

Evidence Transmission - Pre-Registered Tables

Mahvish Shaukat, Andreas Stegmann, and Mattie Toma*

August 2024

This document shows the main tables for “Evidence Transmission Among Policymakers”, which was pre-registered as AEARCTR-0012551. As specified in the pre-registration plan, we are using this document to finalize experimental specifications (including outcome and control variables) using a “fake” randomization procedure (Olken, 2015). In particular, we have now collected all study data and have re-randomized treatment assignments. The tables below use these “fake” treatment assignments; as such, while the specifications are final, we have not yet looked at the treatment data. We will generate real results for the tables shown below using the actual randomization only once this PAP has been filed.

In addition to the tables included in this PAP, we intend to examine heterogeneous treatment effects by the following seed and contact-level characteristics:

- Education (PhD, Masters or Professional Degree)
- Gender
- Baseline ChatGPT use and preferences
- An indicator for a research-oriented position

We also may conduct the following exploratory analyses:

- Long-run follow-up, for contacts in Wave 2 with divisions seeded in Wave 1
- Spillovers, using our saturation design
- Behavior change outcomes, e.g. whether contacts are more likely to use ChatGPT

*Shaukat: mshaukat@worldbank.org, The World Bank; Stegmann: andreas.stegmann@warwick.ac.uk, University of Warwick; Toma: mattie.toma@warwick.ac.uk, University of Warwick.

- Sharing of optimistic/pessimistic information about ChatGPT use at the World Bank (building on the optimistic information treatment for seeds), using contacts' responses to the incentivized questions about ChatGPT use and preferences
- Descriptive analyses, including when and why World Bank employees engage with research evidence and timing of the link engagements.
- Treatment saturation for contacts in divisions seeded once or twice

1 Seed Tables

Table 1: Treatment Effects on Seed Engagement and Sharing

	(1) Sharing Intention	(2) Any sharing	(3) Infographic Engage	(4) Article Engage
Panel A: Full Sample				
High-Ranked	0.028 (0.132)	-0.009 (0.054)	0.060 (0.168)	-0.120 (0.133)
Optimistic Information	0.020 (0.133)	0.016 (0.051)	0.030 (0.167)	0.038 (0.134)
Observations	336	337	335	335
Controls	Yes	Yes	Yes	Yes
Control Mean	2.33	.45	1.49	.93
Panel B: Wave 1				
High-Ranked	0.125 (0.159)	-0.053 (0.074)	-0.052 (0.253)	-0.198 (0.192)
Optimistic Information	-0.142 (0.161)	0.009 (0.071)	0.033 (0.264)	-0.011 (0.198)
Strong Credibility	0.257 (0.160)	-0.070 (0.072)	-0.045 (0.272)	0.178 (0.198)
Observations	204	205	205	205
Controls	Yes	Yes	Yes	Yes
Control Mean	2.48	.58	2.24	1.3
Panel C: Wave 2				
High-Ranked	-0.078 (0.228)	0.080 (0.079)	0.326** (0.161)	0.064 (0.165)
Optimistic Information	0.418* (0.243)	0.026 (0.083)	0.079 (0.157)	0.079 (0.169)
Observations	132	132	130	130
Controls	Yes	Yes	Yes	Yes
Control Mean	2.12	.26	.38	.38

This table shows the impact of optimistic information, sharing with high-ranked seeds, and strong credibility on engagement with evidence, for all seeds. Column 1 reflects self-reported intentions to share evidence with colleagues, on a scale from 0 to 4. Columns 2 is an indicator equal to one if either of the shared links was engaged with more than once. Columns 3 and 4 reflect actual engagement with an infographic and journal article relevant to the evidence, as measured by the number of clicks on links to the materials. Controls include an indicator for whether the seed has a Masters or Professional degree; an indicator for whether the seed has a PhD; an indicator for whether the seed is female; a count of the number of days the seed typically comes into the office; an indicator for whether the seed has ChatGPT; self-reported preference for using ChatGPT, on a likert scale from 1 to 5; standardized priors on the evidence for Waves 1 and 2; an indicator for whether an observation is missing demographic information; an indicator for the wave; and strata fixed effects. Standard errors, clustered at the division level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table 2: Treatment Effects on Seeds' Incentivized Second-Order Beliefs

