botex: Using LLMs as Experimental Participants in oTree

Joachim Gassen and Victor Maas
Humboldt-Universität zu Berlin and University of Amsterdam

May 24, 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.
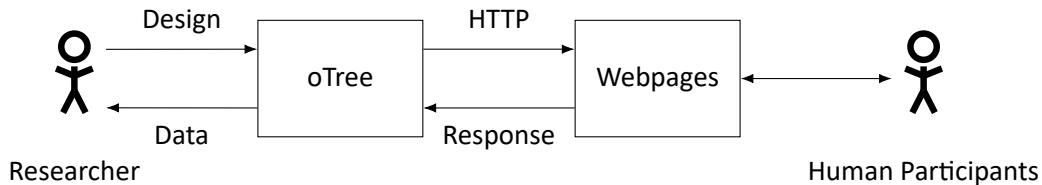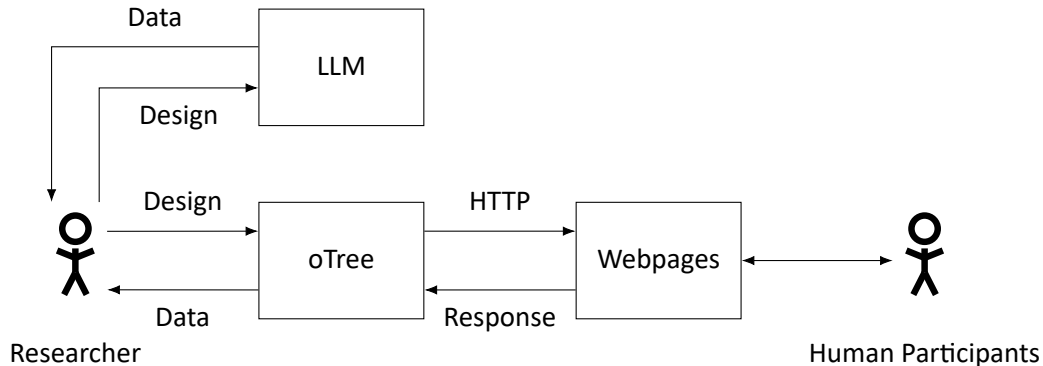
Engel, Grossmann and Ockenfels (2024, SSRN)
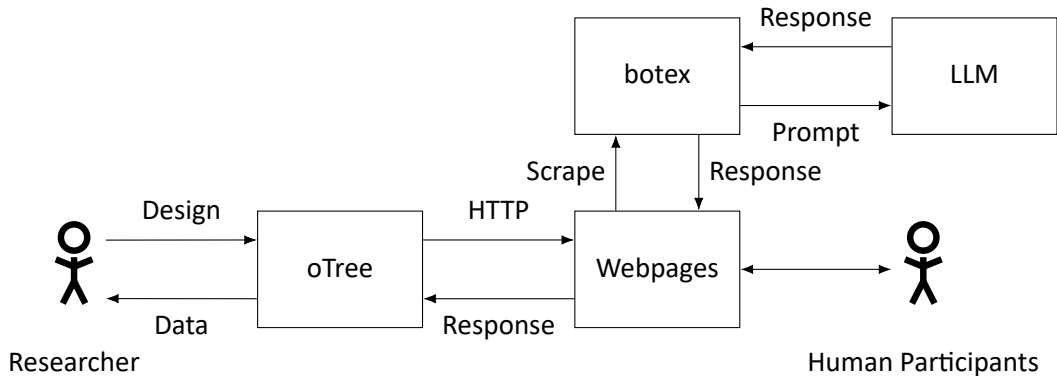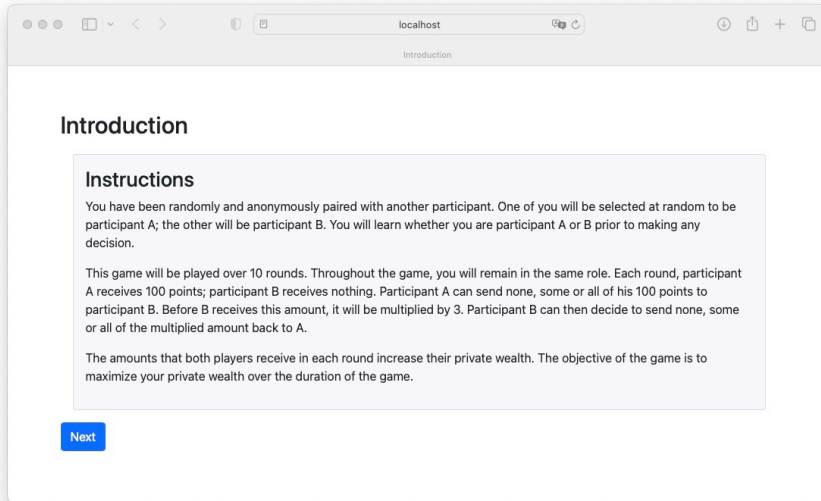
How does business context framing affect the response behavior of LLMs?

