



SFB/Transregio 266

ACCOUNTING FOR
TRANSPARENCY

botex: Using LLMs as Experimental Participants in oTree

Joachim Gassen and Victor Maas

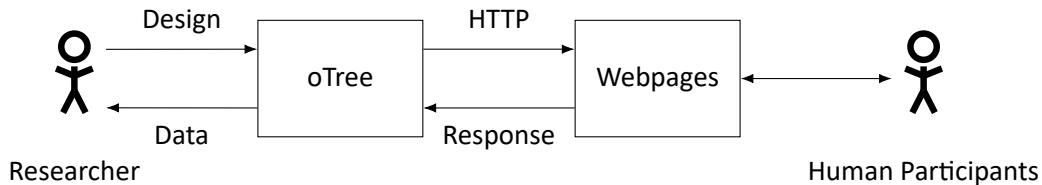
Humboldt-Universität zu Berlin and University of Amsterdam

May 13, 2024

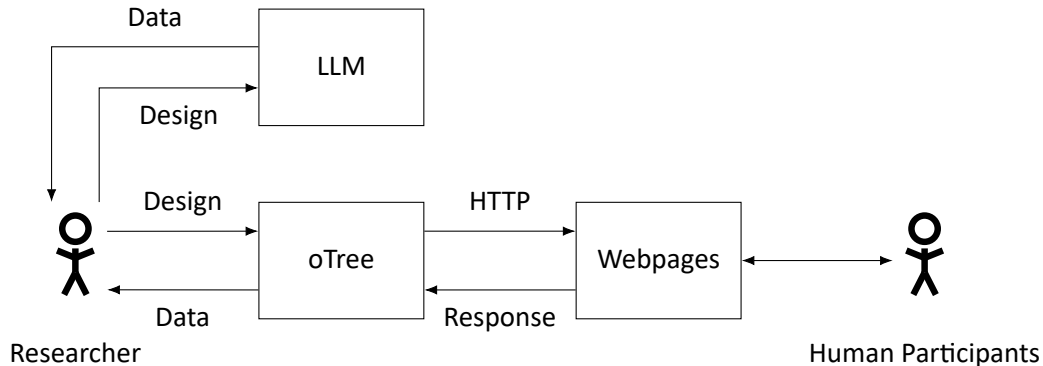
Motivation

- 1 Designing experiments is hard and small differences in experimental materials can have significant impact on experimental findings. Wouldn't it be great to have a "cheap" way to pretest experimental designs without bothering humans?
- 2 Understanding how LLMs act in behavioral experiments - on their own and interacting with humans - is an emergent and rapidly evolving research field. Providing an infrastructure to run such experiments should be beneficial for the profession.
- 3 Applied experimental work often uses context framing in their experimental designs without explicitly hypothesizing and assessing its effect on findings. LLM based experiments might inform priors in that regard.

Traditional oTree Setup

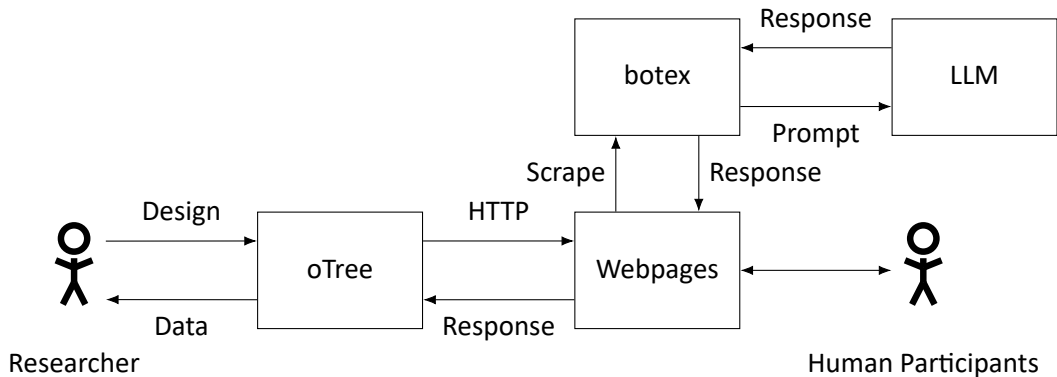


oTree + Alter Ego



Engel, Grossmann and Ockenfels (2024, SSRN)

