Abstract Title Page

Title:

Bound Constrained Optimization of Sample Sizes with Monetary Restrictions in Planning Multilevel Randomized Experiments and Regression Discontinuity Studies

Author(s):

Metin Bulus, *Adiyaman University*
Nianbo Dong, *University of North Carolina at Chapel Hill*

Abstract Body

Background:

Sample size determination in multilevel randomized trials (MRTs) and multilevel regression discontinuity designs (MRDDs) can be complicated due to multilevel structure, monetary restrictions, and other limitations with sample sizes at one or more levels. These issues have sparked a set of studies under optimal design literature (Hedges & Borenstein, 2014; Konstantopoulos, 2011, 2013; Liu, 2003; Moerbeek, & Safarkhani, 2018; Raudenbush, 1997; Raudensbush & Liu, 2000; Rhoads & Dye, 2016). Most of the ideas from these studies have been implemented in CRT-Power (Borenstein, Hedges, & Rothstein, 2012), Optimal Design Plus (Raudenbush et al., 2011), PowerUp! (Dong & Maynard, 2013a, 2013b) and PowerUpR (Bulus, Dong, Kelcey, Spybrook, 2018). The literature on the sample size determination in MRTs and MRDDs and their implementation in software packages has been scarce, scattered, and incomplete. Table 1C summarizes key aspects of the studies and software packages, and contribution of this study.

Focus of study:

Although analytical solutions exist when there are two or less unknown parameters with the fixed budget case, there are some limitations in more complex designs especially when the intention is to minimize the total cost. For example, Figure 1C illustrates a case for a four-level multisite individual-randomized trial where minimum variance does not correspond to the maximum power. More importantly, analytical solution may not be within the feasible region. For example, very few number of classrooms in a three-level cluster randomized trial may pose challenges to the estimation of the variance parameters. In such cases and more complex problems, setting bound constraints allow finding optimal solutions for sample sizes within a reasonable finite range.

This study proposes a general and flexible framework under constrained optimal sample allocation (COSA) term that is applicable to two-group MRTs and MRDDs with continuous outcomes using bound constrained optimization technique and implements the framework in the `cosa` R library. As a result, four distinct design questions can be addressed while sample sizes at one or more levels can be fixed or limited by bounds: (i) Given marginal costs per units and a fixed budget, what is the optimal allocation of subjects/clusters across levels [and/or across treatment condition] to achieve highest level of precision? (ii) Given marginal costs per units and a fixed budget, what is the optimal allocation of subjects/clusters across levels [and/or across treatment condition] to achieve highest level of power? (iii)) Given marginal costs per units, what is the most cost-efficient allocation of subjects/clusters across levels [and/or across treatment condition] given the desired level of power? (iv) Given marginal costs per units, what is the most cost-efficient allocation of subjects/clusters across levels [and/or across treatment condition] given the desired level of MDES?

Research Design:

Both regression discontinuity designs (RDDs) and randomized trials (RTs) rely on allocation of subjects to treatment conditions via an explicit assignment mechanism. Essentially, RDDs resembles RTs because random measurement error dominates the assignment mechanism considering a narrow range around the cutoff score (Boruch, 1975; Campbell & Stanley, 1963; Lee & Lemieux, 2010). In this sense, an RDD converges to RT when the measurement error completely dominates the assignment mechanism across the range of the distribution of the score variable, that is, when there is no relationship between the treatment condition and the score variable. Therefore, when correlation between the score variable and treatment condition is set to zero results pertain to MRTs. Based on this theory, constrained optimization framework unifies MRTs and MRDDs designs. Appendix B provides details of the framework and the optimization technique.

Results:

In a series of figures, we demonstrate the optimization surfaces, primary constraints and bound constraints using perspective and contour plots for a two-level multisite individual-randomized trial. Figure 2C illustrates the bound constrained solution with bound constraints on average number of level 1 units (n_1) and proportion of level 1 units assigned to the treatment condition (p). The inner rectangle with lighter color indicates the region restricted by bound constraints. The crossing points (or lines) within the rectangle indicates the feasible region defined by the primary constraint on the total cost. The larger (red) point indicates the optimal combination of the two parameters (n_1 and p).

