

International collaborations in science, why and how: evidence from scientists' survey in the US and Japan

Sadao NAGAOKA

Institute of Innovation Research, Hitotsubashi University, Japan

Masatsura IGAMI

National Institute of Science and Technology Policy (NISTEP), Japan

John P. WALSH

School of Public Policy, Georgia Institute of Technology, USA

June 2012

Introduction

- International institutional collaborations in science have become important.
 - Institutional collaborations refer to the collaborations between researchers affiliated with different institutions in this paper.
 - one third of the highly cited papers (top 1%) in the US and around 40% in Japan
- Many growing literature analyzing the causes and consequences of institutional collaborations
 - Adams, Black, Clemmons, and Stephan (2005),
 - Wuchty, Jones, and Uzzi (2007) and Jones B. F., S. Wuchy and B. Uzzi, 2008), and
 - Stephan (2011))

Table 1. Organizational structure by focal papers, Number of observations, %

Japan	single	dom_coll	intemat_coll	Total
N	638	491	346	1,475
	43	33	23	100
H	147	183	234	564
	26	32	41	100
Total	785	674	580	2,039
	39	33	28	100

US	single	dom_coll	intemat_coll	Total
N	663	387	343	1,393
	48	28	25	100
H	222	252	243	717
	31	35	34	100
Total	885	639	586	2,110
	42	30	28	100

This paper focuses on international (institutional) collaborations

- They have new dimensions
 - it could promote a research with strong public good nature than a domestic collaboration (access to more number of funding sources)
 - access to foreign talent (combination of foreign and domestic talents)
 - higher cost of coordination

Analytical Framework

- Scientific knowledge produced(→Private benefit of the project leader)

$$Y_j = Aq_j n_j^\beta \quad (1)$$

j : domestic collaboration (D) or international collaboration (INT)

-Team size (n) and quality of the team (q)

$$n_j = R_j + \alpha A \quad (3) \quad 0 < q_D < q_{INT} < 1 \quad (2)$$

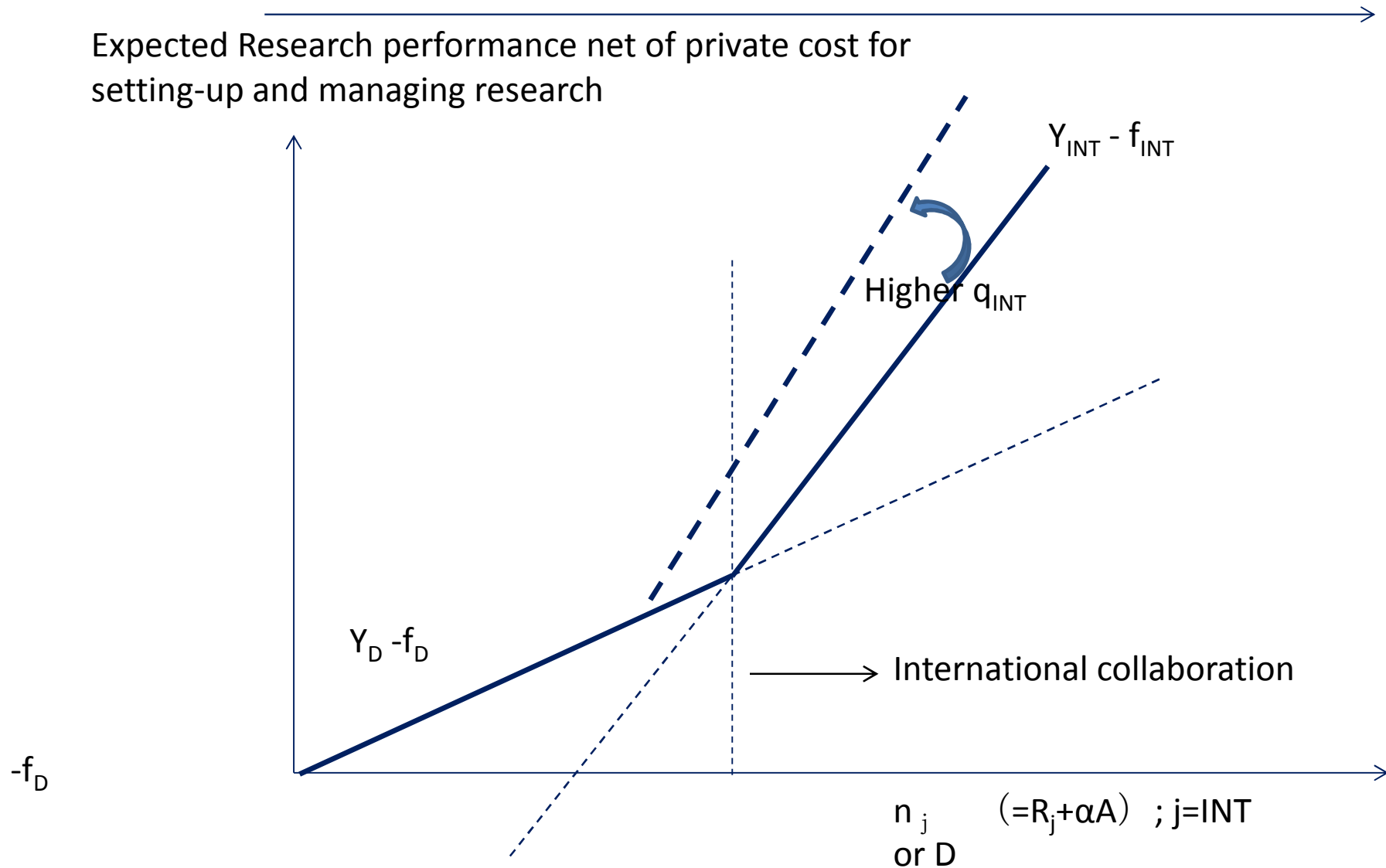
-Private cost of the project leader (fixed cost component)

$$0 < f_D < f_{INT} \quad (6)$$

-Net benefit from the project for the project leader

$$u_j = Y_j - f_j \quad (7)$$

Figure 1 Choice of international collaborations



Why international?

- International research collaboration is more likely to be chosen if
 - The project pursues fundamental research, and/or if it is not strongly driven by practical use considerations ($R_D < R_{INT}$)
 - It requires a large research team (large A),
 - It is triggered by foreign knowledge
 - The lead researcher has the experience of doing research abroad, prior to the project (lower f_{INT}).

Figure 3-1. Importance of fundamental understanding as a research objective (Very important, %)

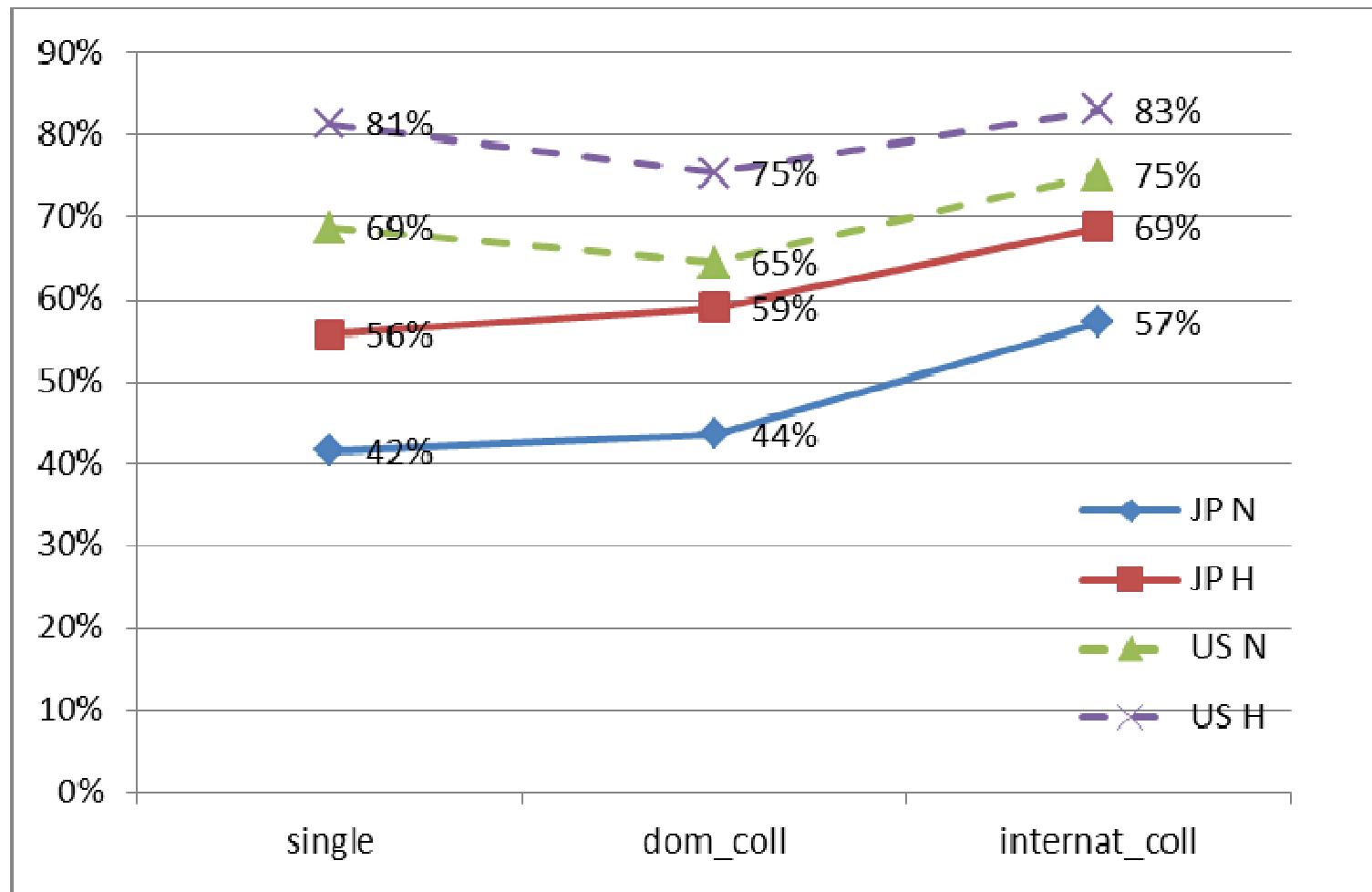


Figure 3-2. Importance of use considerations (Solving specific issues in real life as a research objective , Very important, %)