Motivation

- 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 Pscaol 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 believes and social desirability biases. Thus, they might inform us how a general participant pool might react to contextual frames

The Games

- 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

Introduction

Instructions

You have been randomly and anonymously paired with another participant. One of you will be selected at random to be participant A; the other will be participant B. You will learn whether you are participant A or B prior to making any decision.

This game will be played over 10 rounds. Throughout the game, you will remain in the same role. Each round, participant A receives 100 points; participant B receives nothing. Participant A can send none, some or all of his 100 points to participant B. Before B receives this amount, it will be multiplied by 3. Participant B can then decide to send none, some or all of the multiplied amount back to A.

The amounts that both players receive in each round increase their private wealth. The objective of the game is to maximize your private wealth over the duration of the game.

Next

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 an 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.
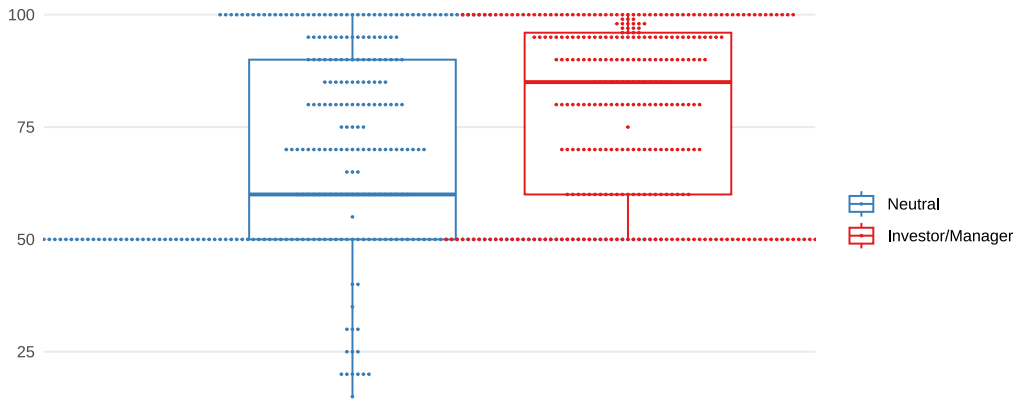
Next

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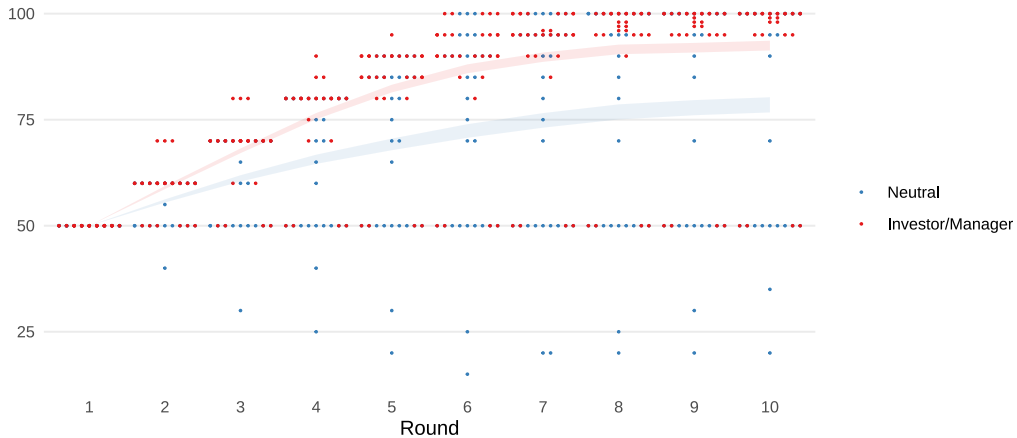
	(1)
(Intercept)	50.900***
	(2.165)
Round	3.148***
	(0.349)
Investor/Management	1.896
	(3.061)
Round × Investor/Management	1.572***
	(0.493)
Num.Obs.	600
R2 Adj.	0.346

