Online Appendix to:

Do Firms Report More University Knowledge Sourcing When Informed About Engagement Channels? Evidence from a Field Experiment

David Angenendt
Economics & Policy
TUM School of Management
and CBR Cambridge

Ammon Salter
Entrepreneurship and Innovation
Warwick Business School

6 May 2026

Contents:

A. Implications for survey design to capture knowledge sourcing from universities by companies

Our study seeks to help inform the design and development of surveys that capture information from practicing managers. First, with respect to innovation surveys, there is a potential to conduct additional question order experiments to learn about whether different aspects of these managerial responses are influenced by the positioning of questions on the survey. To date, we have mixed evidence for such effects, as our study shows an effect for indications of university knowledge sourcing, but Han et al. (2024), which focused the order of questions on R&D and innovation, demonstrated no order effects. It may be the order effects are stronger in parts of the survey that address areas with greater managerial agency and more limited knowledge, such as knowledge sourcing with universities, rather than more concrete and specific questions, such as expenditures on R&D. Future innovation surveys could experiment with which parts of the survey are most susceptible to question order effects, and how different question orders might improve the quality of information collected from the surveys.

Second, innovation surveys could be supplemented with greater information and detail on different aspects of managerial behavior and practice to help inform and educate respondents about practices that are often poorly monitored within the company. In the past, innovation surveys have been adapted and extended to explore a wide range of innovation types, including eco-innovations and organizational innovation. To improve the quality and reliability of information on sourcing knowledge from different information sources, innovation surveys could offer richer guidance to respondents to help them better understand the dimensions and characteristics of each of the potential knowledge sources. It may be that other knowledge sources, like universities, are under reported on innovation surveys due to limited information provided to respondents.

Third, innovation surveys and university-industry link surveys could be enhanced with providing respondents with public information on companies' engagement with universities. Given that university engagement often covers a wide range of different corporate activities and may not be directly managed or mapped by the firm, there is scope to use draw together information from various public sources and records, such as publications, patents, research grants, and educational backgrounds on staff from LinkedIn, to provide detailed and specific information to respondents about their firm's different forms of engagement with universities before asking them about the value of universities as a source of knowledge for innovation and/or about their different channels of university engagement. Such an effort might help to help respondents gain awareness of different ways universities have shaped their innovative efforts that are not always easily observed in management information systems.

Fourth, the innovation survey questions about universities are relatively modest in scale and scope, with only two items to capture a complex set of potential relationships. Moreover, the innovation survey question format asks respondents to report on both "use" and "importance" in the same response. As it stands, respondents are unable to respond that they do not use universities as a source of knowledge but also find them as an important source of knowledge and vice-versa. We appreciate that these cases may be rare, but it may be that some firms acknowledge the value of universities to their firms, but they themselves do not directly draw on university knowledge in their innovation efforts. An example of this would be a firm that hires graduates from universities but does not report the direct use of universities as a source of information or knowledge in its innovation efforts. Given that effects of universities may often be transmitted indirectly to companies, it is likely that the reports from innovation surveys are only capturing the visible and direct part of these effects on companies' activities. A revised question that separates use from importance could help to shed light on the indirect benefits of universities to companies. In addition, future surveys could also provide more detailed anchors for the degree of importance and use, as different respondents may interpret these terms differently. An example of potential new anchors is presented in Figure A.1 below.

Figure A.1 Potential new anchors for questions for university knowledge sourcing

Low importance	Minimal influence on our innovation activities Information from this source is rarely considered when defining or adjusting innovation projects. It has little or no bearing on decisions about new or improved products, services, processes, or markets. Whether or not this source is considered has almost no effect on the direction, scope, or outcomes of our innovation activities.
Medium importance	A supporting influence on innovation activities Information from this source is one of several inputs considered in innovation-related decisions. It helps refine or validate ideas, reduce uncertainty, or adjust project details, but is not usually decisive on its own. If this source were not available, innovation projects could still proceed, but some decisions would be less well-informed or somewhat riskier.
High importance	A major influence on our innovation activities Information from this source is central in shaping the choice, design, or prioritisation of innovation projects. It strongly affects decisions about whether to start, stop, expand or redirect innovation efforts. If this source were not available, many innovation activities would change substantially, be delayed, or might not take place at all.
Low use	Rarely used The source is consulted only in exceptional cases or for a small number of innovation projects. It is not part of our usual innovation-related information gathering.
Medium use	Occasionally used The source is consulted for some innovation projects or specific questions, but not routinely. It plays a role in selected projects rather than across the board.
High use	Regularly used The source is consulted for most innovation projects and is part of our standard innovation routines or processes. It is a normal and expected channel for innovation-related information.

Source: Generated with the help of GPT 5.2.

B. Further details on the dataset construction

B.1 Sampling frame

The sampling frame, including company director-level email addresses, was drawn from Bureau van Dijk's FAME database, which provides financial and contact data for UK companies across all sectors, drawing primarily on official filings from Companies House and other public records. Contact data from FAME were supplemented by email quality checks using Hunter IO; firms with available email addresses tended to be larger than those without. Table B.1 benchmarks both our sampling frame and respondent sample against the ONS 2020 Business Population Estimates.[3] As expected, larger firms and knowledge-intensive sectors are over-represented, while the regional distribution closely mirrors national patterns.

Table B.1 Sample representativeness

	ONS 2020	Respondent sample	FAME data export: Sampling frame	FAME data export: Firms with email addresses
Employment				
1-9	81.9%	7.9%	14.0%	7.6%
10-49	15.0%	63.2%	60.6%	60.2%
50-249	2.6%	21.2%	17.9%	23.5%
250 and more	0.6%	7.7%	7.4%	8.6%
Sector				
ICT, Professional & Scientific	19.9%	25.7%	14.1%	19.3%
Other KI Services	11.0%	18.1%	20.7%	18.4%
Other services & non-manufacturing	62.8%	38.8%	54.6%	45.3%
high-tech manuf.	1.2%	5.9%	2.5%	4.2%
low-tech manuf.	5.0%	11.6%	8.1%	12.8%
Region				
London	17.3%	18.4%	22.7%	19.0%
North	20.2%	18.3%	19.6%	19.9%
Midlands	15.6%	13.5%	14.3%	15.4%
South	33.6%	37.6%	31.5%	34.3%
Scotland, Wales, N. Ireland	13.3%	12.2%	11.8%	11.3%

Notes: This table compares the distribution of firms across employment size, sector, and geographic region between the ONS 2020 Business Population Estimates, the full FAME sampling frame without and with the restriction to firms with contact email addresses, and the final respondent sample used in the analysis. Percentages are based on the total number of firms in each source. "Other KI Services" refers to other knowledge-intensive services.

