# Power Analysis

## 1 Design

### 1.1 Treatments

The experiment has three treatments:

- **T0**: Baseline with high inequality (no redistribution, no mobility).
- **T1**: *Redistribution* (incomes are made more equal within each period).
- **T2**: *Mobility* or rank reversal (within a period, inequality is high as in T0, but people can swap positions between periods, reducing inequality "on average over time").

### 1.2 Roles and outcomes

There are two main roles:

- **Senders**, who choose *In* or *Out*.
- **Receivers**, who choose *Share* or *Take* if the Sender chooses *In*.

We focus on two key outcomes, defined for individuals who are *Poor* in the current treatment:

1. **Trust** (Senders):
    - For each Sender and each round, we look at whether they choose *In* when matched with a *Poor* Receiver.
    - Let
    $$p_T^{P,In} = \Pr(\text{Sender chooses In} \mid \text{partner is Poor, treatment } T),$$
    be the probability of trusting a Poor partner in treatment $T$.
2. **Trustworthiness** (Poor Receivers):

- For each Poor Receiver and each round, we look at whether they choose *Share*.
- Let

$$p_T^{P,Share} = \Pr(\text{Poor Receiver chooses Share} \mid \text{treatment } T),$$

  be the probability that a Poor Receiver behaves trustworthily in treatment $T$.

Power calculations in this document are framed in terms of these probabilities.

# 2 Assumed treatment effects and hypotheses

## 2.1 Effect size assumptions

For both trust and trustworthiness among the Poor, we assume these treatment effects:

$$\begin{aligned}
\Delta_{T1,T0}^{P} &= p_{T1}^{P} - p_{T0}^{P} = 0.30, \\
\Delta_{T2,T0}^{P} &= p_{T2}^{P} - p_{T0}^{P} = 0.15, \\
\Delta_{T1,T2}^{P} &= p_{T1}^{P} - p_{T2}^{P} = 0.15,
\end{aligned}$$

where $p_T^P$ denotes either $p_T^{P,In}$ (trust) or $p_T^{P,Share}$ (trustworthiness).
Our assumed treatment effects are in line with magnitudes found in related experimental work. Trust and reciprocity in standard trust games respond strongly to manipulations of social identity, status, and partner type, with differences of 10–20 percentage points or more across conditions (Tsutsui and Zizzo, 2014; Cettolin and Suetens, 2019). Information and history treatments likewise generate shifts of similar size in trusting and returning behavior (Hofmeyr et al., 2023).

## 2.2 Hypotheses

The Pre-Analysis Plan (PAP) defines six hypotheses for Poor partners/receivers.

**Trust (Senders matched with Poor).**

- H1: $p_{T1}^{P,In} > p_{T0}^{P,In}$, effect $\approx 0.30$ (T1 vs T0).
- H2: $p_{T2}^{P,In} > p_{T0}^{P,In}$, effect $\approx 0.15$ (T2 vs T0).
- H3: $p_{T1}^{P,In} > p_{T2}^{P,In}$, effect $\approx 0.15$ (T1 vs T2).

**Trustworthiness (Poor Receivers).**

- H4: $p_{T1}^{P,Share} > p_{T0}^{P,Share}$, effect $\approx 0.30$ (T1 vs T0).
- H5: $p_{T2}^{P,Share} > p_{T0}^{P,Share}$, effect $\approx 0.15$ (T2 vs T0).
- H6: $p_{T1}^{P,Share} > p_{T2}^{P,Share}$, effect $\approx 0.15$ (T1 vs T2).

All tests are one-sided at significance level $\alpha = 0.05$.

# 3 From effects to required *effective* sample size

We first imagine a simplified world where each subject makes only one binary decision. We compare two treatments, A and B, with equal sample size $n$ per arm. Let $p_A$ and $p_B$ be the outcome probabilities, and $\delta = p_A - p_B$ the effect size. For a one-sided $z$-test at level $\alpha$ with power $1 - \beta$, an approximation for the required (effective) sample size per arm is:

$$n_{\text{eff}} \approx \frac{0.5\,(z_{1-\alpha} + z_{1-\beta})^2}{\delta^2},$$

where $z_{1-\alpha}$ and $z_{1-\beta}$ are standard normal quantiles.

