, two thirds of those inventions cite only unimportant science sources; about 20% of inventions without any reference to scientific literature in the patent document have important science sources

Conclusion (2)

- Decision on using the scientific knowledge in the R&D process and the decision on citing those scientific literature are determined by different factors
  - Former depends on the inventor's scientific absorptive capacity and the difference in technology field
  - Latter depends on the value of the invention and the firm size, which can reflect the firm's application strategy is more important.

Implications

- ■Policy implication
  - – Importance of the development of scientific knowledge as a public good for industrial innovation
    - • Not only publication but also research equipment and materials
- ■Implication for innovation research
  - • Citation is an incomplete and noisy index
    - • Important to control for the value of the invention, firm size, inventor's science activity, and the technology field

Future work

- Difference between the “essential” scientific literature and the literature cited in the patent documents as an index of knowledge flow
  – Results are different?
- Are key science sources really so domestic in Japan? Then, why?
  – Plenty sources exist in Japan?
  – Language constraint on the science absorptive capacity?
  – Geographical closeness matters for the application of scientific knowledge to innovation?