[3] Obtained from https://www.gov.uk/government/statistics/business-population-estimates-2020.

Firms in the sampling frame were randomly assigned to nine invitation waves, each comprising approximately 10,000 companies. Invitations were distributed sequentially over five months. Random allocation of firms from the original sampling frame involved only waves 1-6. Wave 7 instead contains firms from the same original sampling frame whose invitations were delayed due to data issues (e.g., determining appropriate salutations or resolving cases where the same director email address was associated with multiple businesses). The sampling frame for waves 8 and 9 was obtained later, with modified inclusion criteria to address the low response rate during the COVID-19 pandemic. The full survey dataset also includes two smaller waves (10 and 11) of firms whose contact details were obtained manually to increase the sample size among larger firms.

B.2 Detailed list of independent variables

Number of employees: This variable is added (in logs) as a measure of firm size. It is well known from earlier research that firm size influences university interactions and their perceived importance (Fontana et al., 2006). This variable is sourced from the FAME database for the financial year 2019. If missing, we use the respondent's self-reported value in the survey.

Start-up: Although the literature has found mixed results about the importance of firm age on university sourcing (Laursen and Salter, 2004), we control for the age of the firm. Based on the establishment year in FAME, we define a dummy variable taking the value 1 if the firm was established at most seven years prior to the survey start date.

R&D active: Prior research demonstrates that Research and Development (R&D) plays an important role in sourcing knowledge from universities (Cohen et al., 2002; Laursen and Salter, 2004), but it is also well documented that many firms innovate without maintaining dedicated R&D units (Hervas-Oliver et al., 2012; Tether and Tajar, 2008). This dummy takes the value 1 for firms reporting any R&D expenditure during the three years before March 2020.[4]

Asset growth: A firm's growth rate may relate closely to its use of university knowledge, as firms with high rates of growth are liable to have greater slack resources for external engagement (Bruneel et al., 2016). We calculate two-year asset growth to maximize sample size due to data availability, but using three-year asset growth yields comparable results. Two firms with extreme asset growth rates in the range of 500,000 were excluded, as the next-highest value was around 21.

Regional fixed effects: We include five regional dummies to account for differences between the major countries and regions of the UK. We also examined using a finer regional breakdown that slightly reduced the sample size but yielded consistent results.

Industry fixed effects: We include five dummies representing broad sectoral groups based on technological intensity. We allocate firms based on their UK SIC code using UK Office of National Statistics and OECD criteria. Our dummies distinguish two groups within manufacturing

[4] The *amount* of R&D spending would likely also be an important predictor of university importance, but these data are frequently unavailable for UK firms. The survey responses to this question did not appear reliable. We also matched the responding firms to UK patent records, but only a small fraction of firms was active in patenting.

and three within the services and non-manufacturing sectors; see Table 2 for details. Again, a more granular disaggregation reduced the sample size slightly without changing the result of interest.

Innovator. This binary variable is equal to one if the firm reports at least one of a product innovation, process innovation, or new logistics or distribution processes in the reference period. Prior innovation activity is a strong predictor of current innovation activity, which is relevant given that our dependent variable concerns universities' role as information source for innovation.

External search breadth. Defined as the number of external information sources other than universities, this is another important predictor of a firm's use of information from universities (Laursen and Salter, 2004). However, we exclude this variable from our main specification due to potential endogeneity: it is derived from the same survey question as our dependent variable and was likewise moved further to the end in the treatment group survey. We consider this variable separately below.

Response-time fixed effects. While all respondents were asked about their activities in the period ending in March 2020, there is variation in the time of invitation and response submission. This variation may significantly impact our results, given the ongoing COVID-19 pandemic. To account for it, we optionally include dummies indicating the month of response recording. Since the COVID-19 restrictions could change quite drastically over a short time frame, we additionally distinguish between the first and second half of a month. i.e., the first half of a month is represented by a different dummy than the second half of the same month. Accounting for this heterogeneity over time increases the estimated treatment effect, and using a single dummy for each month has a similar effect.

Interactor. A binary variable taking value 1 if the respondent indicated they had used at least one of the university engagement channels detailed in Table 1 of the main text. In addition to that list, the survey contained a question allowing firms to indicate and identify engagement channels not covered by the list. We mapped all responses to the existing channels where this was possible and classified the remaining firms as non-interactors in the context of this survey.

C. Further details on the sample description

C.1 Sample composition and re-invitations

The treatment group in the main analysis is larger than the control group primarily due to survey administration procedures rather than systematic differences in response behavior. All firms associated with survey wave 6 were initially invited to the control version of the questionnaire. When an invitation email bounced, we attempted to obtain an updated address for the corresponding director. These firms were then re-invited at a later date. For administrative simplicity, all re-invitations coincided with invitations for wave 7 and were sent only the treatment version of the questionnaire. No firm ever had access to both versions of the survey, and duplicate responses were not possible.

Because waves 5 and 6 were drawn at random from the same FAME-based sampling frame, we retain all firms from these waves in our baseline sample. Table D.1, column (1), further below shows that excluding re-invited firms leaves the estimated treatment effects largely unchanged, indicating that the administrative re-invitation process does not drive our findings.

C.2 Non-response analysis

C.2.1 Response rates by wave

Table C.1 summarizes invitations and responses by wave. Across the nine waves we invited 90,113 firms and obtained 2,415 fully completed questionnaires, corresponding to an overall completion rate of 2.7%. Wave-specific completion rates lie in a narrow band between 2.3% and 3.1%. A χ^2 test of equal completion rates across all nine waves rejects the null of identical response rates (p=0.013), driven by the slightly higher rate in wave 9. When restricting attention to waves 1–8, the null of equal response rates is not rejected (p=0.125), and the p-values increase further when restricting to waves 1–7 (p=0.157) and waves 1–6 (p=0.343). Quantitatively, however, the differences are small: for our baseline analysis, which uses waves 5 and 6, the completion rates are 2.7% and 2.9%, respectively. If we also count partial responses, which we define as questionnaires in which at least the first block of questions was completed, the overall response rate increases to 4.1%, with wave-specific rates between 4.0% and 4.6%.

