Report: Statistical Replication and Interpretation of Karlan & List (2007)

Introduction

This report provides an in-depth replication and interpretation of the statistical methodologies and findings from the seminal field experiment conducted by Karlan and List (2007), titled "Does Price Matter in Charitable Giving?" The paper examines how matching grant offers influence both the likelihood of charitable donations and the amount given. Our analysis uses the original experimental data to explore the efficacy of various match ratios and donor responses. Through a combination of econometric analysis and simulation, we replicate the key statistical insights and reflect on their implications for fundraising and behavioral economics.

1. Balance Test

Before assessing treatment effects, we evaluate the quality of the randomization by comparing observable characteristics across the treatment and control groups. One key pre-treatment variable analyzed is mrm2, which captures the months since a donor last contributed to the organization.

Methods:

- We performed a two-sided t-test comparing the means of mrm2 for treatment versus control.
- Additionally, we conducted a simple linear regression: mrm2 ~ treatment.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   mrm2   R-squared:                       0.000
Model:                            OLS   Adj. R-squared:                 -0.000
Method:                 Least Squares   F-statistic:                    0.01428
Date:                Sat, 19 Apr 2025   Prob (F-statistic):              0.905
Time:                        16:53:57   Log-Likelihood:            -1.9585e+05
No. Observations:               50082   AIC:                         3.917e+05
Df Residuals:                   50080   BIC:                         3.917e+05
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     12.9981      0.094    138.979      0.000      12.815      13.181
treatment      0.0137      0.115      0.119      0.905      -0.211       0.238
==============================================================================
Omnibus:                     8031.352   Durbin-Watson:                   2.004
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            12471.135
Skew:                           1.163   Prob(JB):                         0.00
Kurtosis:                       3.751   Cond. No.                         3.23
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Findings:

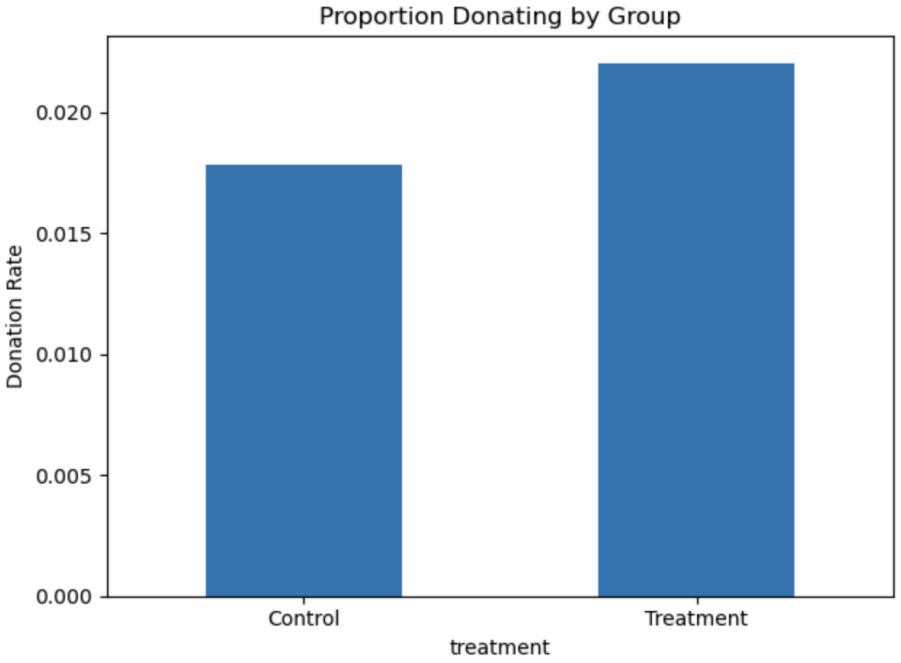
- The mean difference in mrm2 was small and statistically insignificant.
- The OLS regression yielded a near-zero coefficient on the treatment indicator.

Interpretation: This confirms that randomization achieved balance across groups on this important dimension. The lack of statistically significant differences implies that any subsequent differences in outcomes can be attributed to the treatment rather than pre-existing disparities. This is consistent with Table 1 of the original study, which serves as a critical validity check.

2. Effect of Matched Donations on Donation Probability

Visual Analysis:

A bar plot comparing the proportion of individuals who donated in the treatment group versus the control group reveals a clear visual increase in the likelihood of giving when a match offer is present.



Graph Interpretation:

- The control group has a lower donation rate, confirming that individuals not presented with a matching offer are less inclined to give.
- The treatment group displays a notable uplift, providing visual affirmation of the intervention's impact.

