# Pre-registration and Power Analysis: Low Odds, High Hopes: Supporting Carbon Pricing with Cash Rebates

March 31, 2026

## 1 Power Analysis: Omnibus Tests of Treatment Effects

### 1.1 Setup and Notation

The primary outcome variable in our study is the number of units of the virtual good chosen by participant $i$, denoted by $Y_i \in \{0, 1, 2, 3, 4\}$. The power analysis is conducted for the primary endpoint $Y_{i1}$ (Round 1 units purchased). For the purpose of power analysis, the outcome is treated as a continuous variable. All power calculations are conducted for the experimental condition with a single price level $P = 13$.

While Round 1 purchases constitute the primary endpoint for hypothesis testing and power calculations, the experimental design includes four rounds of decisions. As secondary outcomes, we will examine the dynamic treatment effects across the four rounds to assess whether responses evolve over repeated decisions. In particular, we will estimate regression specifications that include round indicators and treatment-by-round interaction terms to test whether the effectiveness of the probabilistic rebate conditions diminishes over time. We will also estimate specifications separately by round. These secondary analyses complement the primary analysis but are not used for power calculations.

The experimental design includes five treatment arms: no rebate (0%), a sure rebate (100%), and three probabilistic rebates (50%, 25%, and 10%). In the regression framework, treatment assignment is represented using four indicator variables, with one treatment serving as the omitted reference category. Let $T_i$ denote the vector of treatment indicators and $X_i$ denote a vector of pre-specified control variables.

The main outcome equation is given by

$$Y_i = \alpha + T_i'\beta + X_i'\gamma + \varepsilon_i, \tag{1}$$

where $\varepsilon_i$ is an idiosyncratic error term $\mathbb{E}[\varepsilon_i] = 0$ and $\text{Var}(\varepsilon_i) = \sigma_\varepsilon^2$.

We implement a new design in order to fully address the comments of reviewers and isolate the effect of probabilistic rebates on consumption and carbon emissions. In the absence of pilot data from the exact current design, we base the variance assumption for the power analysis on data from the closest available previous study implementing the same underlying purchasing and carbon emission tradeoff task. At the same time, the current design differs from the earlier experiment both in the support of the quantity choice and in the nominal payoff scale, so we do not anchor the variance assumption to a single price point. Instead, we normalize quantity by the maximum feasible number of units in each design and use purchase share from the previous study as the benchmark. Specifically, letting $q_i^{\text{old}} \in \{0, 1, 2, 3\}$ denote the number of units chosen in the earlier experiment, we define

$$s_i^{\text{old}} = \frac{q_i^{\text{old}}}{3} \tag{2}$$

To avoid conflating subject-level variability with the price manipulation in the earlier design, we estimate the regression

$$s_i^{\text{old}} = \alpha + \sum_j \gamma_j \mathbf{1}\{p_i = j\} + \varepsilon_i \tag{3}$$

where the indicators $\mathbf{1}\{p_i = j\}$ capture the price levels used in the previous experiment. We then use the regression root mean squared error (RMSE), denoted $\widehat{\sigma}_\varepsilon$, as the pooled within-price residual standard deviation of normalized quantity. This provides a measure of dispersion in purchasing behavior after removing systematic differences across the previously used price levels.

We map this quantity to the current design, where $q_i^{\text{new}} \in \{0, 1, 2, 3, 4\}$, by multiplying by the new maximum number of units:

$$\widehat{\sigma}_Y = 4\widehat{\sigma}_\varepsilon \tag{4}$$

Therefore, because we calculate the Root Mean Square Error (RMSE) as $\widehat{\sigma}_\varepsilon = 0.300$ from the previous experiment, the value of $\widehat{\sigma}_Y$ used in the power analysis is $\widehat{\sigma}_Y = 4 \times 0.300 = 1.200$.

In addition, to benchmark the explanatory power of the control variables, we use data from the closest available previous experiment while netting out the price levels used in that experiment. Specifically, we regress normalized purchase share on price fixed effects and the full set of control variables, and compute the partial coefficient of determination of the controls conditional on price,

$$R^2_{\text{controls|price}} = \frac{R^2_{\text{full}} - R^2_{\text{price}}}{1 - R^2_{\text{price}}}, \tag{5}$$

where $R^2_{\text{price}}$ is obtained from a regression including only price fixed effects and $R^2_{\text{full}}$ from a regression including both price fixed effects and controls. This yields a benchmark for the share of within-price variation explained by the controls, which is more appropriate for the current design than the raw $R^2$ from any single price level. We include price as a full set of fixed effects, rather than as a linear regressor, in order to net out the price manipulation without imposing a linear functional-form restriction on the relationship between price and quantity chosen. The resulting value of $R^2_{\text{controls|price}} = 0.0723$ is treated as a fixed input for all subsequent power calculations.