oTree + botex



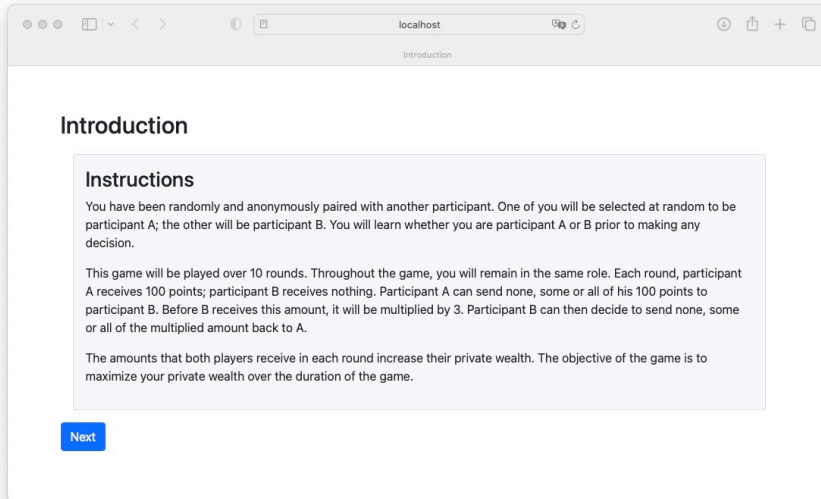
Research Question

How does business context framing affect the response behavior of LLMs?

- Context frames are common in the applied experimental literature
- Prior work has shown that their effect can be significant (e.g., Liberman, Samuels and Ross Pers Soc Pschol Bull 2014, and for investment trust games Al-Ubaydli, Houser, Nye, Paganelli, and Pan, PLOS One 2013; Cronk and Wasielewski, J of Evolutionary Psychology 2008)
- Then again, they are rarely hypothesized or even explored in the accounting literature
- LLMs reflect the priors of the average population in terms of, e.g., first order beliefs and social desirability biases. Thus, they might inform us how a general participant pool might react to contextual frames

- Investment trust game (Berg, Dickhaut and McCabe, Games and Econ Behav 1995)
- Deception (Gneezy, AER 200)
- Honesty in budgeting (Evans et al., TAR 2001)
- Gift exchange (Fehr, Kirchsteiger, and Riedl, QJE 1993)

Trust: The Neutral Frame



Trust: The Investor/Manager Frame

Introduction

Instructions

This is an experimental study to learn how investors and managers interact. You have been randomly selected to act either as an investor or as a manager and anonymously paired with another participant who will be playing the other role. You will learn whether you are an investor or a manager prior to making any decision. Please try to impersonate your role very closely when taking any decision and think about how you would act if you are a real investor or a manager of a publicly-listed firm.

This experiment will be run over 10 rounds. Throughout it, you will remain in your role. Each round, first the investor receives a private wealth of 100 points that he can decide to invest into the firm of the manager or to keep. The firm is very profitable meaning that the invested amount will be multiplied by 3 to create the firm returns. Next, the manager has to decide how much of the firm returns to return to the investor as a dividend. The remaining amount is paid out to the manager as compensation.

The dividends and compensation that both players receive in each round increase their private wealth. In the experiment, please try to maximize your private wealth.

Next

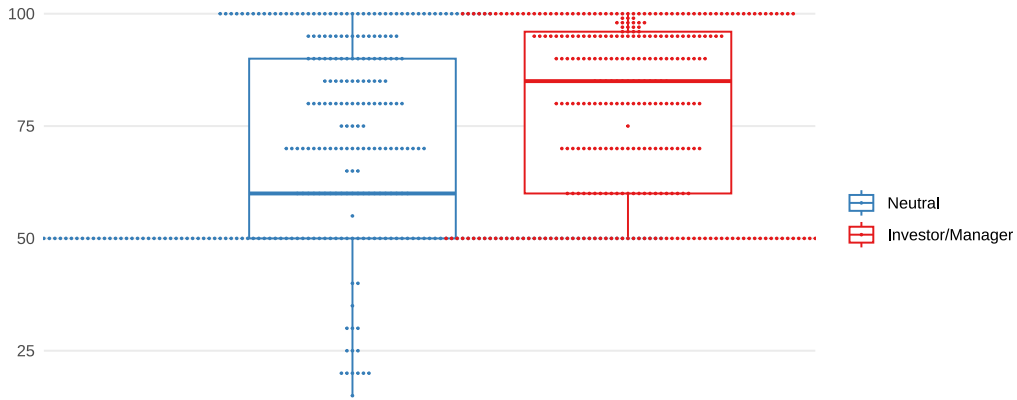
Trust: Sample Composition

	Sender	Receiver
Neutral Framing	30	30
Investor/Management Framing	30	30
Total	60	60

