Abstract

This paper provides a theoretical base for further simulation study to examine the effectiveness of post-hoc adjustment approaches such as propensity score matching in reducing the selection bias of synthetic cohort design (SCD) for casual inference and program evaluation. Compared with the Solomon Four Groups Design (SFGD), the SCD often faces selection bias due to the unbalance of covariates between the two cohorts. The efficiency of SCD is assured by the historical equivalence of groups (HEoG) assumption indicating the comparability between the two cohorts. Structural equation modeling (SEM) framework is used to define the HEoG assumption. The mathematical proof shows that HEoG assures that the use of SCD results an unbiased estimate of schooling effect.
Structural Equation Modeling Based Quasi-Experimental Synthetic Cohort Design

Background

The Synthetic Cohort Design (SCD) was first proposed and used for cross-national comparison of schooling (Wiley and Wolfe, 1992) in the Third International Mathematics and Science Study (TIMSS). Academic Growth is determined by comparing data of adjacent grades. Two Cohorts, such as 7th and 8th graders, measured at the same time point. The contrast between the two cohorts, a quasi-longitudinal growth, is a sufficient measure of the schooling effect under the historical equivalence of groups (HEoG) assumption. This assumption asserts that students in adjacent grades are similar except for the additional year of schooling. This paper contrasts SCD with an optimal longitudinal design - the Solomon Four Group Design (SFGD, Solomon, 1949; Campbell and Stanley, 1963) in order to further delineate and to identify the strengths and weaknesses of the SCD. We begin with a brief introduction of the two designs, then we frame the design in the context of Structural Equations Model so that we can compare and contrast the two designs.

Solomon Four Group Design (Solomon, 1949)

In the SFGD, participants are randomly assigned to one of four different groups (See Figure 1). Treatment "T" can be a particular instructional method. The dependent variable is measured on "O's", administered as pre-test (Time 0, before "T") and post-test (Time 1, after "T"). The SFGD investigates if changes on the dependent variable are due to some interaction between the pre-test effect (τ) and the treatment effect (δ). Randomization, however, is a very powerful requirement because all initial differences among the groups are attributed to sampling variation. Furthermore, randomization also renders the interaction effects to be indistinguishable from the main effects. For example, the interaction between prior learning (α) and τ is confounded with τ, in the sense that τ and α × τ cannot be estimated separately (Solomon, p148).

Synthetic Cohort Design (Wiley and Wolfe, 1992)

Figure 2 depicts the time line of the SCD. Four possible sets of data, two cohorts at two time points, can be collected. But the SCD intends to collect data at only time 1. The hypothetical data at Time 0 are important to the comparison of the SCD with the SFGD. Putting Figure 2 into the context of TIMSS, Cohort 1 corresponds to 7th graders, and Cohort 2 corresponds to 8th graders. Data are collected at only Time 1. The SCD investigates δ_{C2T0—C2T1}, which is the effect of 8th grade instruction on student learning. δ now is the schooling effect. The subscript C2T0—C2T1 indicates that the schooling effect is the academic growth of Cohort 2 from Time 0 to Time 1.

An estimate of δ_{C2T0—C2T1} through the SCD is δ_{C2T1—C1T1}. δ_{C2T1—C1T1} represents what Cohort 1 at Time 1 would learn if they go through the schooling system that students of Cohort 2 at Time 0 had gone through. Under the HEoG assumption, Cohort 1 at Time 1 (Cohort 1 at 7th grade) are comparable with Cohort 2 at Time 0 (Cohort 2 at 7th
Therefore, $\delta_{C_2T1-C_1T1}$ produces an unbiased estimate of $\delta_{C_2T0-C_2T1}$ under the HEoG assumption. That is, mathematically $HEoG \Rightarrow \delta_{C_2T1-C_1T1} = \delta_{C_2T0-C_2T1}$ holds.

**Practical Importance**

The SCD is cost effective because it involves a two-cohorts-one-time-point design for group comparison. Using this design, researchers collect data at one time point such as Time 1, which reduces time and cost as in the case of a longitudinal study. However, due to the impossibility of random assignment, the validity of the inference using the SCD requires the HEoG assumption. Quantifying HEoG allows the examination of how selection bias (Heckman, 1979) is reduced so that the SCD can be a valid and convenient design for researches in evaluating program effectiveness.

