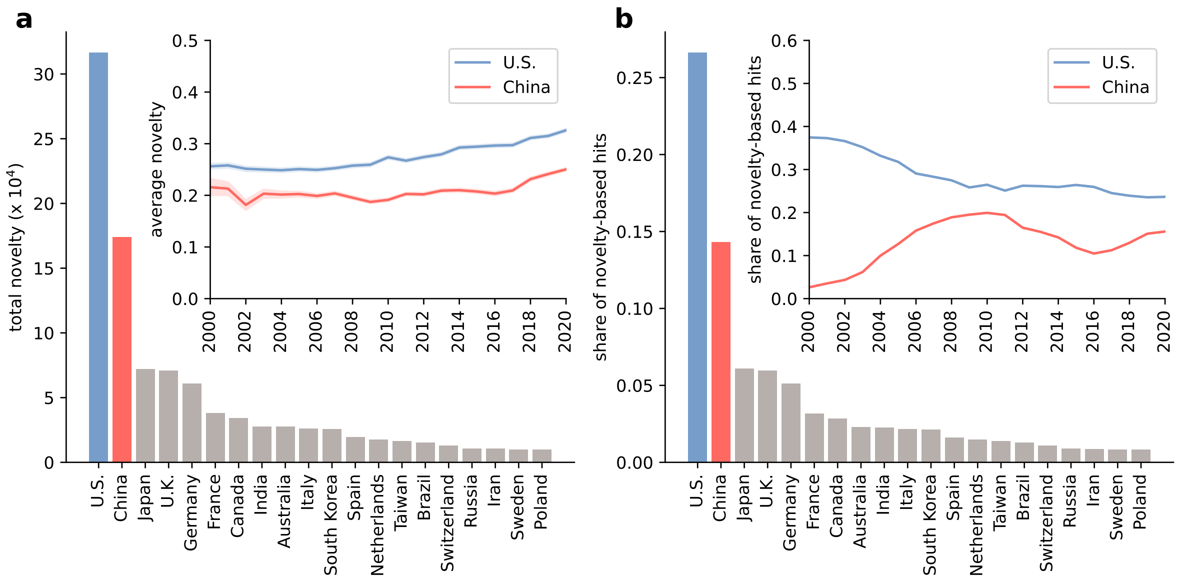
Supplementary Materials for

China and the U.S. produce more impactful AI research when collaborating together

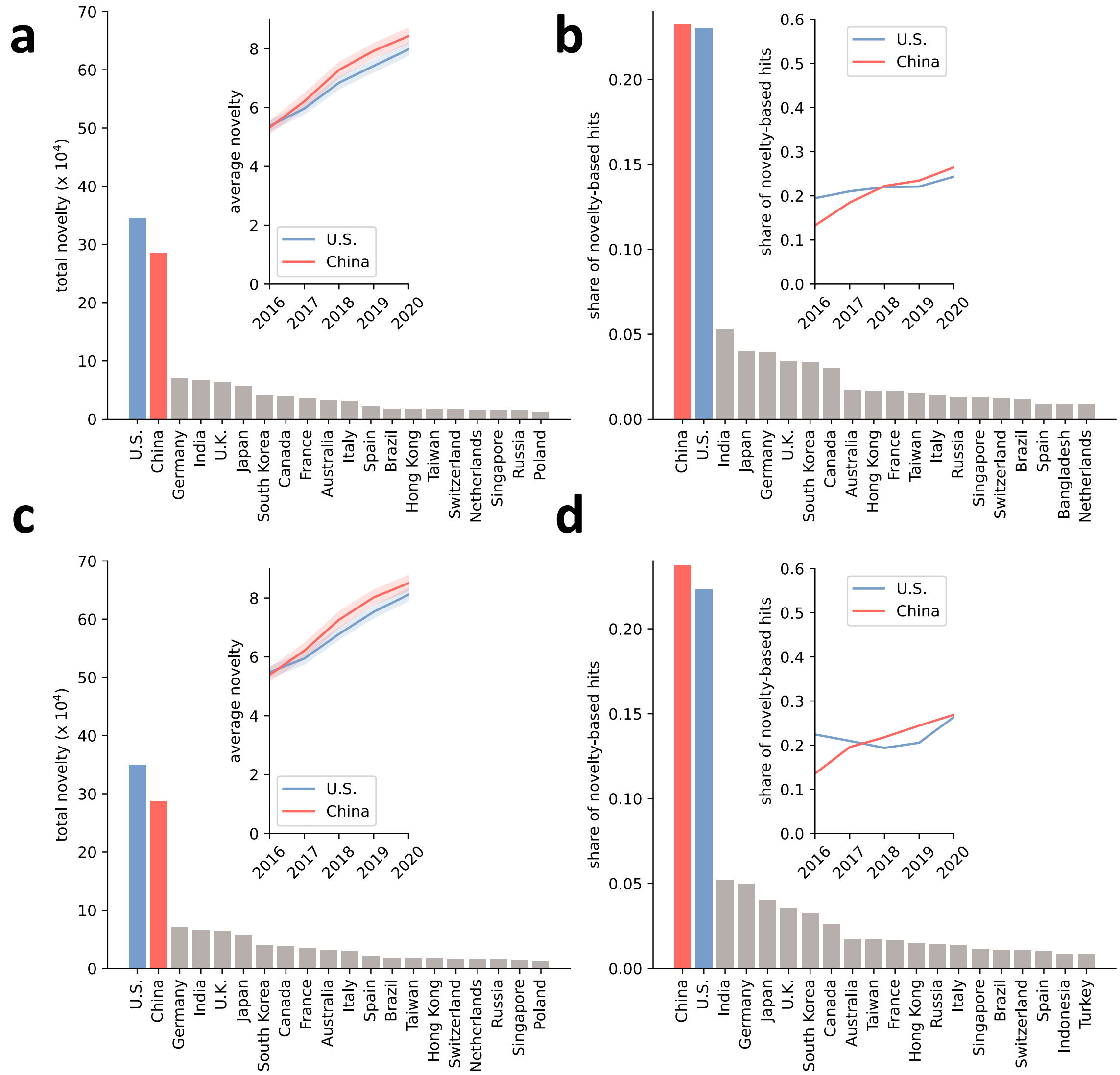
Bedoor AlShebli*, Shahan Ali Memon, James A. Evans, Talal Rahwan*