In our case:

- $\alpha = 0.05$ (one-sided) $\Rightarrow z_{1-\alpha} \approx 1.645$;
- we target about $1 - \beta = 0.80$ power for the primary contrasts $\Rightarrow z_{1-\beta} \approx 0.84$.

Then:

$$(z_{1-\alpha} + z_{1-\beta})^2 \approx (1.645 + 0.84)^2 \approx 6.18,$$

so that

$$n_{\text{eff}} \approx \frac{3.09}{\delta^2}. \tag{1}$$

Applying this to the two effect sizes of interest:

- For $\delta = 0.30$ (T1 vs T0): $n_{\text{eff}} \approx 3.09/0.09 \approx 34$.
- For $\delta = 0.15$ (T1 vs T2 and T2 vs T0): $n_{\text{eff}} \approx 3.09/0.0225 \approx 137$.

The **most demanding** case is the 15 percentage point difference (T1 vs T2 and T2 vs T0). We therefore take as a target:

$$n_{\text{eff,target}} \approx 137 \text{ effective observations per treatment arm for } \delta = 0.15.$$

# 4 Panel structure and correlation between rounds

In the experiment, each subject makes multiple decisions in Block 1. This increases the amount of information, but decisions from the same person are correlated.

## 4.1 Decisions per subject

In Block 1 (rounds 1–10):

- Each Poor Receiver makes $m_R = 10$ Share/Take decisions.
- Each Sender plays 10 rounds and is matched with a Poor or Rich partner with equal probability. On average, each Sender therefore has about $m_S = 5$ "trust toward Poor" decisions (In/Out when the partner is Poor).

## 4.2 Intra-person correlation

Decisions by the same subject across rounds are likely to be similar (for instance, a "trusting" Sender may often choose In). We capture this with an *intra-person correlation*:

$$\rho = 0.10.$$

A common approximation is that a subject with $m$ decisions and correlation $\rho$ contributes:

$$m_{\text{eff}} = \frac{m}{1 + (m-1)\rho}$$

*effective* independent observations.[1]

**Poor Receivers.** For Poor Receivers ($m_R = 10$),

$$m_{\text{eff,recv}} = \frac{10}{1 + 9 \times 0.10} = \frac{10}{1.9} \approx 5.26.$$

Each Poor Receiver contributes about 5.26 effective observations.

[1] The term $1 + (m-1)\rho$ is the usual *design effect*: when $\rho > 0$, repeated decisions from the same subject are correlated, so the variance of estimators (e.g. sample means, regression coefficients) is larger than it would be with the same number of independent observations, and the $m$ repeated outcomes behave like only $m_{\text{eff}} = \dfrac{m}{1 + (m-1)\rho}$ independent observations.

**Senders.** For Senders and their trust toward Poor partners ($m_S = 5$),

$$m_{\text{eff,send}} = \frac{5}{1 + 4 \times 0.10} = \frac{5}{1.4} \approx 3.57.$$

Each Sender contributes about 3.57 effective observations for trust toward Poor.

# 5 Role composition constraints

Since treatments (T0, T1, T2) are mean preserving, the experiment is designed to satisfy:

1. The number of Poor Receivers equals the number of Rich Receivers:

$$N_{\text{Poor}} = N_{\text{Rich}} = R.$$

2. The number of Senders equals the total number of Receivers:

$$N_{\text{Senders}} = N_{\text{Poor}} + N_{\text{Rich}} = 2R.$$

Therefore, per treatment:

- Poor Receivers: $R$,
- Rich Receivers: $R$,
- Senders: $2R$.

Total subjects per treatment:

$$N_{\text{per treatment}} = R + R + 2R = 4R.$$

With three treatments (T0, T1, T2), the total sample size is:

$$N_{\text{total}} = 3 \times 4R = 12R.$$

Thus, once we choose $R$ (the number of Poor and Rich Receivers per treatment), the entire sample size is determined.

