Money Talks: When and Why Real Incentives Matter in Experiments

Jack Fitzgerald

Vrije Universiteit Amsterdam

Introduction

RQ: **When do real incentives matter for accurate experimental results?** I make three contributions:

1. I build a theoretical and statistical framework that reveals when real incentives matter in experiments.
2. I identify an oft-observed behavioral effect where real incentives likely matter: the foreign language effect (FLE).
3. I design an experiment to assess whether observed FLEs are biased by hypothetical incentives.

Terminology

- **Elicitation experiments**: The researcher is interested in measuring outcome Y_i (e.g., choice experiments).
- **Intervention experiments**: The researcher is interested in the treatment effect (TE) of some intervention D_i on an outcome Y_i (e.g., randomized controlled trials).

Let H_i be an indicator for hypothetical incentives and consider the data-generating process

$$Y_i = \alpha + \delta H_i + \epsilon_i.$$

δ represents **classical hypothetical bias (CHB)**.

Framework

- **Elicitation experiments**: CHB is fully informative regarding the necessity of real incentives. Thus real incentives likely matter for most elicitation experiments since CHB is widely documented (e.g., Camerer and Hogarth 1999).
- **Intervention experiments**: CHB is uninformative. Real incentives matter if and only if hypothetical incentives cause heterogeneous treatment effects.

To see why, consider the relevant data-generating process:

$$Y_i = \alpha + \beta_1 D_i + \beta_2 H_i + \beta_3 (D_i \times H_i) + \mu_i.$$

Suppose $\mathbb{E}\left[\mu_i | D_i, H_i\right] = 0$ due to randomization. Then

$$\underbrace{\frac{\partial Y_i}{\partial D_i} = \beta_1 + \beta_3 H_i}_{\text{intervention TE}} \quad \underbrace{\frac{\partial Y_i}{\partial H_i} = \beta_2 + \beta_3 D_i}_{\text{CHB}} \quad \underbrace{\frac{\partial^2 Y_i}{\partial D_i \partial H_i} = \beta_3}_{\text{hypothetical bias on TE}}.$$

Thus in intervention experiments, β_3 determines whether real incentives matter, not $\delta = \beta_2 + \beta_3 D_i$.

Application: Holt and Laury (2005)

Do people become more risk-averse when making high-stakes choices? The answer you get depends on whether hypothetical or real incentives are used.

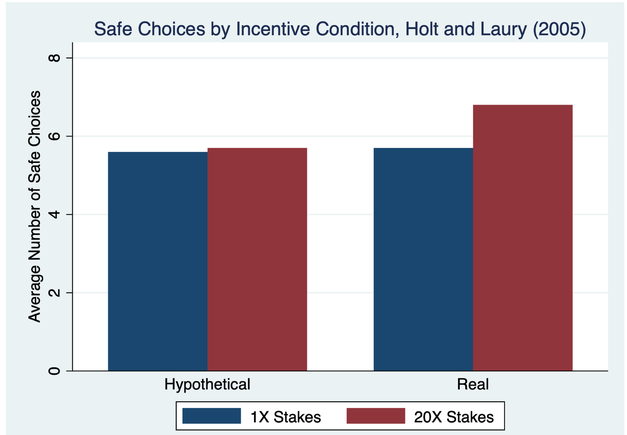

Figure 1. An Intervention Experiment Where Real Incentives Matter

While there is no significant CHB, hypothetical incentives cause considerable treatment effect heterogeneity (even changing conclusions), so real incentives matter in this setting.

The Foreign Language Effect

Dozens of psychological studies (almost all of which are unincentivized) find one of three types of FLEs (Circi et al. 2021):

1. **Risk FLE**: FL usage reduces risk aversion.
2. **Rationality FLE**: FL usage increases rationality.
3. **Moral FLE**: FL usage decreases deontological reasoning.

The hypothetical mechanism is *dual processing*: FL usage increases the likelihood of System 2 usage. Thus if real incentives motivate participants to use System 2 *ceteris paribus*, real incentives will attenuate observed FLEs by reducing System 1 $\rightarrow$ System 2 switching. I.e., real incentives likely matter here.

Meta-analytic evidence is consistent with this notion. Risk FLE experiments with real incentives have 53.3% smaller effect sizes, 88.6% larger TE variances, and 78.1% smaller test statistics than those with hypothetical incentives (see Circi et al. 2021).

The Experimental Proposal

I propose a 2x2 experiment with two between-subject treatment arms: FL (vs. NL control) and hypothetical incentives (vs. real incentives control). I use three business-framed games:

1. **Investment game**: Participants must choose whether to invest in a safe or risky franchise across ten different success probabilities. This is effectively a multiple price list (Holt and Laury 2005). *Outcomes*: Safe choices (risk aversion) and preference consistency (rationality).
2. **Embezzlement game**: Participants must report their franchise's earnings to the franchisee across ten different earnings levels, where 10% of reported earnings are taken as a franchise fee. *Outcomes*: Underreporting (i.e., embezzlement), both on the extensive and intensive margins.
3. **Modified embezzlement game**: Identical to the embezzlement game, but each earnings report comes with a 5% chance of a random audit. If embezzlement is detected, the participant earns nothing for that earnings report. *Outcomes*: Underreporting (i.e., embezzlement), both on the extensive and intensive margins.

Hypotheses

Earnings levels and audit risk are **incentive-relevant interventions** while the FL condition is an **incentive-irrelevant intervention**. For each outcome, I hypothesize:

1. Hypothetical incentives will *attenuate* TE estimates for incentive-*relevant* interventions.
2. Hypothetical incentives will *amplify* TE estimates for incentive-*irrelevant* interventions.
3. Study conclusions regarding the TE of each intervention will differ between real- and hypothetical-incentive conditions.

References

[1] Colin F Camerer and Robin M Hogarth. The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19:7–42, 1999.

[2] Riccardo Circi, Daniele Gatti, Vincenzo Russo, and Tomaso Vecchi. The foreign language effect on decision-making: A meta-analysis. *Psychonomic Bulletin & Review*, 28(4):1131–1141, 2021.

[3] Charles A Holt and Susan K Laury. Risk aversion and incentive effects: New data without order effects. *American Economic Review*, 95(3):902–904, 2005.