On equivalencies between design-based and regression-based variance estimators for randomized experiments

Cyrus Samii [a,*], Peter M. Aronow [b]

[a] Department of Politics, New York University, 19 West 4th St. 2nd Floor, New York, NY 10012, United States
[b] Department of Political Science, Yale University, Rosenkranz Hall, 115 Prospect Street, New Haven, CT 06520, United States

A B S T R A C T

This paper demonstrates that the randomization-based "Neyman" and constant-effects estimators for the variance of estimated average treatment effects are equivalent to a variant of the White "heteroskedasticity-robust" estimator and the homoskedastic ordinary least squares (OLS) estimator, respectively.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Design-based methods for analyzing experiments use the known treatment assignment distribution as the basis for inference on causal quantities such as the average treatment effect. Such methods for "randomization inference" have recently been presented as more reasoned alternatives to more commonly used regression-based methods that rely on seemingly arbitrary assumptions about the distribution of outcomes in order to derive properties for estimators. Freedman (2008b); Rosenbaum (2002a,b, 2009) develop this point, often with particular reference to the estimated uncertainty of regression estimators. Regression models have been the standard mode of analysis for experiments. If the assumptions embedded in regression analysis lead to misestimates of the true estimator variance, the reliability of many reported results would be called into question.

At the same time, there may be no cause for alarm, as there is a difference between the manner in which an estimator is derived and the way the estimator operates in practice. Indeed, it is straightforward to demonstrate that regression analysis reproduces common randomization-based estimators. We show that the conservative randomization-based Neyman (1923) variance estimator is exactly reproduced by a variant of the commonly-used White heteroskedasticity-robust variance estimator (Horn et al., 1975; White, 1980; MacKinnon and White, 1985), and that the randomization variance of the difference-in-means under a constant effects assumption is reproduced (subject to a finite-sample scaling coefficient) by the ordinary least squares (OLS) homoskedastic variance estimator. We also show that all four of these estimators are equivalent under a balanced design, suggesting that, in many cases, the choice of variance estimator is irrelevant. The paper presents each of these results in turn. We conclude with a note on the use of regression for the analysis of randomized experiments.

---

* Corresponding author. Tel.: +1 917 301 6421; fax: +1 212 995 4184.
  E-mail addresses: cdsamii@gmail.com (C. Samii), peter.aronow@yale.edu (P.M. Aronow).

2. The setting

Consider an experiment on a finite population U of N units in which $1 < M < N - 1$ units are randomly assigned to treatment and the remaining units are assigned to control. Denote control or treatment status, respectively, for unit j with the random variable $X_j \in \{0, 1\}$ for which $\Pr(X_j = 1) = M/N$. Reassign an index ordering, $i = 1, \ldots, N$, such that those assigned to treatment come first, $X_1, \ldots, X_M = 1$ and those assigned to control come after, $X_{M+1}, \ldots, X_N = 0$. We observe

$$Y_i = X_i y_{1i} + (1 - X_i) y_{0i}, \tag{1}$$

where y_{1i} and y_{0i} are unit i's fixed "potential outcomes" under treatment and control, respectively.

Given a feature v_i for $i \in 1, \ldots, N$, we define the population mean, $\overline{v}$, and variance, $\sigma^2(v)$, as,

$$\overline{v} = \frac{1}{N} \sum_{i=1}^{N} v_i$$

$$\sigma^2(v) = \frac{1}{N} \sum_{i=1}^{N} (v_i - \overline{v})^2.$$

For features v_i and w_i for $i \in 1, \ldots, N$, we define the population covariance, $\sigma(v, w)$ as,

$$\sigma(v, w) = \frac{1}{N} \sum_{i=1}^{N} (v_i - \overline{v})(w_i - \overline{w}).$$

We wish to estimate the average treatment effect, β, for the finite population U,

$$\beta = \overline{y_1} - \overline{y_0} = \frac{1}{N} \sum_{i=1}^{N} (y_{1i} - y_{0i}).$$

It is well known that by random assignment, we can estimate β without bias via the simple difference in treated versus control means,

$$\hat{\beta} = \frac{1}{M} \sum_{i=1}^{M} Y_i - \frac{1}{N - M} \sum_{i=M+1}^{N} Y_i, \tag{2}$$

with $E\left[\hat{\beta}\right] = \beta$, where the $E[\cdot]$ operator averages over all $\binom{N}{M}$ treatment assignments, each being equally likely.