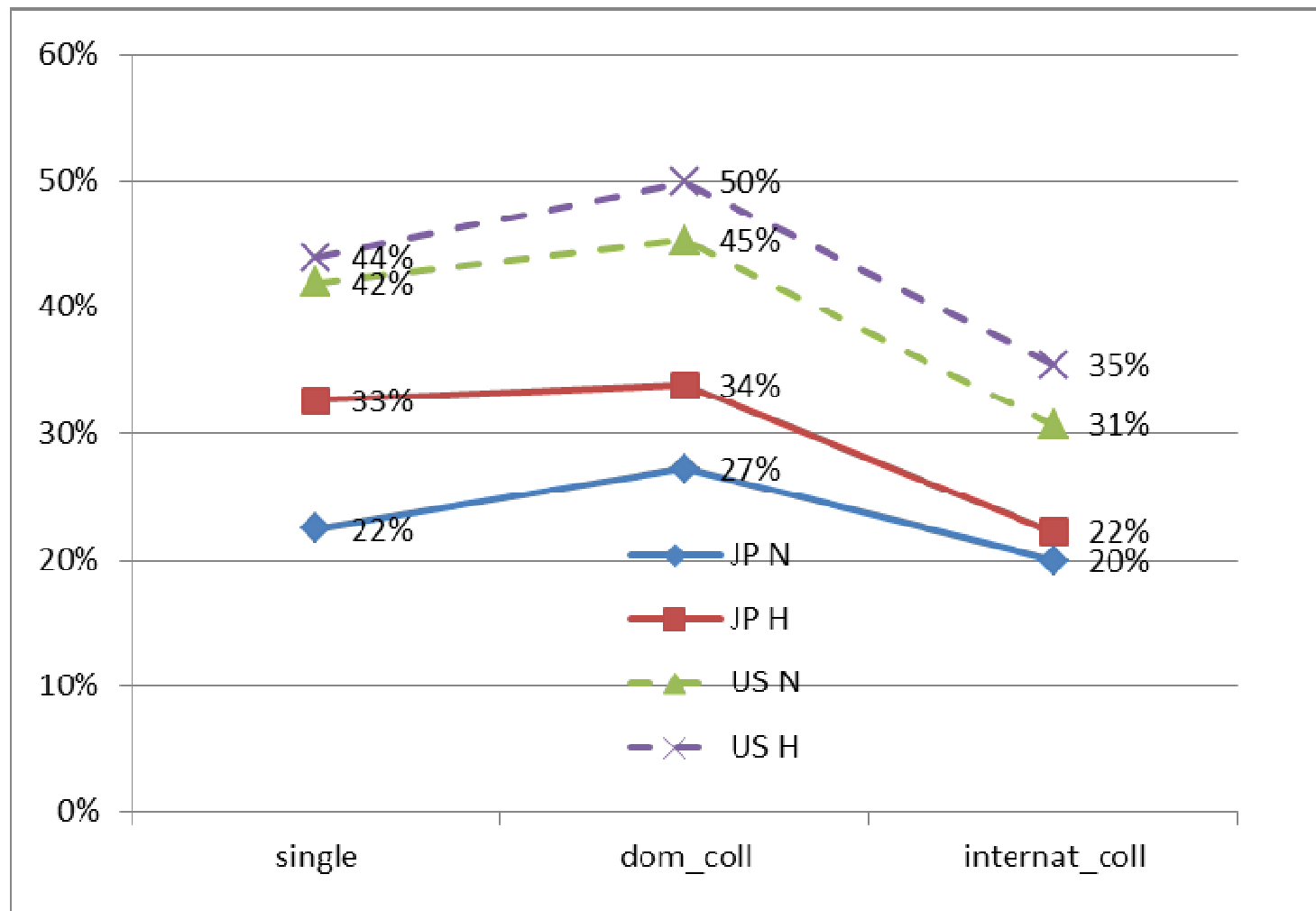


Figure 4-1. Team size and the incidence of international collaborations (JP, H paper)

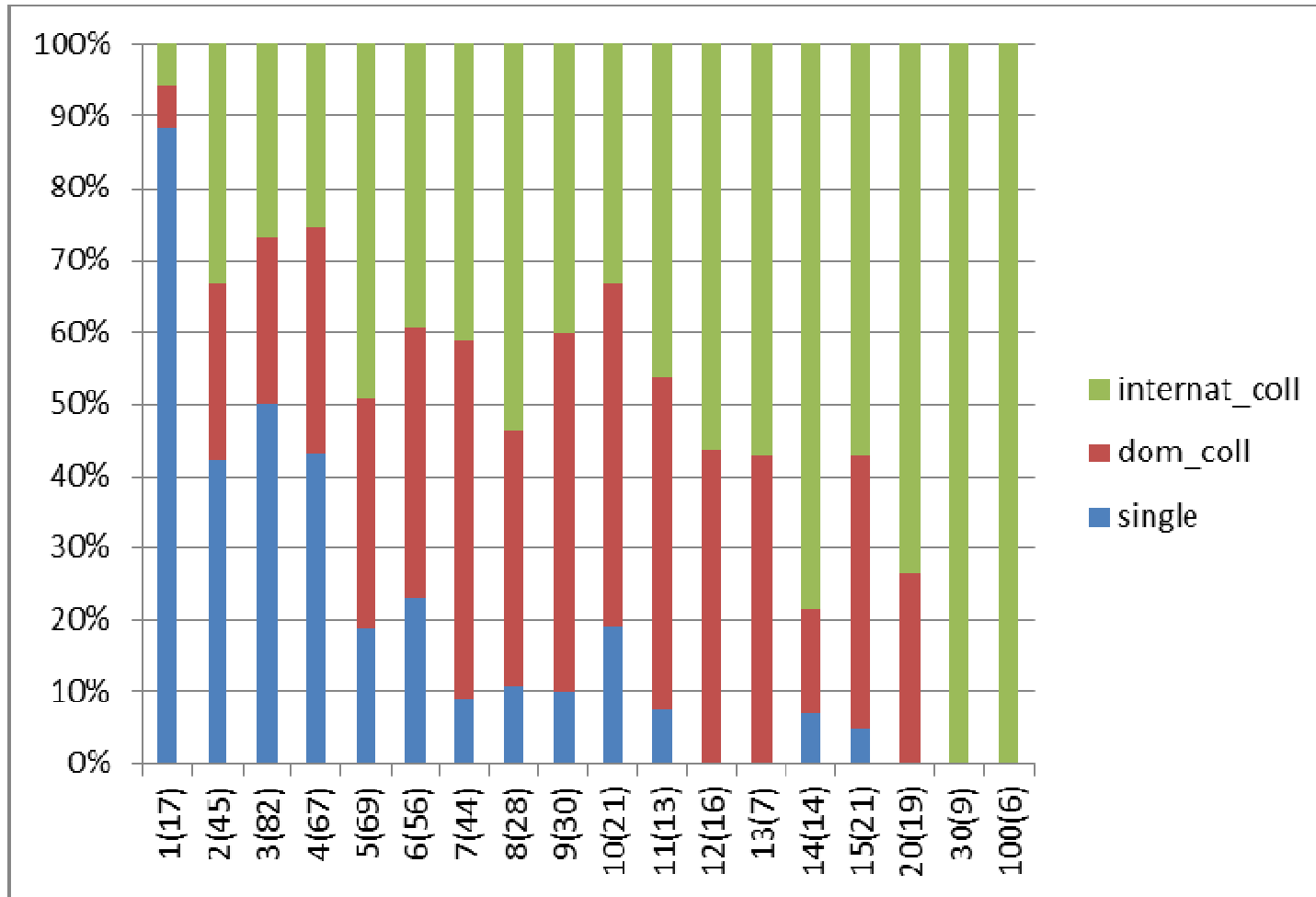


Figure 4-2. Team size and the incidence of international collaborations (US,H paper)

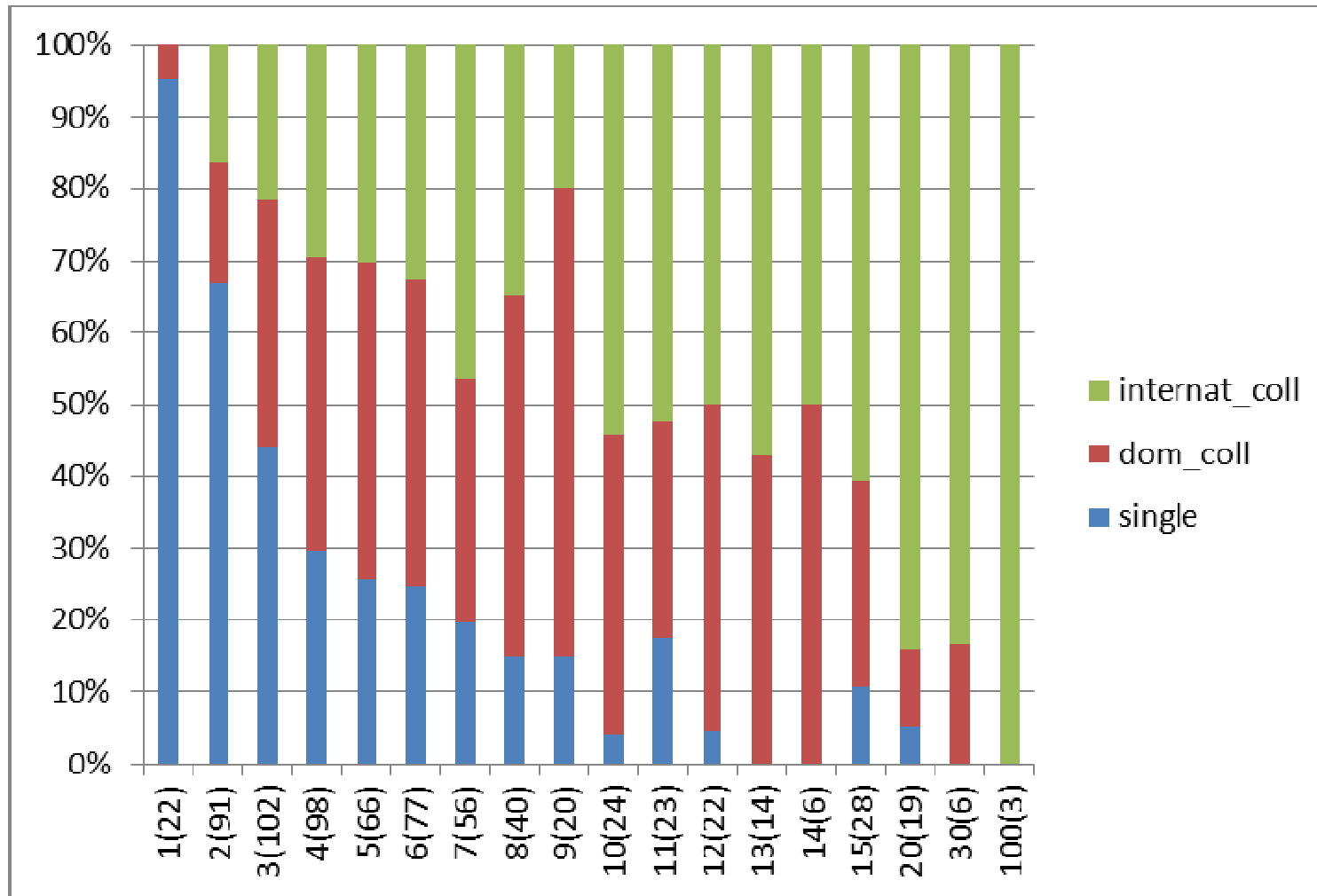
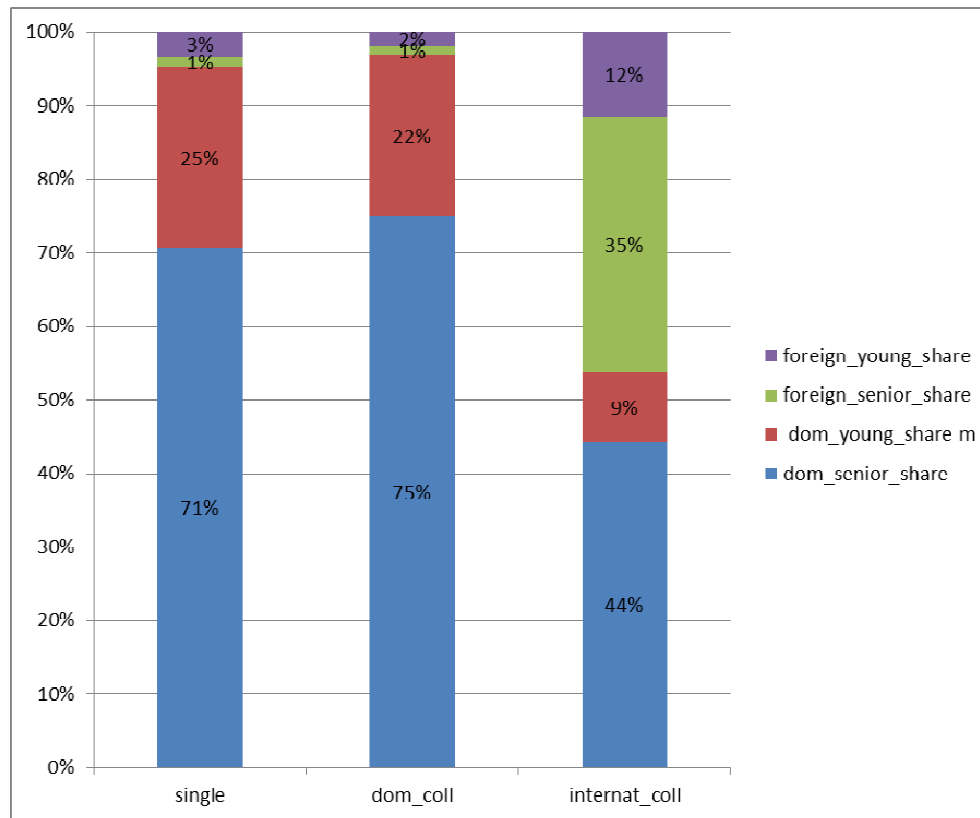
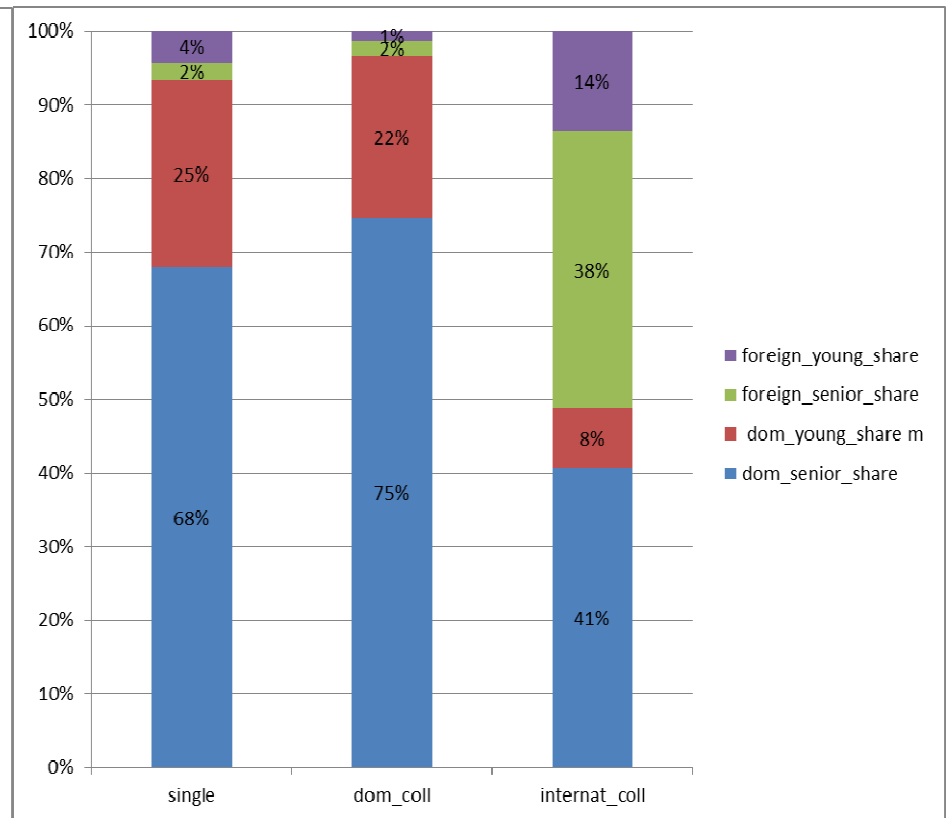