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

Neutral Framing

Investor/Management Framing

		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

	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

	Neutral Framing				Business Framing			
	N	Mean	S.D.	Median	N	Mean	S.D.	Median
Absolute slack	100	749.50	565.57	550.00	100	297.50	361.84	200.00
% slack claimed	98	0.77	0.31	1.00	98	0.37	0.34	0.24

Absolute Slack: -6.73***

Relative Slack: -8.51***

	Neutral Framing	Business Framing
Strategy	N	N
All slack	4	1
No slack	0	1
Some slack	6	8

Chi-Square: 3.09
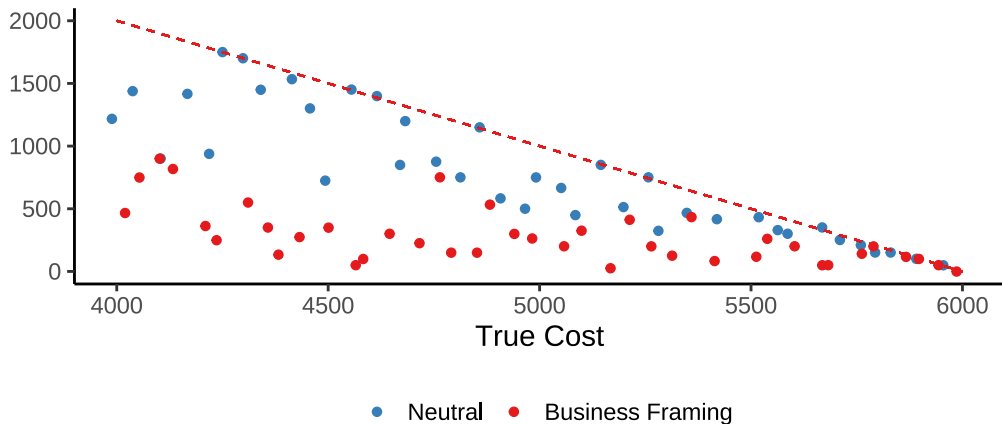
Honesty: Absolute Regression Results

	Slack
Intercept	4327.724***
	(295.058)
True Cost	-0.745***
	(0.058)
Business Framing	-2855.743***
	(417.275)
Round	26.798**
	(12.072)
True Cost × Business Framing	0.509***
	(0.083)
Round × Business Framing	-25.661
	(17.072)
Adjusted R²	0.568
Number of observations	200

	% Claimed Slack
Intercept	0.659***
	(0.097)
% Availabe Slack	-0.068
	(0.111)
Business Framing	-0.091
	(0.137)
Round	0.026**
	(0.011)
% Availabe Slack × Business Framing	-0.345**
	(0.157)
Round × Business Framing	-0.024
	(0.016)
Adjusted R^2	0.325
Number of observations	196

Neutral Framing

Business Framing

maximize
actual reporting
payoff

payoff predicted
actual costs
budget request

	Neutral Framing				Business Framing			
	N	Mean	S.D.	Median	N	Mean	S.D.	Median
Wage offered	150	59.60	20.72	52.50	150	68.54	13.80	70.00
Effort Level	150	0.64	0.10	0.60	150	0.62	0.09	0.60