Following Neyman (1923) and Freedman et al. (1998, A32–A34), elementary sampling theory allows us to derive the expression for the exact variance of $\hat{\beta}$ as,

$$V(\hat{\beta}) = \frac{N}{N - 1} \left[\frac{\sigma^2(y_1)}{M} + \frac{\sigma^2(y_0)}{N - M} \right] + \frac{1}{N - 1} \left[2\sigma(y_1, y_0) - \sigma^2(y_1) - \sigma^2(y_0) \right]. \tag{3}$$

The unknown quantities in this expression are $\sigma^2(y_1), \sigma^2(y_0)$ and $\sigma(y_1, y_0)$. With estimators of each of these quantities, we could construct an unbiased estimator of $V(\hat{\beta})$. By Cochran (1977, Theorem 2.4), unbiased estimators of $\sigma^2(y_1)$ and $\sigma^2(y_0)$ are simple to construct:

$$\hat{\sigma}^2_{FP}(y_1) = \frac{N - 1}{N} \frac{1}{M - 1} \left[\sum_{i=1}^{M} \left(Y_i - \frac{1}{M} \sum_{i=1}^{M} Y_i \right)^2 \right]$$

$$= \frac{N - 1}{N} \hat{\sigma}^2(y_1)$$

and

$$\hat{\sigma}^2_{FP}(y_0) = \frac{N - 1}{N} \frac{1}{N - M - 1} \left[\sum_{i=M+1}^{N} \left(Y_i - \frac{1}{N - M} \sum_{i=M+1}^{N} Y_i \right)^2 \right]$$

$$= \frac{N - 1}{N} \hat{\sigma}^2(y_0),$$

where $\hat{\sigma}^2(y_1)$ and $\hat{\sigma}^2(y_0)$ are the traditional sample variance estimators. Since both potential outcomes for the same unit can never be observed at the same time, the term $\sigma(y_0, y_1)$ cannot generally be estimated without bias. However, we can bound this quantity: $2\sigma(y_0, y_1) - \sigma^2(y_0) - \sigma^2(y_1) \leq 0$.[1]

---

[1] This is due to $\sigma^2(y_0) + \sigma^2(y_1) - 2\sqrt{\sigma^2(y_0)}\sqrt{\sigma^2(y_1)} = (\sqrt{\sigma^2(y_0)} - \sqrt{\sigma^2(y_1)})^2 \geq 0$, and $2\sqrt{\sigma^2(y_0)}\sqrt{\sigma^2(y_1)} \geq 2\sigma(y_0, y_1)$ by the Cauchy–Schwartz inequality.

Now, let us turn to re-expressing this setting in regression form. Define

$$
Z = \begin{pmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_N \end{pmatrix} = \begin{pmatrix} Z_1 \\ \vdots \\ Z_N \end{pmatrix} \qquad Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_N \end{pmatrix}.
$$

Then the regression estimator,

$$
(Z'Z)^{-1}Z'Y = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta}_{OLS} \end{pmatrix} = \begin{pmatrix} \dfrac{1}{N-M} \displaystyle\sum_{i=M+1}^{N} Y_i \\ \dfrac{1}{M} \displaystyle\sum_{i=1}^{M} Y_i - \dfrac{1}{N-M} \displaystyle\sum_{i=M+1}^{N} Y_i \end{pmatrix}, \tag{4}
$$

and so $\hat{\beta}_{OLS} = \hat{\beta}$ (Freedman, 2008b, Theorem 1). For future reference, define $\hat{e}_i = Y_i - Z_i(Z'Z)^{-1}Z'Y$, the regression residual.