Figure 5-1. Foreign born scholars (JP)

(N paper)



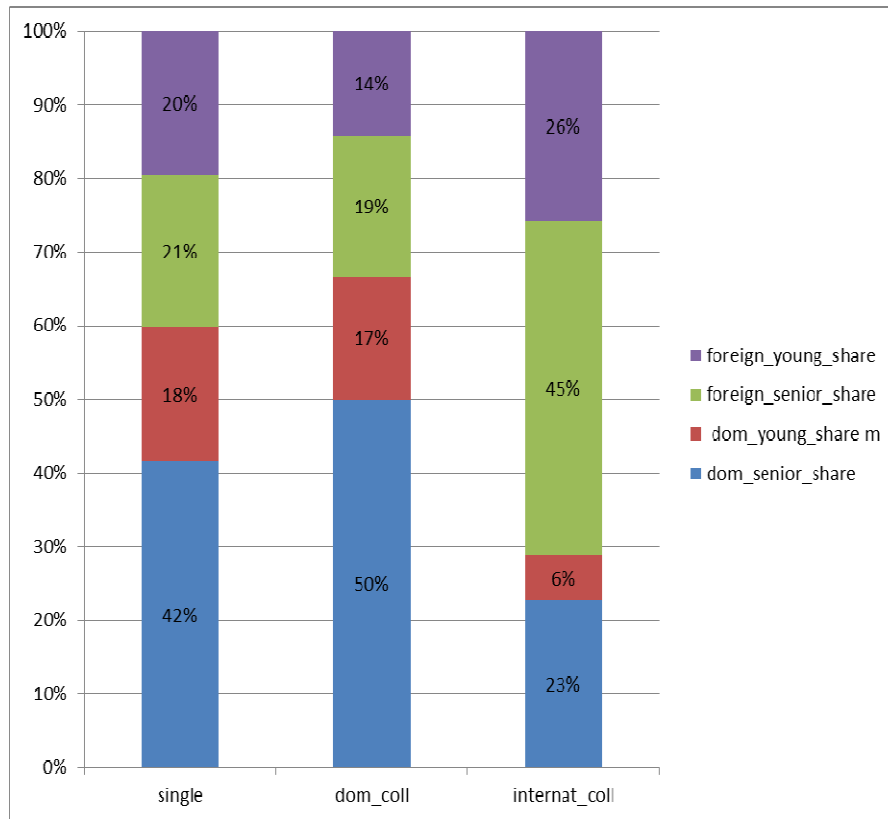
(H paper)



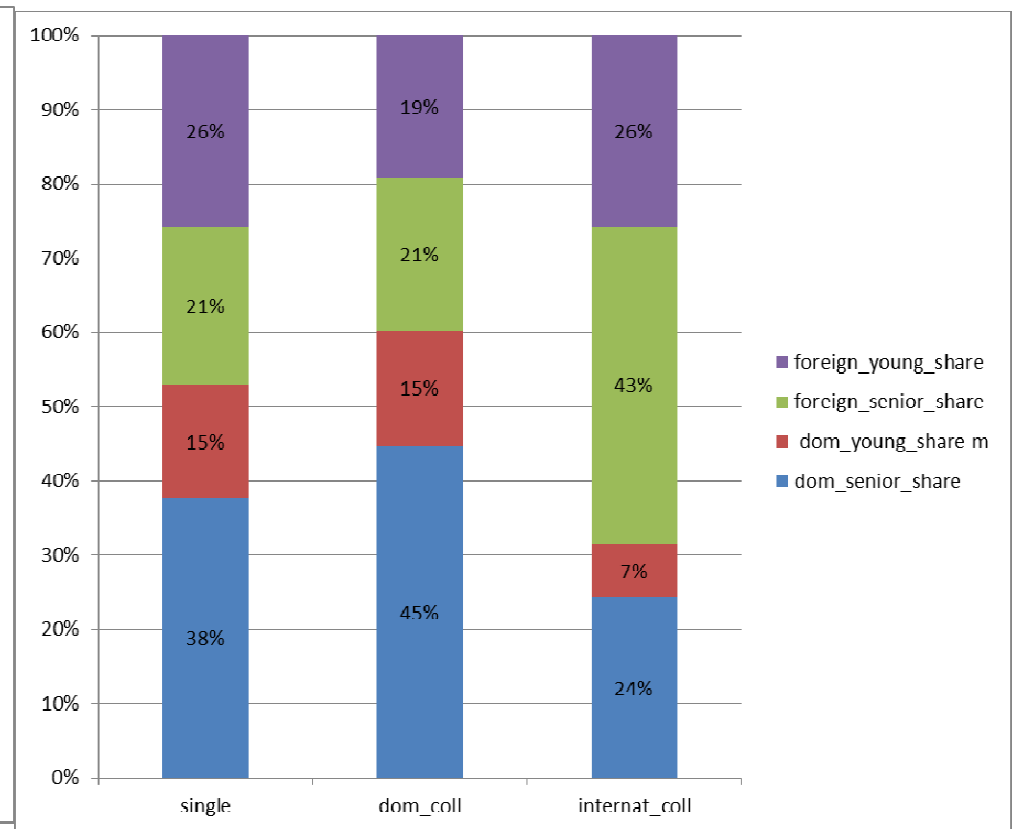
Note. The sample is restricted to those with 6 or less number of authors.

Figure 5-2. Foreign born scholars (US)

(N paper)



(H paper)



Note. The sample is restricted to those with 6 or less number of authors.

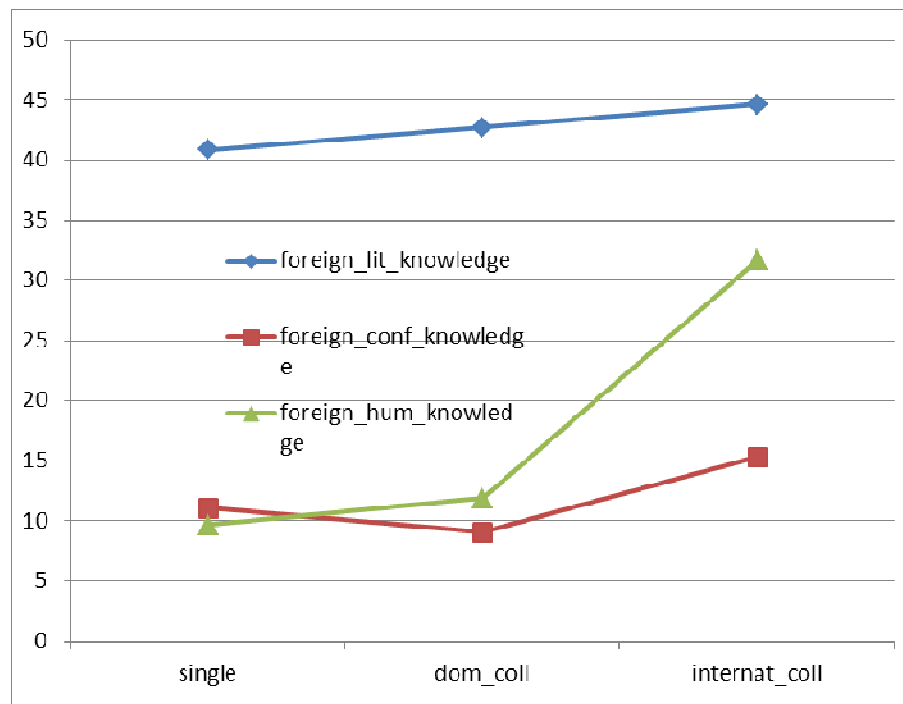
Table 3. Field and skill diversity of research team (US and Japan)

	JP	single	dom_coll	intemat_coll		US	single	dom_coll	intemat_coll
N	variety_field	1.32	1.61	1.54	N	variety_field	1.41	1.73	1.54
	variety_skill	1.18	1.31	1.41		variety_skill	1.22	1.35	1.34
	n. of obs.	589	350	216		n. of obs.	635	311	257
H	variety_field	1.40	1.78	1.51	H	variety_field	1.51	1.80	1.57
	variety_skill	1.13	1.39	1.30		variety_skill	1.27	1.37	1.36
	n. of obs.	130	95	111		n. of obs.	192	153	111

Note. The sample is restricted to those with 6 or less number of authors.