* Joint corresponding authors. E-mail: bedoor@nyu.edu; talal.rahwan@nyu.edu;

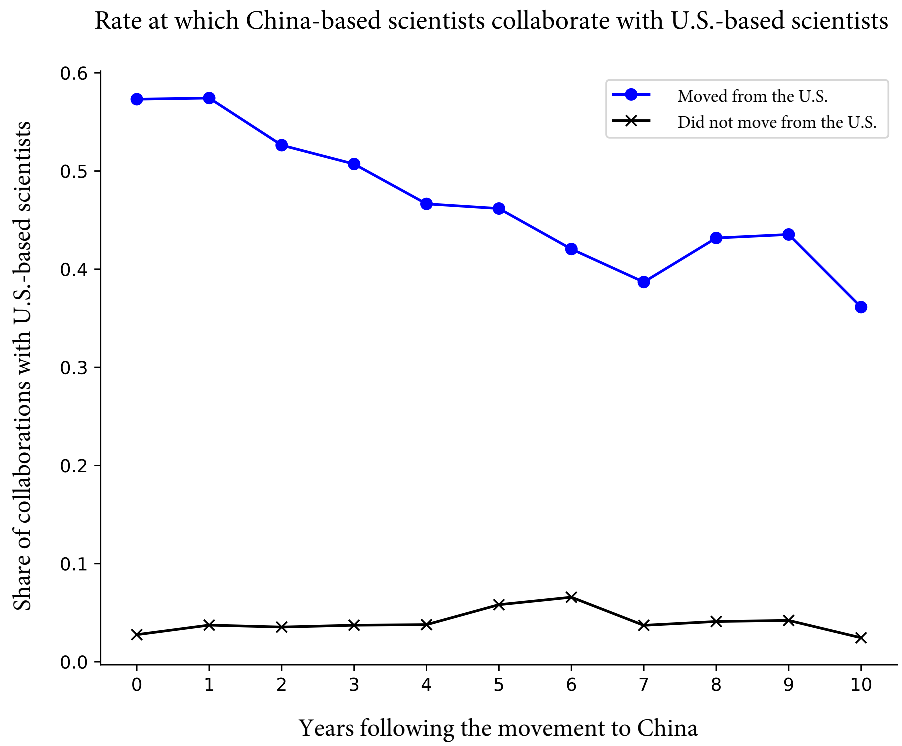
Supplementary Figures

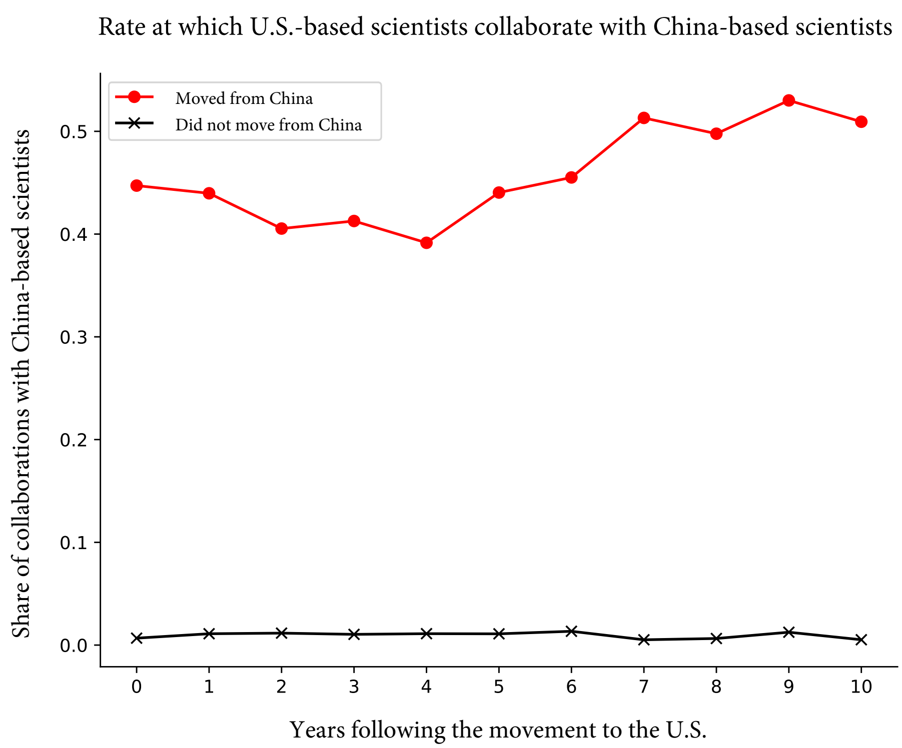


Supplementary Figure 1: **Content novelty analysis. a,** Main plot: The 20 countries that garner the highest total content novelty in AI research. Inset: The average content novelty of AI papers in the leading two countries (U.S. and China) over time. **b,** Main plot: The 20 countries with the greatest share of hits based on content novelty. Inset: The share of hit AI papers based on content novelty in the leading two countries (U.S. and China) over time.