3. The Neyman conservative randomization-based estimator and the White heteroskedasticity-robust estimator

The Neyman estimator of $V(\hat{\beta})$ substitutes $\hat{\sigma}^2_{FP}(y_1)$ and $\hat{\sigma}^2_{FP}(y_0)$ for $\sigma^2(y_1)$ and $\sigma^2(y_0)$, and assumes that $2\sigma(y_0, y_1) - \sigma^2(y_0) - \sigma^2(y_1) = 0$, yielding,

$$
\hat{V}_N(\hat{\beta}) = \frac{N}{N-1} \left[\frac{\hat{\sigma}^2_{FP}(y_1)}{M} + \frac{\hat{\sigma}^2_{FP}(y_0)}{N-M} \right] = \frac{\hat{\sigma}^2(y_1)}{M} + \frac{\hat{\sigma}^2(y_0)}{N-M}.
$$

Since $\hat{\sigma}^2_{FP}(y_1)$ and $\hat{\sigma}^2_{FP}(y_0)$ are unbiased,

$$
E\left[\hat{V}_N(\hat{\beta}) \right] = \frac{N}{N-1} \left[\frac{\sigma^2(y_1)}{M} + \frac{\sigma^2(y_0)}{N-M} \right].
$$

Subtracting off $V(\hat{\beta})$, as in (3), the bias of $\hat{V}_N(\hat{\beta})$,

$$
E\left[\hat{V}_N(\hat{\beta}) \right] - V(\hat{\beta}) = -\frac{1}{N-1} \left[2\sigma(y_1, y_0) - \sigma^2(y_1) - \sigma^2(y_0) \right] \geq 0.
$$

The Neyman estimator is therefore conservative: its bias is always nonnegative. As demonstrated in Freedman (2008a, Theorem 1), asymptotic normality of $\hat{\beta}$ implies that large-sample confidence intervals, using $\hat{V}_N(\hat{\beta})$ in a normal approximation will have either approximately correct or conservative coverage.

In the traditional regression framework, the $\hat{e}_i$ residual approximates the "error" term of a linear model. For the case of heteroskedastic errors, White (1980) proposed a consistent estimator for the covariance of $(\hat{\alpha} \ \hat{\beta})'$,

$$
\hat{V}_{HR} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = N(Z'Z)^{-1} \left(\frac{1}{N} \sum_{i=1}^{N} Z_i'Z_i \hat{e}_i^2 \right) (Z'Z)^{-1}.
$$

MacKinnon and White (1985) and Angrist and Pischke (2009, 299–308) review three amended versions of $\hat{V}_{HR}$, constituting a class of "heteroskedasticity-robust" variance estimators. Each amendment reduces some form of finite sample bias. Among these, a commonly used version is the "HC2" estimator, in which $\frac{\hat{e}_i^2}{1-h_{ii}}$ is substituted for $\hat{e}_i^2$ in the previous equation, where $h_{ii} = Z_i(Z'Z)^{-1}Z_i'$ is the leverage of observation i. In this case,

$$
h_{ii} = \begin{cases} \dfrac{1}{M} & \text{if } X_i = 1 \\ \dfrac{1}{N-M} & \text{if } X_i = 0. \end{cases}
$$

The HC2 estimator for the variance of the average treatment effect, $\hat{V}_{HC2}(\hat{\beta})$, is the [2, 2] element of the resulting covariance matrix:

$$
\hat{V}_{HC2} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = N(Z'Z)^{-1} \left(\frac{1}{N} \sum_{i=1}^{N} Z_i'Z_i \frac{\hat{e}_i^2}{1-h_{ii}} \right) (Z'Z)^{-1}.
$$

We now present our first result.

Theorem 1. $\hat{V}_{HC2}(\hat{\beta}) = \hat{V}_N(\hat{\beta})$.

Proof. The HC2 estimator for the variance of $\hat{\beta}$ reduces algebraically to,

$$\hat{V}_{HC2}(\hat{\beta}) = \frac{1}{M}\frac{1}{M-1}\sum_{i=1}^{M}\hat{e}_i^2 + \frac{1}{N-M}\frac{1}{N-M-1}\sum_{i=M+1}^{N}\hat{e}_i^2.$$

The regression residual, $\hat{e}_i$, is algebraically equivalent to $Y_i - \frac{1}{M}\sum_{i=1}^{M}Y_i$ if $i \leq M$ (i.e., treated), and $Y_i - \frac{1}{N-M}\sum_{i=M+1}^{N}Y_i$ if $i > M$ (i.e., control). Therefore,

$$\frac{1}{M-1}\sum_{i=1}^{M}\hat{e}_i^2 = \hat{\sigma}^2(y_1) \quad \text{and} \quad \frac{1}{N-M-1}\sum_{i=M+1}^{N}\hat{e}_i^2 = \hat{\sigma}^2(y_0).$$