### 1.2 Omnibus Test of Treatment Effects Without Controls

We first consider an omnibus test of whether treatment assignment affects the outcome in the absence of additional covariates. This corresponds to an $F$-test of the joint null hypothesis

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0, \tag{6}$$

where the four coefficients correspond to the treatment indicators included in the regression model (the no rebate treatment is used as the base).

This hypothesis can equivalently be stated in terms of the coefficient of determination as

$$H_0 : R_F^2 = R_R^2, \tag{7}$$

where $R_F^2$ denotes the $R^2$ of the full model including treatment indicators and $R_R^2$ denotes the $R^2$ of the reduced model. In this specification, the reduced model includes only a constant, implying

$$R_R^2 = 0. \tag{8}$$

Power calculations are conducted for an $F$-test of the incremental explanatory power attributable to the treatment indicators. Let

$$\Delta R^2 = R_F^2 - R_R^2 = R_F^2. \tag{9}$$

We consider a grid of minimally detectable effect sizes in terms of explained variance, with $\Delta R^2 \in \{0.01,\ 0.03,\ 0.06\}$.[1]

[1]To aid interpretation, the incremental coefficients of determination considered in the power analysis, $\Delta R^2 \in \{0.01,\ 0.03,\ 0.06\}$, are translated into standardized effect sizes using Cohen's $f^2$ metric for multiple linear regression.

In a test comparing a full model to a reduced model, Cohen's effect size is defined as

$$f^2 = \frac{R_F^2 - R_R^2}{1 - R_F^2} = \frac{\Delta R^2}{1 - R_F^2}, \tag{10}$$

where $R_F^2$ is the coefficient of determination of the full model and $R_R^2$ is that of the reduced model. For the omnibus tests without controls, the reduced model contains only an intercept, implying $R_R^2 = 0$ and $R_F^2 = \Delta R^2$. Substituting each value of $\Delta R^2$ yields

$$f^2(\Delta R^2 = 0.01) = \frac{0.01}{1 - 0.01} \approx 0.010,$$
$$f^2(\Delta R^2 = 0.03) = \frac{0.03}{1 - 0.03} \approx 0.031,$$
$$f^2(\Delta R^2 = 0.06) = \frac{0.06}{1 - 0.06} \approx 0.064.$$

For the omnibus tests conditional on controls, the reduced-model coefficient of determination is fixed at $R_R^2 = 0.0723$. The corresponding values of $f^2$ are therefore

$$f^2(\Delta R^2 = 0.01) = \frac{0.01}{1 - (0.0723 + 0.01)} = \frac{0.01}{0.9177} \approx 0.0109,$$
$$f^2(\Delta R^2 = 0.03) = \frac{0.03}{1 - (0.0723 + 0.03)} = \frac{0.03}{0.8977} \approx 0.0334,$$
$$f^2(\Delta R^2 = 0.06) = \frac{0.06}{1 - (0.0723 + 0.06)} = \frac{0.06}{0.8677} \approx 0.0691.$$

The significance level is fixed at $\alpha = 0.05$, corresponding to a Type I error rate of 5%. The target power is set to $1 - \beta = 0.80$, implying a Type II error rate of $\beta = 0.20$. The required total sample size is computed using the noncentral $F$-distribution associated with the test of joint significance in a multiple linear regression.

```
* Constants
local sd        = 0.29931507*4
local r2_ctrl   = 0.07227477

* (A) Treatments only: reduced model is constant-only => R2_R = 0
power rsquared 0, ntested(4) ncontrol(1) diff(0.01 0.03 0.06) ///
    power(0.8) alpha(0.05) parallel
```

where `ntested(4)` specifies the four treatment coefficients being jointly tested and `ncontrol(1)` indicates that the reduced model includes only an intercept.[2]