Table C.1 Response rates by wave

		Complete responses		Partial responses	
Wave	Invited	Number	rate	Number	rate
1	9599	233	2.4%	405	4.2%
2	9636	249	2.6%	393	4.1%
3	9692	265	2.7%	425	4.4%
4	9764	240	2.5%	391	4.0%
5	9697	265	2.7%	389	4.0%
6	9883	283	2.9%	451	4.6%
7	9389	216	2.3%	392	4.2%
8	11262	318	2.8%	517	4.6%
9	11191	346	3.1%	519	4.6%

A response is coded as partial when at least the first block of questions is filled in. For all waves except wave 6 this first block consists of the list of engagement channels, whereas for wave 6 it contains the CIS-style question on sources of information for innovation. Capturing the full range of engagement channels was the main focus of the original project, and partial responses are analyzed in detail in Hughes et al. (2022). In the present paper we restrict attention to complete responses. This avoids imposing different completeness thresholds on treatment and control waves (engagement vs. innovation sections appearing first) and ensures that the analysis is based on a consistent set of answered questions across all respondents.

C.2.2 Comparing early to late responders

We assess potential non-response bias by comparing early and late responders, treating late responders as proxies for non-respondents (Armstrong and Overton, 1977). We focus on the baseline sample from waves 5 and 6. Hughes et al. (2022) also contains detailed checks of non-response bias in the full survey sample.

Table C.2 compares firms that respond to the initial invitation with those that respond only after one of the four reminders. Early responders are slightly smaller and have lower asset growth, although these differences are not statistically significant. They are somewhat less likely to be start-ups and more likely to conduct R&D, with p-values around the 10% level. There are also some differences in regional and sector composition (for example, the North of England and low-tech manufacturing are marginally over-represented among early responders, while “other KI services” are under-represented). Crucially, however, there are no significant differences between early and late responders in the probability of using universities as an information source or in the distribution over low, medium, and high importance.

Given that only about 18% of firms respond to the invitation and the remaining 82% respond after at least one reminder, we also consider a broader definition of “early” response: firms that respond either to the invitation or to the first reminder. Table C.3 shows that this leads to more balanced group sizes and attenuates several of the earlier differences. Under this definition, early responders are modestly smaller and have lower asset growth ($p \approx 0.07$), and there are some differences in the composition of service industries and in the share of London-based firms. Again, we find no statistically significant differences between early and late responders in university use or in the relative importance assigned to universities as a knowledge source.

The more important question is whether early and late responders react differently to the treatment. Table C.4 reports logit regressions of university information sourcing that add an indicator for early response and its interaction with the treatment dummy to our preferred specification. Columns (1) and (2) use the narrow definition of early responders (invitation only), while columns (3) and (4) use the broader definition (invitation or first reminder) and include response-time fixed effects to capture general trends in response behavior over the field period.

Table C.2 Comparing early and late responders (1): reminder 1 counts as late response

Variable	Early responders (invitation)				Late responders (reminders 1-4)				Two-tailed t-test	
	Mean	p50	SD	N	Mean	p50	SD	N	p-val.	sign.
Dummy: univ. used	0.41	0	0.49	101	0.44	0	0.50	447	0.579	
Unis not used	0.59	1	0.49	101	0.56	1	0.50	447	0.579	
Unis low importance	0.22	0	0.41	101	0.23	0	0.42	447	0.860	
Unis medium importance	0.14	0	0.35	101	0.16	0	0.37	447	0.613	
Unis high importance	0.05	0	0.22	101	0.05	0	0.22	447	0.936	
Employees (log.)	3.40	3.18	1.06	101	3.50	3.26	1.33	447	0.414	
Asset growth (2-year)	0.29	0.14	0.63	101	0.43	0.10	2.24	447	0.239	
Start-up dummy	0.01	0	0.10	101	0.05	0	0.21	447	0.087	*
R&D dummy	0.54	1	0.50	101	0.45	0	0.50	447	0.071	*
Innovator dummy	0.67	1	0.47	101	0.57	1	0.50	447	0.048	**
Interactor dummy	0.63	1	0.48	101	0.58	1	0.49	447	0.338	
No. of sources except unis	7.06	7	2.04	101	6.79	7	2.53	447	0.250	
Sources with high importance except unis	1.58	1	1.37	101	1.75	1	1.69	447	0.285	
Regional dummies:										
London	0.14	0	0.35	101	0.19	0	0.39	447	0.264	
North of England	0.22	0	0.41	101	0.14	0	0.35	447	0.063	*
Midlands	0.15	0	0.36	101	0.15	0	0.36	447	0.883	
South of England	0.39	0	0.49	101	0.39	0	0.49	447	0.987	
Scotland, Wales, N. Ireland	0.11	0	0.31	101	0.13	0	0.34	447	0.569	
Industry dummies:										
ICT, Professional & Scientific	0.28	0	0.45	101	0.25	0	0.43	447	0.547	
Other KI Services	0.10	0	0.30	101	0.19	0	0.39	447	0.029	**
Other Service and Non-Mfg. Sectors	0.40	0	0.49	101	0.40	0	0.49	447	0.999	
High-Tech Manuf.	0.04	0	0.20	101	0.06	0	0.23	447	0.460	
Low-Tech Manuf.	0.19	0	0.39	101	0.11	0	0.31	447	0.025	**

Table C.3 Comparing early and late responders (2): reminder 1 counts as "early" response