Figure 6-2. Location of three types of external knowledge sources initiating the research (Foreign location of important knowledge, %)

JP



US

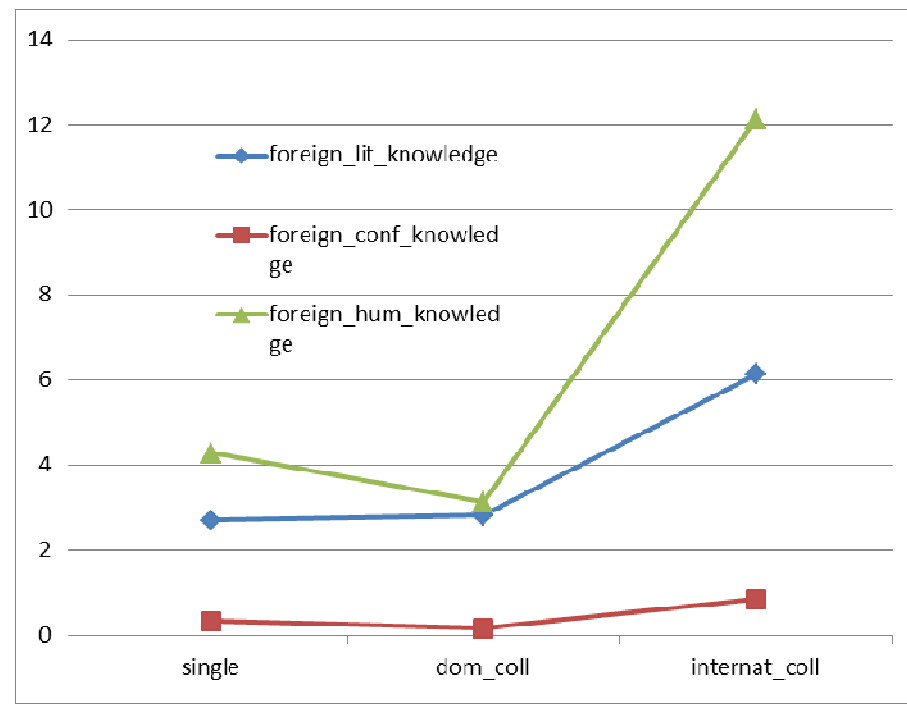


Table 5. Likelihood of an international research (the linear probability model : 1 international coll, 0 domestic coll)

	Japan				US			
	international	international	international	hlog for	international	international	international	hlog for
	Model1	Model2	Model3	Model4	Model5	Model6	Model7	Model8
fundamental_d	0.0838*** (0.0296)	0.0873*** (0.0306)	0.0796*** (0.0299)	0.0871*** (0.0322)	0.0721** (0.0340)	0.0763** (0.0352)	0.0752** (0.0353)	0.0372 (0.0379)
used	-0.0702** (0.0327)	-0.0725** (0.0338)	-0.0770** (0.0327)	-0.111*** (0.0371)	-0.106*** (0.0318)	-0.105*** (0.0336)	-0.106*** (0.0336)	-0.131*** (0.0349)
foreign_knowledge	0.0627** (0.0281)	0.0616** (0.0287)			0.264*** (0.0425)	0.264*** (0.0452)		
foreign_hum_knowledge			0.261*** (0.0351)	0.245*** (0.0405)			0.265*** (0.0503)	0.234*** (0.0537)
foreign_lit_knowledge			-0.0475 (0.0289)	-0.0338 (0.0321)			0.164** (0.0791)	0.226*** (0.0862)
foreign_conf_knowledge			0.0550 (0.0458)	0.0891 (0.0581)			0.212** (0.0869)	0.283* (0.148)
hauthors		0.0885*** (0.0208)	0.0813*** (0.0205)	0.412*** (0.0372)		0.165*** (0.0256)	0.167*** (0.0257)	0.362*** (0.0360)
hmanmonths		-0.00494 (0.0125)	-0.00620 (0.0120)	-0.0104 (0.0141)		0.0206 (0.0145)	0.0200 (0.0145)	0.0331* (0.0172)
hres_fund		-0.00641 (0.00815)	-0.0122 (0.00788)	-0.0115 (0.00875)		-0.0188* (0.00961)	-0.0184* (0.00961)	-0.0361*** (0.0117)
needs_adv equip_ext	-0.0238 (0.0326)	-0.0228 (0.0337)	-0.0280 (0.0329)	-0.0121 (0.0365)	-0.0203 (0.0333)	-0.0364 (0.0348)	-0.0397 (0.0350)	-0.0262 (0.0358)
needs_materials	-0.0200 (0.0347)	-0.0286 (0.0354)	-0.0304 (0.0345)	-0.0559 (0.0395)	0.0813** (0.0316)	0.0539 (0.0330)	0.0527 (0.0331)	0.0602* (0.0340)
abroad	0.121*** (0.0282)	0.123*** (0.0288)	0.106*** (0.0282)	0.111*** (0.0316)	0.148*** (0.0321)	0.129*** (0.0330)	0.127*** (0.0330)	0.183*** (0.0363)
award	0.0266 (0.0282)	0.0334 (0.0287)	0.0219 (0.0282)	-0.0102 (0.0314)	0.0246 (0.0332)	0.0108 (0.0346)	0.00927 (0.0346)	-0.00627 (0.0356)
citedness	0.126*** (0.0299)	0.106*** (0.0323)	0.107*** (0.0316)	0.134*** (0.0379)	0.00284 (0.0300)	-0.0494 (0.0327)	-0.0506 (0.0327)	-0.0563 (0.0345)
Observations	1,244	1,188	1,188	1,188	1,086	987	987	987
R-squared	0.090	0.101	0.142	0.324	0.113	0.142	0.143	0.260
Log Likelihood	-840.7	-795.4	-768.1	-900.4	-721.9	-640.6	-640.1	-683.1
Robust standard errors in parentheses. The coefficients for science fields and publication years are not presented.								
*** p<0.01, ** p<0.05, * p<0.1								

How do international collaborations contribute ?

- Selection effect.
- Larger team size
- Combining domestic and foreign talents for higher quality team.
- Effects would be larger in Japan than in the US, due to larger and more globalized scientists' labor market in the US

Estimation of knowledge production functions

- Performance
 - the level of forward citations of the focal paper
 - a number of papers from the project
- strength of two basic motivations for research
- Levels of domestic collaboration and international collaboration (*lnorg_dom*, *ln1org_for*)
- scale variables : the logarithm of the number of authors of the focal paper (*lnauthors*) , the logarithm of the man months spent for the project (*lnmanmonths*) and the logarithm of research money (*lnres_fund*)
- the composition of the team: measured by the shares of foreign born senior and junior scientists (*for_senior_share* and *for_young_share* respectively, *dom_young_share* represents the share of domestic young share, with the domestic senior share as a control)
- Fields diversity and research infrastructure
- 9 science field dummies and publication year dummies

Table 6-1 Knowledge production estimates
(dependent variable: forward citations)

		JP					US				
	VARIABLES	ln1fwdcitation_2009					ln1fwdcitation_2009				
		Model1	Model2	Model3	Model4	Model5	Model6	Model7	Model8	Model9	Model10
Objective	fundamental	0.123***	0.103***	0.0989***	0.101***	0.101***	0.109***	0.115***	0.101***	0.104***	0.104***
		(0.0279)	(0.0279)	(0.0280)	(0.0285)	(0.0286)	(0.0290)	(0.0286)	(0.0300)	(0.0313)	(0.0313)
	use	0.0186	-0.0154	-0.0138	-0.0131	-0.0110	-0.00331	-0.0320*	-0.0234	-0.0195	-0.0177
		(0.0176)	(0.0181)	(0.0182)	(0.0186)	(0.0186)	(0.0173)	(0.0173)	(0.0182)	(0.0186)	(0.0186)
Organizations	horg_dom	0.204***	0.0735	0.124**	0.131**	0.143***	0.163***	-0.0296	-0.00126	0.00148	-0.00548
		(0.0429)	(0.0490)	(0.0502)	(0.0508)	(0.0512)	(0.0409)	(0.0467)	(0.0498)	(0.0517)	(0.0521)
	h1org_for	0.360***	0.196***	0.0939	0.0878	-0.0315	0.142***	-0.0451	-0.0529	-0.0486	0.00731
		(0.0402)	(0.0471)	(0.0666)	(0.0673)	(0.0901)	(0.0441)	(0.0490)	(0.0553)	(0.0569)	(0.112)
Number of authors	hauthors		0.164***	0.146***	0.145***	0.160***		0.313***	0.285***	0.298***	0.277***
			(0.0456)	(0.0471)	(0.0485)	(0.0492)		(0.0451)	(0.0479)	(0.0523)	(0.0532)
Man months	hman months										
Money	hres_fund		0.0785***	0.0772***	0.0763***	0.0787***		0.0656***	0.0677***	0.0686***	0.0681***
			(0.0111)	(0.0111)	(0.0112)	(0.0112)		(0.0119)	(0.0125)	(0.0128)	(0.0127)

Note. 9 science field dummies and publication year dummies are introduced as controls.
Robust standard errors. Ln(2)=0.693

Table6-2 Knowledge production estimates (dependent variable: forward citations)

[illegible]

Figure 7-1. Size of the coefficients of the variable for international collaborations (citation performance)

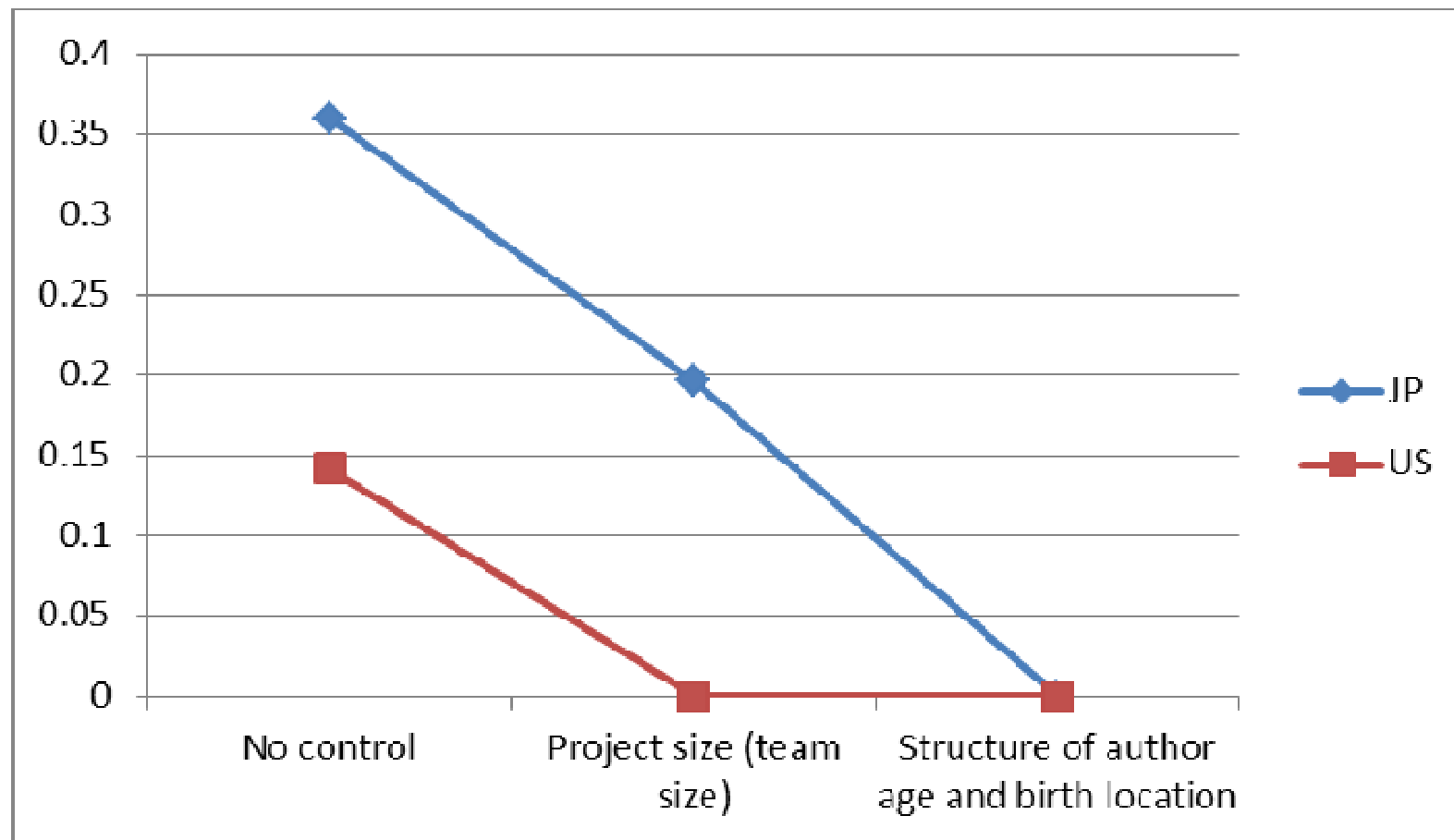


Table 7-1. Knowledge production estimates
(dependent variable: the number of papers from the
project)