### 1.2.1 Stata output

```
Performing iteration ...

Estimated sample size for multiple linear regression
F test for R2 testing subset of coefficients
H0: R2_F = R2_R  versus  Ha: R2_F != R2_R

  +-------------------------------------------------------------------------+
  |   alpha   power       N   delta    R2_R    R2_F R2_diff ntested ncontrol |
  |-------------------------------------------------------------------------|
  |     .05      .8   1,187   .0101       0     .01     .01       4        1 |
  |     .05      .8     391  .03093       0     .03     .03       4        1 |
  |     .05      .8     192  .06383       0     .06     .06       4        1 |
  +-------------------------------------------------------------------------+
```

Cohen's conventional benchmarks for $f^2$ are $0.1^2 = 0.01$ (small), $0.25^2 = 0.0625$ (medium), and $0.4^2 = 0.16$ (large). Accordingly, the values implied by the chosen grid of $\Delta R^2$ correspond to small/medium effect sizes. The use of multiple values of $\Delta R^2$ in the power analysis is intended to provide a sensitivity range rather than to assert a specific expected effect size.

[2] Power calculations using `power rsquared` are based on the noncentral $F$ distribution for testing linear restrictions in multiple regression, following Cohen (1988). Consider a regression of $y_i$ on $\kappa+\mu$ covariates with i.i.d. normal errors. For a test of a subset of $\mu$ coefficients, the null hypothesis $H_0 : \beta_{\kappa+1} = \cdots = \beta_{\kappa+\mu} = 0$ can be equivalently stated as $H_0 : R_F^2 = R_R^2$, where $R_F^2$ and $R_R^2$ denote the coefficients of determination of the full and reduced models, respectively. The effect size is defined as $\delta = (R_F^2 - R_R^2)/(1 - R_F^2)$. Under the alternative, the corresponding $F$ statistic follows a noncentral $F$ distribution with numerator degrees of freedom $\nu_1 = \mu$, denominator degrees of freedom $\nu_2 = n - \kappa - \mu - 1$, and noncentrality parameter $\lambda = n\delta$. Power is computed as $\pi = 1 - F_{\nu_1,\nu_2,\lambda}(F_{\nu_1,\nu_2,1-\alpha})$, where $F_{\nu_1,\nu_2,\lambda}(\cdot)$ denotes the cumulative noncentral $F$ distribution. Sample size is obtained by numerically solving this expression for $n$ given target power and significance level, and the resulting value is rounded up to the nearest integer.

## 1.3 Omnibus Test of Treatment Effects With Controls

We next consider an omnibus test of treatment effects conditional on a rich set of pre-specified control variables. The reduced model includes a constant and the full vector of controls $X_i$, while the full model additionally includes the four treatment indicators.

The null hypothesis remains

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0, \tag{11}$$

but the explanatory power of the reduced model is no longer zero. Based on pilot data and the analysis shown in Section 1.1, we set $R^2_R = R^2_{\text{controls|price}} = 0.0723$. The reduced model contains 32 parameters, corresponding to one intercept and 31 control variables. Power calculations are conducted for detecting an incremental increase in explained variance attributable to treatment assignment of $\Delta R^2 = R^2_F - R^2_R \in \{0.01,\ 0.03,\ 0.06\}$.

As before, the target power is 0.80 and the significance level is $\alpha = 0.05$. The required total sample size is derived from the noncentral $F$-distribution for testing a subset of regression coefficients in a multiple linear regression model.

These calculations are implemented in Stata using the following command:

```
* (B) Treatments + 31 controls: reduced model includes controls only
*     R2_R = 0.07227477, ncontrol(32) (constant + 31 controls)
power rsquared `r2_ctrl', ntested(4) ncontrol(32) diff(0.01 0.03 0.06) ///
    power(0.8) alpha(0.05) parallel
```

### 1.3.1 Stata output

```
Performing iteration ...

Estimated sample size for multiple linear regression
F test for R2 testing subset of coefficients
H0: R2_F = R2_R  versus  Ha: R2_F != R2_R

  +-------------------------------------------------------------------------+
  |   alpha   power       N   delta    R2_R    R2_F R2_diff ntested ncontrol |
  |-------------------------------------------------------------------------|
  |     .05      .8   1,101   .0109  .07227  .08227     .01       4       32 |
  |     .05      .8     363  .03342  .07227   .1023     .03       4       32 |
  |     .05      .8     179  .06915  .07227   .1323     .06       4       32 |
  +-------------------------------------------------------------------------+
```

## 1.4 Interpretation

In both specifications, the `power rsquared` command reports the total sample size required to achieve the specified power for the omnibus $F$-test of joint treatment significance. These calculations

provide benchmark estimates for the detectability of aggregate treatment effects and serve as a complement to the pairwise and trend-based power analyses that focus on specific contrasts of substantive interest.