Variable	Early responders (invitation and reminder 1)				Late responders (reminders 2-4)				Two-tailed t-test	
	Mean	p50	SD	N	Mean	p50	SD	N	p-val.	sign.
Dummy: univ. used	0.46	0	0.50	262	0.41	0	0.49	286	0.216	
Unis not used	0.54	1	0.50	262	0.59	1	0.49	286	0.216	
Unis low importance	0.23	0	0.42	262	0.22	0	0.42	286	0.807	
Unis medium importance	0.16	0	0.37	262	0.15	0	0.35	286	0.578	
Unis high importance	0.06	0	0.25	262	0.04	0	0.19	286	0.161	
Employees (log.)	3.38	3.18	1.14	262	3.58	3.24	1.39	286	0.070	*
Asset growth (2-year)	0.25	0.12	0.61	262	0.55	0.12	2.76	286	0.071	*
Start-up dummy	0.03	0	0.18	262	0.05	0	0.21	286	0.509	
R&D dummy	0.49	0	0.50	262	0.44	0	0.50	286	0.195	
Innovator dummy	0.60	1	0.49	262	0.57	1	0.50	286	0.541	
Interactor dummy	0.60	1	0.49	262	0.58	1	0.49	286	0.591	
No. of sources except unis	7.01	7	2.35	262	6.68	7	2.52	286	0.119	
Sources with high importance except unis	1.79	1	1.76	262	1.66	1	1.52	286	0.333	
Regional dummies:										
London	0.15	0	0.36	262	0.20	0	0.40	286	0.099	*
North of England	0.16	0	0.37	262	0.15	0	0.36	286	0.659	
Midlands	0.16	0	0.37	262	0.15	0	0.35	286	0.663	
South of England	0.39	0	0.49	262	0.38	0	0.49	286	0.910	
Scotland, Wales, N. Ireland	0.14	0	0.34	262	0.12	0	0.32	286	0.439	
Industry dummies:										
ICT, Professional & Scientific	0.25	0	0.43	262	0.26	0	0.44	286	0.775	
Other KI Services	0.12	0	0.32	262	0.22	0	0.42	286	0.001	***
Other Service and Non-Mfg. Sectors	0.45	0	0.50	262	0.35	0	0.48	286	0.020	**
High-Tech Manuf.	0.06	0	0.23	262	0.05	0	0.22	286	0.805	
Low-Tech Manuf.	0.13	0	0.34	262	0.12	0	0.32	286	0.608	

Table C.4 Treatment effect among early and late responders

	(1)	(2)	(3)	(4)
Treatment	0.712***	1.149***	0.613**	0.869**
	(0.222)	(0.330)	(0.262)	(0.396)
Early responder	0.345	0.639	0.247	0.322
	(0.358)	(0.492)	(0.344)	(0.483)
Treatment × early responder	-0.837*	-0.728	-0.244	0.013
	(0.493)	(0.757)	(0.416)	(0.652)
Employees (log.)	0.123	0.117	0.123*	0.112
	(0.075)	(0.076)	(0.074)	(0.076)
Dummy: young firm	-0.324	-0.251	-0.297	-0.251
	(0.441)	(0.454)	(0.426)	(0.454)
Dummy: R&D	0.818***	0.873***	0.795***	0.868***
	(0.193)	(0.198)	(0.192)	(0.198)
Asset growth (2-yr.)	0.017	-0.122	0.015	-0.124
	(0.049)	(0.092)	(0.049)	(0.092)
Constant	-1.606***	-2.517***	-1.597***	-2.650***
	(0.456)	(0.891)	(0.448)	(0.641)
Reminder 1 included in:	early	early	late	late
Response-time fixed effects		yes		yes
Observations	548	544	548	544
Log lik.	-349.685	-342.177	-351.081	-341.844

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Across the specifications in Table C.4, the main treatment effect remains positive and significant. The interaction term between treatment and early response is negative in all four columns, indicating that the treatment effect is, if anything, stronger among late responders. This difference is statistically significant only in the specification that uses the narrow early-responder definition without response-time fixed effects (column 1); it becomes smaller and statistically insignificant once we broaden the early-responder group and include response-time fixed effects.

Overall, these analyses provide only limited evidence of systematic differences between early and late responders, and we find no indication that early responders are more responsive to the treatment than late responders. If anything, the pattern suggests a somewhat stronger treatment effect among late responders, implying that any non-response bias would tend to attenuate rather than inflate our estimated treatment effect.

C.2.3 Comparing responders to non-responders

We also compare survey responders to complete non-responders in the sampling frame. We define non-responders as firms that were invited (and for which we did not receive an "email bounced" or similar delivery failure) but did not return a complete or partial questionnaire, including the small number of firms that replied only to request removal from the database. In this exercise we can only use variables available in the FAME sampling frame: firm size (log employees), two-year asset growth, region, and industry.

Table C.5 reports these comparisons. Responders and non-responders are very similar in terms of firm size (p = 0.159) but differ in asset growth: non-responders exhibit a much higher mean growth rate (2.38 vs 0.41) and substantially larger dispersion. This reflects a small number of extreme values in the administrative data—indeed, median asset growth is slightly lower among non-responders—so the mean difference should be interpreted with caution (we also dropped two firms with extreme growth from our analysis sample). Regionally, firms in the North of England are somewhat under-represented among responders, while firms in the South are over-represented; this pattern is consistent with the particularly severe COVID-19 outbreaks in the Northern regions during our fieldwork (Morrissey et al., 2021; Munford et al., 2021). Industry-wise, firms in "ICT, Professional & Scientific" are over-represented among responders, and firms in "Other Service and Non-Manufacturing Sectors" are under-represented, which is plausible given that knowledge-intensive services are more likely to engage with universities and to respond to a survey on university links.

To gauge the impact of these differences on our main estimates, we follow the standard inverse-probability weighting approach to adjust for unit nonresponse (e.g. Little and Rubin (2019) and Wooldridge (2007)) by estimating a logistic regression for the response indicator and using the inverse of the predicted response probability as survey weights. Our concern is that survey response may not be random with respect to firm characteristics that are also related to university use (e.g. size, growth, sector, region). If, for example, faster-growing or more knowledge-intensive firms are more likely to respond, naive estimates based only on respondents may be biased. The inverse-probability weighting approach reweights respondents so that their distribution of observables matches that of the full sampling frame, under the assumption that, conditional on these observables, response is ignorable.

Specifically, we estimate a response-propensity model for all firms in waves 5 and 6 using a logit regression of a response indicator (1 = complete questionnaire, 0 = no response) on firm size (log employees), two-year asset growth, region dummies, and industry dummies (Table C.6, column 1). We then construct inverse-probability weights as the reciprocal of the predicted response probabilities, $w_i = 1/\hat{p}_i$, and use these as sampling weights in a second-stage logit regression of university information sourcing on the treatment indicator and the same set of controls, restricted to responding firms in waves 5 and 6 (Table C.6, columns 2 and 3).