Figure 3C illustrates the solution when treatment and control units are associated with differing marginal costs and the primary constraint is on the statistical power. However, as the feasible region approaches to relatively flat area on the optimization surface there might be other approximate solutions within the algorithm tolerance. In such cases, although optimal solution slightly deviates from analytic solutions the minimum variance or maximum statistical power does not deviate substantially.

Figure 4C illustrates the case where treatment and control units are not associated with marginal costs and the primary equality constraint is on the statistical power. Unlike previous cases, dark points (line) crossing the inner rectangle indicates both feasible region and optimal solutions, however, the algorithm picks the solution that is closest to the starting values within the algorithm tolerance. For this reason, marginal cost information should be provided or researcher should be aware of the local solution. In such cases, experimenting different starting values and algorithms would assist with decision regarding sample size allocation.

Conclusions:

Bound constrained optimization framework provides flexibility in the planning of MRTs and MRDDs. Different from other software packages, there are two types of constraints; the primary constraint can be placed on either statistical power, MDES or total cost, along with bound constraints on the sample sizes for one or more levels. In addition, marginal cost per treatment and control units does not need to be equal where proportion of units assigned to

treatment condition can also be optimized in MRTs. This potentially will assist researchers to make informed decisions regarding sample sizes and whether unbalanced treatment allocation will result in expected cost efficiency as defined by Liu (2003). In cases where cost reduction is trivial researcher should favor the balanced design for attrition and crossover reasons. Bound constrained optimization is available in `cosa` R library.

Appendixes

Not included in page count.

Appendix A. References

Birgin, E. G., Martínez, J. M. (2008). Improving Ultimate Convergence of an Augmented Lagrangian method. *Optimization Methods and Software*, *23*(2), 177-195. https://doi.org/10.1080/10556780701577730

Bloom, H. S. (1995). Minimum detectable effects a simple way to report the statistical power of experimental designs. *Evaluation Review, 19*(5), 547-556. https://doi.org/10.1177/0193841X9501900504

Bloom, H. S. (2006). The core analytics of randomized experiments for social research. MDRC Working Papers on Research Methodology. New York, NY: MDRC. Retrieved from https://www.mdrc.org/sites/default/files/full_533.pdf

Borenstein, M., Hedges, L. V., & Rothstein, H. (2012). CRT-Power [Software]. Teaneck, NJ: Biostat.

Boruch, R. F. (1975). Coupling Randomized Experiments and Approximations to Experiments in Social Program Evaluation. *Sociological Methods & Research, 4*(1), 31-53. https://doi.org/10.1177/004912417500400103

Bulus, M., Dong, N., Kelcey, B., & Spybrook, J. (2018). PowerUpR: Power Analysis Tools for Multilevel Randomized Experiments. R package version 1.0.0. https://CRAN.R-project.org/package=PowerUpR

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. In N.L. Gage (Ed.), *Handbook of Research on Teaching* (pp. 1-76). Chicago, IL: Rand McNally.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York, NY: Wiley.

Conn, A. R., Gould, N. I., & Toint, P. (1991). A Globally Convergent Augmented Lagrangian Algorithm for Optimization with General Constraints and Simple Bounds. *SIAM Journal on Numerical Analysis*, *28*(2), 545-572. Retrieved from http://www.jstor.org/stable/2157828

Dong, N., & Maynard, R. (2013a). *PowerUp!*: A Tool for Calculating Minimum Detectable Effect Sizes and Minimum Required Sample Sizes for Experimental and Quasi-experimental Design Studies. *Journal of Research on Educational Effectiveness, 6*(1), 24-67. https://doi.org/10.1080/19345747.2012.673143

Dong, N., & Maynard, R. A. (2013b). PowerUp!: A tool for calculating minimum detectable effect sizes and minimum required sample sizes for experimental and quasi-experimental design studies [Software]. http://www.causalevaluation.org/

Hedges, L. V., & Borenstein, M. (2014). Conditional Optimal Design in Three- and Four-Level Experiments. *Journal of Educational and Behavioral Statistics*, *39*(4), 257-281. https://doi.org/10.3102/1076998614534897

Johnson, S. G. (n.d.). The NLopt nonlinear-optimization package. http://ab-initio.mit.edu/nlopt

Konstantopoulos, S. (2011). Optimal Sampling of Units in Three-Level Cluster Randomized Designs: An ANCOVA Framework. *Educational and Psychological Measurement*, *71*(5), 798-813. https://doi.org/10.1177/0013164410397186