The HC2 version of the White estimator may be represented as,

$$\hat{V}_{HC2}(\hat{\beta}) = \frac{\hat{\sigma}^2(y_1)}{M} + \frac{\hat{\sigma}^2(y_0)}{N-M}.$$

Therefore, $\hat{V}_{HC2}(\hat{\beta}) = \hat{V}_N(\hat{\beta})$, and the Neyman estimator is equivalent to the White HC2 estimator. □

Corollary 1. $E\left[\hat{V}_{HC2}(\hat{\beta})\right] \geq V(\hat{\beta})$.

Proof. Since $E\left[\hat{V}_N(\hat{\beta})\right] \geq V(\hat{\beta})$ and $\hat{V}_{HC2}(\hat{\beta}) = \hat{V}_N(\hat{\beta})$, $E\left[\hat{V}_{HC2}(\hat{\beta})\right] \geq V(\hat{\beta})$. The HC2 estimator is therefore generally conservative. □

Above, we noted that large-sample confidence intervals based on $\hat{V}_N(\hat{\beta})$ had either approximately correct or conservative coverage. Clearly, this extends to $\hat{V}_{HC2}(\hat{\beta})$.[2]

4. Constant effects randomization variance and homoskedastic OLS estimator

Rosenbaum (2002a,b, 2009) studies randomization-based inference by maintaining constant additive effects hypotheses. Under the maintained hypothesis, the treatment effect is assumed to be the same scalar quantity for all units. Then, inferences about the average treatment effect are based on studying the distribution of estimated treatment effects as one permutes the treatment assignment. The general approach has attracted a fair number of adherents among applied researchers (e.g., Bowers, 2011). Here we show that the variance of the estimated treatment effects that obtains, through such permutation under a maintained hypothesis of constant effects equal to $\hat{\beta}$, reproduces the variance estimate from ordinary least squares (OLS) regression analysis under the homoskedasticity assumption, subject to a slight scaling factor.

By maintaining the hypothesis of constant additive effects equal to $\hat{\beta}$, we can use expressions (1) and (2) to construct a hypothesized set of potential outcomes, $\tilde{y}_{0i}$ and $\tilde{y}_{1i}$, for all $i = 1, \ldots, N$ as follows:

$$\tilde{y}_{0i} = Y_i - X_i\hat{\beta} = Y_i - X_i\left(\frac{1}{M}\sum_{i=1}^{M}Y_i - \frac{1}{N-M}\sum_{i=M+1}^{N}Y_i\right)$$

and

$$\tilde{y}_{1i} = Y_i + (1 - X_i)\hat{\beta} = Y_i + (1 - X_i)\left(\frac{1}{M}\sum_{i=1}^{M}Y_i - \frac{1}{N-M}\sum_{i=M+1}^{N}Y_i\right).$$

By (3), the hypothesis of constant effects implies that the randomization variance for the estimated treatment effect should be estimated as,

$$\hat{V}_{CE}(\hat{\beta}) = \frac{N}{N-1}\left[\frac{\sigma^2(\tilde{y}_1)}{M} + \frac{\sigma^2(\tilde{y}_0)}{N-M}\right] + \frac{1}{N-1}\left[2\sigma(\tilde{y}_1, \tilde{y}_0) - \sigma^2(\tilde{y}_1) - \sigma^2(\tilde{y}_0)\right]. \tag{5}$$

Note that there are no unknown terms in (5): the $\sigma^2(\tilde{y}_1)$, $\sigma^2(\tilde{y}_0)$ and $\sigma(\tilde{y}_1, \tilde{y}_0)$ values in (5) may be computed using the hypothesized potential outcomes that have been defined for all N units.

---

[2] Implications for the other heteroskedasticity-robust estimators (HR, HC1, and HC3) studied in MacKinnon and White (1985) and Angrist and Pischke (2009, 299–308) are straightforward. The other estimators are all asymptotically equivalent to HC2 and, thus, large-sample confidence intervals based on any of the heteroskedasticity-robust estimators have either approximately correct or conservative coverage.

