Populated Pre-analysis plan for "Strategic Use of Unfriendly Leadership and Labor Market Competition: An Experimental Analysis"

Danilov, Hong, Schöttner

Humboldt-Universität zu Berlin

September 24, 2024

This document reports additional of the results of the experiment pre-registered as AEARCTR-0009365 in connection to the pre-analysis plan submitted at the time of pre-registration. Here, we also identify and specify any deviations from the original pre-analysis plan. The complete analysis is available in our research paper.

1. Research questions

As stated in the pre-analysis plan, our primary research question was to examine how managers adjust their leadership styles based on labor market conditions. Additionally, we aimed to analyze the impact of these leadership styles on workers' behaviors, such as contract acceptance and effort levels. To address these questions, we conducted experiments at the Laboratory for Experiments in Economics at the Technische Universität Berlin in October and December 2022.

2. Treatments

In the pre-analysis plan, we describe an experimental design with variations in three following dimensions:

    (i)   the choice set of leadership styles for the managers ('*friendly and unfriendly*' versus '*unfriendly and no leadership style*'),

    (ii)   market conditions concerning competitiveness (*ELD* versus *ELS*, standing for Excess Labor Demand and Excess Labor Supply),

    (iii)   the source of leadership behavior (*Manager-Message, Computer-Message*, and *No-Message*).

As of September 2024, the following experimental conditions were conducted:

		Leadership style	
		Unfriendly	Friendly and unfriendly
Market condition	Excess Labor Demand	x	x
	Excess Labor Supply	x	x

The third dimension (the source of leadership behavior) was kept constant in all sessions conducted so far: the source of leadership behavior was the human managers' choice. The main reason for not conducting these sessions was the empty subject pool mainly due to the COVID-19 pandemic. Thus, the final paper focuses on the research question concerning the impact of market competition on managers' use of leadership styles. Manipulation of the third dimension to shed light on the interplay between the two mechanisms underlying the effectiveness of leadership styles—incentive and reciprocal effects—is left for future research.

Procedural details of the experiment

We conducted two sets of sessions, each varying in the first dimension—the choice set of leadership styles for managers. Each set contained 16 experimental sessions, as pre-registered. These sets are

referred to as the first and second experiments in our research paper. The second dimension, market competition conditions, varied within each session.

3. Data analysis plan, populated

3.1 Definitions of variables

Our primary outcome variable was the managers' choice of leadership style, which we denoted by m_i, for a subject i in the role of a manager. The number of unfriendly (or friendly) messages in our research paper refers to this variable.

As we stated in Section 2 of our pre-analysis plan, we also analyze managers' adoption of friendly (unfriendly) leadership. To address this point, in the research paper, we use the binary variable 'Adoption of unfriendly leadership' for the first experiment and the categorical variable 'Adoption of leadership style' for the second experiment.

Regarding the variables marking the workers' decisions, $accept_j$ in the pre-analysis plan is denoted with 'Contract acceptance (0/1)' in the research paper, and the variable $effort_j$ remains the same.

The treatment dummy for the market conditions is defined as we planned; $ELS_i = 1$ if a subject i is assigned to the ELS treatment, $ELS_i = 0$ the ELD treatment.

The variables $belief\_accept_i$ and $belief\_effort_i$ in the pre-analysis plan were defined for the beliefs in the first experiment. We report the analysis in supplementary material C4 of our research paper. We also report the beliefs in the second experiment as exploratory results in Section 7.1 because the detailed method of belief elicitation in the second experiment was not pre-specified in our pre-registration.

In our pre-analysis plan, we stated that we would include the following covariates in our analysis and present a balance test to see whether some of the measures tend to be relevant for the use of leadership style or reaction to it: perceived social appropriateness of the leadership style $socially\_appropriate_i$, perceived fairness of the leadership style $fair_i$, the tendency for positive and negative reciprocity $positive\_reciprocity_i$, $negative\_reciprocity_i$, $trust_i$, $risk\_taking_i$, Big 5 personality traits, Machiavellianism and Psychopathy traits, gender, age, education, and work experience. All these variables were measured in the questionnaire at the end of the first experiment and reported in the paper. Due to restrictions regarding the duration of experimental sessions and other resources, the variables $socially\_appropriate_i$ and $fair_i$ were not collected in the second experiment, and hence, their analysis is not reported in the research paper. Ex-post, we considered it to be useful to add also specific belief elicitation questions in the second experiment. Tables A3 and A8 in the appendix A2 report the analysis.