# 2 Power Analysis: Primary Pairwise Contrast (Sure vs. 10%)

## 2.1 Hypothesis testing

The primary confirmatory comparison evaluates whether a sure rebate (treatment 2) produces different behavior than a low-probability rebate (treatment 5, 10%). Let $\mu_2 = \mathbb{E}[Y_i \mid \text{treat} = 2]$ denote the mean number of units chosen under the sure rebate and $\mu_5 = \mathbb{E}[Y_i \mid \text{treat} = 5]$ denote the corresponding mean under the 10% rebate. The primary estimand is the mean difference

$$\delta = \mu_2 - \mu_5. \tag{12}$$

The null hypothesis for the primary pairwise test is

$$H_0 : \delta = 0, \tag{13}$$

against a two-sided alternative $H_A : \delta \neq 0$.

## 2.2 Power Calculation

Power for the primary confirmatory contrast (sure rebate versus 10% probabilistic rebate) is based on a two-sample test of equality of means. The outcome variable $Y_i$ is treated as continuous, and we assume independent samples and equal variances across the two treatment groups. Let $\sigma_Y$ denote the common within-group standard deviation of $Y_i$. As shown in Section 1.1, using data from a previous experiment, we set $\widehat{\sigma} = \widehat{\sigma}_Y = 4\widehat{\sigma}_\varepsilon = 1.200$.

For a target mean difference $\delta = \mu_2 - \mu_1$, significance level $\alpha$, and power $1 - \beta$, a commonly used large-sample approximation for the required per-group sample size under equal allocation is given by (Kupper and Hafner, 1989):

$$n(\delta) = \frac{2\big(z_{1-\alpha/2} + z_{1-\beta}\big)^2 \sigma^2}{\delta^2}, \tag{14}$$

where $z_q$ denotes the $q$-quantile of the standard normal distribution. For $\alpha = 0.05$ and $\beta = 0.20$, the corresponding quantiles are $z_{1-\alpha/2} = 1.96$ and $z_{1-\beta} = 0.84$. This expression provides a useful benchmark and serves as an initial approximation. In implementation, however, sample sizes are computed using the exact two-sided power equation based on the noncentral Student's $t$ distribution. Under the alternative hypothesis, the test statistic follows a noncentral $t$ distribution with degrees of freedom $\nu = 2n - 2$ and noncentrality parameter $\lambda = \frac{\delta}{\sigma_Y \sqrt{2/n}}$.

Because both $\nu$ and $\lambda$ depend on the unknown sample size $n$, the power equation has no closed-form solution. Required sample sizes are therefore obtained by numerically solving this equation for $n$ given target power and significance level.

We conduct a sensitivity analysis over minimally detectable mean differences $\delta \in \{0.15,\ 0.20,\ 0.25\}$, measured in units purchased, with target power $1-\beta = 0.80$ and significance level $\alpha = 0.05$. These calculations are implemented in Stata using the `power twomeans` command as follows:

```
foreach diff of numlist 0.15 0.20 0.25 {
    power twomeans 1.09, diff(‘diff’) sd(‘sd’) power(0.8) alpha(0.05)
}
```

In this command, the value `1.09` specifies an anchor mean for one group. Although this anchor corresponds to the mean outcome in the prior experiment, it does not affect the required sample size, because power depends on the mean difference $\delta$ and the standard deviation $\sigma_Y$, not on the absolute level of the group means. The option `diff(‘diff’)` supplies the target mean difference $\delta$, and `sd(‘sd’)` supplies the estimate $\widehat{\sigma}_Y$.

Using $\widehat{\sigma}_Y = 1.200$, target power $1-\beta = 0.80$, and significance level $\alpha = 0.05$, the required per-group sample sizes are 1,002 for $\delta = 0.15$, 564 for $\delta = 0.20$, and 361 for $\delta = 0.25$. Full Stata output is reported in Appendix A.