Konstantopoulos, S. (2013). Optimal Design in Three-Level Block Randomized Designs with Two Levels of Nesting: An ANOVA Framework with Random Effects. *Educational and Psychological Measurement*, *73*(5), 784-802. https://doi.org/10.1177/0013164413485752

Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, *48*(2), 281-355. https://doi.org/10.1257/jel.48.2.281

Liu, X. (2003). Statistical Power and Optimum Sample Allocation Ratio for Treatment and Control Having Unequal Costs per Unit of Randomization. *Journal of Educational and Behavioral Statistics*, *28*(3), 231-248. https://doi.org/10.3102/10769986028003231

Moerbeek, M., & Safarkhani, M. (2018). The Design of Cluster Randomized Trials With Random Cross-Classifications. *Journal of Educational and Behavioral Statistics*, *43*(2), 159-181. https://doi.org/10.3102/1076998617730303

Raudenbush, S. W. (1997). Statistical Analysis and Optimal Design for Cluster Randomized Trials. *Psychological Methods*, *2*(2), 173. http://dx.doi.org/10.1037/1082-989X.2.2.173

Raudenbush, S. W., & Liu, X. (2000). Statistical Power and Optimal Design for Multisite Trials. *Psychological Methods*, *5*, 199-213. http://dx.doi.org/10.1037/1082-989X.5.2.199

Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X. F., Martinez, A., & Bloom, H. (2011). Optimal design software for multi-level and longitudinal research (Version 3.01) [Software].

Rhoads, C. H., & Dye, C. (2016). Optimal Design for Two-Level Random Assignment and Regression Discontinuity Studies. *The Journal of Experimental Education*, *84*(3), 421-448. https://doi.org/10.1080/00220973.2015.1048846

Ypma, J. (2017). nloptr: R Interface to NLopt. R package version 1.0.4 [Software]. Retrieved from https://CRAN.R-project.org/package=nloptr

Appendix B. Bound Constrained Optimization Framework

Let $\boldsymbol{n}$ be a vector of sample sizes for L number of levels in a hierarchical structure, consisting of $n_1, n_2, \ldots, n_L$, let p be the proportion of units in the treatment group, and δ the estimand of the treatment effect of interest. The sampling variance of the standardized treatment effect (δ), will be stated as a function of $\boldsymbol{n}$ and p as

$$f(\boldsymbol{n}, p) = var(\delta|\boldsymbol{n}, p). \quad (1)$$

The MDES given power can be calculated as

$$MDES(\boldsymbol{n}, p|Power) = M_v\sqrt{f(\boldsymbol{n}, p)}. \quad (2)$$

where $M_v = c(\alpha, v) + c(\beta, v)$ for one tailed test and $M_v = c(\alpha/2, v) + c(\beta, v)$ for a two-tailed test, c is the quantile of *t*-distribution associated with probability of α, $\alpha/2$, or β and degrees of freedom v (Bloom, 1995, 2006). Inversely, the non-centrality parameter is

$$\lambda = \frac{\delta}{\sqrt{f(\boldsymbol{n}, p)}}. \quad (3)$$

Knowing the non-centrality parameter, the power of the test given an effect size (ES) can be calculated from *t* distribution as

$$Power(\boldsymbol{n}, p|ES) = \begin{cases} 1 - H(c(\alpha, v), v, \lambda), & one\ tailed \\ 1 - H\left(c\left(\frac{\alpha}{2}, v\right), v, \lambda\right) + H\left(-c\left(\frac{\alpha}{2}, v\right), v, \lambda\right), & two\ tailed \end{cases}, \quad (4)$$

where *H* is cumulative distribution function of non-central *t*-distribution given the quantile $c(\alpha, v)$ or $c(\alpha/2, v)$, degrees of freedom v, and non-centrality parameter λ ; and c is the quantile of *t*-distribution associated with probability of α or $\alpha/2$, and degrees of freedom v (Hedges & Rhoads, 2010, p. 18). Assuming sampling of each unit is associated with marginal cost, the total cost function can be defined as

$$g(\boldsymbol{n}, p) = p\left(\sum_{j=1}^{L} c_{n_{jt}} \prod_{i=j}^{L} n_i\right) + (1-p)\left(\sum_{j=1}^{L} c_{n_{jc}} \prod_{i=j}^{L} n_i\right), \quad (5)$$

where $c_{n_{jt}}$ is marginal cost per treatment unit for j^{th} level, and $c_{n_{jc}}$ is marginal cost per control unit for j^{th} level, n_i is the average sample size for i^{th} level. This is an extension of linear cost function defined by Raudenbush (1997), and Raudenbush and Liu (2000), allowing unequal marginal cost per treatment and control conditions. Unequal costs only applies to randomization level or below, and the rest of the levels are forced to have equal costs per unit so that Equation (5) accommodates all the designs.