Columns (2) and (3) then report weighted versions of our main logit specifications for university information sourcing: column (2) corresponds to main text Table 3, column (3), while column (3) corresponds to the specification with response-time fixed effects in Table D.1, column (2), included below. The weighted treatment coefficient is slightly larger than in the unweighted re-

Table C.5 Comparing responders and non-responders in the sampling frame

Variable	Responders				Non-responders				Two-tailed t-test	
	Mean	p50	SD	N	Mean	p50	SD	N	p-val.	sign.
Employees (log.)	3.48	3.22	1.28	548	3.56	3.33	1.27	18369	0.159	
Asset growth (2-year)	0.41	0.12	2.04	548	2.38	0.10	44.91	19011	0.000	***
Regional dummies:										
London	0.18	0	0.38	548	0.17	0	0.38	19130	0.894	
North of England	0.16	0	0.36	548	0.21	0	0.41	19130	0.003	***
Midlands	0.15	0	0.36	548	0.16	0	0.37	19130	0.691	
South of England	0.39	0	0.49	548	0.34	0	0.47	19130	0.014	**
Scotland, Wales, N. Ireland	0.13	0	0.33	548	0.12	0	0.32	19130	0.643	
Industry dummies:										
ICT, Professional & Scientific	0.25	0	0.44	548	0.17	0	0.38	19130	0	***
Other KI Services	0.17	0	0.38	548	0.17	0	0.38	19130	0.991	
Other Service and Non-Mfg. Sectors	0.40	0	0.49	548	0.46	0	0.50	19130	0.001	***
High-Tech Manuf.	0.05	0	0.23	548	0.05	0	0.21	19130	0.321	
Low-Tech Manuf.	0.12	0	0.33	548	0.14	0	0.35	19130	0.143	

gression without response-time fixed effects and slightly smaller with those included. The corresponding average marginal effect is 12.8 percentage points, compared with 12.1 percentage points in the unweighted baseline. Thus, accounting for non-response via inverse-probability weighting does not materially change the estimated treatment effect and, if anything, suggests that non-response bias would attenuate rather than inflate our baseline estimate, in line with our finding using early and late responders above.

Table C.6 Regression weighted by inverse propensity to respond

	(1)	(2)	(3)
	Propensity score estimation	Treatment effect estimation with weights	
Treatment		0.578^{***}	0.836^{***}
		(0.203)	(0.262)
Employees (log.)	-0.045	0.119	0.105
	(0.035)	(0.076)	(0.077)
Asset growth	-0.003	0.026	-0.134
	(0.004)	(0.046)	(0.109)
Start-up		-0.314	-0.223
		(0.475)	(0.498)
R&D active		0.862^{***}	0.930^{***}
		(0.197)	(0.204)
Constant	-3.381^{***}	-1.672^{***}	-2.209^{***}
	(0.186)	(0.446)	(0.642)
Region and industry fixed effects:	yes	yes	yes
Response-time fixed effects			yes
Observations	18811	548	544
Log lik.	-2456.655	-11934.189	-11665.537

Robust standard errors in parentheses. Column (1) reports the estimation of a propensity score for survey response, while columns (2) and (3) use the inverted propensity score as survey weights. Column (2) corresponds to the weighted version of Table 3, column (3), our preferred specification; column (3) corresponds to the weighted version of Table D.1, column (2). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Because the IPW weights are themselves generated from the first-stage logit, standard model-based standard errors from the second-stage weighted logit treat the weights as fixed and understate uncertainty. We therefore implement a nonparametric bootstrap that resamples firms in the sampling frame with replacement. For each of 500 bootstrap replications we re-estimate the propensity model, recompute the weights, and re-estimate the weighted logit among respondents, thus propagating first-stage estimation uncertainty into the sampling distribution of the treatment effect. The resulting bootstrapped standard error for the weighted treatment effect is 0.208, implying a p-value of 0.005, which confirms that the treatment effect remains statistically significant after adjusting for differential response propensities.

C.3 Standardized mean differences and equivalence tests

Non-rejection of the null hypothesis of equal means in t-tests does not imply that the treatment and control groups are well balanced, especially with a modest sample size (Altman and Bland, 1995). We therefore complement the t-tests reported in Table 2 of the main text with additional balance diagnostics in Table C.7 below: absolute standardized mean differences (SMDs) for each covariate as well as two one-sided equivalence tests (TOST).

Following common guidelines (Austin, 2011; Steiner and Cook, 2013), we view $|SMD| < 0.10$ as indicating negligible imbalance. For the TOST procedure we adopt the "small difference" bounds proposed by Cohen (1988): ±0.20 pooled standard deviations for continuous variables and the corresponding margin based on Cohen's $h = 0.20$ for binary variables, in line with Fitzgerald (2025) and Rainey (2014). Equivalence is concluded if the 90% confidence interval for the mean difference lies entirely within these bounds.

The diagnostics reveal mixed balance in the baseline sample. Most covariates fall within the conventional thresholds, but employees (log), asset growth, two regional dummies and three sector dummies exceed at least one of the thresholds and thus cannot be regarded as fully balanced. In the main text we address these residual imbalances by (i) including the full set of covariates, regional and sectoral fixed effects in our regression models (Section 4.2) and (ii) implementing propensity-score matching to improve covariate balance between groups prior to estimation (Section 4.3).

Table C.7 Balance diagnostics between treatment and control groups (survey waves 5 and 6)

Variable	|SMD|	TOST equivalence test: 90% CI (|diff|)	Equiv. bound	p-val.
Univ. information sourcing	0.19	0.17	0.10	0.47
Unis low importance	0.11	0.11	0.08	0.16
Unis medium importance	0.16	0.11	0.07	0.30
Unis high importance	**0.03**	0.04	0.04	**0.03**
Employees (log.)	0.34	0.62	0.26	0.95
Start-up dummy	**0.00**	0.03	0.04	**0.01**
R&D active dummy	**0.08**	0.11	0.10	**0.09**
Asset growth (2-year)	0.13	0.59	0.45	0.27
Innovator dummy	0.19	0.16	0.10	0.45
Interactor dummy	**0.04**	0.09	0.10	**0.04**
External search breadth	**0.03**	0.42	0.48	**0.03**
Regional dummies:				
London	0.15	0.12	0.08	0.31
North of England	**0.06**	0.07	0.07	**0.05**
Midlands	0.26	0.14	0.07	0.73
South of England	**0.06**	0.10	0.10	**0.06**
Scotland, Wales, N. Ireland	**0.05**	0.06	0.07	**0.05**
Industry dummies:				
ICT, Professional & Scientific	0.18	0.15	0.09	0.44
Other KI Services	**0.01**	0.06	0.08	**0.01**
Other Service and Non-Mfg. Sectors	**0.04**	0.09	0.10	**0.03**
High-Tech Manuf.	0.14	0.06	0.05	0.24
Low-Tech Manuf.	0.11	0.08	0.07	0.15

Variables and sample as in Table 2 of the main text. All variables except employees (log), asset growth and external search breadth are binary. |SMD| = absolute standardized mean difference; values of $|\text{SMD}| < 0.10$ appear in **bold**. Equiv. bound is the equivalence margin: ± 0.20 pooled SD for continuous variables, $\pm\Delta(h = 0.20)$ for binary variables (Cohen, 1988). Equivalence is concluded when the 90% CI for the mean difference lies entirely within $\pm$ the equivalence bound, which corresponds to both one-sided TOST p values being below the 0.1 threshold (represented in **bold**; for simplicity, only the larger p-value is reported). "Other KI Services" refers to other knowledge-intensive services.