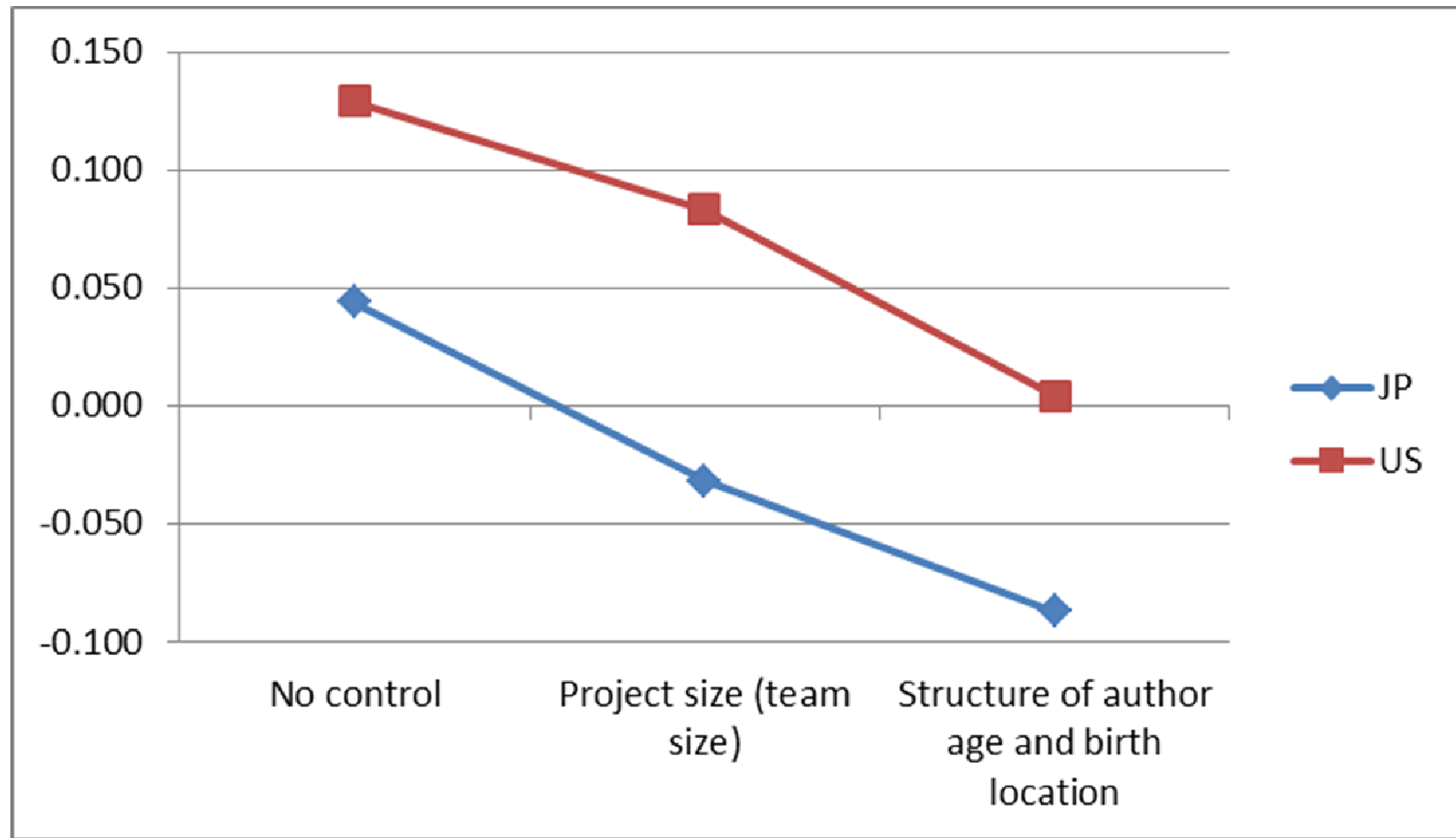
		JP					US				
	VARIABLES	h1papers_en					h1papers_en				
		Model11	Model12	Model13	Model14	Model15	Model16	Model17	Model18	Model19	Model20
Objective	fundamental	0.231***	0.127***	0.119***	0.111***	0.112***	0.0845***	0.0608***	0.0587**	0.0525**	0.0532**
		(0.0299)	(0.0279)	(0.0276)	(0.0276)	(0.0277)	(0.0259)	(0.0233)	(0.0235)	(0.0245)	(0.0247)
	use	0.0725***	-0.0125	-0.0149	-0.0136	-0.0147	0.0683***	0.0300*	0.0253	0.0199	0.0210
		(0.0209)	(0.0185)	(0.0187)	(0.0190)	(0.0191)	(0.0172)	(0.0161)	(0.0168)	(0.0170)	(0.0170)
Organizations	horg_dom	0.114**	-0.00379	0.00433	0.000225	0.00645	0.129**	0.0244	0.0309	0.0369	0.0350
		(0.0483)	(0.0492)	(0.0504)	(0.0510)	(0.0512)	(0.0503)	(0.0441)	(0.0459)	(0.0465)	(0.0471)
	horg_for	0.0439	-0.0319	-0.0869	-0.0826	4.43e-05	0.129***	0.0833*	0.00430	0.00742	0.238*
		(0.0580)	(0.0554)	(0.0885)	(0.0886)	(0.117)	(0.0484)	(0.0475)	(0.0549)	(0.0564)	(0.126)
Number of authors	hauthors		-0.0427	-0.0410	-0.0428	-0.0498		-0.0121	0.0178	0.0105	0.00148
			(0.0484)	(0.0499)	(0.0521)	(0.0526)		(0.0387)	(0.0420)	(0.0458)	(0.0465)
Man months	hman months		0.167***	0.167***	0.158***	0.156***		0.194***	0.192***	0.190***	0.188***
			(0.0245)	(0.0246)	(0.0235)	(0.0234)		(0.0207)	(0.0218)	(0.0224)	(0.0221)
Money	hres_fund		0.204***	0.209***	0.212***	0.211***		0.120***	0.113***	0.108***	0.108***
			(0.0144)	(0.0145)	(0.0143)	(0.0143)		(0.0135)	(0.0141)	(0.0142)	(0.0140)

Note. 9 science field dummies and publication year dummies are introduced as controls.
Robust standard errors.

Table7-2 Knowledge production estimates (dependent variable: the number of papers from the project)

		JP					US				
	VARIABLES	h1papers_en					h1papers_en				
		Model11	Model12	Model13	Model14	Model15	Model16	Model17	Model18	Model19	Model20
Structure of authors in age and birth countries	dom_young_share			0.0263 (0.107)	0.0303 (0.108)	0.0290 (0.108)			-0.0426 (0.0991)	-0.0371 (0.104)	-0.0485 (0.104)
	intn'lcoll					-0.218* (0.120)					-0.223 (0.166)
	for_senior_share			0.188 (0.210)	0.177 (0.209)	0.281 (0.217)			0.207*** (0.0744)	0.194** (0.0770)	0.231*** (0.0847)
	intn'lcoll										-0.0733 (0.180)
	for_young_share			0.0595 (0.185)	0.0436 (0.185)	0.110 (0.183)			0.0102 (0.0896)	-0.00765 (0.0908)	-0.0254 (0.0986)
	intn'lcoll										0.103 (0.217)
Diversity of authors	hdiv_field				-0.0342 (0.0549)	-0.0553 (0.0633)				0.0149 (0.0491)	0.0273 (0.0569)
	intn'lcoll					0.0903 (0.115)					-0.00386 (0.104)
	hdiv_skill				0.0295 (0.0729)	-0.00468 (0.0860)				0.0200 (0.0642)	0.0137 (0.0746)
	intn'lcoll					0.127 (0.155)					0.0287 (0.137)
Research infra	p_adv equip_ext				0.141** (0.0628)	0.136* (0.0742)				0.126*** (0.0463)	0.101* (0.0545)
	intn'lcoll					0.00253 (0.137)					0.0694 (0.101)
	p_materials				-0.0166 (0.0657)	-0.00134 (0.0772)				-0.00709 (0.0425)	0.0257 (0.0480)
	intn'lcoll					-0.0490 (0.133)					-0.122 (0.0930)
sum of size coefficients			0.33	0.34	0.33	0.32		0.30	0.32	0.31	0.30
Sample	citedness	0.635*** (0.0609)	0.373*** (0.0575)	0.371*** (0.0578)	0.370*** (0.0576)	0.371*** (0.0575)	0.384*** (0.0453)	0.238*** (0.0425)	0.233*** (0.0438)	0.236*** (0.0442)	0.237*** (0.0443)
	Observations	2,023	1,947	1,937	1,915	1,915	1,987	1,843	1,711	1,654	1,654
	R-squared	0.149	0.367	0.374	0.376	0.378	0.093	0.319	0.318	0.318	0.322
	Log Likelihood	-3069	-2660	-2636	-2596	-2593	-2574	-2121	-1950	-1888	-1883
Robust standard errors in parentheses											
*** p<0.01, ** p<0.05, * p<0.1											

Figure 8-1. Size of the coefficients of the variable for international collaborations(the number of papers from the project)



Conclusions

- International institutional collaborations are more engaged, when
 - (1) the project pursues fundamental research, and/or if it is not strongly driven by practical use considerations,
 - (2) it requires a large research team,
 - (3) it is triggered by foreign knowledge and
 - (4) the lead researcher has the experience of doing research abroad prior to the project.

Conclusions (continued)

- Papers from international collaborations are significantly more cited than those from domestic collaborations in Japan (not in the US) and a project from an international collaboration produces significantly more papers in the US (not in Japan)
- An international collaboration increases the level of forward citations both by expanding the team size and by combining domestic and foreign talents in Japan but only by expanding the team size in the US.
- International collaborations results in more papers both by expanding the team size and by combining the domestic and foreign talents (especially foreign senior talents) in the US.

Appendix. Hitotsubashi-NISTEP-Georgia Tech scientist's survey : Purpose of the research

- Developing systematic and objective data on the knowledge creation process in science at the project level (Japan and the US).
 - Motivations of the research projects
 - Knowledge sources that inspired the project
 - Uncertainty in the knowledge creation process
 - Research competition
 - Composition of the research team
 - Sources of funds
 - Research outputs, e.g., research papers, students, patents, etc.