	(1)	(2)
	Predict % Use Chat	Predict ChatPref
Panel A: Full Sample		
High-Ranked	-0.933	1.030
	(2.553)	(2.678)
Optimistic Information	0.658	0.738
	(2.471)	(2.611)
Observations	337	337
Controls	Yes	Yes
Control Mean	41.69	49.14
Panel B: Wave 1		
High-Ranked	3.022	2.152
	(3.319)	(3.629)
Optimistic Information	3.389	4.263
	(3.148)	(3.379)
Strong Credibility	1.108	-4.198
	(3.126)	(3.416)
Observations	205	205
Controls	Yes	Yes
Control Mean	38.74	47.2
Panel C: Wave 2		
High-Ranked	-6.760	-0.369
	(4.174)	(4.282)
Optimistic Information	-3.864	-3.678
	(4.318)	(4.601)
Observations	132	132
Controls	Yes	Yes
Control Mean	46.26	52.14

This table shows the impact of optimistic information, sharing with high-ranked seeds, and strong credibility on second-order beliefs, for all seeds. Column 1 reflects incentivized predictions about the percent of one's colleagues who use ChatGPT. Column 2 reflects incentivized predictions about the percent of one's colleagues who belief employees *should* use ChatGPT. Controls include an indicator for whether the seed has a Masters or Professional degree; an indicator for whether the seed has a PhD; an indicator for whether the seed is female; a count of the number of days the seed typically comes into the office; an indicator for whether the seed has ChatGPT; self-reported preference for using ChatGPT, on a likert scale from 1 to 5; standardized priors on the evidence for Waves 1 and 2; an indicator for whether an observation is missing demographic information; an indicator for the wave; and strata fixed effects. Standard errors, clustered at the division level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table 3: Treatment Effects on Seeds' Evidence Perceptions

	(1) Credible	(2) Interesting	(3) Valuable Colleagues	(4) Valuable Personally
Panel A: Full Sample				
High-Ranked	-0.018 (0.110)	0.253** (0.105)	0.152 (0.101)	-0.066 (0.105)
Optimistic Information	0.050 (0.108)	0.008 (0.107)	0.087 (0.106)	0.075 (0.110)
Observations	337	337	337	337
Controls	Yes	Yes	Yes	Yes
Panel B: Wave 1				
High-Ranked	0.080 (0.134)	0.332** (0.131)	0.207 (0.140)	0.062 (0.144)
Optimistic Information	-0.116 (0.132)	-0.150 (0.135)	-0.055 (0.140)	-0.117 (0.141)
Strong Credibility	-0.165 (0.143)	0.012 (0.138)	0.065 (0.137)	-0.019 (0.140)
Observations	205	205	205	205
Controls	Yes	Yes	Yes	Yes
Panel C: Wave 2				
High-Ranked	-0.059 (0.174)	0.191 (0.173)	0.208 (0.162)	-0.158 (0.166)
Optimistic Information	0.357** (0.175)	0.211 (0.165)	0.269 (0.173)	0.300 (0.182)
Observations	132	132	132	132
Controls	Yes	Yes	Yes	Yes

This table shows the impact of optimistic information, sharing with high-ranked seeds, and strong credibility on perceptions about evidence, for all seeds. Column 1 looks at a standardized measure of beliefs about the *credibility* of the evidence in the experiment. Column 2 looks at a standardized measure of beliefs about how *interesting* the evidence in the experiment is. Column 3 looks at a standardized measure of beliefs about how *valuable* the evidence in the experiment is for one's colleagues. Column 4 looks at a standardized measure of beliefs about how *valuable* the evidence in the experiment is for one personally. Controls include an indicator for whether the seed has a Masters or Professional degree; an indicator for whether the seed has a PhD; an indicator for whether the seed is female; a count of the number of days the seed typically comes into the office; an indicator for whether the seed has ChatGPT; self-reported preference for using ChatGPT, on a likert scale from 1 to 5; standardized priors on the evidence for Waves 1 and 2; an indicator for whether an observation is missing demographic information; an indicator for the wave; and strata fixed effects. Standard errors, clustered at the division level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