Returning to the regression formulation, under the assumption of homoskedastic errors, the OLS coefficient covariance matrix is given by,

$$\hat{V}_{OLS}\begin{pmatrix}\hat{\alpha}\\\hat{\beta}\end{pmatrix} = \frac{\sum\limits_{i=1}^{N}\hat{e}_i^2}{N-2}(Z'Z)^{-1},$$

with $\hat{V}_{OLS}(\hat{\beta})$ denoting the [2, 2] element of this covariance matrix. Freedman (2008b, Theorem 5) demonstrates that for $N \neq 2M$, $\hat{V}_{OLS}(\hat{\beta})$ converges asymptotically to a limit that may be too big or too small relative to $V(\hat{\beta})$.

We now demonstrate that the constant effects randomization variance is equivalent to a scaled homoskedastic OLS estimator.

Theorem 2. $\frac{N-2}{N-1}\hat{V}_{OLS}(\hat{\beta}) = \hat{V}_{CE}(\hat{\beta})$.

Proof. Since $\tilde{y}_0 = \tilde{y}_1 - \hat{\beta}, \sigma^2(\tilde{y}_0) = \sigma^2(\tilde{y}_1) = \sigma(y_0, y_1)$. Thus, substituting this into expression (5), randomization yields,

$$\hat{V}_{CE}(\hat{\beta}) = \frac{N}{N-1}\frac{N}{M(N-M)}\sigma^2(\tilde{y}_0).$$

Now, taking $\hat{\alpha}$ and $\hat{\beta}$ as fixed,

$$\frac{\sum\limits_{i=1}^{N}\hat{e}_i^2}{N} = \frac{\sum\limits_{i=1}^{N}\left(Y_i - \hat{\alpha} - X_i\hat{\beta}\right)^2}{N} = \frac{\sum\limits_{i=1}^{N}\left(\tilde{y}_{0i} - \hat{\alpha}\right)^2}{N} = \sigma^2(\tilde{y}_0),$$

by (4) and simple algebraic manipulations. Therefore,

$$\begin{aligned}\frac{N-2}{N-1}\hat{V}_{OLS}(\hat{\beta}) &= \frac{N-2}{N-1}\frac{\sum\limits_{i=1}^{N}\hat{e}^2}{N-2}\frac{N}{M(N-M)}\\ &= \frac{N}{N-1}\frac{\sum\limits_{i=1}^{N}\hat{e}^2}{N}\frac{N}{M(N-M)}\\ &= \frac{N}{N-1}\frac{N}{M(N-M)}\sigma^2(\tilde{y}_0) = \hat{V}_{CE}(\hat{\beta}). \quad \square\end{aligned}$$

As a result, the fact noted above that when $N \neq 2M$, $\hat{V}_{OLS}(\hat{\beta})$ may be either too large or too small relative to $V(\hat{\beta})$ carries over to $V_{CE}(\hat{\beta})$.[3] Then, in large samples, confidence intervals based on $\hat{V}_{OLS}(\hat{\beta})$ or $\hat{V}_{CE}(\hat{\beta})$ are not assured to be approximately correct or conservative, in contrast to $\hat{V}_N(\hat{\beta})$ and $\hat{V}_{HC2}(\hat{\beta})$.

Concerns about outliers or other irregularities may motivate the analyst to conduct inference with statistics other than the difference-in-means, such as rank sums. Alternative procedures based on inverted tests and constant effects hypotheses are also possible.[4] The implications of these results for such tests are not entirely clear, as they would likely depend on distributional features of the data at hand.

5. The balanced design

In a balanced design, equal numbers of units are assigned to the treatment and control groups. Calculations are simplified greatly with a balanced design and, accordingly, many results about the properties of variance estimators assume a balanced design (Gadbury, 2001).

Theorem 3. *Under a balanced design such that* $N = 2M$, $\frac{N-1}{N-2}\hat{V}_{CE}(\hat{\beta}) = \hat{V}_{OLS}(\hat{\beta}) = \hat{V}_N(\hat{\beta}) = \hat{V}_{HC2}(\hat{\beta})$.

---

[3] Freedman (2008b, p. 190) gives the exact asymptotic relation between $\hat{V}_{OLS}(\hat{\beta})$ and $V(\hat{\beta})$, which depends on the size of the treatment versus control group and the limiting values of the control group mean and average treatment effect.

