Covariate Adjustment in Randomized Experiments

Kosuke Imai
Harvard University

STAT186/GOV2002 CAUSAL INFERENCE

Fall 2019
Motivation

- Treatment randomization $\leadsto$ unbiased causal estimates
- We adjust covariates for improved efficiency *before* randomization via blocking/stratification

In some cases, we cannot perform pre-randomization adjustment

1. post-stratification
2. regression adjustment

What are the statistical properties of such adjustments?

Danger of post-randomization covariate adjustment

1. specification search, *p*-hacking $\leadsto$ pre-registration
2. may lose some of the theoretical properties of randomization inference
Let’s assume the following LSM:

\[ Y_i(t) = \alpha + \beta \cdot t + \gamma^\top X_i + \delta^\top U_i + \epsilon_i(t) \]

where \( \mathbb{E}(\epsilon_i(t) \mid X_i, U_i) = 0 \)

and \( U_i \) is an unobserved pre-treatment variable correlated with \( X_i \)

Decomposition via projection:

\[ U_i = \hat{\lambda} + \hat{\xi}^\top T_i + \hat{\zeta}^\top X_i + \hat{\eta}_i \]

Omitted variable bias formula:

\[ \hat{\beta} \xrightarrow{p} \beta + \delta^\top \xi \]

where \( (\xi, \delta) \) are probability limits of \( (\hat{\xi}, \hat{\delta}) \)

Randomization of \( T_i \) implies \( \xi = 0 \) though \( \delta \neq 0_K \)

Random assignment protects you against bias

Does the result hold if we do not assume the LSM?
Covariate Adjustment via Linear Regression

- Multiple linear regression model with centered covariates:
  \[ Y_i = \alpha + \beta T_i + \gamma^\top \tilde{X}_i + \epsilon_i \quad \text{for } i = 1, \ldots, n \]
  where \( \tilde{X}_i = X_i - \bar{X}_n \)

- Ordinary least squares estimator:
  \[
  (\hat{\alpha}, \hat{\beta}, \hat{\gamma}) = \arg\min_{(\alpha, \beta, \gamma)} \sum_{i=1}^{n} (Y_i - \alpha - \beta T_i - \gamma^\top \tilde{X}_i)^2
  \]
  where the right hand side converges to,
  \[
  \mathbb{E}\{( Y_i - \alpha - \beta T_i)^2\} + \mathbb{E}\{(\gamma^\top \tilde{X}_i)^2\} - 2 \cdot \mathbb{E}\{( Y_i - \alpha - \beta T_i)(\gamma^\top \tilde{X}_i)\}
  \]
  \[
  \mathbb{E}( Y_i \cdot \gamma^\top \tilde{X}_i)
  \]
  Thus, \((\hat{\alpha}, \hat{\beta})\) converge to \((\mathbb{E}\{ Y_i(0)\}, \tau = \mathbb{E}\{ Y_i(1) - Y_i(0)\})\)

- OLS estimator is consistent even if the model is incorrect
Efficiency Gain from Regression

- **Asymptotic variance** (Imbens and Rubin, Theorem 7.1):

  \[ \sqrt{n}(\hat{\beta} - \tau) \rightsquigarrow N \left( 0, \frac{\mathbb{E}\{(T_i - k)^2(Y_i - \alpha^* - \beta^* T_i - \gamma^* \tilde{X}_i)^2\}}{k^2(1 - k)^2} \right) \]

  where \( * \) represents a limiting value and \( k = n_1/n \)

- Consistently estimated by the heteroskedasticity-robust variance estimator regardless of model misspecification

- If the linear model is correct,

  \[ \mathbb{V}(\hat{\beta}) \approx \frac{\mathbb{E}\{\mathbb{V}(Y_i(1) | \tilde{X})\}}{n_1} + \frac{\mathbb{E}\{\mathbb{V}(Y_i(0) | \tilde{X})\}}{n_0} \leq \frac{\sigma_1^2}{n_1} + \frac{\sigma_0^2}{n_0} \]

- If the linear model is incorrect, the efficiency might be lost

Regression with Interactions

- Fully interacted model:
  \[ Y_i = \alpha + \beta T_i + \gamma^\top \tilde{X}_i + \delta^\top T_i \tilde{X}_i + \epsilon_i \]
  same as fitting a separate model to each group

- **Imputation** interpretation:
  \[ \hat{\beta} = \frac{1}{n} \sum_{i=1}^{n} \{ T_i (Y_i - \hat{Y}_i(0)) + (1 - T_i)(\hat{Y}_i(1) - Y_i) \} \]
  which is different from
  \[ \frac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i(1) - \hat{Y}_i(0)) \]
  \[ \hat{\beta} \] is consistent for the PATE and asymptotically normal
  Unlike the no interaction model, \( \hat{\beta} \) is at least as efficient as the difference-in-means estimator (Linn. 2013. *Ann. Appl. Stat.*)
Reducing Transphobia (Brookman and Kalla. 2016. Science)

- Can perspective-taking – “imagining the world from another’s perspective” – reduce transphobia?