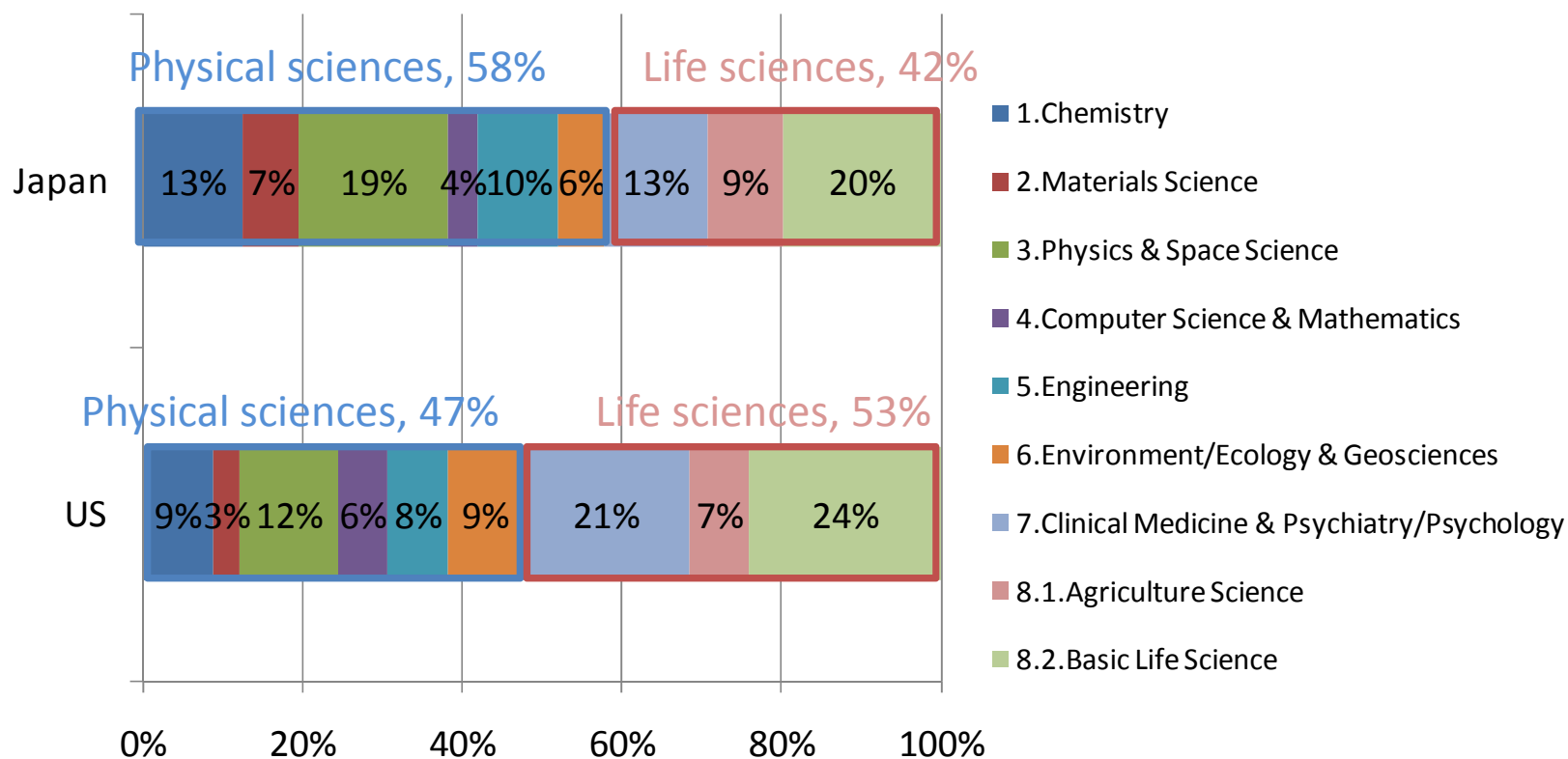
The population of the survey

- Articles and letters in the Science Citation Indexes-Expanded (Thomson Reuters)
- Time window: 2001 – 2006 (database year)
- Forward citation counts were retrieved on December 31, 2006
- 22 fields in the ESI were adopted

Identification of possible focal papers

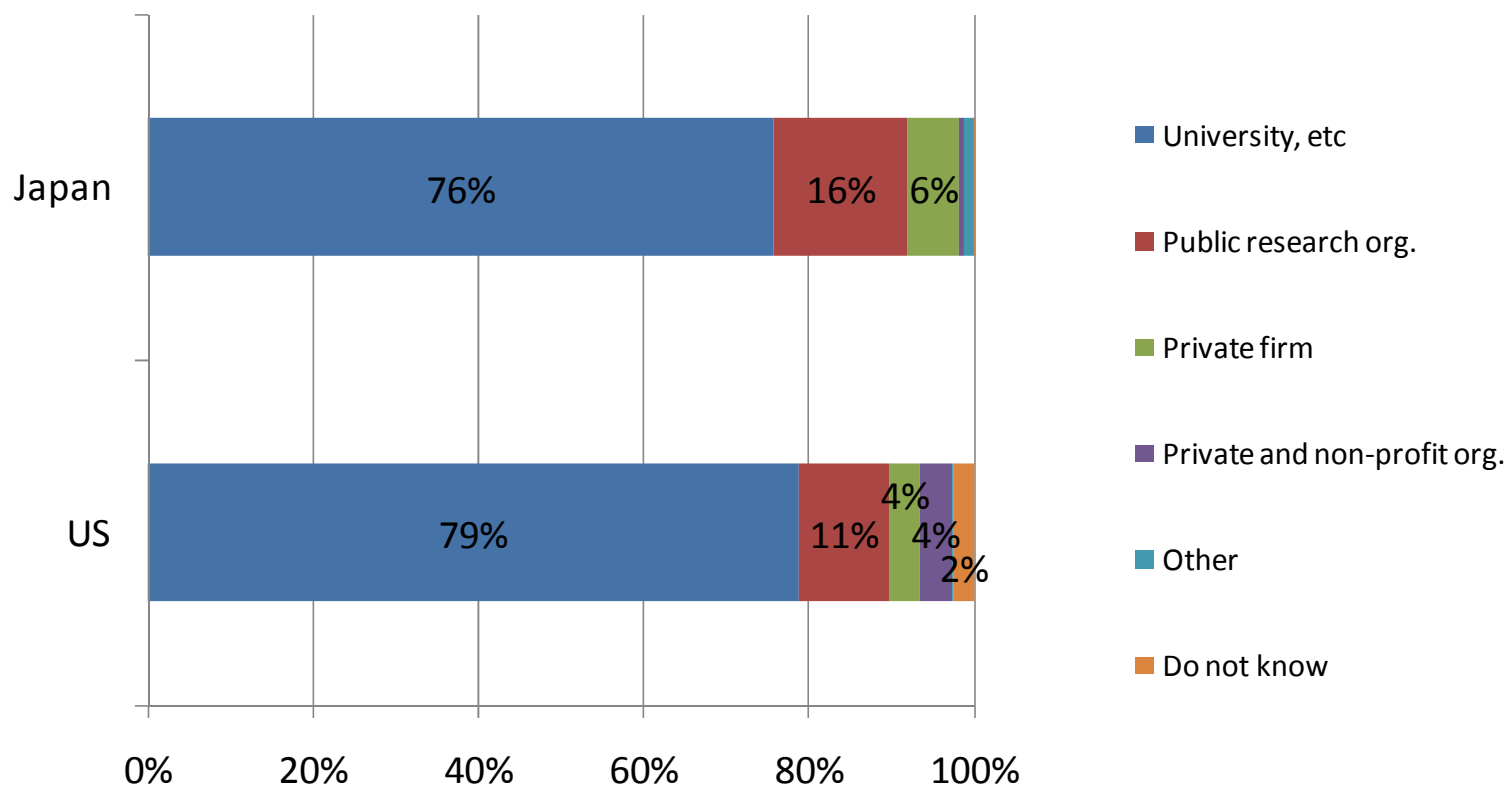
- Highly Cited Papers (H papers)
 - Top 1% highly cited papers in each journal field (22 fields in total) and in each database year; at least one organization of authors should be located in the target country (approximately 3,000 in total).
- Normal Papers (N papers)
 - Randomly selected papers in each journal field and in each database year from the population of the survey, excluding the above highly cited papers; at least one organization of authors should be located in the target country (approximately 7,000).

Composition of the field of science



Note1: Results of all respondents

Sector of respondents



Note1: Results of all respondents

Table A-1 Descriptive statistics (based on Model 1 and 5)

Variable	JP					US				
	Obs	Mean	Std. Dev	Min	Max	Obs	Mean	Std. Dev	Min	Max
international	1244	0.46	0.50	0	1	1086	0.48	0.50	0	1
lnlog_for	1244	0.48	0.65	0	4.29	1086	0.47	0.57	0	3.85
fundamental_d	1244	0.54	0.50	0	1	1086	0.73	0.44	0	1
use_d	1244	0.25	0.44	0	1	1086	0.40	0.49	0	1
foreign_knowledge	1244	0.54	0.50	0	1	1086	0.11	0.32	0	1
foreign_hum_knowledge	1244	0.21	0.41	0	1	1086	0.08	0.27	0	1
foreign_lit_knowledge	1244	0.44	0.50	0	1	1086	0.04	0.20	0	1
foreign_conf_knowledge	1244	0.12	0.32	0	1	1086	0.01	0.07	0	1
lnauthors	1244	1.72	0.68	0	5.79	1086	1.64	0.66	0	5.93
lnmanonths	1224	4.46	1.39	0	11.41	1002	3.48	1.38	0	11.70
lnres_fund	1194	11.73	2.26	4.61	19.52	1047	11.90	2.10	8.52	18.42
needs_adv equip_ext	1244	0.27	0.44	0	1	1086	0.31	0.46	0	1
needs_materials	1244	0.27	0.44	0	1	1086	0.53	0.50	0	1
abroad	1244	0.52	0.50	0	1	1086	0.31	0.46	0	1
award	1244	0.48	0.50	0	1	1086	0.28	0.45	0	1
citedness	1244	0.33	0.47	0	1	1086	0.41	0.49	0	1
publication_year	1244	2003.4	1.68	2000	2007	1086	2003.5	1.71	2000	2006

Figure 2-1 Collaboration structure by field, % (JP, H papers)

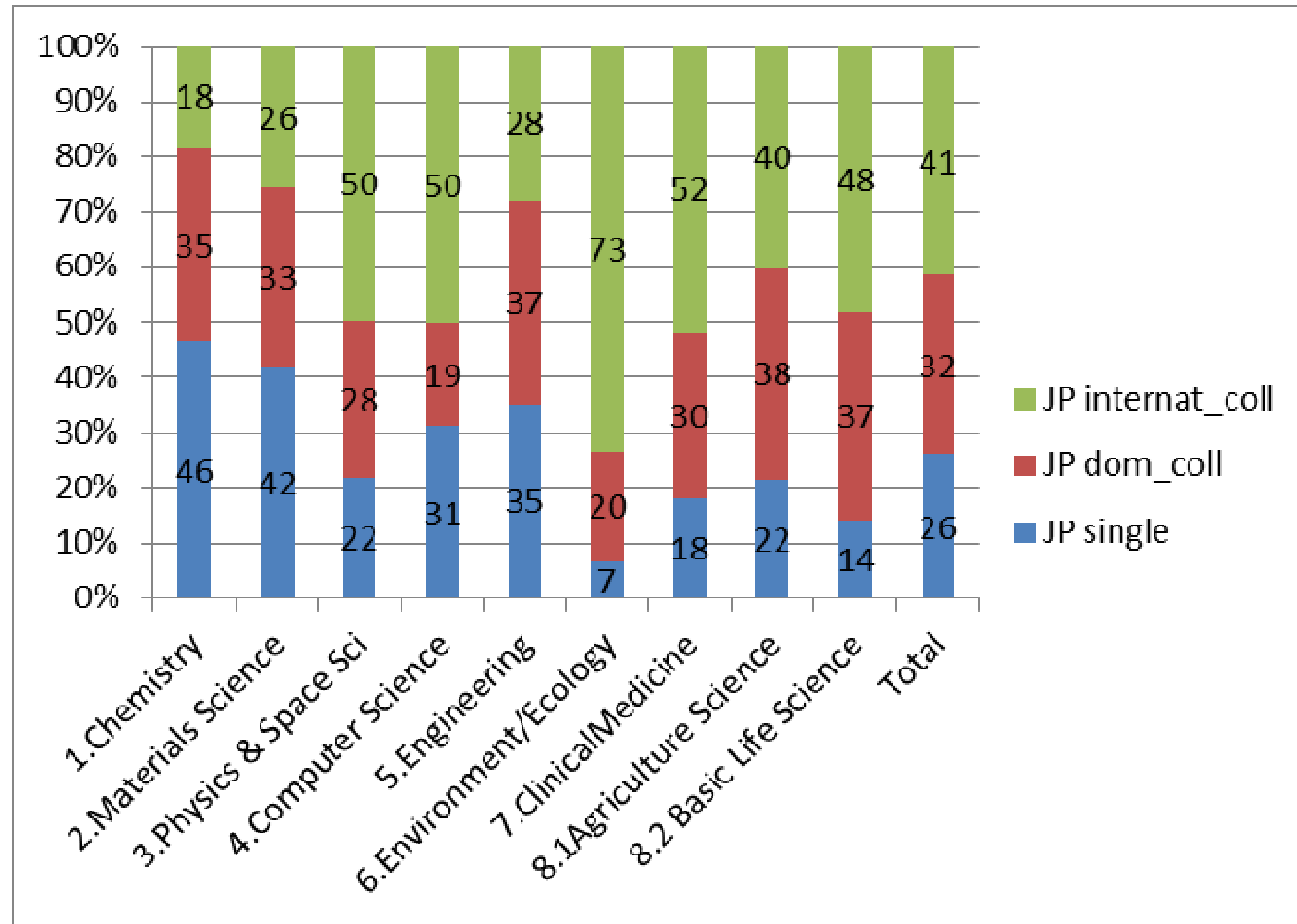


Figure 2-2. Collaboration structure by field, % (US, H papers)