Difference in wage offered (t-stat): 4.40***
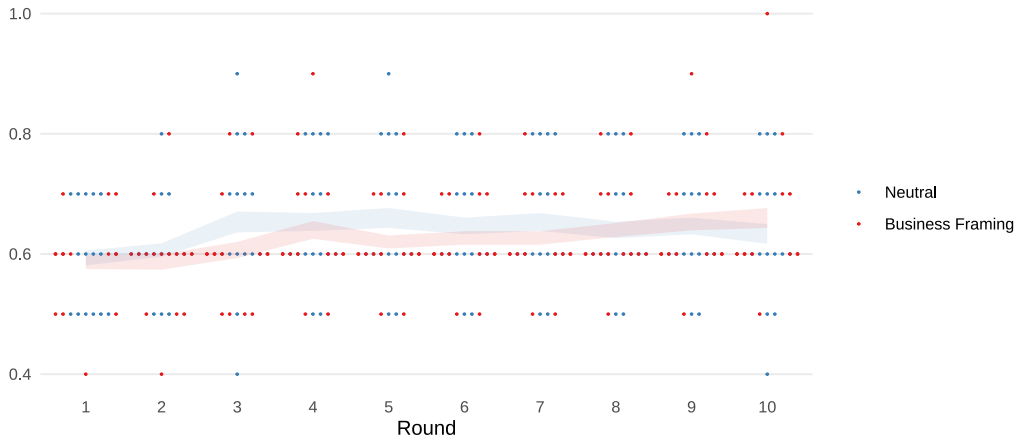
Difference in effort level (t-stat): -1.23

giftex: Wage by Round
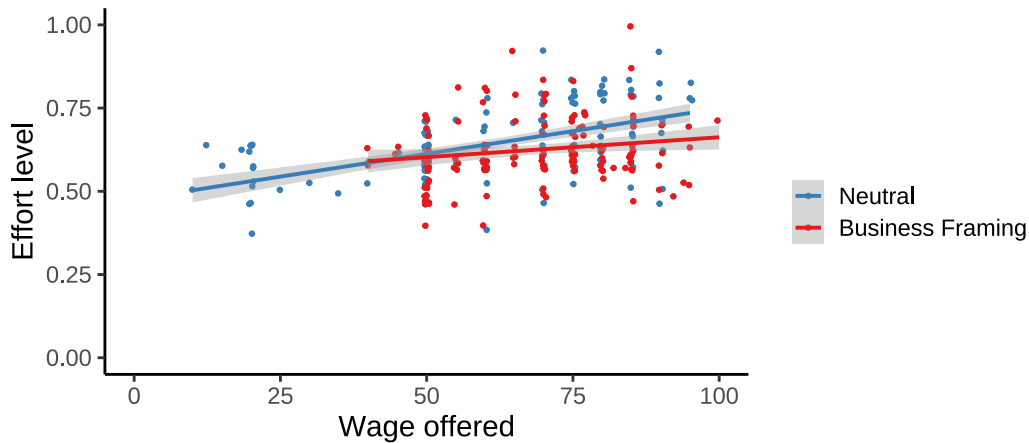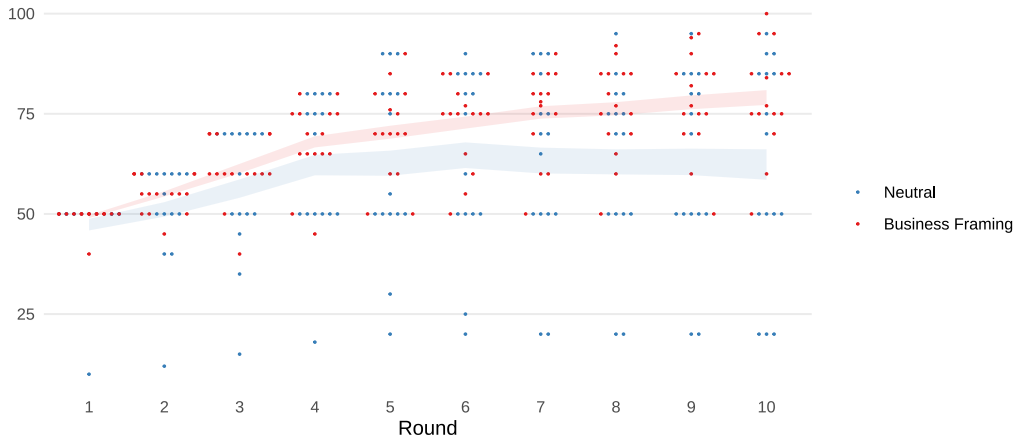
Neutral Framing
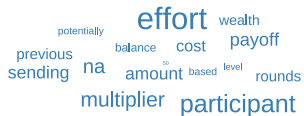
Business Framing

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