- Treatment group ($n_1 = 912$): canvasser asking people to think about a time when they were judged unfairly and were guided to translate that experience to a transgender individual’s experience

- Control group ($n_0 = 913$): canvasser asking people to recycle

- Outcome variable: feeling thermometer towards transgender people (0 – 100)
  1. baseline
  2. 3 day
  3. 3 week
  4. 6 week
  5. 3 month
Durable Effects

- 3 month effect
- Dropping additional observations with missing covariates
  \[ n_1 = 266 \text{ and } n_0 = 274 \]

- Neyman: est. = 7.99, s.e. = 2.38, 95% CI = [3.33, 12.65]
- Simple Regression: est. = 7.99, s.e. = 2.38 (HC2)

- Regression with covariates (age, gender, party ID, baseline thermometer rating)
  - no interaction: est. = 4.28, s.e. = 1.65 (HC2)
  - with interaction: est. = 4.28, s.e. = 1.65 (HC2)
Regression Adjustment in Stratified Designs

- Neyman’s analysis of stratified designs \(\leadsto\) aggregate within-strata estimates across strata
- Linear regression with fixed effects does within-strata comparison

\[
Y_{ij} = \alpha_j + \beta T_{ij} + \epsilon_{ij}
\]

\[
\hat{\beta} \xrightarrow{p} \frac{\sum_{j=1}^{J} w_j \cdot k_j (1 - k_j) \mathbb{E}(Y_{ij}(1) - Y_{ij}(0))}{\sum_{j=1}^{J} w_j \cdot k_j (1 - k_j)} \neq \tau
\]

where \(w_j = n_j / n\)

- \(\hat{\beta}\) is consistent for PATE if:
  1. treatment assignment probability is identical across strata
  2. average treatment effect is identical across strata

- Weighted fixed effects with \(n_j/n_{j1} = 1/k_j\) as regression weights equals the Neyman’s estimator, removing the bias (Imai and Kim. 2019. *Am. J. Political Sci*)
STAR Project Revisited

- Pre-randomization stratification by schools
- Covariates to be considered: gender, race, birth year

Effect of kindergarten class size on high school graduation:
1. Neyman: est. = 0.018, s.e. = 0.017
2. Linear regression with strata fixed effects
   - No covariate: est. = 0.004, s.e. = 0.016 (HC2)
   - With covariates (gender, birth year, race):
     est. = -0.005, s.e. = 0.016 (HC2)

Effect of kindergarten class size on 8th grade reading score
1. Neyman: est. = 2.76, se. = 1.73
2. Linear regression with strata fixed effects
   - No covariate: est. = 2.58, s.e. = 1.72 (HC2)
   - With covariates: est. = 2.45, s.e. = 1.71 (HC2)
Matched-Pair Designs

- Linear regression with pair fixed effects

\[ Y_{ij} = \alpha_j + \beta T_{ij} + \epsilon_{ij} \quad \text{where} \quad T_1 + T_2 = 1 \quad \text{for all} \quad j \]

- identical to the average pairwise difference estimator
- homoskedastic (or HC2) variance estimator is also identical
- covariates can be added, equivalent to the first-difference model

\[
(Y_{1j} - Y_{2j}) = \beta (T_{1j} - T_{2j}) + \gamma^\top (X_{1j} - X_{2j}) + (\epsilon_{1j} - \epsilon_{2j})
\]

An alternative model:

\[
(Y_{1j} - Y_{2j}) = \beta W_j + \gamma^\top (X_{1j} + X_{2j}) + \eta_j
\]

- Both differencing models yield conservative variances but they are smaller than the standard variance (Fogarty. 2018. *Biometrika*)
Seguro Popular Revisited

- Outcome: health cluster-level Seguro Popular enrollment rates
- Cluster-level covariates: average assets, education, urban/rural

- Neyman: est. = 0.374, s.e. = 0.036

- Linear regression with fixed effects:
  - No covariates: est. = 0.374, s.e. = 0.036
  - With covariates: est. = 0.372, s.e. = 0.035

- First difference regression:
  - No covariates: est. = 0.374, s.e. = 0.036
  - With covariates: est. = 0.390, s.e. = 0.035
Summary

- Covariate adjustment can be used to improve efficiency in randomized experiments
- Under various experimental designs, linear regression models are useful methods for this purpose
- Randomization of treatment assignment protects researchers from misspecification
  - independence between treatment and covariates
  - linear regression estimators are often consistent even when the model is incorrect
- Danger of \( p \)-hacking \( \leadsto \) pre-randomization blocking is preferred whenever possible

Readings:
- Imbens and Rubin, Chapters 7, 9, and 10
- Angrist and Pischke, Chapters 3 and 8