Equations (1), (2), (4) and (5) constitutes the core of the COSA framework. Combination of (1), (2), (4) and (5) allows four distinct albeit related COSA routines each answering a particular research question. In the following section we will define primary constraints, bound constraints and the optimization technique.

Primary constraint on the total cost. To answer the question "Given marginal cost per units and a fixed budget, what is the optimal allocation of subjects/clusters across levels [and/or

across treatment condition in MRTs] to achieve highest level of precision?" Minimize sampling variance of the treatment

$$\underset{N,p\in \mathbb{R}^+ \; and \; 0<p<1}{min} f(\boldsymbol{n}, p), \tag{6}$$

subject to primary equality constraint on total cost

$$g(\boldsymbol{n}, p) = Budget, \tag{7}$$

where $\boldsymbol{n}, p \in \mathbb{R}^+$ and $0 < p < 1$. Alternatively to answer the question "Given marginal costs per units and a fixed budget, what is the optimal allocation of subjects/clusters across levels [and/or across treatment condition in MRTs] to achieve highest level of power?" Maximize the power

$$\underset{N,p\in \mathbb{R}^+ \; and \; 0<p<1}{max} Power(\boldsymbol{n}, p|ES) \tag{8}$$

subject to primary equality constraint on total cost where $\boldsymbol{n}, p \in \mathbb{R}^+$ and $0 < p < 1$.

Primary constraint on the statistical power or MDES. To answer the question "Given marginal costs per units, what is the most cost-efficient allocation of subjects/clusters across levels [and/or across treatment condition in MRTs] given the desired level of power?" Minimize total cost

$$\underset{N,p\in \mathbb{R}^+ \; and \; 0<p<1}{min} g(\boldsymbol{n}, p), \tag{9}$$

subject to primary equality constraint

$$Power(\boldsymbol{n}, p|ES) = Power, \tag{10}$$

where $\boldsymbol{n}, p \in \mathbb{R}^+$ and $0 < p < 1$. Alternatively, To answer the question "Given marginal costs per units, what is the most cost-efficient allocation of subjects/clusters across levels [and/or across treatment condition in MRTs] given the desired level of MDES?" Minimize the total cost subject to primary equality constraint

$$MDES(\boldsymbol{n}, p|Power) = ES, \tag{11}$$

where $\boldsymbol{n}, p \in \mathbb{R}^+$ and $0 < p < 1$.

Bound constraints. Along with primary constraints (constraints on total cost, statistical power or MDES), bound constraints can be placed on one or more levels by re-defining bounds, in this case, $\boldsymbol{n_{LB}} < \exists \boldsymbol{n} < \boldsymbol{n_{UB}}$, where $\boldsymbol{n_{LB}}, \boldsymbol{n_{UB}} \in \mathbb{R}^+$ where the corresponding member in the set $\boldsymbol{n_{LB}}$ does not violate minimum degrees of freedom.

Optimization. $f(\boldsymbol{n}, p)$ and $g(\boldsymbol{n}, p)$ equations can be combined into single form $h(\boldsymbol{n}, p)$ using Lagrangian function as described in Cochran (1977, p. 289) as

$$h(\boldsymbol{n}, p) = f(\boldsymbol{n}, p) + \lambda_{LM} g(\boldsymbol{n}, p), \tag{12}$$

when the primary constraint is on the total cost, and

$$h(\boldsymbol{n}, p) = g(\boldsymbol{n}, p) + \lambda_{LM} f(\boldsymbol{n}, p), \tag{13}$$

when the primary constraint is on the statistical power or MDES, where λ_{LM} is Lagrange multiplier. We use Augmented Lagrangian algorithms developed by Birgin & Martines (2008) and Conn, Gould, & Toint (1991) that are implemented in NLopt (Johnson, n.d.) and `nloptr` (Ypma, 2017) libraries. Optionally, solutions are rounded in a top-down fashion. As the top-level sample size must be an integer, post-optimization procedure re-optimizes while rounding the top-

level. Furthermore, product of a higher-level sample size with a lower sample sizes should be integers (as sample sizes for levels other than the top-level are averages or harmonic means). Instead of rounding the averages, the algorithm rounds the products. p also slightly changes depending on whether the sample size at the treatment level (as a product with top-levels) is an odd or even number, therefore, p is modified to produce integer numbers for treatment and control conditions.