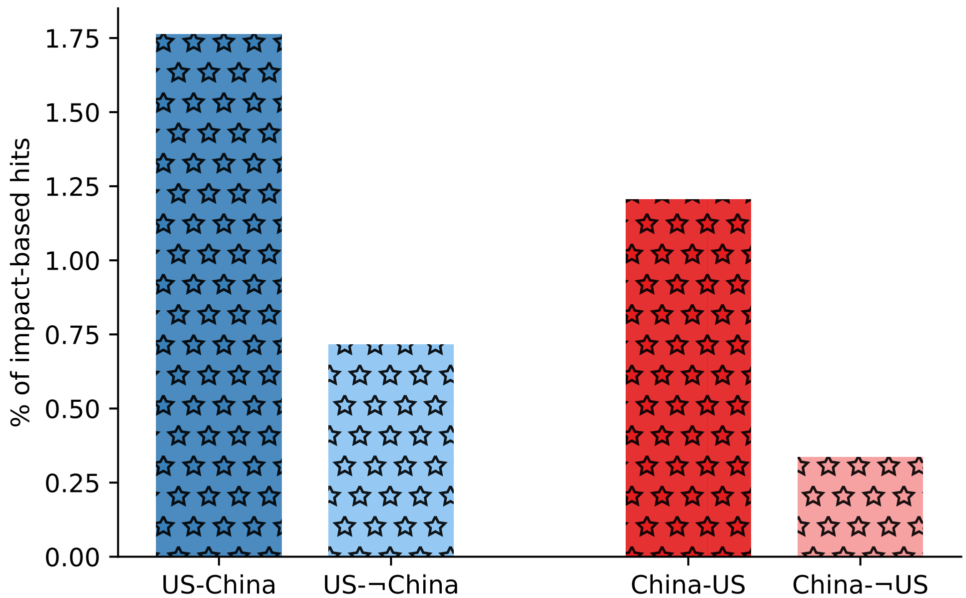


Supplementary Figure 2: **Context novelty analysis after excluding relatively new venues.** Analyzing how context novelty in AI papers published in [2016,2020] would change as a result of excluding references published in journals that are established after 2015. **a**, Main plot: the 20 countries with the highest context novelty in AI papers published in [2016,2020] (before excluding references). Inset: The average result over time for the two leading countries, U.S. and China. **b**, Main plot: The 20 countries with the greatest share of AI hits based on impact in [2016,2020] (before excluding references). Inset: The average result over time for the two leading countries, U.S. and China. **c** and **d**, The same as **(a)** and **(b)**, respectively, but after excluding references published in journals that are established after 2015.