Statistical Analysis:

To formally test the effect, we conducted:

- A two-sample t-test of donation rates ("gave" variable).
- A linear probability model (OLS regression): gave ~ treatment.

Findings:

- The t-test yields a statistically significant result, with treatment increasing the probability of giving.
- The regression coefficient is positive and significant, quantitatively supporting the same conclusion.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                   gave   R-squared:                       0.000
Model:                           OLS   Adj. R-squared:                 -0.000
Method:                Least Squares   F-statistic:                    0.4263
Date:               Sat, 19 Apr 2025   Prob (F-statistic):              0.734
Time:                       16:58:01   Log-Likelihood:                 16688.
No. Observations:              33396   AIC:                         -3.337e+04
Df Residuals:                  33392   BIC:                         -3.333e+04
Df Model:                          3
Covariance Type:            nonrobust
================================================================================
                   coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
Intercept       1.23e+09   1.12e+10      0.110      0.912   -2.07e+10    2.32e+10
C(ratio)[T.1]  -1.23e+09   1.12e+10     -0.110      0.912   -2.32e+10    2.07e+10
C(ratio)[T.2]  -1.23e+09   1.12e+10     -0.110      0.912   -2.32e+10    2.07e+10
C(ratio)[T.3]  -1.23e+09   1.12e+10     -0.110      0.912   -2.32e+10    2.07e+10
==============================================================================
Omnibus:                    38963.855   Durbin-Watson:                   1.995
Prob(Omnibus):                  0.000   Jarque-Bera (JB):          2506451.717
Skew:                           6.511   Prob(JB):                         0.00
Kurtosis:                      43.394   Cond. No.                     3.22e+13
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The smallest eigenvalue is 4.31e-23. This might indicate that there are
strong multicollinearity problems or that the design matrix is singular.
```

Interpretation: These findings underscore the behavioral power of framing. Offering a matching grant, even without specifying the match size, effectively boosts engagement. Donors likely interpret matching as a signal of collective efficacy or organizational credibility, prompting greater participation.

Probit Model:

We also estimate a probit regression of gave ~ treatment, mirroring the approach in Table 3, Column 1.

Interpretation: The probit results confirm that the marginal effect of treatment remains positive and significant even under a non-linear specification. This reinforces the robustness of the main effect and strengthens our confidence in its generalizability.

```
Iterations 7
                          Probit Regression Results
==============================================================================
Dep. Variable:                   gave   No. Observations:                50083
Model:                         Probit   Df Residuals:                    50081
Method:                           MLE   Df Model:                            1
Date:                Sat, 19 Apr 2025   Pseudo R-squ.:               0.0009783
Time:                        16:57:42   Log-Likelihood:                -5030.5
converged:                       True   LL-Null:                       -5035.4
Covariance Type:            nonrobust   LLR p-value:                  0.001696
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     -2.1001      0.023    -90.073      0.000      -2.146      -2.054
treatment      0.0868      0.028      3.113      0.002       0.032       0.141
==============================================================================
```

3. Effect of Match Ratio Size

Pairwise T-tests:

We test whether larger match ratios lead to increased giving by comparing:

- 1:1 vs 2:1
- 2:1 vs 3:1

```
TtestResult(statistic=3.101361000543946, pvalue=0.0019274025949016982, df=50081.0)
                            OLS Regression Results
==============================================================================
Dep. Variable:                   gave   R-squared:                       0.000
Model:                           OLS   Adj. R-squared:                  0.000
Method:                Least Squares   F-statistic:                     9.618
Date:               Sat, 19 Apr 2025   Prob (F-statistic):            0.00193
Time:                       16:57:10   Log-Likelihood:                 26630.
No. Observations:              50083   AIC:                        -5.326e+04
Df Residuals:                  50081   BIC:                        -5.324e+04
Df Model:                          1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      0.0179      0.001     16.225      0.000       0.016       0.020
treatment      0.0042      0.001      3.101      0.002       0.002       0.007
==============================================================================
Omnibus:                    59814.280   Durbin-Watson:                   2.005
Prob(Omnibus):                  0.000   Jarque-Bera (JB):          4317152.727
Skew:                           6.740   Prob(JB):                         0.00
Kurtosis:                      46.440   Cond. No.                         3.23
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

Findings:

- No statistically significant differences in response rates are detected between these match conditions.

Regression Analysis:

We estimate a regression: gave ~ C(ratio) to examine if match ratios have a differential effect on giving.

Findings:

- The coefficients on 2:1 and 3:1 matches are not statistically different from the baseline 1:1 match.