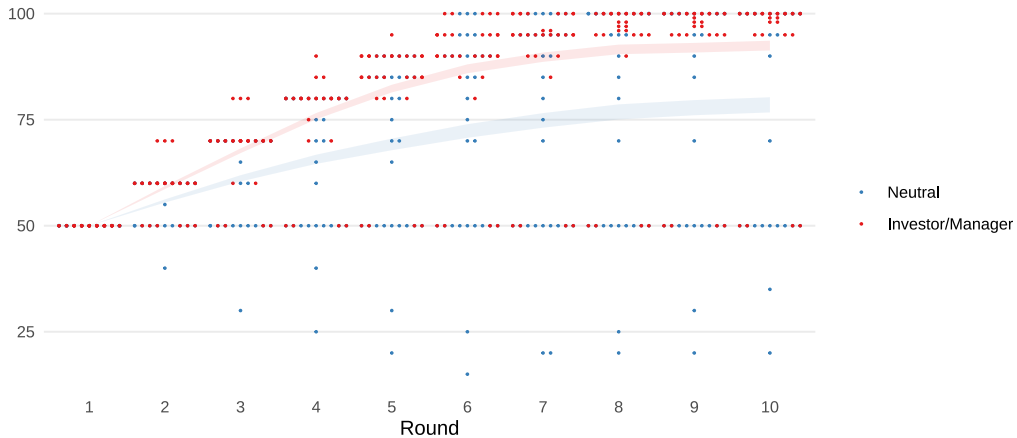
Trust: Descriptive Statistics by Framing

	Neutral Framing				Investor/Management Framing			
	N	Mean	S.D.	Median	N	Mean	S.D.	Median
Amount sent	300	68.22	22.10	60.00	300	78.76	19.46	85.00
Amount returned	300	93.96	39.24	90.00	300	115.95	31.23	120.00
% returned	300	0.45	0.08	0.50	300	0.49	0.04	0.50

Trust: Sent Amount by Framing



Trust: Sent Amount by Round



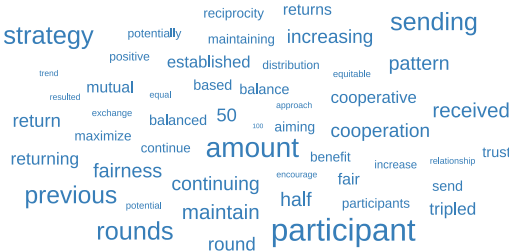
Trust: Regression Results

	(1)
(Intercept)	50.900*** (2.165)
Round	3.148*** (0.349)
Investor/Management	1.896 (3.061)
Round × Investor/Management	1.572*** (0.493)
Num.Obs.	600
R2 Adj.	0.346

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Trust: Some Verbal Response

Neutral Framing



Investor/Management Framing



Trust: Manipulation Checks ;-)

		Neutral	Inv/Manag
Answers comprehension question correctly	Yes	60	60
	No	0	0
Remembers role	Yes	59	59
	No	1	1
Characterizes as ...	Human	48	39
	Bot	12	21

Deception: Descriptives by Framing

	Neutral Framing		Accountant/HQ Framing	
	N	Mean	N	Mean
% lied	30	0.77	30	0.80
% followed	30	1.00	30	1.00

Chi-Square: 0.00

Honesty: Descriptives

	Neutral Framing				Business Framing			
	N	Mean	S.D.	Median	N	Mean	S.D.	Median
Absolute slack	300	500.00	628.10	100.00	300	641.50	655.81	350.00
% slack claimed	297	0.52	0.50	1.00	292	0.63	0.44	1.00