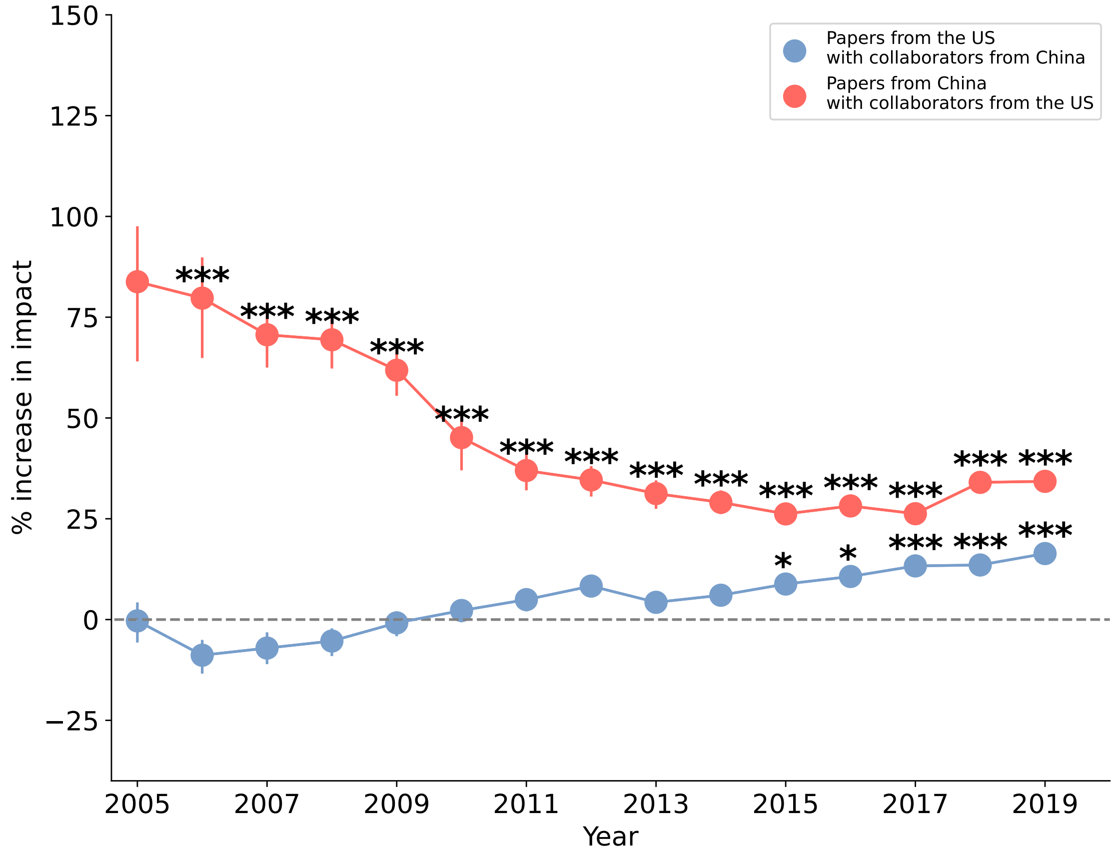




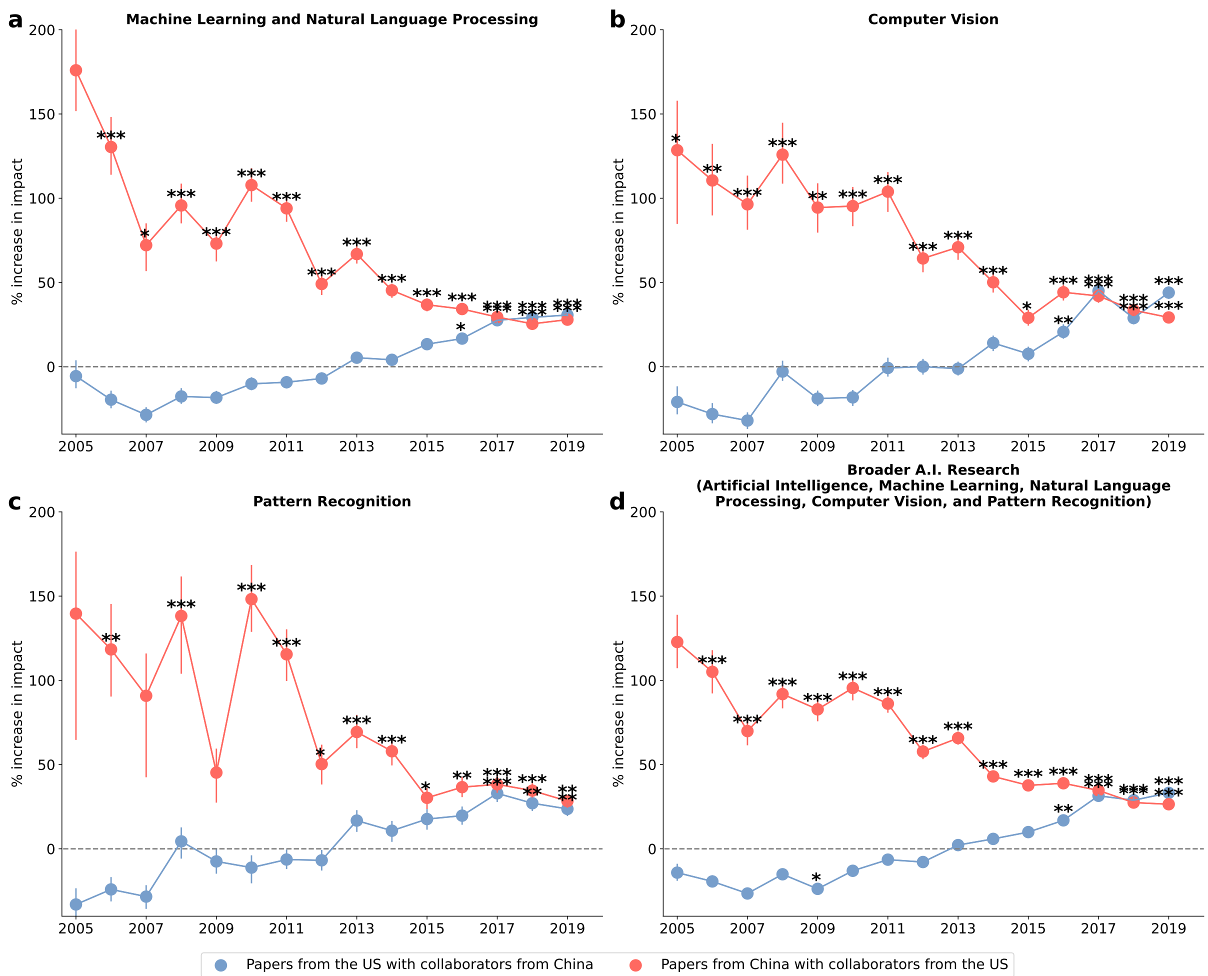
Supplementary Figure 3: **Collaboration rate with the country of origin over time.** The same as Figure 2f, but over the 10 years that followed the move.



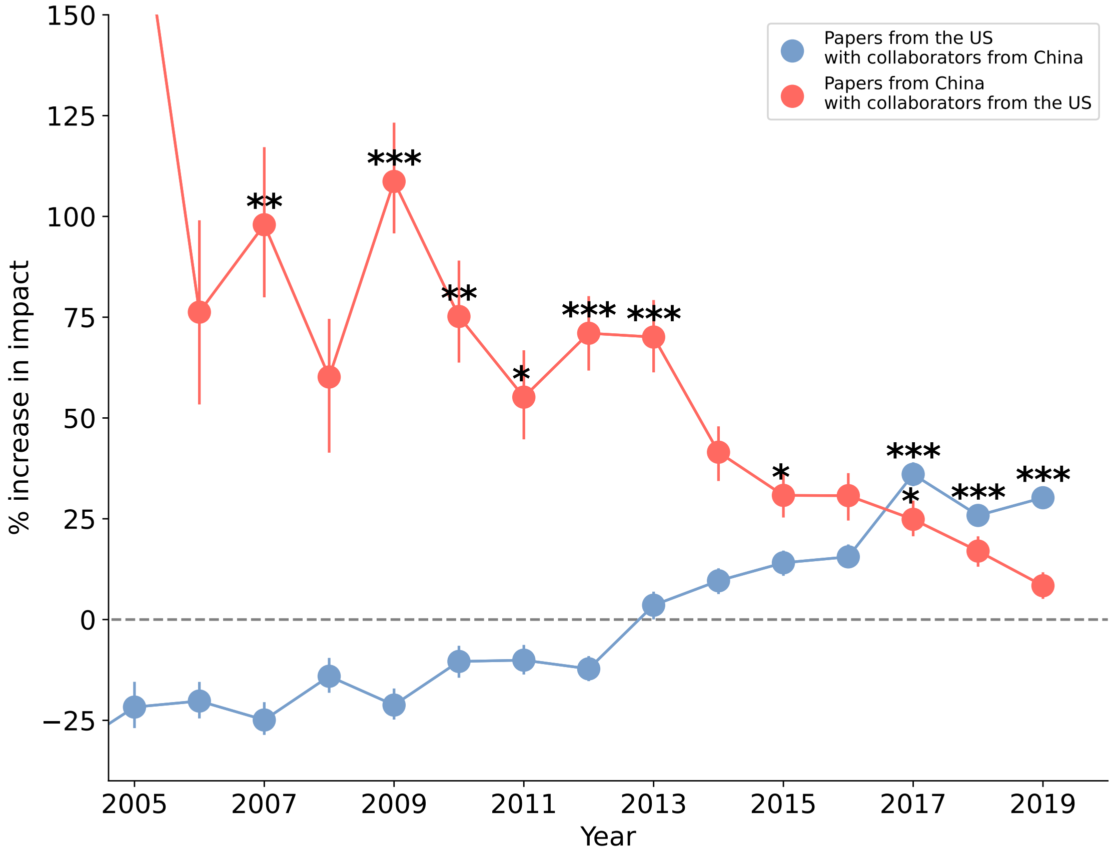
Supplementary Figure 4: **Impact-based AI hit rates.** Comparing the rate of impact-based AI hits for the four types of AI papers: (i) U.S.-based papers produced in collaboration with China, $(US, China)$; (ii) U.S.-based papers produced without China, $(US, \neg China)$; (iii) China-based papers produced in collaboration with the U.S., $(China, US)$; (iv) China-based papers produced without the U.S., $(China, \neg US)$.