[4] Indeed, Rosenbaum (2002a,b, 2009) develops his analysis of randomization inference for the constant effects model using a rank sum test and Hodges–Lehmann confidence intervals.

Proof. With $N = 2M$ and by Theorems 1 and 2,

$$\frac{N-1}{N-2} V_{CE}(\hat{\beta}) = \hat{V}_{OLS}(\hat{\beta}) = \frac{\sum_{i=1}^{N} \hat{e}_i^2}{2(M-1)} \frac{2M}{M^2} = \frac{\frac{\sum_{i=1}^{M} \hat{e}_i^2}{M-1}}{M} + \frac{\frac{\sum_{i=M+1}^{N} \hat{e}_i^2}{M-1}}{M}$$

$$= \frac{\hat{V}(y_1)}{M} + \frac{\hat{V}(y_0)}{M} = \hat{V}_N(\hat{\beta}) = \hat{V}_{HC2}(\hat{\beta}). \quad \square$$

While this result is heartening, it nevertheless represents a special case. For imbalanced designs, the choice of variance estimator is nontrivial.

6. Discussion

The results allow us to qualify Freedman (2008b)'s critique of the use of regression models to analyze randomized experiments. In the critique, Freedman claims that variance estimation based on the "usual" homoskedastic error model is, in general, inconsistent for the randomization variance of the estimated average treatment effect. While technically correct, it is of little practical significance for current practice in econometrics and associated social sciences. Standard estimation practice today, as presented for example in Angrist and Pischke (2009), is to use a heteroskedastic model and some variant of the White (1980) "robust" estimator. Theorem 1 shows that randomization justifies the HC2 estimator and Theorem 2 shows that the randomization-based constant effects estimator is equivalent (minus a scaling factor) to the OLS estimator that Freedman criticizes.

Regression may not be the correct tool for analyzing data from experiments and we do not advocate its use over design-based estimators. In small experiments, covariate adjustment introduces bias that may be severe, and clustered random assignment may render simple difference-in-means and related regression estimators biased (Middleton, 2008). Nonetheless, while randomization may not justify regression assumptions, randomization does justify some modern econometric practice, including heteroskedasticity-robust variance estimation. In addition, design-based estimators that exploit the randomization distribution while eschewing regression assumptions may not be as different from classical regression estimators as may seem at first glance.

Acknowledgments

References

Angrist, J.D., Pischke, J.S., 2009. Mostly Harmless Econometrics. Princeton University Press, Princeton, NJ.
Bowers, J., 2011. Making effects manifest in randomized experiments. In: Druckman, J.N., Green, D.P., Kuklinski, J.H., Lupia, A. (Eds.), Cambridge Handbook of Experimental Political Science. Cambridge University Press, (chapter 32).
Cochran, W.G., 1977. Sampling Techniques, third ed.. John Wiley, New York.
Freedman, D., Pisani, R., Purves, R., 1998. Statistics, third ed.. W.W. Norton, New York.
Freedman, D.A., 2008a. On regression adjustments in experiments with several treatments. The Annals of Applied Statistics 2, 176–196.
Freedman, D.A., 2008b. On regression adjustments to experimental data. Advances in Applied Mathematics 40, 180–193.
Gadbury, G.L., 2001. Randomization inference and the bias of standard errors. The American Statistician 55, 310–313.
Horn, S.D., Horn, R.A., Duncan, D.B., 1975. Estimating heteroskedastic variances in linear models. Journal of the American Statistical Association 70, 380–385.
MacKinnon, J.G., White, H., 1985. Some heteroskedastiticy-consistent covariance matrix estimators with improved finite sample properties. Journal of Econometrics 29, 305–325.
Middleton, J.A., 2008. Bias of the regression estimator for experiments using cluster random assignment. Statistics and Probability Letters 78, 2654–2659.
Neyman, J.S., 1923. On the application of probability theory to agricultural experiments, essay on principles, Section 9, 1990, reprint. Statistical Science 5, pp. 465–472.
Rosenbaum, P.R., 2002a. Covariance adjustment in randomized experiments and observational studies. Statistical Science 17, 286–327.
Rosenbaum, P.R., 2002b. Observational Studies, second ed.. Springer, New York.
Rosenbaum, P.R., 2009. Design of Observational Studies. Springer, New York.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.