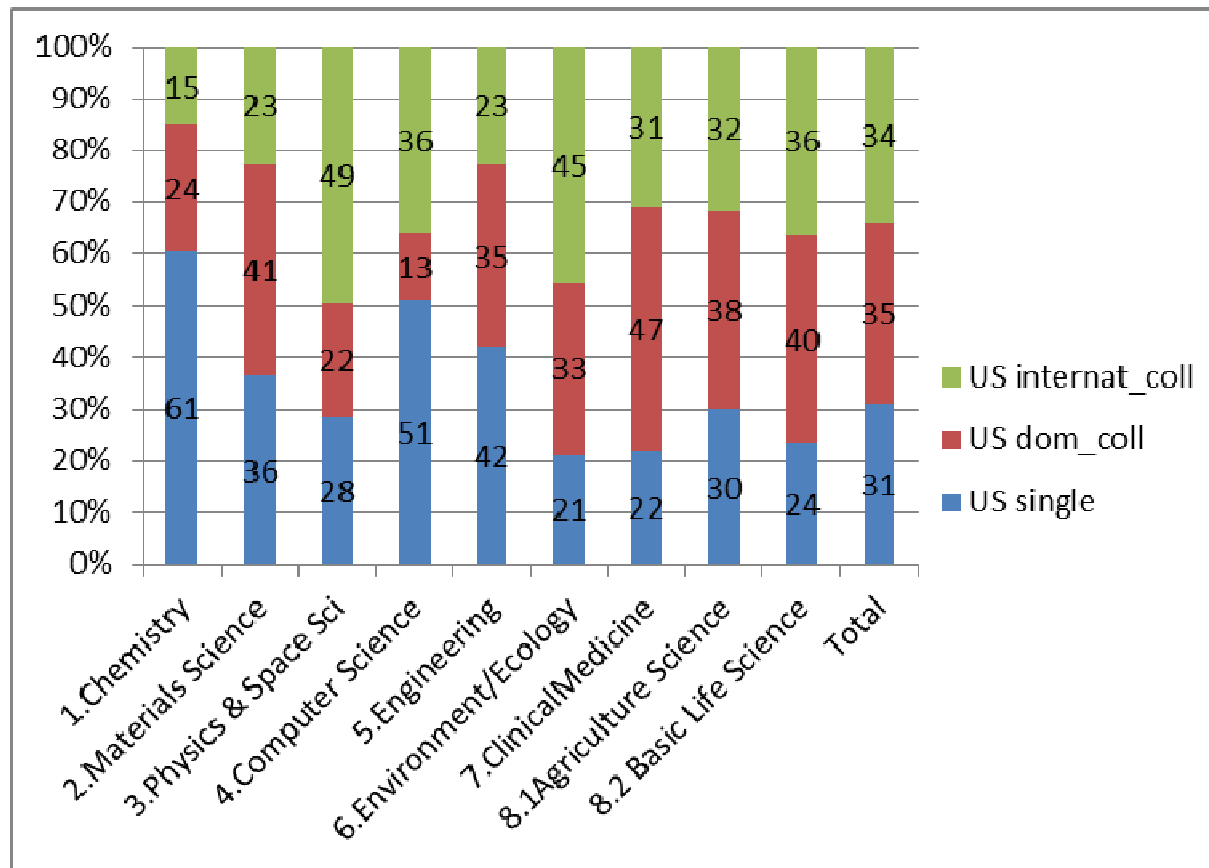


Table 2-1. Scale of a team (Japan)

JP, N	authors, average	authors, m ed	N
single	3.5	3	638
dom_coll	5.3	5	491
intemat_coll	7.5	5	346

JP, H	authors, average	authors, m ed	N
single	4.0	3	147
dom_coll	7.3	6	183
intemat_coll	16.0	7	234

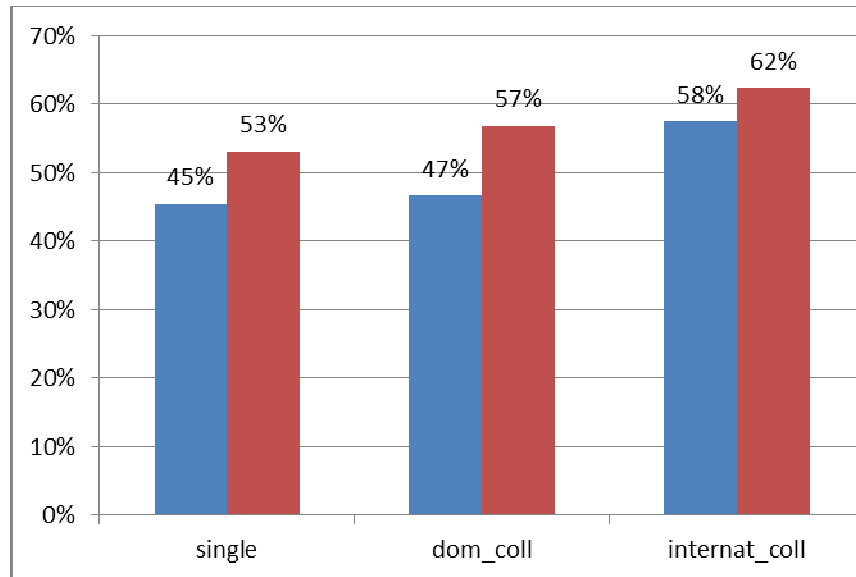
Table 2-2. Scale of a team (US)

US, N	authors, average	authors, m ed	N
single	3.1	3	663
dom_coll	4.8	4	387
intemat_coll	5.6	4	343

US, H	authors, average	authors, m ed	N
single	4.0	3	222
dom_coll	6.7	6	252
intemat_coll	12.3	7	243

Figure 6-1. Location of external knowledge sources
(foreign location of important knowledge, %)

JP



US

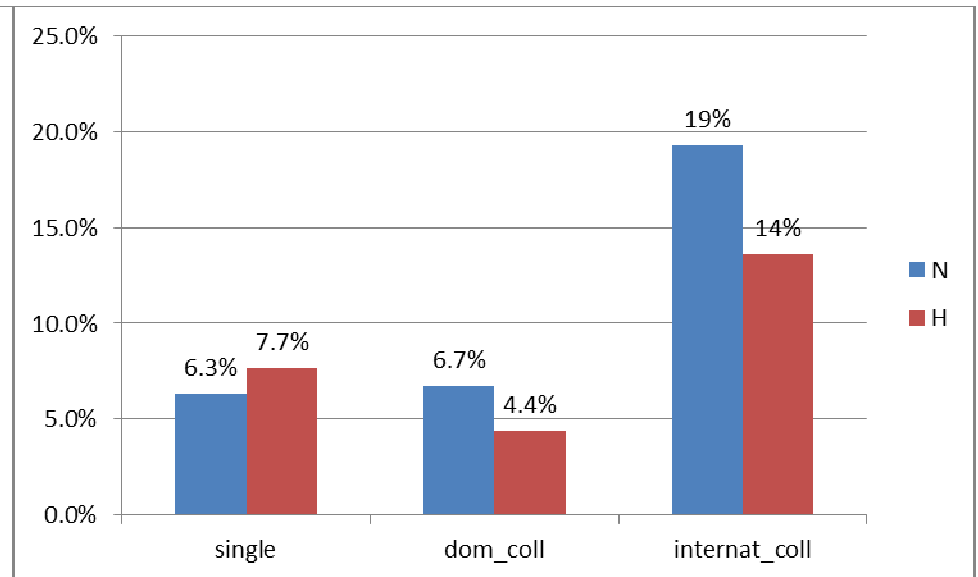


Figure 7-2. Size of the coefficients of the variable for domestic collaborations (citation performance)

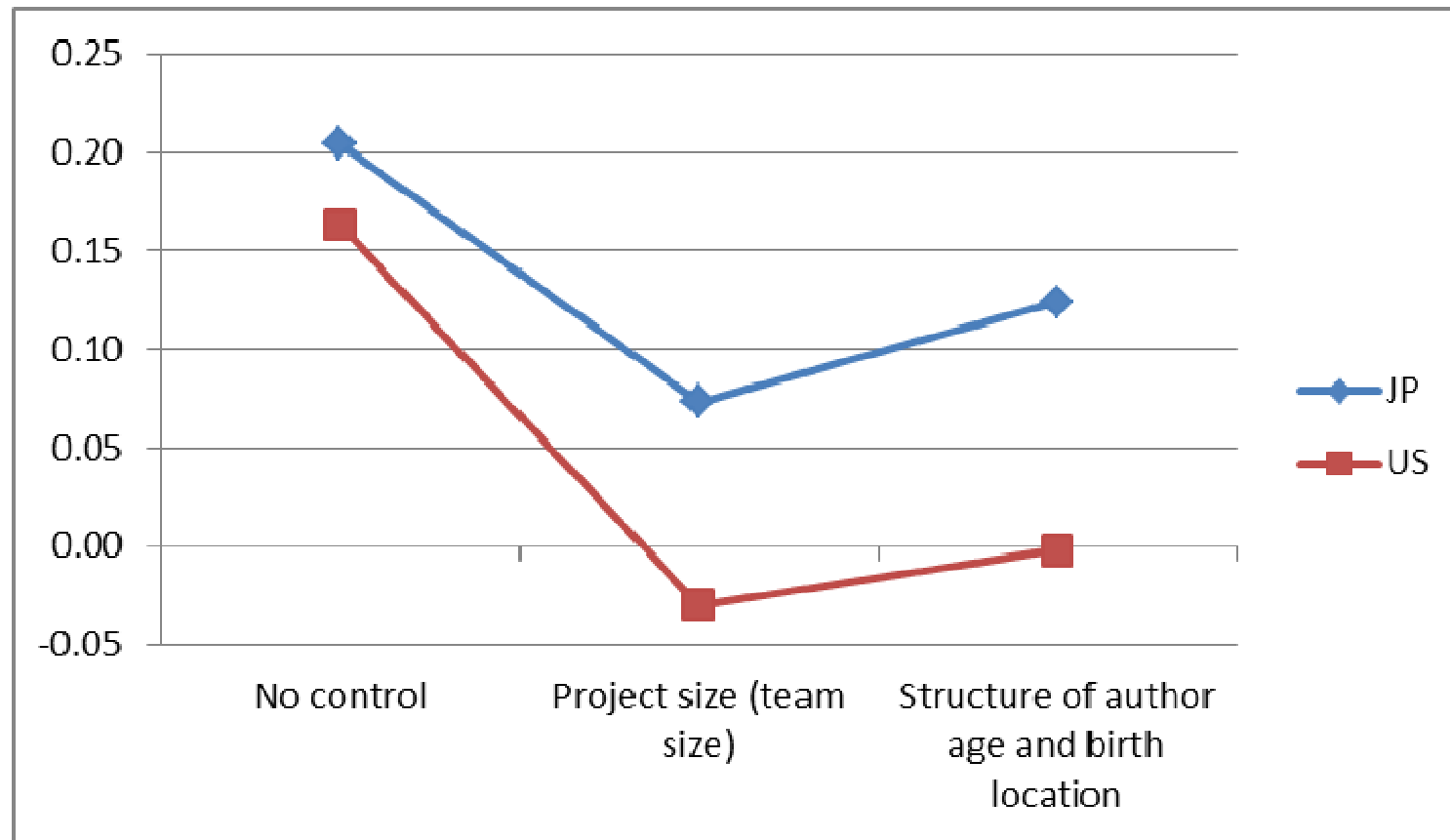


Figure 8-2 Size of the coefficient of the variable for domestic collaborations (the number of papers from the project)