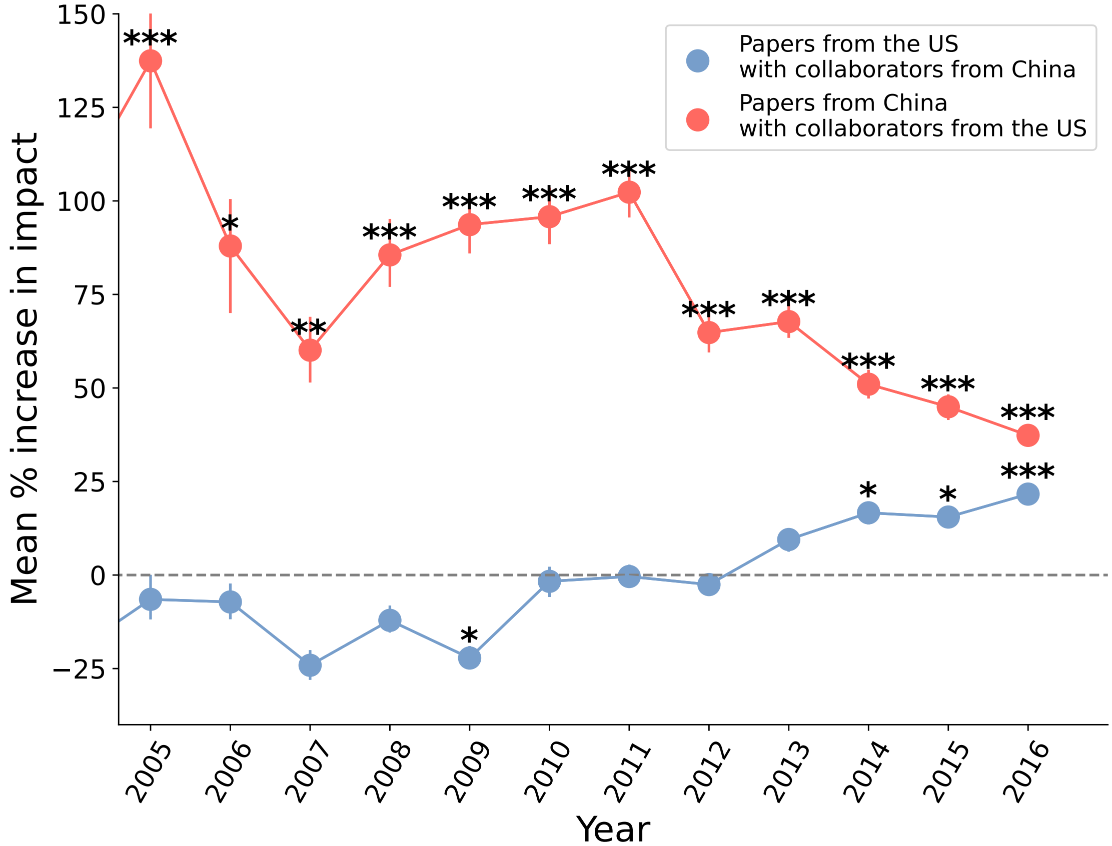
Supplementary Figure 5: **Within-subject analysis of U.S.-China collaborations.** The same as Figure 3d, but controlling for the last author. That is, we compare papers that involve U.S.-China collaborations to those produced by one country without the other, but the comparison is now performed among papers that have the same last-author. When performing this comparison, we allow for up to two years difference in publication year. Moreover, we bin the sizes of teams that involve five or more authors, using the following bins: [5,6], [7,9], [10,∞].



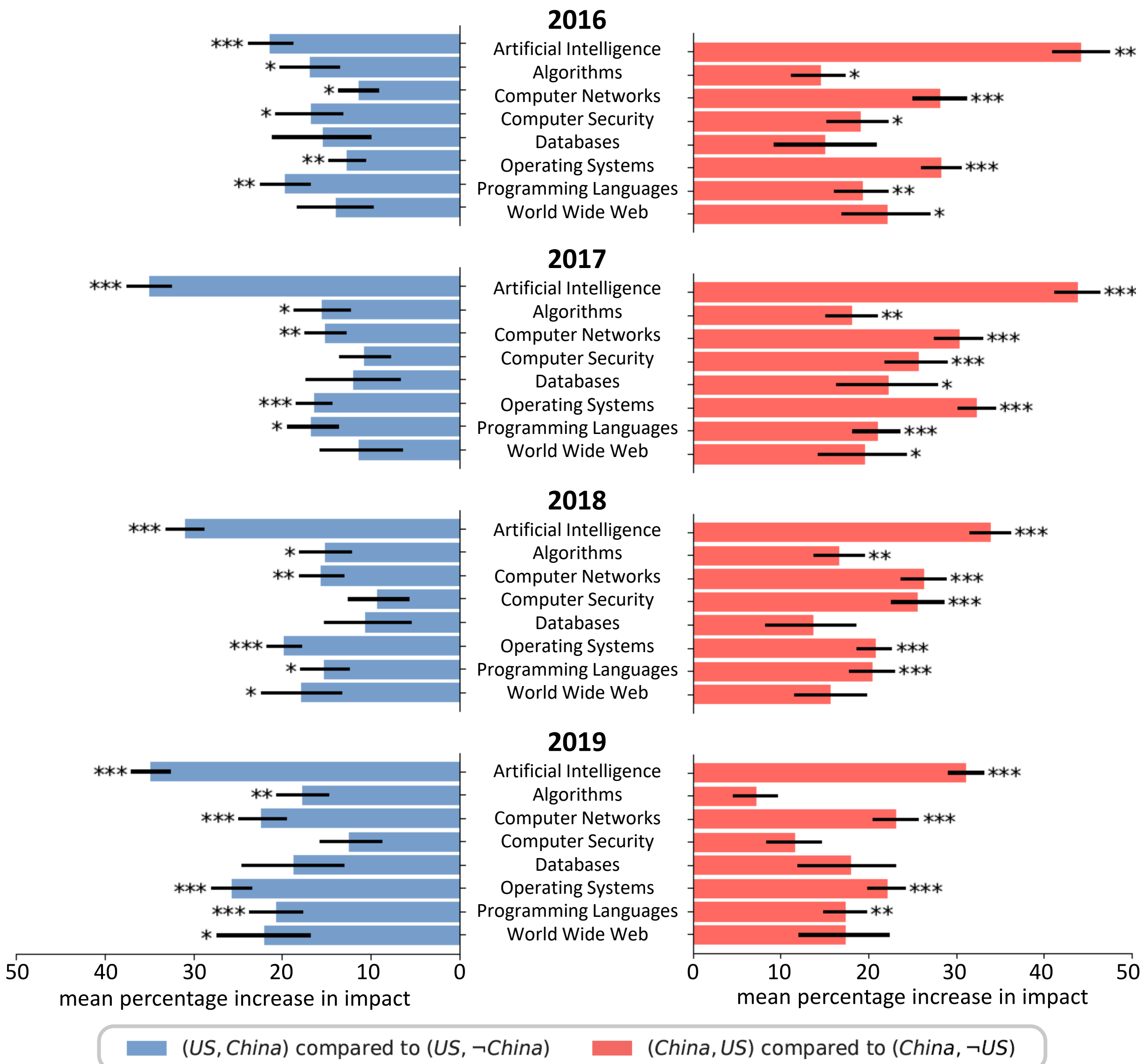
Supplementary Figure 6: **Other AI-related subfields** The same as Figure 3d, but examining the following AI-related areas: (**a**) Machine Learning and Natural Language Processing, (**b**) Computer Vision, (**c**) Pattern Recognition, and (**d**) a broader definition of AI research that encompasses the three aforementioned areas (**a**, **b**, and **c**) in addition to the area of "AI".