# 3 Power Analysis: Secondary Confirmatory Test (Trend Across Rebate Probabilities)

The secondary confirmatory analysis evaluates whether outcomes vary systematically with the probability that the rebate is realized. This analysis is conducted within the rebate treatments only, i.e., among participants assigned to treatments 2, 3, 4, and 5. Let $p_i$ denote the rebate probability assigned to participant $i$, coded as

$$p_i = \begin{cases} 1.00 & \text{if treat} = 2 \text{ (sure rebate)}, \\ 0.50 & \text{if treat} = 3, \\ 0.25 & \text{if treat} = 4, \\ 0.10 & \text{if treat} = 5. \end{cases} \tag{15}$$

Within the rebate subsample, the trend specification is:

$$Y_i = \alpha + \tau p_i + X_i'\gamma + \varepsilon_i, \tag{16}$$

where $X_i$ denotes the vector of pre-specified control variables (included in one specification and omitted in another), and $\tau$ captures the linear association between the rebate probability and the outcome.

The secondary confirmatory null hypothesis is

$$H_0 : \tau = 0, \tag{17}$$

tested against a two-sided alternative $H_A : \tau \neq 0$ at significance level $\alpha = 0.05$.

### 3.1 Power Calculation via an Incremental $R^2$ Test

Power for the trend test uses Stata's `power rsquared` command, treating the test of $H_0 : \tau = 0$ as a one-degree-of-freedom $F$-test for the incremental contribution of the single tested covariate $p_i$ to the coefficient of determination. In this framework, the null hypothesis can be equivalently expressed as

$$H_0 : R_F^2 = R_R^2, \tag{18}$$

where $R_F^2$ is the $R^2$ from the full model including $p_i$, and $R_R^2$ is the $R^2$ from the reduced model excluding $p_i$. The effect size is specified in terms of the incremental explained variance $\Delta R^2 = R_F^2 - R_R^2$.

We conduct a sensitivity analysis over the grid $\Delta R^2 \in \{0.01,\ 0.03,\ 0.06\}$ with target power $1 - \beta = 0.80$ and significance level $\alpha = 0.05$. The resulting sample size requirements are computed using the noncentral $F$-distribution underlying the test of a subset of coefficients in a multiple linear regression model.

### 3.2 Implementation in Stata and Interpretation of the Reported Sample Size

The trend test is planned in two specifications. In the first specification, the reduced model contains only a constant (so $R_R^2 = 0$); in the second specification, the reduced model contains the constant and the full set of pre-specified controls, with $R_R^2$ fixed at the estimate $R_R^2 = R^2_{\text{controls|price}} = 0.0723$ from a previous experiment and with 32 reduced-model parameters (intercept plus 31 controls). These calculations are implemented in Stata as follows:

```
* Trend test, treatments only (rebate arms): reduced model constant-only
power rsquared 0, ntested(1) ncontrol(1) diff(0.01 0.03 0.06) ///
    power(0.8) alpha(0.05) parallel

* Trend test with 31 controls (rebate arms): reduced model includes controls
power rsquared 'r2_ctrl', ntested(1) ncontrol(32) diff(0.01 0.03 0.06) ///
    power(0.8) alpha(0.05) parallel
```

where `ntested(1)` specifies that the trend coefficient $\tau$ (associated with the single regressor $p_i$) is the only tested covariate.

Because the trend regression is estimated on the rebate arms only, the sample size $N_{\text{rebate}}$ reported by `power rsquared` should be interpreted as the required number of observations in the rebate subsample. Under the planned allocation of subjects across all treatment arms, let $s_{\text{rebate}}$

denote the share of the total sample assigned to the rebate arms (treatments 2, 3, 4, and 5): $s_{\text{rebate}} = s_2 + s_3 + s_4 + s_5$. Hence, $s_{\text{norebate}} = 1 - s_{\text{rebate}}$.

The implied total sample size required to achieve $N_{\text{rebate}}$ observations within the rebate subsample is therefore

$$N_{\text{total}} \geq \left\lceil \frac{N_{\text{rebate}}}{s_{\text{rebate}}} \right\rceil, \tag{19}$$

where $\lceil \cdot \rceil$ denotes the ceiling operator.