D. Further results

D.1 Robustness of results across survey waves

The baseline regressions in Table 3 of the main text use only responses from survey waves 5 and 6. These waves were drawn at random from the same FAME-based sampling frame and were fielded in close temporal proximity, under the same level of national COVID-19 restrictions (see Figure 1 in the main text). To assess whether our results depend on this sample choice, we re-estimate the baseline specification including additional waves. Table D.1 reports estimates when (i) expanding the treatment group to include waves 1-4 and (ii) expanding it to include waves 7-9.

Waves 1-4 are indistinguishable from waves 5 and 6 in the sampling frame except for employment numbers (in the work for Hughes et al., 2022, employment was only considered in four ordinal categories). However, Figure 1 shows that responses in waves 1-3 were collected considerably closer to the onset of the COVID-19 pandemic than waves 4-6. This temporal difference may have affected which firms were willing and able to respond to a survey on their interactions with universities. When waves 1-4 are added to the treatment group, the estimated treatment effect decreases slightly but remains positive and statistically significant.

Waves 7-9 were fielded in close temporal proximity to wave 6, which reduces concerns about changing pandemic conditions. However, these waves were sampled from a revised FAME extract obtained after it became clear that contact details in the original extract provided low coverage, particularly for smaller firms. The updated extract had better coverage of email addresses but a somewhat different composition, with relatively fewer small firms. As a result, including waves 7–9 increases the mean size in the treatment group by around two employees, although the median size is unchanged.[5] When adding these later waves to the treatment group, the estimated treatment effect increases slightly. Overall, across all definitions of the treatment sample, the estimated treatment effects are of similar magnitude and remain statistically significant, indicating that our main results are not driven by the choice of survey waves.

D.2 Controlling for external search breadth

Column (5) of Table 3 in the main text adds a measure of external search breadth, defined as the number of external information sources other than universities that the firm reports as being of low, medium or high importance for innovation , as used in Laursen and Salter (2004). Including this variable yields a slightly smaller treatment effect estimate closer to the unconditional mean difference reported in Table 2. However, caution is required when interpreting this result. As is apparent from the substantial increase in the calculated log-likelihood and pseudo-R^2, external search breadth is a strong predictor of university importance. This high explanatory power is to be expected, considering that both variables derive from mutually exclusive parts of the same survey question about the importance of various external information sources for innovation. A firm considering a wider range of external information sources is also more likely to identify universities as one such source. Due to the high correlation of 0.47 between external search breadth and the university dummy, we do not include external search breadth as a control variable in subsequent analysis.

[5] Consequently, when expanding the dataset to include later survey waves, the difference in mean employment between control and treatment groups is no longer significant.

Table D.1 Regressions in different samples

Regression sample:	Waves 5-6		Waves 1-6		Waves 6-7	Waves 6-9	All observations	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.584***	0.884***	0.336**	0.653***	0.808***	0.441***	0.370**	0.444***
	(0.218)	(0.256)	(0.163)	(0.235)	(0.201)	(0.165)	(0.156)	(0.162)
Employees (log.)	0.068	0.109	0.117**	0.113**	0.139*	0.082*	0.105***	0.101***
	(0.082)	(0.076)	(0.046)	(0.046)	(0.072)	(0.044)	(0.033)	(0.033)
Start-up	-0.720	-0.264	-0.566*	-0.535*	0.048	-0.303	-0.481**	-0.463**
	(0.520)	(0.450)	(0.307)	(0.309)	(0.452)	(0.314)	(0.233)	(0.235)
R&D active	0.918***	0.879***	0.667***	0.690***	0.459**	0.440***	0.566***	0.570***
	(0.221)	(0.197)	(0.114)	(0.115)	(0.200)	(0.128)	(0.089)	(0.090)
Asset growth	0.025	-0.125	0.030	-0.016	0.057	0.037	0.031	0.005
	(0.046)	(0.089)	(0.036)	(0.048)	(0.046)	(0.036)	(0.032)	(0.038)
Constant	-1.532***	-2.344***	-1.417***	-1.794***	-1.412***	-1.106***	-1.302***	-1.463***
	(0.469)	(0.613)	(0.290)	(0.365)	(0.528)	(0.324)	(0.244)	(0.284)
Response-time fixed effects		yes		yes				yes
Observations	413	544	1535	1531	499	1163	2432	2429
Log lik.	-261.848	-343.058	-999.485	-992.270	-319.679	-762.526	-1600.642	-1593.047
Pseudo-R2	0.069	0.076	0.043	0.047	0.072	0.045	0.038	0.042
Average Marginal Effect of Treatment	0.130***	0.190***	0.076**	0.141***	0.183***	0.100***	0.084**	0.100***
	(0.047)	(0.051)	(0.036)	(0.047)	(0.044)	(0.036)	(0.034)	(0.035)

Note: in all columns, the control group is identical and consists of firms allocated to wave 6, as only wave 6 received the control version of the questionnaire. The respective treatment group is constructed as described in the top row: wave 5 in columns (1) and (2), waves 1-5 in columns (3) and (4), wave 7 in column 5, waves (7)-(9) in column 6, and all waves except wave 6 in columns (7) and (8). In column (1), we drop all observations from the baseline regression sample for which we were unable to deliver the original survey invitation and that we reassigned to wave 7 with an updated email address. The balance checks in this restricted sample look very similar to those reported in Table 2. All regressions contain the set of region and industry dummies. Response-time fixed effects are achieved by including dummies indicating the month in which the response was recorded and whether it was the first or second half of the month (e.g., "March, 1st half", "March, 2nd half", "April, 1st half", and so on). The joint significance of these dummies is rejected in all regressions. Parentheses report robust standard errors.