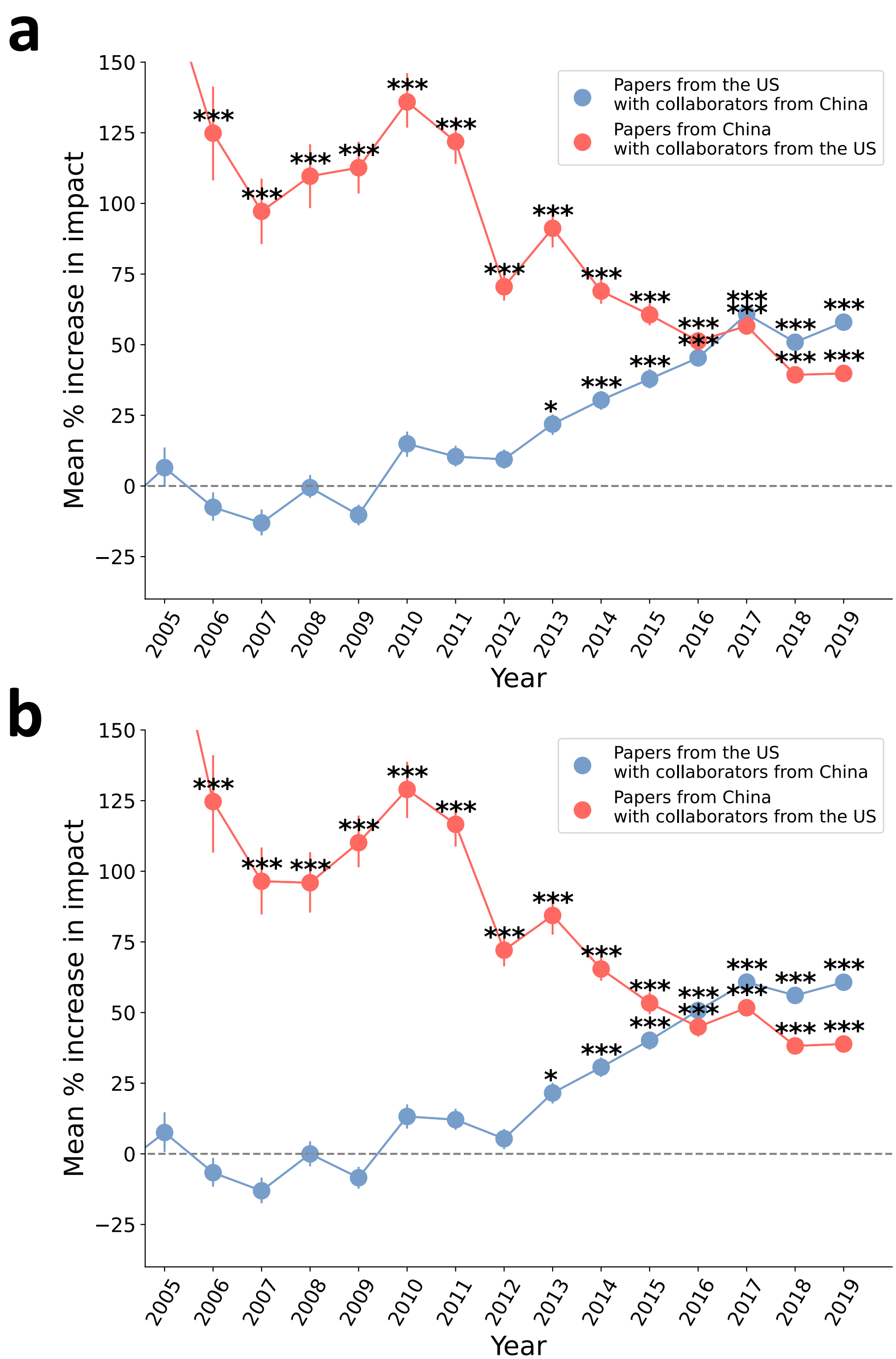
Supplementary Figure 7: **Alternative Collaboration Criteria.** The same as Figure 3d, but only focusing on papers where the last author is in one country and the first author is in the other.