Absolute Slack: 2.70***

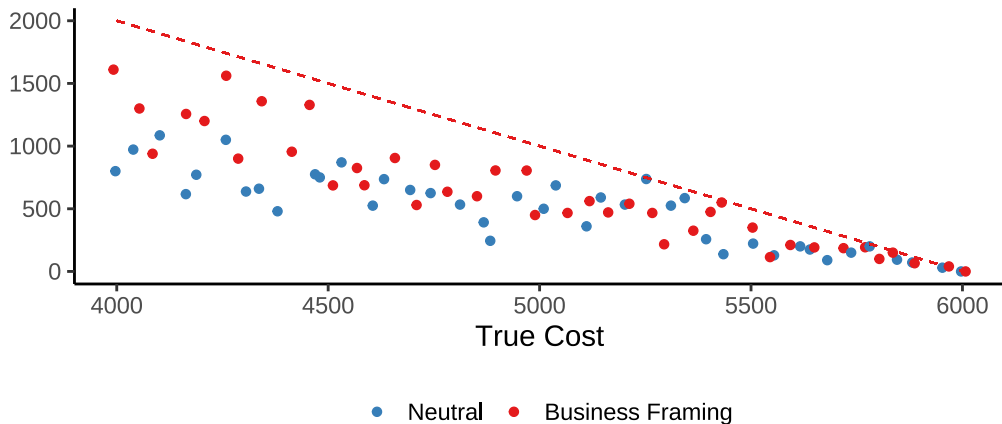
Relative Slack: 3.00***

Honesty: Strategies

	Neutral Framing	Business Framing
Strategy	N	N
All slack	9	11
No slack	13	4
Some slack	8	15

Chi-Square: 7.10**

Honesty: Average Claimed Slack by True Cost



Honesty: Absolute Regression Results

	Slack
Intercept	2574.321*** (279.770)
True Cost	-0.432*** (0.056)
Business Framing	1263.912*** (391.465)
Round	14.020 (11.047)
True Cost \times Business Framing	-0.241*** (0.077)
Round \times Business Framing	15.939 (15.567)
Adjusted R ²	0.289
Number of observations	600

Honesty: Relative Regression Results

	% Claimed Slack
Intercept	0.584*** (0.082)
% Availabe Slack	-0.183* (0.098)
Business Framing	-0.085 (0.111)
Round	0.005 (0.010)
% Availabe Slack × Business Framing	0.204 (0.135)
Round × Business Framing	0.017 (0.013)
Adjusted R ²	0.022
Number of observations	589

Honesty: Verbal Response

Neutral Framing



Business Framing



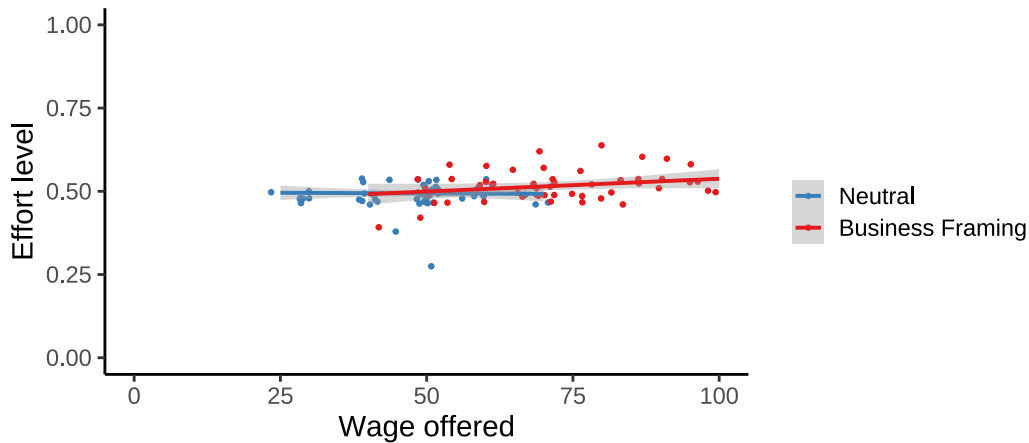
Gift: Descriptives

	Neutral Framing				Business Framing			
	N	Mean	S.D.	Median	N	Mean	S.D.	Median
Wage offered	50	47.10	10.31	50.00	50	71.80	14.63	70.00
Effort Level	50	0.49	0.03	0.50	50	0.52	0.05	0.50