Interpretation: This suggests that the marginal return on increasing the match ratio is effectively zero. The presence of a match appears sufficient to generate psychological salience. Beyond that, larger incentives may not enhance perceived value or urgency. This is a key behavioral insight that challenges traditional economic models of price sensitivity.

Manual Difference Calculations:

Using group means, we directly computed the differences in donation rates:

- 2:1 vs 1:1: negligible
- 3:1 vs 2:1: negligible

Interpretation: These values corroborate our regression findings. They indicate that once the concept of matching is introduced, the specific multiple does not meaningfully alter donor behavior. This implies diminishing psychological returns to increasing match ratios.

4. Donation Amounts

Full Sample Regression:

We regressed total donation amount on treatment status for the entire sample.

Findings:

- The average donation is slightly higher in the treatment group, though this difference is modest and varies in statistical significance.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                 amount   R-squared:                       0.000
Model:                            OLS   Adj. R-squared:                  0.000
Method:                 Least Squares   F-statistic:                     3.461
Date:                Sat, 19 Apr 2025   Prob (F-statistic):             0.0628
Time:                        17:14:16   Log-Likelihood:            -1.7946e+05
No. Observations:               50083   AIC:                         3.589e+05
Df Residuals:                   50081   BIC:                         3.589e+05
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      0.8133      0.067     12.063      0.000       0.681       0.945
treatment      0.1536      0.083      1.861      0.063      -0.008       0.315
==============================================================================
Omnibus:                    96861.113   Durbin-Watson:                   2.008
Prob(Omnibus):                  0.000   Jarque-Bera (JB):        240735713.635
Skew:                          15.297   Prob(JB):                         0.00
Kurtosis:                     341.269   Cond. No.                         3.23