Table 1 below reports the mean and standard deviations of the variables $socially\_appropriate_i$ and $fair_i$ by treatment, in the first experiment, regarding the use of unfriendly leadership. Overall, subjects find it very socially inappropriate and very unfair to send 10 unfriendly messages, while sending only one message is rated as more acceptable. We do not find any treatment effect on the ratings (see P-values resulting from the two-sided t-test, reported in Table 1).

Table 1. **Social appropriateness and fairness of unfriendly leadership, by treatment**

Question	Number of messages	Treatment		P-value
		ELD	ELS	(t-test)
How socially appropriate do you consider sending [0, 1, 10] statements of Style B to a worker who fails on a project?	0	0.38	0.40	0.799
		(0.70)	(0.66)	
	1	0.10	0.13	0.659
		(0.63)	(0.63)	
	10	-0.81	-0.85	0.526
		(0.49)	(0.44)	
What do you think was the most common response to the same question as above among 10 randomly selected managers in this experiment?	0	0.28	0.33	0.639
		(0.66)	(0.67)	
	1	0.14	0.19	0.518
		(0.56)	(0.59)	
	10	-0.76	-0.79	0.720
		(0.51)	(0.51)	
How fair do you consider sending [0, 1, 10] statements of Style B to a worker who fails on a project?	0	0.28	0.20	0.371
		(0.68)	(0.71)	
	1	0.14	0.15	0.883
		(0.57)	(0.59)	
	10	-0.82	-0.70	0.055*
		(0.39)	(0.50)	
What do you think was the most common response to the same question as above among 10 randomly selected managers in this experiment?	0	0.14	0.07	0.460
		(0.65)	(0.68)	
	1	0.13	0.18	0.485
		(0.51)	(0.58)	
	10	-0.75	-0.68	0.286
		(0.47)	(0.54)	

Mean and standard deviations (in parentheses below means) by treatments. The number of observations, N=96 in the ELD treatment and N=160 in the ELS treatment. For social appropriateness questions, subjects rate each possible case of sending 0, 1, or 10 statements on 4-point Likert scale: (1) "very socially inappropriate", (2) "somewhat socially inappropriate", (3) "somewhat socially appropriate", or (4) "very socially appropriate". For fairness questions, we used the following 4-point Likert scale: (1) "very unfair", (2) "somewhat unfair", (3) "somewhat fair", or (4) "very fair". Following Krupka and Weber (2013), we centered the ratings to the scores to zero and converted them to the values -1, -1/3, 1/3, and 1. The last column reports p-values from two-sided t-test by treatment. $^{*} p < 0.10$, $^{**} p < 0.05$, $^{***} p < 0.01$.

3.2 Model specifications

We proposed the following OLS model to test the effect of market conditions on managers' strategic use of leadership styles:

$$m = X\beta + \delta \cdot ELS + session\ FE + \varepsilon,$$

where the matrix X contains covariates and ε stands for the error term. The pre-registered dependent variable is m, manager's leadership choice. In our analysis, we report different estimations of the chosen leadership style (see Tables 2 and 5 of the paper): (i) as a categorical (or binary) variable indicating the chosen leadership style, (ii) as the number of chosen messages of a certain leadership style. Where appropriate, we rely on the Probit, ordered Probit, and Tobit estimations, as we believe

these methods are better suited for the binary, censored, and categorical variables. Estimating the same specifications with the method of Ordinary Least Squares virtually does not change the results. In the research paper, we chose to estimate with the standard errors clustered at the session level instead of including the session-fixed effects. Also here, the results remain virtually the same. Controlling for social appropriateness and fairness measures does not change the estimated treatment effect. See Tables 2 and 3 below.