Appendix C. Tables and Figures

Table 1C
Studies and Software Packages (in bold) with a Focus on Sample Allocation in MRTs and MRDDs (Chronological Order)

Study/Software	Treatment Assignment	Marginal Cost	Minimize - Constrain	Constraints on Sample Sizes
Raudenbush (1997)	Random	Equal	Variance - Cost	*NA*
Raudensbush & Liu (2000)	Random	Equal	Variance - Cost	*NA*
Liu (2003)	Random	Equal Unequal	Variance - Cost	*NA*
OD+ (Raudenbush et al., 2011)	Random	Equal	Variance – Cost β - Cost	*NA*
Konstantopoulos (2011)	Random	Equal	Variance - Cost	*NA*
CRT-Power (Borenstein, Hedges, & Rothstein, 2012)	Random	Equal	Cost - Power	Any level(s) up to L-1 levels fixed
Konstantopoulos (2013)	Random	Equal	Variance - Cost	*NA*
PowerUp! (Dong & Maynard, 2013)	Random	*NA*	*NA* - ES	L-1 levels fixed
Hedges & Borenstein (2014)	Random	Equal	Cost - Power	L-1 or L-2 levels fixed
Rhoades and Dye (2016)	Cutoff	Equal	Variance - Cost	NA
cosa (This study)	Random Cutoff	Equal Unequal	Variance - Cost β - Cost Cost - ES Cost - Power	Any level(s) up to L-1 fixed or bounded

Note. L: Total number of levels (maximum: four levels). ES: Effect Size. β: Type II error rate. OD+: Optimal Design Plus. OD+ only allows optimal design of two-level cluster randomized trials with monetary constraints. Minimizing β is equivalent to maximizing power.

(a) (Optimization Surface)

(b) (Contour Plot)