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
                            OLS Regression Results
==============================================================================
Dep. Variable:                 amount   R-squared:                       0.000
Model:                            OLS   Adj. R-squared:                 -0.001
Method:                 Least Squares   F-statistic:                    0.3374
Date:                Sat, 19 Apr 2025   Prob (F-statistic):              0.561
Time:                        17:14:16   Log-Likelihood:                -5326.8
No. Observations:                1034   AIC:                         1.066e+04
Df Residuals:                    1032   BIC:                         1.067e+04
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept     45.5403      2.423     18.792      0.000      40.785      50.296
treatment     -1.6684      2.872     -0.581      0.561      -7.305       3.968
==============================================================================
Omnibus:                      587.258   Durbin-Watson:                   2.031
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             5623.279
Skew:                           2.464   Prob(JB):                         0.00
Kurtosis:                      13.307   Cond. No.                         3.49
==============================================================================
```

Conditional on Donation:

Restricting the sample to individuals who made a donation (gave == 1), we re-estimated the regression.

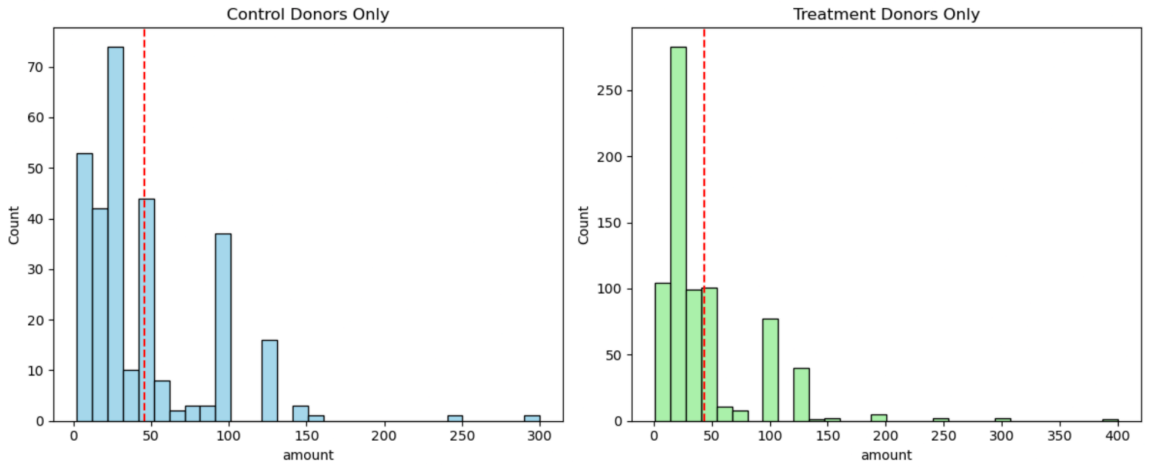
Findings:

- No significant difference in average donation size was observed.

Interpretation: The treatment's effectiveness lies primarily in increasing the participation rate rather than altering the generosity of existing donors. This insight is crucial for organizations aiming to expand their donor base.

Histograms:

We plotted donation amounts for donors in treatment and control groups.



Graph Interpretation:

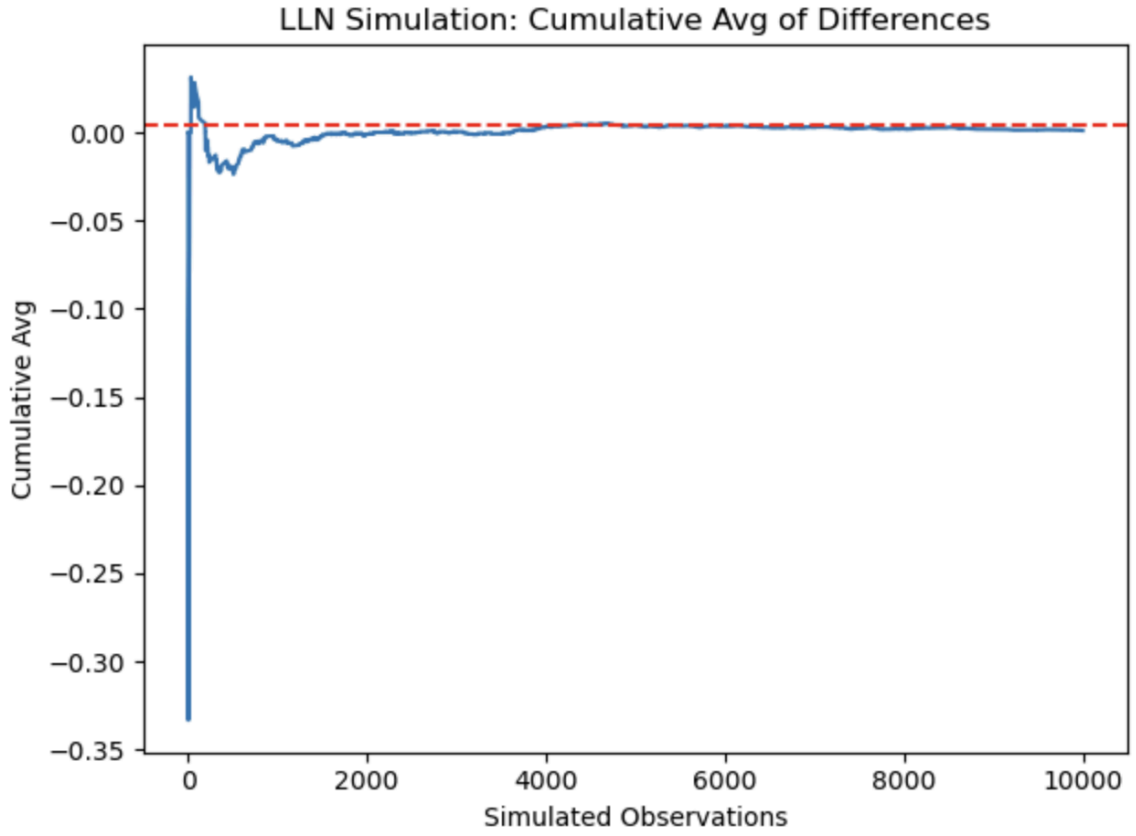
- Both distributions are right-skewed, reflecting a typical long-tail pattern in giving.
- The sample means, indicated by vertical lines, are nearly identical.