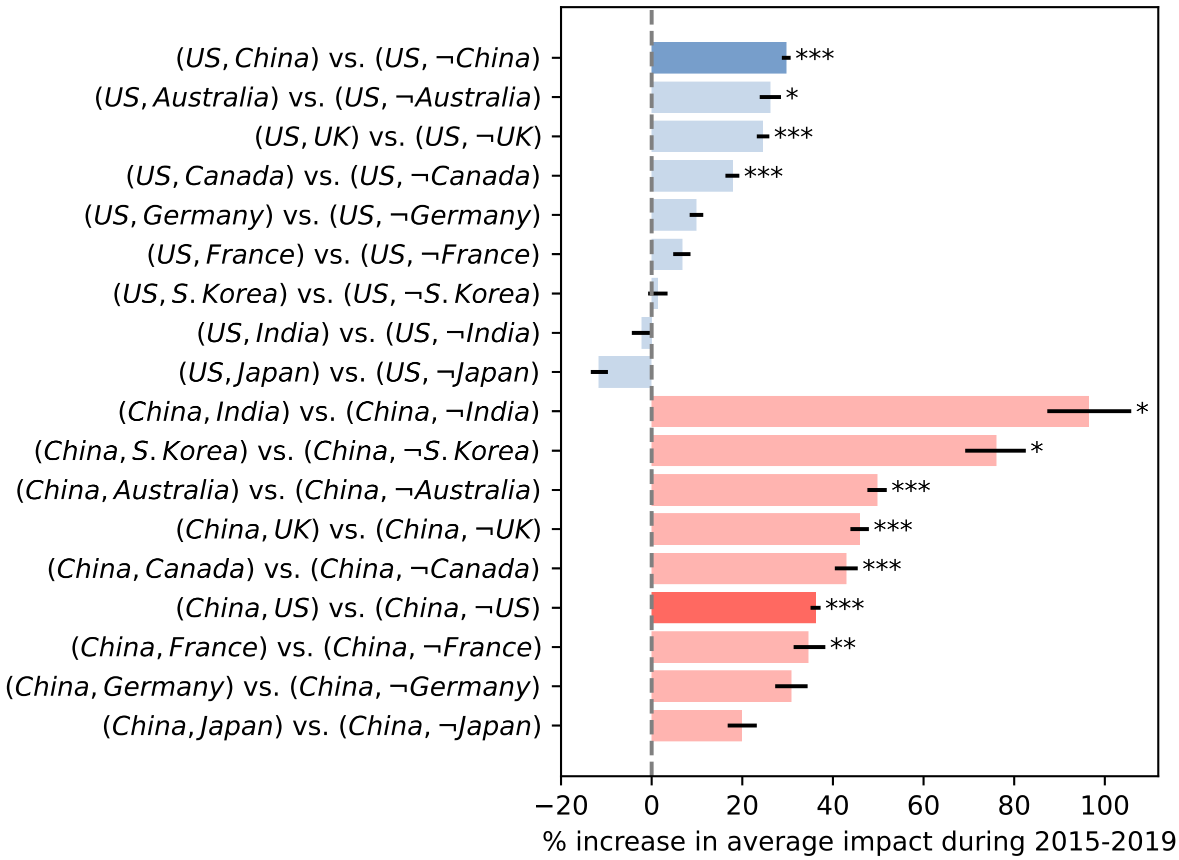
Supplementary Figure 8: **Citations within five years.** The same as Figure 3d, but for the number of citations received within five years, instead of two years, post publication.



Supplementary Figure 9: **Examining U.S.-China collaborations across Computer Science subfields.** Analyzing the four types of papers used in Figure 3a, i.e., $(US, China)$, $(US, \neg China)$, $(China, US)$, and $(China, \neg US)$ using Coarsened Exact Matching (CEM) to quantify the percentage increase in impact of $(US, China)$ compared to $(US, \neg China)$, as well as $(China, US)$ compared to $(China, \neg US)$. However, unlike Figure 3a, the comparison is now done in different fields of Computer Science (rather than AI only), for the years 2016 to 2019 (inclusive).



Supplementary Figure 10: **Excluding home citation bias.**The same as Figure 3d, but after removing "home citations" based on last author affiliation (**a**), and based on any-author affiliation (**b**).



Supplementary Figure 11: **Exploring the impact when the U.S. and China collaborate with other countries.** The other countries considered in this analysis are those appearing among the 10 most productive countries as per Figure 1a. For each country, $X \neq US$, the figure depicts (as a blue bar) the difference in impact between the papers in (US, X) and those in $(US, \neg X)$ over the last five years in our dataset (2015-2019). Similarly, for each country $Y \neq China$, the figure depicts (as a red bar) the difference in impact between the papers in $(China, Y)$ and those in $(China, \neg Y)$. The comparison is performed using CEM as per Figure 3d. P values are calculated using t-tests; * $p < .05$; ** $p < .01$; *** $p < .001$.

Supplementary Tables

Supplementary Table 1: **Coarsened Exact Matching results for the migration analysis of Figure 2f.** T' and C' are the populations of matched treatment and matched control papers, respectively; $\mathcal{L}_1$ is the multivariate imbalance statistic [1]; $\mu_{T'}$ is the percentage of papers that include a China-based collaborator (first row) or a U.S.-based collaborator (second row) in T'; $\mu_{C'}$ is the percentage of papers that include a China-based collaborator (first row) or a U.S.-based collaborator (second row) in C'; a bootstrap of 95% confidence interval ($CI_{95\%}$) is provided for $\mu_{T'}$ and $\mu_{C'}$; a t-test shows which δ values are statistically significant; see the resulting p-values.

	n	$\|T'\|$	$\|C'\|$	$\mathcal{L}_1$	$\mu_{T'}$	$\mu_{C'}$	$CI_{95\%,T'}$	$CI_{95\%,C'}$	p
T: U.S.-based scientists who migrated from China C: U.S.-based scientists who did not migrate from China	$4,740$	$1,958$	$1,856$	0.24	45.83	1.24	[44.22, 47.44]	[0.92, 1.55]	$< .001$
T: China-based scientists who migrated from the U.S. C: China-based scientists who did not migrate from the U.S.	$2,466$	700	765	0.19	55.09	3.01	[52.67, 57.52]	[2.21, 3.82]	$< .001$