2 Contact Tables

Table 4: Treatment Effects on Contact Evidence Awareness

	(1)	(2)
	Recognize Study (Count)	Recognize Study (≥ 1)
Treatment vs. Control		
Treatment	0.046* (0.025)	0.022 (0.014)
Observations	380	380
Controls	Yes	Yes
Control Mean	.02	.02
Within-Treatment		
Optimistic Information	-0.023 (0.029)	-0.016 (0.021)
High-Ranked	0.118** (0.044)	0.076** (0.027)
Strong Credibility	-0.059 (0.033)	-0.037 (0.022)
Observations	292	292
Controls	Yes	Yes
Control Mean	.05	.03
Long-run Follow-up - Treatment vs. Control		
Treatment	-0.082 (0.055)	-0.055 (0.035)
Observations	204	204
Controls	Yes	Yes
Control Mean	.06	.04
Long-run Follow-up - Within Treatment		
Optimistic Information	0.010 (0.020)	0.005 (0.010)
High-Ranked	-0.018 (0.024)	-0.009 (0.012)
Strong Credibility	-0.022 (0.022)	-0.011 (0.011)
Observations	153	153
Controls	Yes	Yes
Control Mean	.01	.01

This table shows the impact of having a treated seed as well as seeds in the optimistic information, sharing with high-ranked seeds, and strong credibility treatments on awareness of the evidence shared with a seed. Panel A compares contacts in treated divisions with those in control divisions in Wave 1. Panel B compares contacts across seeds treatments in Wave 1. Panel C looks at long-run outcomes for contacts in Wave 2 with treated divisions in Wave 1. Column 1 reflects a hand-coded count of the number of study features recognized out of: author, institutional affiliation, journal, and content in the study. Column 2 is an indicator for whether at least one study feature is recognized. Controls include an indicator for whether the contact has a Masters or Professional degree; an indicator for whether the contact has a PhD; an indicator for whether the contact is female; a count of the number of days the contact typically comes into the office; an indicator for whether an observation is missing demographic information; the seed-level controls included in the tables above; and strata fixed effects. Standard errors, clustered at the division level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

Table 5: Treatment Effects on Contact Evidence Awareness - Secondary Outcomes

	(1)	(2)	(3)	(4)
	Confidence Impact	Discuss AI	Receive Email	Recognize Email
Wave 1				
Treatment vs. Control				
Treatment	-3.806 (3.903)	-0.006 (0.056)	0.023 (0.055)	
Observations	370	380	380	
Controls	Yes	Yes	Yes	
Control Mean	65.33	.75	.16	
Within-Treatment				
Optimistic Information	1.572 (4.424)	-0.164** (0.052)	-0.100 (0.056)	
High-Ranked	-1.433 (5.223)	0.020 (0.062)	-0.031 (0.060)	
Strong Credibility	-2.927 (4.977)	-0.040 (0.054)	0.044 (0.054)	
Observations	284	292	292	
Controls	Yes	Yes	Yes	
Control Mean	62.4	.72	.18	
Wave 2				
Treatment vs. Control				
Treatment	-11.339 (8.733)	-0.037 (0.079)	-0.016 (0.039)	0.005 (0.005)
Observations	257	257	257	257
Controls	Yes	Yes	Yes	Yes
Control Mean	37.73	.55	.08	0
Within-Treatment				
Optimistic Information	7.349 (9.959)	-0.016 (0.092)	-0.006 (0.045)	0.015 (0.012)
High-Ranked	3.296 (8.615)	-0.028 (0.073)	-0.083* (0.041)	-0.011 (0.010)
Observations	193	193	193	193
Controls	Yes	Yes	Yes	Yes
Control Mean	35.02	.53	.09	.01

This table shows the impact of having a treated seed as well as seeds in the optimistic information, sharing with high-ranked seeds, and strong credibility treatments on secondary measures of awareness of the evidence shared with a seed. Panel 1 shows contacts surveyed in Wave 1 with seeded divisions in Wave 1. Panel 2 shows contacts surveyed in Wave 2 with seeded divisions in Wave 2. For all outcomes, Wave 1 outcomes are relevant to ChatGPT impacts on productivity while Wave 2 outcomes are relevant to ChatGPT impacts on creativity. Column 1 is the confidence in the perceived impact of ChatGPT on productivity/creativity from 0 to 100, when the sign of the outcome is flipped if the perceived impact is negative and zero if the perceived impact is zero. Column 2 is an indicator for whether the contact has discussed the impacts of AI on productivity/creativity with colleagues. Column 3 is an indicator for whether a contact self-reports having received an email about the impacts of AI on productivity/creativity. Column 4 is an indicator for whether a contact self-reports recognizing the infographic was shared. Controls are the same as those denoted in Table 4. Standard errors, clustered at the division level, are reported in parentheses. ***, **, * denote that estimates are statistically significant at the 1%, 5%, and 10% levels, respectively.

References

Olken, Benjamin, “Promises and Perils of Pre-Analysis Plans,” *Journal of Economic Perspectives*, 2015, 29 (3), 61–80.