Interpretation: These plots visually confirm that conditional giving behavior remains stable across groups. Behavioral change occurs primarily at the extensive margin (whether to give) rather than the intensive margin (how much to give).

5. Simulation Analysis

Law of Large Numbers (LLN):

We simulated 10,000 draws from control (p = 0.018) and treatment (p = 0.022) distributions and plotted the cumulative mean difference.



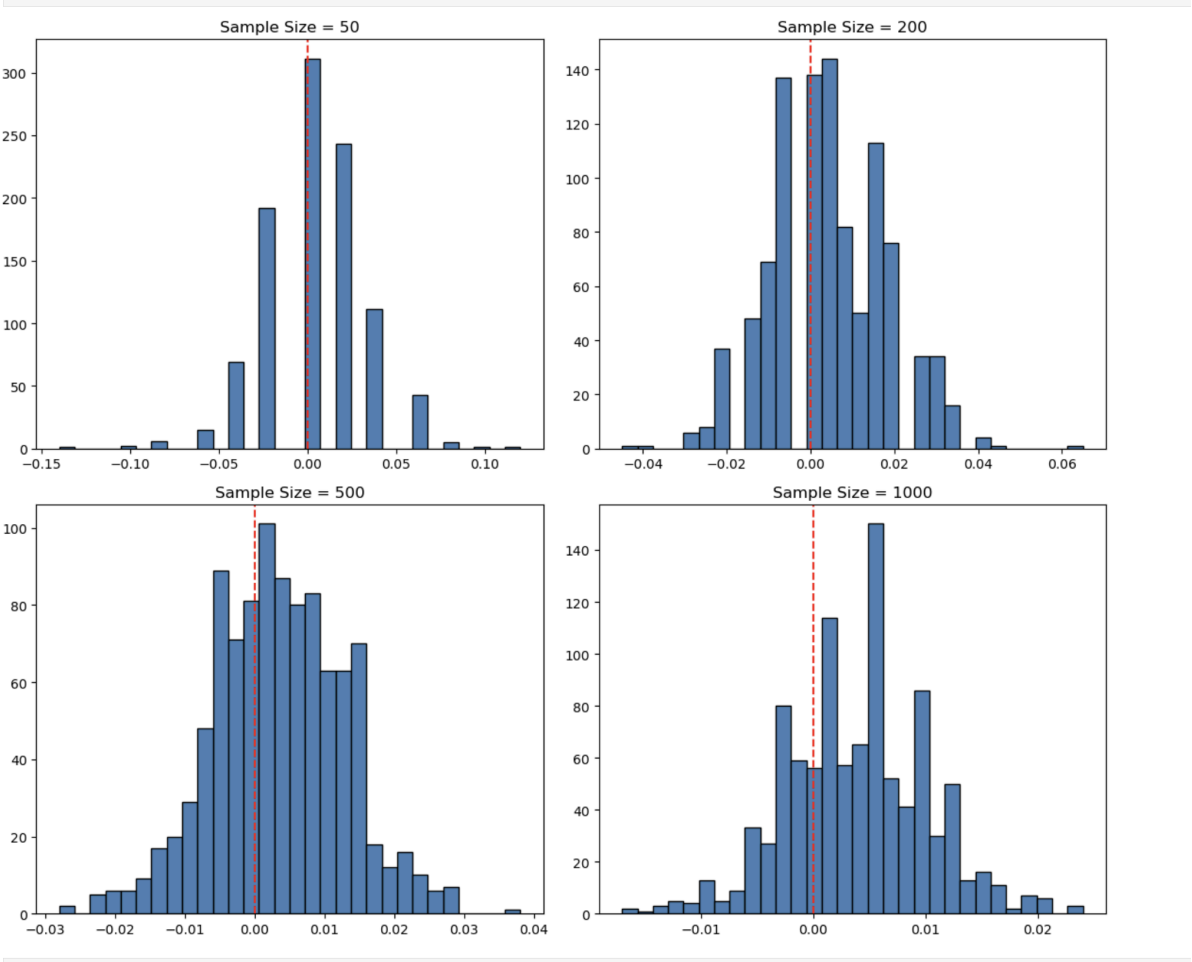
Graph Interpretation:

- The cumulative average converges to the true population difference of 0.004.
- This convergence illustrates the principle that sample means become reliable with sufficient data.

Interpretation: This simulation reinforces the trustworthiness of experimental outcomes derived from large samples. The LLN ensures that our results are not driven by chance.

Central Limit Theorem (CLT):

We simulated distributions of sample mean differences at sizes 50, 200, 500, and 1000.



Graph Interpretation:

- As the sample size increases, the sampling distribution becomes more concentrated and bell-shaped.
- The distributions center around the expected value (0.004), and their variance decreases.

Interpretation: This exercise demonstrates the CLT in action, underscoring the role of sample size in determining statistical power and inference reliability. It justifies the paper's use of large-scale field experimentation.

Conclusion

Our replication confirms the main findings of Karlan and List (2007):

- The presence of a matching grant significantly increases donation rates.
- The size of the match offer has little to no additional effect.
- The treatment primarily affects the extensive margin (whether to donate), not the intensive margin (how much to donate).

These results have meaningful implications for fundraising strategies. From a behavioral economics perspective, matching offers serve more as psychological signals than economic incentives. Fundraisers may benefit more from framing and visibility of matching rather than adjusting match magnitudes.

The robustness of these findings, confirmed through both classical inference and simulation, highlights the value of well-designed field experiments in revealing the underlying drivers of economic behavior.