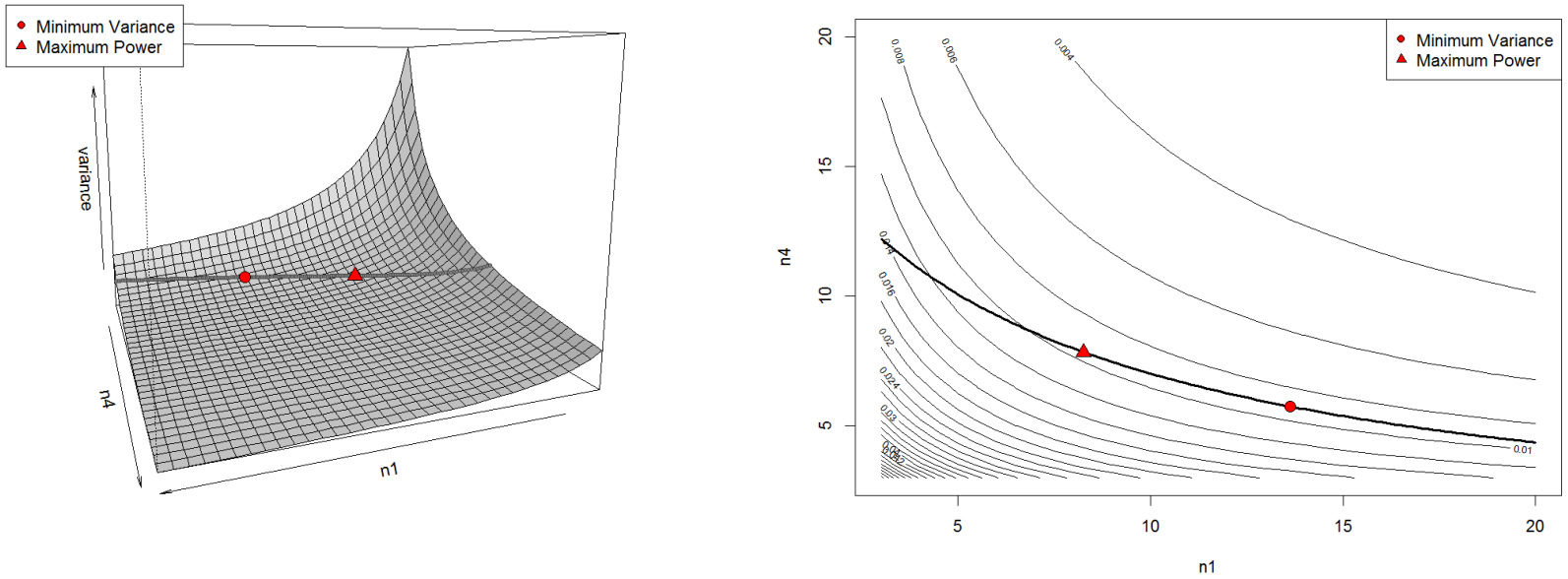

Figure 1C. *(a) demonstrates the optimization surface for n_1 and n_4 when treatment and control units are associated with marginal costs and the primary equality constraint is on the total cost. Both in (a) and (b) the dark line is the feasible region that is restricted by primary equality constraint on the total cost. The circle (red) indicates the optimal solution when variance is minimized and the triangle (red) indicates the optimal solution when power is maximized. Parameters used in the optimization procedure: marginal cost per level 1 treatment and control units are 15 and 5 correspondingly, marginal cost per level 2 unit is 30, marginal cost per level 3 unit is 60, and marginal cost per level 4 unit is 90, total cost is fixed at 5000,* $n_2 = 2$, $n_3 = 6$, $n_4 = 5$, $\rho_{TS} = 0$, $R_1^2 = 0$, $R_{T2}^2 = 0$, $R_{T3}^2 = 0$, $R_{t4}^2 = 0$, $\rho_2 = .20$, $\rho_3 = .10$, $\rho_4 = .05$, $\omega_2 = .20$, $\omega_3 = .20$, *and* $\omega_4 = .20$.

(a) (Optimization Surface)

(b) (Contour Plot)