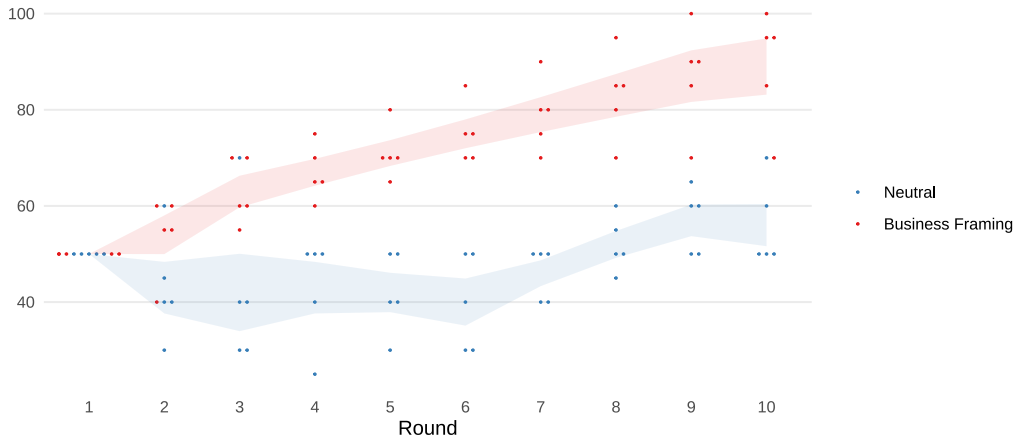
Difference in wage offered (t-stat): 9.76***

Difference in effort level (t-stat): 2.76***

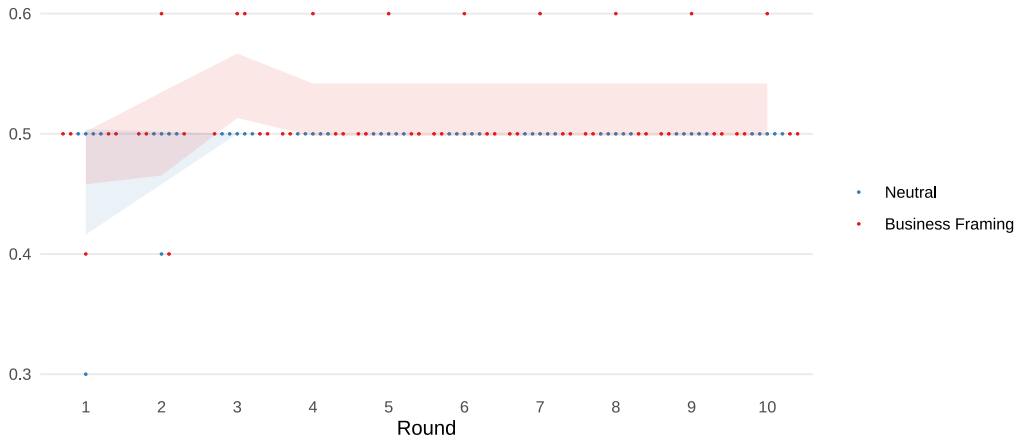
Gift: Association between Wage and Effort



Gift: Wage by Round



Gift: Effort Level by Round



Gift: Verbal Response

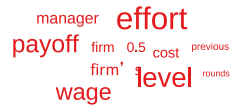
Neutral Framing



A word cloud for the 'Neutral Framing' condition. The words are in shades of blue. 'multiplier' is the largest word at the top right. 'participant' is at the bottom left. 'payoff' is in the middle left. 'cost' is in the middle right. '0.5' is a small word in the center. 'balance' is a very small word to the left of '0.5'.

multiplier
payoff
participant
cost
0.5
balance

Business Framing



A word cloud for the 'Business Framing' condition. The words are in shades of red. 'effort' is the largest word at the top right. 'level' is below it. 'payoff' is on the left. 'wage' is at the bottom left. 'manager' is at the top left. 'firm' and 'firm'' are in the center. '0.5' is in the center. 'cost' is to the right of '0.5'. 'previous' is at the top right. 'rounds' is at the bottom right.

effort
level
payoff
wage
manager
firm
firm'
0.5
cost
previous
rounds

Next Steps

- Decide on experiments to include in first draft
- Finalize experimental materials for those
- Run experiments with reasonable power
- Write first draft
- Later: Compare (selected) findings with results for human participants