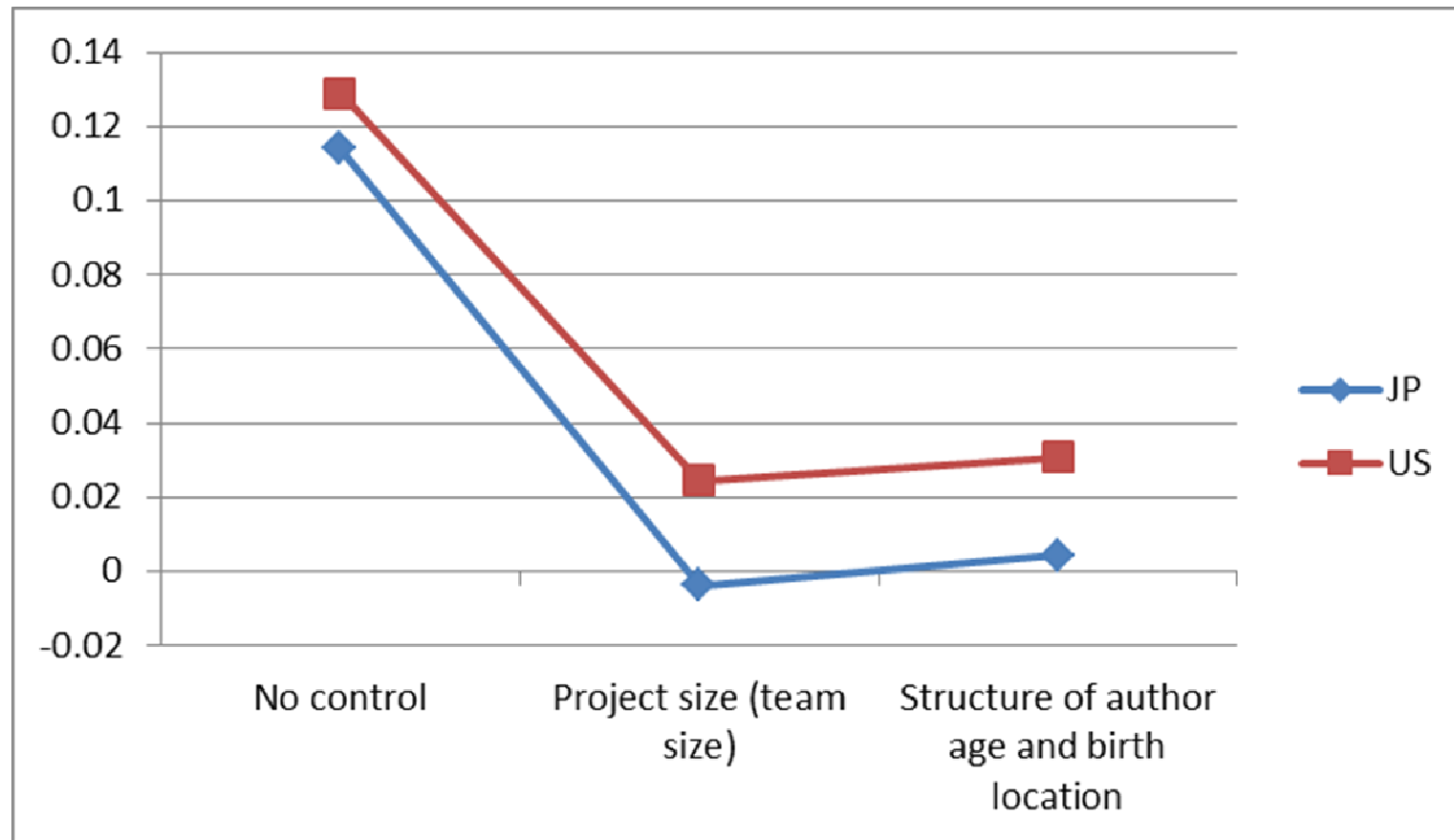


Table 4-1. Research money and infrastructure used for a project (JP)

JP, N	res_fund, million, average	res_fund, million, median	advanced external equipments	materials DB	remote researchers
single	0.5	0.05	19%	14%	12%
dom_coll	1.3	0.08	23%	15%	22%
intemat_coll	0.9	0.08	27%	21%	54%

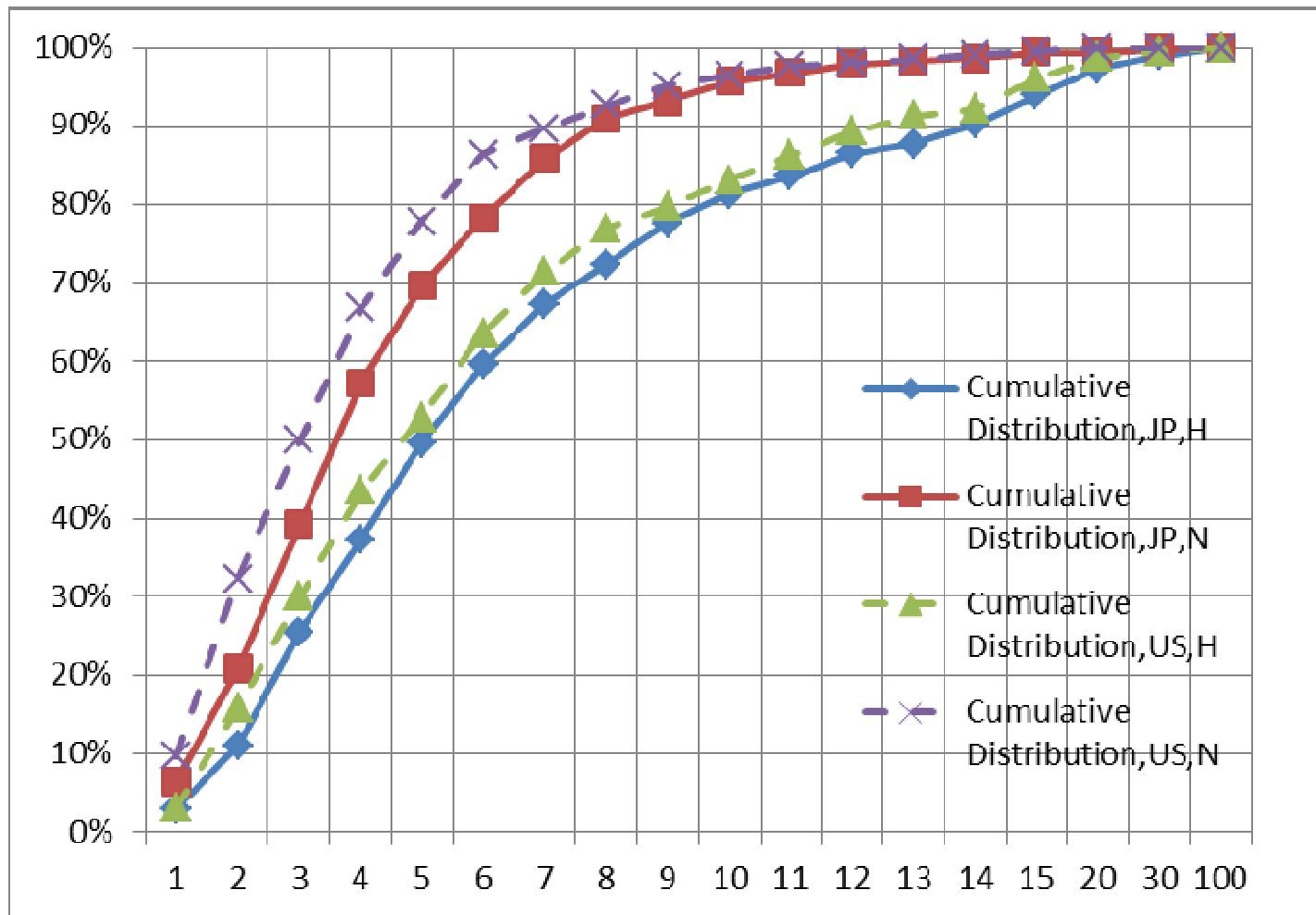
JP, H	res_fund, million, average	res_fund, million, median	advanced external equipments	materials DB	remote researchers
single	1.7	0.30	16%	15%	16%
dom_coll	2.2	0.40	27%	26%	25%
intemat_coll	3.3	0.20	24%	23%	58%

Table 4-2. Research money and infrastructure used for a project (US)

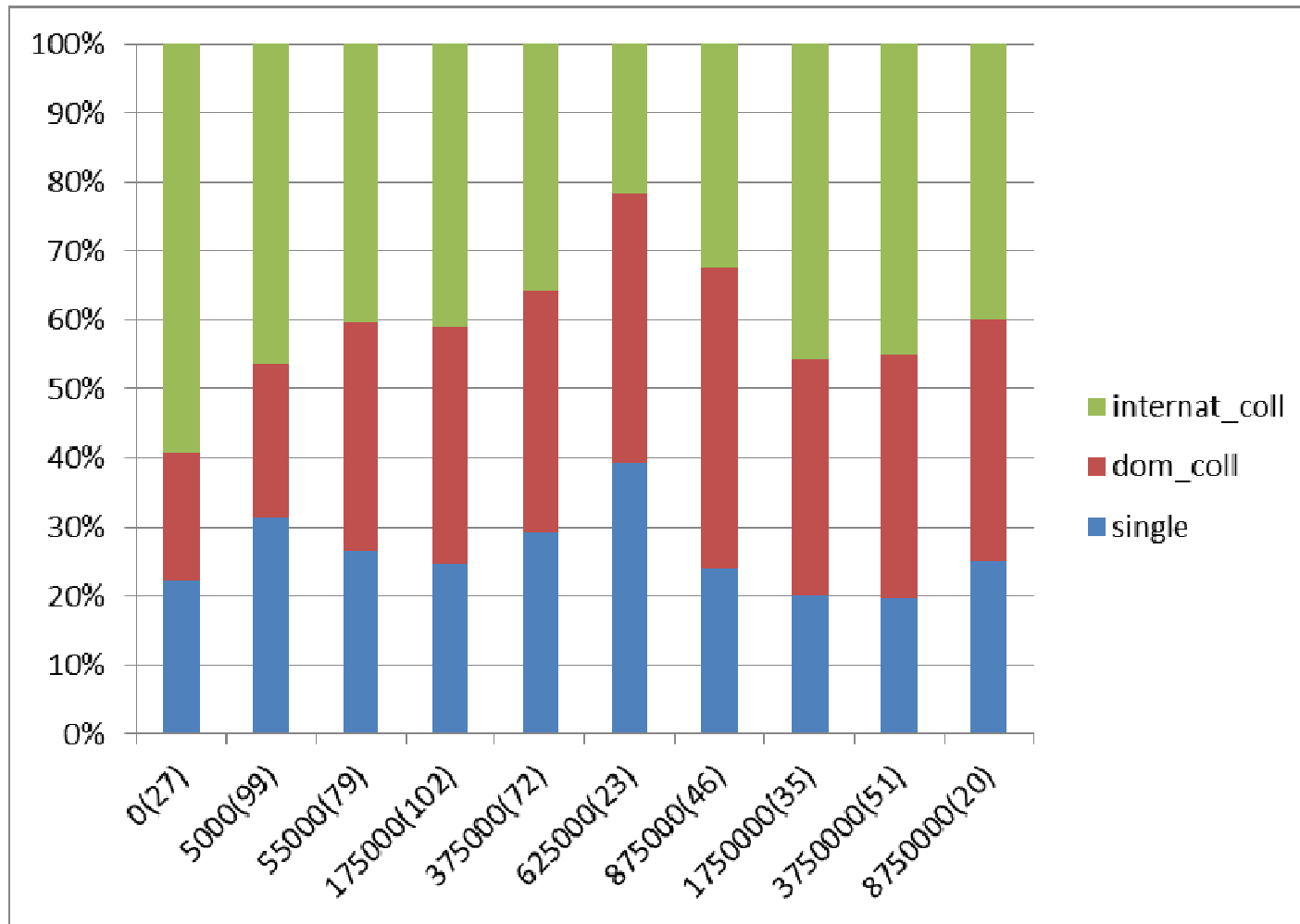
US, N	res_fund, million, average	res_fund, million, median	advanced external equipments	materials DB	remote researchers
single	1.6	0.06	26%	48%	53%
dom_coll	1.6	0.06	28%	47%	59%
intemat_coll	2.2	0.06	33%	53%	68%

US, H	res_fund, million, average	res_fund, million, median	advanced external equipments	materials DB	remote researchers
single	2.4	0.18	23%	53%	60%
dom_coll	4.7	0.18	33%	54%	66%
intemat_coll	4.5	0.18	36%	62%	75%

Distribution of team size (US and Japan)



Research money size and the incidence of international collaborations (JP, H paper)



Research money size and the incidence of international collaborations (US, H paper)