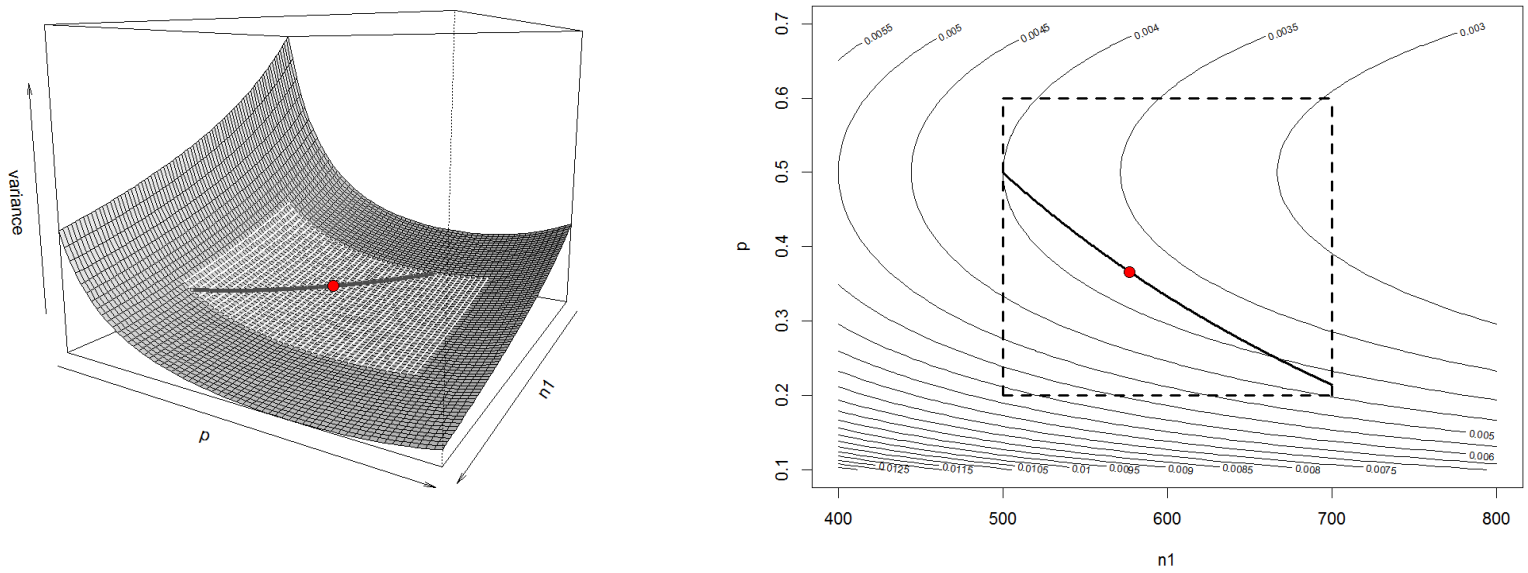

Figure 2C. *(a) demonstrates the optimization surface for p and n_1 when treatment and control units are associated with marginal costs and the primary inequality constraint is on total cost. Both in (a) and (b) the inner rectangles indicates the region restricted by bound constraints. The dark points within the inner rectangle (line) is the feasible region that is restricted by primary equality constraint on the total cost. The larger (red) point within the rectangle indicates the optimal solution. Parameters used in the optimization procedure: costs per treatment and control units are 15 and 5 correspondingly, total cost is fixed at 5000, the correlation between the score variable and the treatment condition is zero ($\rho_{TS} = 0$), proportion of variance in the outcome explained by level 1 covariates are assumed to be zero ($R_1^2 = 0$), and $\alpha = .05$ for a two-tailed hypothesis test.*

(a) (Optimization Surface)

(b) (Contour Plot)