# 6 Choosing $R$ using the target effective sample size

## Effective sample size as a function of $R$

With $R$ Poor Receivers and $2R$ Senders per treatment:

**Poor Receivers (trustworthiness).**

$$n_{\text{eff,recv}}(R) = R \times m_{\text{eff,recv}} \approx R \times 5.26.$$

**Senders (trust toward Poor).**

$$n_{\text{eff,send}}(R) = 2R \times m_{\text{eff,send}} \approx 2R \times 3.57 = 7.14R.$$

## Target: about 137 effective observations per arm for $\delta = 0.15$

As shown in Section 3, detecting a 15 percentage point effect ($\delta = 0.15$) with a one-sided test at $\alpha = 0.05$ and power $1 - \beta = 0.80$ requires approximately $n_{\text{eff,target}} \approx 137$ effective independent observations per treatment arm. In what follows, we translate this target into a condition on $R$ using the effective sample size expressions $n_{\text{eff,send}}(R)$ (for senders) and $n_{\text{eff,recv}}(R)$ (for receivers) derived above.

**Trust (Senders, H3 and H2).** For Senders,

$$n_{\text{eff,send}}(R) \approx 7.14R.$$

Setting $7.14R \approx 137$ gives

$$R \approx \frac{137}{7.14} \approx 19.2.$$

We choose the convenient round number $R = 20$. This yields slightly more effective observations than needed for trust hypotheses involving $\delta = 0.15$.

At $R = 20$:

$$n_{\text{eff,send}}(20) \approx 7.14 \times 20 \approx 143.$$

**Trustworthiness (Poor, H6 and H5).** For Poor Receivers,

$$n_{\text{eff,recv}}(R) \approx 5.26R.$$

At $R = 20$:

$$n_{\text{eff,recv}}(20) \approx 5.26 \times 20 \approx 105.$$

This is below the ideal 137 effective observations for 80% power at $\delta = 0.15$. We accept this as a reasonable compromise for a lab experiment.

### Resulting sample size

With $R = 20$ per treatment:

- Poor Receivers: 20,
- Rich Receivers: 20,
- Senders: 40,
- Total per treatment: $4R = 80$ participants.

Over the three treatments (T0, T1, T2):

$$N_{\text{total}} = 3 \times 80 = 240.$$

## 7 Resulting power for each hypothesis

Given $R = 20$, the effective sample sizes per arm are approximately:

- For trust (Senders, Poor partners): $n_{\text{eff,send}} \approx 143$,
- For trustworthiness (Poor Receivers): $n_{\text{eff,recv}} \approx 105$.

Using a normal approximation for one-sided tests at $\alpha = 0.05$ with baseline probability around 0.5, we obtain the approximate powers shown in Table 1.

**Table 1:** Approximate power for each hypothesis under $N = 240$.

| Hypothesis | Outcome | Comparison | Effect size $\delta$ | Power |
|---|---|---|---|---|
| H1 | Trust (Senders) | T1 vs T0 | 0.30 | $\approx 0.99$ |
| H2 | Trust (Senders) | T2 vs T0 | 0.15 | $\approx 0.81$ |
| H3 | Trust (Senders) | T1 vs T2 | 0.15 | $\approx 0.81$ |
| H4 | Trustworthiness (Poor) | T1 vs T0 | 0.30 | $\approx 0.99$ |
| H5 | Trustworthiness (Poor) | T2 vs T0 | 0.15 | $\approx 0.70$ |
| H6 | Trustworthiness (Poor) | T1 vs T2 | 0.15 | $\approx 0.70$ |

## References


Cettolin, E. and S. Suetens (2019). Return on trust is lower for immigrants. *The Economic Journal 129*(621), 1992–2009.

Hofmeyr, A., H. Kincaid, and B. Monroe (2023). The trust game: Salience, beliefs, and socia l history. Technical report, Working Paper.

Tsutsui, K. and D. J. Zizzo (2014). Group status, minorities and trust. *Experimental Economics 17*(2), 215–244.