D.3 Further details on the re-weighting approaches

Table D.2 reports descriptive statistics for the main sample after re-weighting by covariate-balancing propensity scores (CBPS; (Imai and Ratkovic, 2014). The weights are constructed using log employment, two-year asset growth, the start-up and R&D dummies, and the full sets of regional and industry dummies as balancing covariates. Under these weights, all covariates achieve excellent balance between treatment and control firms, with absolute standardized mean differences below 0.01 and all two one-sided equivalence tests (TOST) indicating equivalence within the ±0.20 SD bounds used in Section C.3 above.

For comparison, we also implemented propensity-score matching (PSM) and entropy balancing (EB; (Hainmueller, 2012). Both achieve substantial improvements in balance relative to the unweighted sample; EB reaches a balance profile similar to CBPS but does so by up-weighting only the smaller control group, which reduces its effective sample size. In contrast, CBPS reweights both groups toward a common target, leading to a higher effective sample size and more stable estimates. As reported in Table 4 of the main text, all three sets of weights yield very similar estimated treatment effects (12.2–12.5 percentage points), reinforcing the conclusion that our findings are not driven by residual covariate imbalance. Detailed balance tables for PSM and EB are available from the authors on request.

D.4 Relative importance

The dependent variable in Table 5 has four ordered categories (“not used”, “low”, “medium”, “high” importance). A natural benchmark is therefore the ordered logit model. However, Brant tests indicate that the parallel-regressions assumption is violated for several covariates, including the treatment indicator, so we do not rely on ordered logit in the main text.

To verify the robustness of our findings, we estimate two sets of alternative models. First, we run a series of independent binary logistic regressions that compare each positive importance level (low, medium, high) separately against the base category “not used”. Second, we estimate a generalized ordered logit model (Williams, 2006, 2016), which relaxes the parallel-slopes assumption for selected covariates. In both approaches, the estimated treatment effects display the same qualitative pattern as in Table 5: the treatment significantly increases the probability of reporting low or medium importance relative to “not used”, while the effect on reporting high importance is small and statistically insignificant.

We also replicate the multinomial and generalized ordered logit models on the full sample, including firms that do not report any university interaction, and report the results from the multinomial regressions in Table D.3 below. The sign pattern of the treatment coefficients remains unchanged: the treatment shifts probability mass from “not used” to “low” and “medium” importance, with no systematic effect on “high” importance. Effect sizes are smaller and less precisely estimated, reflecting that the treatment effect is concentrated among firms that interact with universities, while we now also include firms with no reported engagement.

Table D.2 Summary statistics of the preferred re-weighted sample

	Control group			Treatment group					TOST equivalence test		
	N=200			N=348			Mean		90% CI	Equiv.	
Variable	mean	SD	med.	mean	SD	med.	|diff|	|SMD|	(|diff|)	bound	p-val.
Univ. information sourcing	0.35			0.47			0.12	0.25	0.20	0.10	0.70
Unis low importance	0.18			0.24			0.06	0.14	0.12	0.08	0.26
Unis medium importance	0.11			0.18			0.06	0.14	0.12	0.07	0.40
Unis high importance	0.05			0.05			0.00	**0.01**	0.04	0.04	**0.02**
Employees (log.)	3.47	1.23	3.22	3.47	1.22	3.26	0.00	**0.00**	0.16	0.23	**0.01**
Start-up dummy	0.04			0.04			0.00	**0.00**	0.03	0.04	**0.01**
R&D active dummy	0.46			0.46			0.00	**0.00**	0.08	0.10	**0.02**
Asset growth (2-year)	0.38	1.98	0.16	0.38	1.71	0.11	0.00	**0.00**	0.16	0.27	**0.00**
Innovator dummy	0.62			0.57			0.05	0.10	0.12	0.10	0.13
Interactor dummy	0.57			0.60			0.03	**0.06**	0.10	0.10	**0.06**
External search breadth	6.74	2.26	7	6.88	2.52	7	0.14	**0.06**	0.52	0.48	**0.07**
Regional dummies:											
London	0.18			0.18			0.00	**0.00**	0.06	0.08	**0.01**
North of England	0.16			0.16			0.00	**0.00**	0.06	0.07	**0.03**
Midlands	0.16			0.16			0.00	**0.00**	0.06	0.07	**0.03**
South of England	0.38			0.38			0.00	**0.00**	0.07	0.10	**0.02**
Scotland, Wales, N. Ireland	0.13			0.13			0.00	**0.00**	0.05	0.07	**0.02**
Industry dummies:											
ICT, Professional & Scientific	0.24			0.24			0.00	**0.00**	0.06	0.09	**0.01**
Other KI Services	0.18			0.18			0.00	**0.00**	0.06	0.08	**0.02**
Other Service and Non-Mfg. Sectors	0.40			0.40			0.00	**0.00**	0.08	0.10	**0.02**
High-Tech Manuf.	0.06			0.06			0.00	**0.00**	0.04	0.05	**0.03**
Low-Tech Manuf.	0.12			0.12			0.00	**0.00**	0.05	0.07	**0.02**

The notes below Table 2 apply. Beyond those, in this table, statistics are calculated using weights from covariate-balancing propensity score matching (CBPS) (Imai and Ratkovic, 2014; Kranker et al., 2021) on all covariates. Entropy balancing (EB) yields a very similar table. Standard propensity score matching (PSM) performs less well on equalizing some of the region and sector dummies. The three matching approaches yield Kish effective sample sizes of 172.18 (PSM), 177.60 (CBPS), and 157.09 (EB).

This pattern is consistent with the design of the innovation knowledge sourcing question, which follows the CIS format and asks respondents to classify universities as "not used" or as a source of "low", "medium", or "high" importance for their innovation activities. Firms that did not use universities as an information source in the reference period can only choose "not used", even if they nonetheless regard universities as generally important (for example, as providers of skilled labor or interaction partners of other firms in their supply chain). By contrast, firms that select an importance level must have used universities as a knowledge source to some extent, which makes it likely that they would identify at least one engagement channel from Table 1 as applicable to their case.[6]

Overall, these alternative specifications confirm that the main impact of the information treatment is to shift firms from reporting "no use" of universities to acknowledging low or medium importance, with little effect on the high-importance category.