Table 2. **Managers' adoption of unfriendly leadership**

	(1) OLS	(2) OLS	(3) Probit	(4) Probit	(5) OLS	(6) OLS
ELS	0.164**	0.170*	0.162**	0.162**	0.171**	0.171*
	(0.076)	(0.086)	(0.074)	(0.078)	(0.085)	(0.087)
Male	0.212**	0.240**	0.211**	0.254***	0.225**	0.200**
	(0.089)	(0.094)	(0.086)	(0.083)	(0.095)	(0.098)
Social appropriateness, 1 message					0.094*	
					(0.047)	
Fairness, 1 message					0.010	
					(0.039)	
Social appropriateness, 10 messages						0.008
						(0.025)
Fairness, 10 messages						0.032
						(0.022)
Constant	0.371**	0.350*			0.447**	0.428**
	(0.139)	(0.200)			(0.206)	(0.206)
Observations	128	128	128	128	128	128
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Session fixed effects	No	Yes	No	Yes	Yes	Yes
Clustered SE	Yes	No	Yes	No	No	No
Adjusted R^2	0.039	0.071			0.088	0.073
Pseudo R^2			0.063	0.186		

The dependent variable equals one if a manager incorporates at least one unfriendly message into her contract and zero otherwise. *ELS* is a treatment dummy that equals one in the ELS treatment and zero in the ELD treatment. *Male* equals one if the subject's gender is male and zero otherwise. All specifications include demographics such as gender, age, education, years of work experience, and experience in management positions as covariates. For models (1) and (3), the standard errors are clustered at the session level. Models (2) and (4) control for session fixed effects. Models (3) and (4) present marginal effects resulting from Probit regressions. Models (5) and (6) control for measures of social norms, regarding the changes in perception for one message in model (5) and 10 messages in model (6) compared to zero message. Specifically, the variables *Social appropriateness* and *Fairness* are computed by subtracting and dividing by the ratings for zero message, regarding the subjects' guess about 10 other managers' ratings for one and 10 messages, respectively. See the table note in Table 1 for the details of norms elicitation. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

***Table 3.* Managers' choice of the number of unfriendly messages**

	(1) OLS	(2) OLS	(3) Tobit	(4) Tobit	(5) OLS	(6) OLS
ELS	1.025***	1.042**	1.943***	1.904**	1.042**	1.080***
	(0.339)	(0.398)	(0.726)	(0.762)	(0.401)	(0.404)
Male	0.778**	0.756*	1.898**	2.128**	0.740	0.693
	(0.329)	(0.435)	(0.776)	(0.874)	(0.448)	(0.455)
Social appropriateness, 1 message					0.157	
					(0.223)	
Fairness, 1 message					0.039	
					(0.182)	
Social appropriateness, 10 messages						0.091
						(0.118)
Fairness, 10 messages						0.039
						(0.102)
Constant	0.850*	0.937	-1.636	-1.383	1.120	1.105
	(0.479)	(0.926)	(1.247)	(1.718)	(0.970)	(0.963)
Observations	128	128	128	128	128	128
Demographic controls	Yes	Yes	Yes	Yes	Yes	Yes
Session fixed effects	No	Yes	No	Yes	Yes	Yes
Clustered SE	Yes	No	Yes	No	No	No
Adjusted R^2	0.037	0.099			0.087	0.089
Pseudo R^2			0.026	0.083		

The dependent variable is the number of unfriendly messages included in the contract. In models (3) and (4), it is with a lower limit of zero and an upper limit of 10. *ELS* is a treatment dummy that equals one in the ELS treatment and zero in the ELD treatment. *Male* equals one if the subject's gender is male and zero otherwise. All specifications include demographics such as gender, age, education, years of work experience, and experience in management positions as covariates. For models (1) and (3), the standard errors are clustered at the session level. Models (2) and (4) control for session fixed effects. Models (3) and (4) present marginal effects resulting from Probit regressions. Models (5) and (6) control for measures of social norms, regarding the changes in perception for one message in model (5) and 10 messages in model (6) compared to zero message. Specifically, the variables *Social appropriateness* and *Fairness* are computed by subtracting and dividing by the ratings for zero message, regarding the subjects' guess about 10 other managers' ratings for one and 10 messages, respectively. See the table note in Table 1 for the details of norms elicitation. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Regarding the effect of leadership styles on the workers' behavior, we proposed the following models in the original analysis plan.