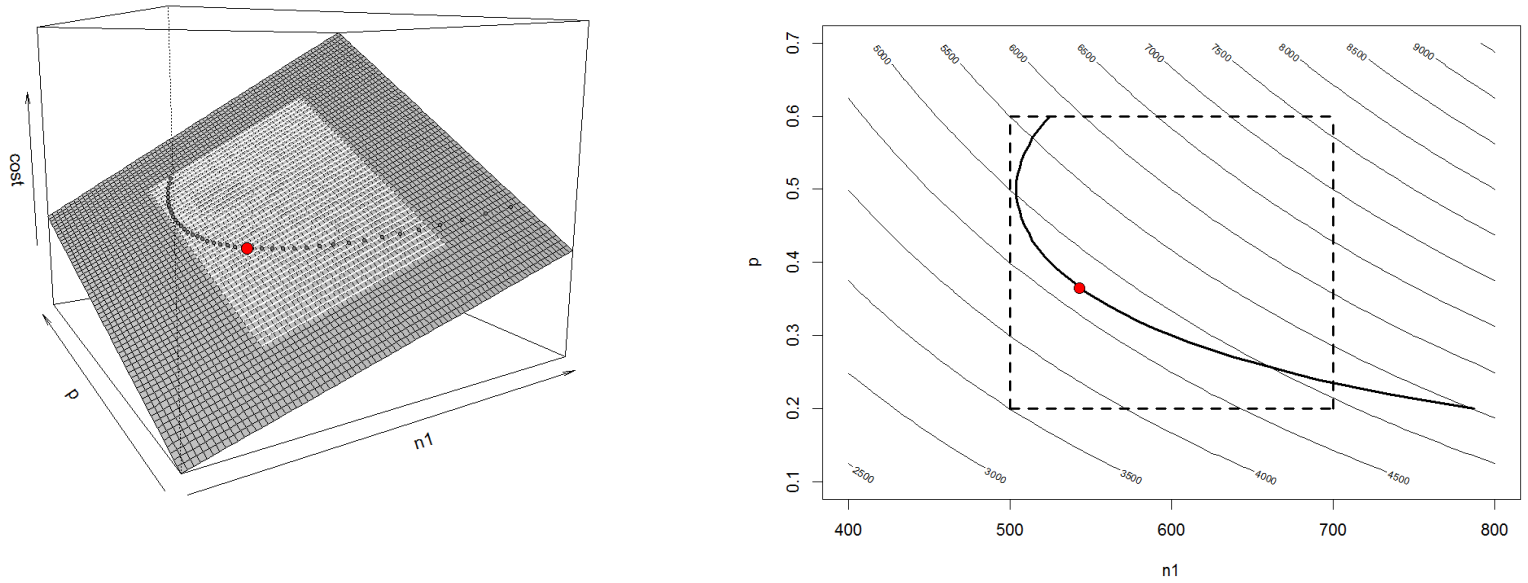

Figure 3C. *(a) demonstrates the optimization surface for p and n_1 when treatment and control units are associated with marginal costs and the primary equality constraint is on the statistical power. Both in (a) and (b) the inner rectangles indicates the region restricted by bound constraints. The dark points within the inner rectangle (line) is the feasible region that is restricted by primary equality constraint on the statistical power. The larger (red) point within the rectangle indicates the optimal solution. Parameters used in the optimization procedure: costs per treatment and control units are 15 and 5 correspondingly, power is fixed at 80%, the correlation between the score variable and the treatment condition is zero ($\rho_{TS} = 0$), proportion of variance in the outcome explained by level 1 covariates are assumed to be zero ($R_1^2 = 0$), and $\alpha = .05$ for a two-tailed hypothesis test.*

(a) **(**Optimization Surface) (b) (Contour Plot)

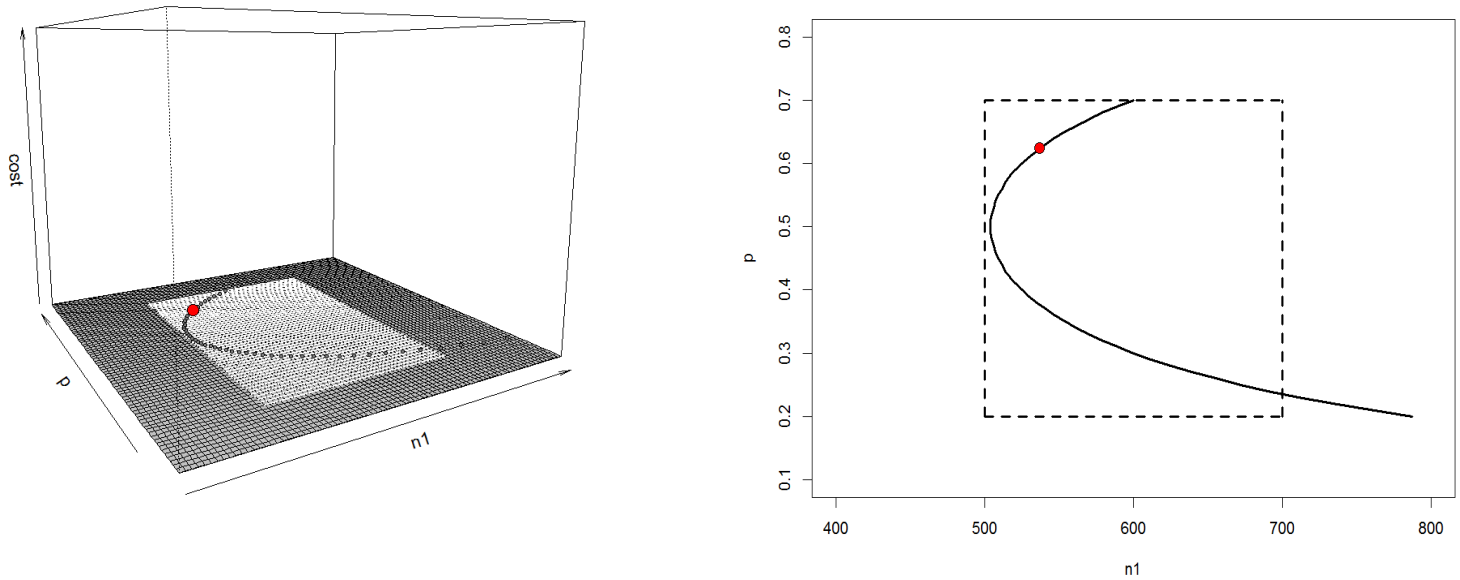

Figure 4C. *(a) demonstrates the optimization surface for p and n_1 when treatment and control units are not associated with marginal costs and the primary equality constraint is on the statistical power. Both in (a) and (b) the inner rectangles indicates the region restricted by bound constraints. The dark points within the inner rectangle (line) is the feasible region that is restricted by primary equality constraint on the statistical power. The larger (red) point within the rectangle indicates the optimal solution. Parameters used in the optimization procedure: there are no cost associated with treatment and control units, power is fixed at 80%, the correlation between the score variable and the treatment condition is zero ($\rho_{TS} = 0$), proportion of variance in the outcome explained by level 1 covariates are assumed to be zero ($R_1^2 = 0$), and $\alpha = .05$ for a two-tailed hypothesis test.*