Supplementary Table 2: **Coarsened Exact Matching results for U.S.-based papers produced in collaboration with China.** T and C are the treatment and control populations, respectively; T' and C' are the populations of matched treatment and matched control papers, respectively; $\mathcal{L}_1$ is the multivariate imbalance statistic [1]; $\mu_{T'}$ is the mean impact (i.e., mean c_2) of T'; $\mu_{C'}$ is the mean impact of C'; δ is the relative impact gain of T' over C', i.e., $\delta = 100 \times (\mu_{T'} - \mu_{C'})/\mu_{C'}$; a bootstrap of 95% confidence interval ($CI_{95\%}$) is provided; a t-test shows which δ values are statistically significant; see the resulting p-values.

year	$\|T\|$	$\|C\|$	$\|T'\|$	$\|C'\|$	$\mathcal{L}_1$	$\mu_{T'}$	$\mu_{C'}$	δ	$CI_{95\%}$	p
2005	345	35,925	314	7,165	0.44	5.06	5.79	−12.73	[-17.74, -6.81]	0.441
2006	556	39,016	512	10,753	0.48	4.15	5.02	−17.42	[-21.46, -13.06]	0.201
2007	605	41,961	531	11,600	0.45	3.97	5.35	−25.86	[-29.50, -22.18]	0.074
2008	820	43,795	747	15,791	0.44	4.31	4.97	−13.21	[-16.70, -9.41]	0.248
2009	1,070	44,801	983	17,373	0.45	4.20	5.34	−21.38	[-24.41, -17.98]	0.042
2010	1,280	48,163	1,169	22,790	0.45	4.69	4.93	−4.93	[-8.98, -1.51]	0.697
2011	1,420	49,730	1,301	23,540	0.42	5.02	5.21	−3.78	[-6.67, -0.58]	0.673
2012	1,661	53,169	1,542	24,835	0.43	5.22	5.52	−5.39	[-8.17, -2.76]	0.524
2013	1,855	54,976	1,716	25,685	0.43	5.80	5.62	3.16	[0.29, 5.74]	0.738
2014	2,206	56,469	2,037	27,656	0.41	6.15	5.57	10.42	[7.64, 13.07]	0.155
2015	2,577	57,276	2,405	30,590	0.43	6.27	5.47	14.60	[12.07, 17.05]	0.043
2016	3,021	56,829	2,796	32,596	0.41	6.49	5.34	21.39	[18.92, 23.84]	$< .001$
2017	3,613	57,234	3,384	34,828	0.42	7.79	5.77	35.02	[32.52, 37.57]	$< .001$
2018	4,436	61,290	4,106	38,511	0.42	8.15	6.22	30.93	[28.73, 32.87]	$< .001$
2019	5,382	69,110	5,036	43,877	0.41	7.57	5.61	34.87	[32.76, 37.10]	$< .001$

Supplementary Table 3: **Coarsened Exact Matching results for China-based papers produced in collaboration with the U.S.** T and C are the treatment and control populations, respectively; T' and C' are the populations of matched treatment and matched control papers, respectively; $\mathcal{L}_1$ is the multivariate imbalance statistic [1]; $\mu_{T'}$ is the mean impact (i.e., mean c_2) of T'; $\mu_{C'}$ is the mean impact of C'; δ is the relative impact gain of T' over C', i.e., $\delta = 100 \times (\mu_{T'} - \mu_{C'})/\mu_{C'}$; a bootstrap of 95% confidence interval ($CI_{95\%}$) is provided; a t-test shows which δ values are statistically significant; see the resulting p-values.

year	$\lvert T\rvert$	$\lvert C\rvert$	$\lvert T'\rvert$	$\lvert C'\rvert$	$\mathcal{L}_1$	$\mu_{T'}$	$\mu_{C'}$	δ	$CI_{95\%}$	p
2005	222	8,262	199	4,039	0.37	3.95	1.85	113.88	[93.45, 130.29]	$< .001$
2006	343	13,607	319	7,094	0.44	3.11	1.59	95.93	[81.92, 109.73]	$< .001$
2007	461	13,857	434	7,404	0.42	3.67	2.12	73.64	[62.68, 83.32]	$< .001$
2008	581	20,437	547	9,647	0.37	3.90	1.95	100.14	[90.24, 110.57]	$< .001$
2009	770	27,657	721	14,979	0.37	3.82	1.98	93.31	[84.22, 101.51]	$< .001$
2010	832	33,184	776	15,831	0.36	4.06	1.97	106.38	[94.05, 115.64]	$< .001$
2011	1,042	32,086	992	16,629	0.36	4.39	2.11	107.75	[100.39, 115.50]	$< .001$
2012	1,209	31,045	1,147	18,016	0.35	4.30	2.58	66.17	[61.16, 71.53]	$< .001$
2013	1,356	31,093	1,294	18,698	0.34	4.99	2.80	77.83	[72.46, 83.26]	$< .001$
2014	1,770	33,997	1,671	22,467	0.34	4.81	3.11	54.60	[50.68, 58.67]	$< .001$
2015	1,980	31,635	1,846	21,733	0.34	5.72	3.90	46.60	[43.29, 50.16]	$< .001$
2016	2,339	36,236	2,213	25,520	0.34	5.57	3.86	44.18	[40.97, 47.26]	$< .001$
2017	2,673	41,097	2,540	29,483	0.35	6.27	4.36	43.75	[40.86, 46.34]	$< .001$
2018	3,402	49,554	3,215	35,635	0.34	6.66	4.97	33.88	[31.63, 36.14]	$< .001$
2019	4,209	64,023	4,024	48,540	0.35	6.20	4.73	31.10	[29.13, 33.03]	$< .001$

References

[1] Iacus, S. M., King, G. & Porro, G. Causal inference without balance checking: Coarsened exact matching. *Political analysis* **20**, 1–24 (2012).