### 3.2.1 Stata output

```
. * Trend test, treatments only (rebate arms): reduced model constant-only
. power rsquared 0, ntested(1) ncontrol(1) diff(0.01 0.03 0.06) ///
>     power(0.8) alpha(0.05) parallel

Performing iteration ...

Estimated sample size for multiple linear regression
F test for R2 testing subset of coefficients
H0: R2_F = R2_R  versus  Ha: R2_F != R2_R

  +------------------------------------------------------------------------------+
  |   alpha   power    N   delta    R2_R    R2_F R2_diff ntested ncontrol |
  |------------------------------------------------------------------------------|
  |     .05      .8    779   .0101       0     .01     .01       1        1 |
  |     .05      .8    256  .03093       0     .03     .03       1        1 |
  |     .05      .8    125  .06383       0     .06     .06       1        1 |
  +------------------------------------------------------------------------------+
```

```
. * Trend test with 31 controls (rebate arms): reduced model includes controls
. power rsquared 'r2_ctrl', ntested(1) ncontrol(32) diff(0.01 0.03 0.06) ///
>     power(0.8) alpha(0.05) parallel

Performing iteration ...

Estimated sample size for multiple linear regression
F test for R2 testing subset of coefficients
H0: R2_F = R2_R  versus  Ha: R2_F != R2_R

  +------------------------------------------------------------------------------+
  |   alpha   power       N   delta    R2_R    R2_F R2_diff ntested ncontrol |
  |------------------------------------------------------------------------------|
  |     .05      .8     722  .01091   .0737   .0837     .01       1       32 |
  |     .05      .8     237  .03347   .0737   .1037     .03       1       32 |
  |     .05      .8     117  .06926   .0737   .1337     .06       1       32 |
  +------------------------------------------------------------------------------+
```

### 3.3 Interpretation

The trend test evaluates whether choices vary monotonically with the probability of receiving a rebate among rebate treatments. In the specification without additional covariates, detecting an incremental contribution of $\Delta R^2 = 0.01$ requires 779 observations in the rebate subsample, while larger effects of $\Delta R^2 = 0.03$ and $\Delta R^2 = 0.06$ require 256 and 125 observations, respectively. When conditioning on the full set of 31 pre-specified control variables, the required rebate-subsample sizes decrease modestly to 722, 237, and 117 observations for the same effect sizes. All calculations target 80% power at a 5% significance level.

## 4 Conclusion

We plan to recruit 362 subjects in each of the two focal rebate treatments, namely the sure rebate treatment (100%) and the low-probability rebate treatment (10%), and 104 subjects in each of the remaining three treatment arms. This allocation yields a total planned sample of $N = 1{,}036$. The allocation is chosen to ensure adequate power for the primary pairwise comparison in Section 2, which requires 361 subjects per group to detect a mean difference of $\delta = 0.25$ between the 100% and 10% rebate treatments with 80% power at the 5% significance level. This analysis also yields 932 observations across the rebate treatments, which exceeds the 779 observations required for the secondary trend analysis in Section 3. This allocation satisfies the omnibus treatment-effects analysis in Section 1.1 for the case of an incremental difference in explained variance of $\Delta R^2 = 0.03$. Accordingly, the planned sample is sufficient for the primary pairwise and rebate-trend analyses and for omnibus tests targeting small-to-moderate effects, though not for the smallest omnibus effect considered in the sensitivity analysis.

## References


Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2 ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Kupper, L. L. and K. B. Hafner (1989). How appropriate are popular sample size formulas? *The American Statistician 43*(2), 101–105.


## A Stata Output for Pairwise Power Calculations

```
Performing iteration ...

Estimated sample sizes for a two-sample means test
t test assuming sd1 = sd2 = sd
H0: m2 = m1  versus  Ha: m2 != m1
```

```
Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =    0.1500
           m1 =    1.0900
           m2 =    1.2400
         diff =    0.1500
           sd =    1.1973

Estimated sample sizes:

            N =     2,004
  N per group =     1,002

Performing iteration ...

Estimated sample sizes for a two-sample means test
t test assuming sd1 = sd2 = sd
H0: m2 = m1  versus  Ha: m2 != m1

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =    0.2000
           m1 =    1.0900
           m2 =    1.2900
         diff =    0.2000
           sd =    1.1973

Estimated sample sizes:

            N =     1,128
  N per group =       564

Performing iteration ...

Estimated sample sizes for a two-sample means test
t test assuming sd1 = sd2 = sd
H0: m2 = m1  versus  Ha: m2 != m1

Study parameters:

        alpha =    0.0500
        power =    0.8000
        delta =    0.2500
```

```
           m1 =    1.0900
           m2 =    1.3400
         diff =    0.2500
           sd =    1.1973

Estimated sample sizes:

            N =       722
  N per group =       361
```