Table D.3 Multinomial regression of relative university importance including firms without university engagement

	(1)	(2)	(3)
Low importance (variable = 1)			
Treatment=1	0.378*	0.465**	0.479**
	(0.226)	(0.229)	(0.234)
Medium importance (variable = 2)			
Treatment=1	0.543**	0.690**	0.704**
	(0.267)	(0.283)	(0.292)
High importance (variable = 3)			
Treatment=1	0.046	0.232	0.404
	(0.404)	(0.421)	(0.451)
Control variables		yes	Yes
Region and industry fixed effects			Yes
Observations	548	548	548
Log lik.	-598.224	-579.459	-562.083

Parentheses report robust standard errors. The specifications are identical to those reported in Table 5 in the main paper, except that the sample here is not restricted to "interactors" but also includes firms that did not select even a single form of university engagement they were active in during the three years prior to the survey. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

[6] Of the 236 firms in our baseline sample that report using universities as a source of knowledge for innovation, 205 (≈ 87%) also select at least one engagement channel from Table 1. Conversely, of the 312 firms that report not using universities as a knowledge source, 128 (41%) nonetheless report some form of university engagement. Hence, while a sizeable share of firms engages with universities without perceiving them as an information source for innovation, the opposite case is (naturally) very rare.

D.5 Placebo tests

Columns (1)–(3) of Table D.4 analyze an indicator for having used at least one university engagement channel and show no robust treatment effect independent of which controls are included. Column (4) examines the number of engagement channels used; again, the treatment coefficient is small and statistically insignificant. These results suggest that moving the engagement section from the first to the second page of the questionnaire does not materially affect how often firms report engaging with universities or how many channels they select.

Finally, column (5) estimates the treatment effect separately for firms with and without prior university engagement and finds a significant effect only for the former group of firms (row 1 + row 3). This finding supports our view of firms re-assessing the importance of universities based on their actual engagement history.

Table D.4 Regression of university engagement channel variables

	(1)	(2)	(3)	(4)	(5)
Dependent variable:	Interactor dummy			Number of engagement channels	University dummy
Treatment	-0.090	0.054	0.147	0.108	0.113
	(0.181)	(0.193)	(0.201)	(0.111)	(0.418)
Interactor					1.959^{***}
					(0.409)
Treatment × Interactor					0.645
					(0.486)
Control variables		yes	yes	yes	yes
Region and industry fixed effects			yes	yes	yes
Observations	548	548	548	548	548
Log lik.	-370.546	-343.385	-328.213	-1544.483	-290.096
Pseudo-R2	0.000	0.074	0.115	0.146	0.225

Parentheses report robust standard errors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

References

Altman, D.G., Bland, J.M., 1995. Statistics notes: Absence of evidence is not evidence of absence. British Medical Journal 311.

Armstrong, J.S., Overton, T.S., 1977. Estimating nonresponse bias in mail surveys. Journal of Marketing Research 14, 396-402.

Austin, P.C., 2011. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behavioral Research 46, 399-424.

Bruneel, J., D'Este, P., Salter, A., 2016. The impact of financial slack on explorative and exploitative knowledge sourcing from universities: evidence from the UK. Industrial and Corporate Change 25, 689-706.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences, 2nd ed. Lawrence Erlbaum Associates.

Cohen, W.M., Nelson, R.R., Walsh, J.P., 2002. Links and impacts: the influence of public research on industrial R&D. Management Science 48, 1-23.

Fitzgerald, J., 2025. The Need for Equivalence Testing in Economics, I4R Discussion Paper Series.

Fontana, R., Geuna, A., Matt, M., 2006. Factors affecting university–industry R&D projects: The importance of searching, screening and signalling. Research Policy 35, 309-323.

Hainmueller, J., 2012. Entropy Balancing for Causal Effects: A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. Political Analysis 20, 25-46.

Han, L., Tian, Z., Wojan, T., Goetz, S., 2024. Testing biasedness of self-reported microbusiness innovation in the annual business survey. PLoS ONE 19.

Hervas-Oliver, J.-L., Albors-Garrigos, J., Baixauli, J.-J., 2012. Beyond R&D activities: the determinants of firms' absorptive capacity explaining the access to scientific institutes in low–medium-tech contexts. Economics of Innovation and New Technology 21, 55-81.

Hughes, A., Kitson, M., Salter, A., Angenendt, D., Hughes, R., 2022. The Changing State of Business-University Interactions in the UK, 2005-2021. National Centre for Universities & Business (NCUB) and Centre for Business research (CBR).

Imai, K., Ratkovic, M., 2014. Covariate Balancing Propensity Score. Journal of the Royal Statistical Society Series B: Statistical Methodology 78, 243–263.

Kranker, K., Blue, L., Vollmer Forrow, L., 2021. Improving Effect Estimates by Limiting the Variability in Inverse Propensity Score Weights. The American Statistician 75.

Laursen, K., Salter, A., 2004. Searching high and low: what types of firms use universities as a source of innovation? Research Policy 33, 1201-1215.

Little, R.J.A., Rubin, D.B., 2019. Statistical Analysis with Missing Data 3rd ed. John Wiley & Sons, Ltd.

Morrissey, K., Spooner, F., Salter, J., Shaddick, G., 2021. Area level deprivation and monthly COVID-19 cases: The impact of government policy in England. Social Science & Medicine 289, 114413.

Munford, L., Khavandi, S., Bambra, C., Barr, B., Davies, H., Doran, T., Kontopantelis, E., Norman, P., Pickett, K., Sutton, M., Taylor-Robinson, D., Wickham, S., 2021. A year of COVID-19 in the North: Regional inequalities in health and economic outcomes. Northern Health Sciences Alliance, Newcastle upon Tyne, p. 44.

Rainey, C., 2014. Arguing for a Negligible Effect. American Journal of Political Science 58, 1083–1091.

Steiner, P.M., Cook, D., 2013. Matching and Propensity Scores, in: Little, T., D, (Ed.), The Oxford Handbook of Quantitative Methods in Psychology. Oxford University Press.

Tether, B.S., Tajar, A., 2008. Beyond industry-university links: Sourcing knowledge for innovation from consultants, private research organisations and the public science-base. Research Policy 37, 1079-1095.

Williams, R., 2006. Generalized Ordered Logit/Partial Proportional Odds Models for Ordinal Dependent Variables. The Stata Journal: Promoting communications on statistics and Stata 6, 58-82.

Williams, R., 2016. Understanding and interpreting generalized ordered logit models. The Journal of Mathematical Sociology 40, 7-20.

Wooldridge, J.M., 2007. Inverse probability weighted estimation for general missing data problems. Journal of Econometrics 141, 1281–1301.