$$accept = \alpha_0 + \alpha_1 \cdot m + individual\ FE + \varepsilon,$$
$$effort = \gamma_0 + \gamma_1 \cdot m + individual\ FE + \varepsilon.$$

The results in Table 4 below, for acceptance decisions, and Table 5, for effort level, report the OLS, Probit and Tobit estimation of the above models. OLS results are very similar to the results obtained with the Probit and Tobit estimations. The difference in the coefficients of the effect size is mainly due to the censored data. Especially for the effort level, the proportion of data points censored at the lower limit of zero amounts to 30%, driving the differences between OLS estimation and Tobit estimation, making Tobit specifications ex-post more suitable for consistent estimation of the coefficients (Wooldridge 2016, Ch.17).

Table 4. **Workers' contract acceptance**

	(1) OLS	(2) OLS	(3) Probit	(4) Probit
ELS	0.094	0.142***	0.101*	0.156***
	(0.060)	(0.027)	(0.058)	(0.028)
Number of unfriendly messages (m)	-0.049***	-0.049***	-0.047***	-0.048***
	(0.005)	(0.003)	(0.005)	(0.003)
Male	0.022	0.018	0.021	0.016
	(0.049)	(0.022)	(0.047)	(0.021)
Constant	0.939***	0.883***		
	(0.137)	(0.078)		
Observations	1408	1408	1408	1408
Demographic controls	Yes	Yes	Yes	Yes
Individual fixed effects	No	Yes	No	Yes
Clustered SE	Yes	No	Yes	No
Adjusted R^2	0.141	0.159		
Pseudo R^2			0.142	0.168

The dependent variable is contract acceptance (equal to one if a contract with m unfriendly messages is accepted and zero otherwise). *ELS* is a treatment dummy that equals one in the ELS treatment and zero in the ELD treatment. *Male* equals one if the subject's gender is male and zero otherwise. All specifications include demographics such as gender, age, education, years of work experience, and experience in management positions as covariates. For models (1) and (3), the standard errors are clustered at the individual worker level. Models (2) and (4) control for individual fixed effects. Models (3) and (4) present the marginal effects resulting from Probit regressions. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5. **Workers' effort level**

	(1) OLS	(2) OLS	(3) Tobit	(4) Tobit
ELS	0.647	0.556***	2.026**	1.821***
	(0.498)	(0.178)	(0.884)	(0.300)
Number of unfriendly messages (m)	-0.191***	-0.191***	-0.295***	-0.286***
	(0.030)	(0.022)	(0.057)	(0.037)
Male	0.232	0.285**	0.054	0.118
	(0.420)	(0.145)	(0.681)	(0.236)
Constant	2.091	1.108**	-0.338	-2.121**
	(1.354)	(0.522)	(2.246)	(0.849)
Observations	1408	1408	1073	1073
Demographic controls	Yes	Yes	Yes	Yes
Individual fixed effects	No	Yes	No	Yes
Clustered SE	Yes	No	Yes	No
Adjusted R^2	0.087	0.167		
Pseudo R^2			0.030	0.060

The dependent variable is the effort choice. In models (3) and (4) of Tobit estimations, it is with a lower limit of zero and an upper limit of 10. _ELS_ is a treatment dummy that equals one in the ELS treatment and zero in the ELD treatment. _Male_ equals one if the subject's gender is male and zero otherwise. All specifications include demographics such as gender, age, education, years of work experience, and experience in management positions as covariates. For models (1) and (3), the standard errors are clustered at the individual worker level. Models (2) and (4) control for individual fixed effects. Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Regarding subgroup analysis or the heterogeneity of treatment effect, we report the gender difference in the use of leadership styles and reactions to them in Section 7.2 of the research paper.