**Theoretical Framework**

The HEoG assumption is mathematically defined through the structural equation modeling framework. Structural equation modeling (SEM, Bollen, 1989) takes into account measurement errors and reduce the estimation bias. SEM depicts the measurement relationship between the surrogate variables and their latent variables, whose relationships are captured by the structural model (Jöreskog, & Sörbom, 1996).

**SFGD in SEM Framework**

The SFGD includes two Experimental groups (denoted as $E_1, E_2$) and two Control groups (denoted as $C_1, C_2$). It also involves two testing point, pre-test at Time 0 (denoted as 0) and post-test at Time 1 (denoted as 1). Using randomization, the SFGD assumes that the four groups are comparable (Solomon, 1949). Table 1 displays the SEM equations of the SFGD. Each group involves two measurement models and one structural model capturing the latent growth relationship of pre-test and post-test. Let $Y_i^t$ be outcome variable in group $i$ at time $t$, with $i = E_1, C_1, E_2, C_2; t = 0, 1$. Let $\eta_0$ and $\eta_1$ represent latent mathematics proficiency at pre-test and post-test time points, respectively. $\eta_0$ is measured by $k_0$ surrogate variables, which are denoted in vector $Y_0 = [Y_1, Y_2, \cdots, Y_{k_0}]$. $\eta_1$ is measured by $k_1$ surrogate variables, which are denoted in vector $Y_1 = [Y_1', Y_2', \cdots, Y_{k_1}']$.

The measurement equation for Experimental Group at pre-test time is

$$Y_0^{E_1} = \delta_0 + \Lambda_0 \eta_0 + \varepsilon_0. \tag{1}$$

The measurement equation for Experimental Group at post-test time is

$$Y_1^{E_1} = (\delta_0 + \delta) + \Lambda_1 \eta_1 + \varepsilon_1. \tag{2}$$

The extra term $\delta$ in the intercept of the post-test measurement model indicates the intervention effect. The structural equation,

$$\eta_1 = \tau + \gamma + \beta \eta_0. \tag{3}$$

It reveals the latent mathematics proficiency growth between two time points. $\gamma$ and $\tau$ indicate the maturation effect and learning effect due to taking pre-test, respectively. The
latent growth rate, namely the acceleration effect of intervention, is captured by the slope \( \beta \). \( \beta \) is unity indicating a "flat" effect in Control Group 1 due to the lack of intervention. Group 2, Experimental and Control, only observes post-test data. Pre-test measurement model and structural model are not observable and are displayed in the text-boxes. Without pre-test in Experimental Group 2, \( \tau \) is zero in the implicit structural model. Further, \( \beta \) is unity in Control Group 2. Covariates are not explicit in the model, because their effects are evened out due to randomization.

**Extended Solomon Four Groups Design in SEM Framework**

The SFGD is extended by including covariates \( X \) in the SEM framework. After including covariates \( X \) in the SFGD, the general measurement models\(^1\) are:

\[
\begin{align*}
Y &= \delta + \lambda \eta + e_1; \\
X &= v + g \xi + e_2.
\end{align*}
\]

\( e_1 \sim N(0, \Theta_{e_1}^T) \) is independent of \( \eta, \xi \) and \( e_2, e_2 \sim N(0, \Theta_{e_2}^T) \) is independent of \( \eta, \xi \) and \( e_1 \). The structural model in LISREL notation (Jöreskog & Sörbom, 1996) is

\[
\eta = A + B \xi + U. \quad (5)
\]

\( U \sim N(0, \Theta_E^T) \) is independent of \( \xi, e_1 \) and \( e_2 \). Intercept \( A \) is generally set zero for the purpose of model identification (Lee, 2007). Table 2 displays the models for both pre-test at Time 0 and post-test at Time 1. The two measurement models of \( Y \) are:

\[
\begin{pmatrix}
Y_0 \\
Y_1
\end{pmatrix} =
\begin{pmatrix}
\delta_0 \\
\delta_1
\end{pmatrix} +
\begin{pmatrix}
\lambda_0 & 0 \\
0 & \lambda_1
\end{pmatrix}
\begin{pmatrix}
\eta_0 \\
\eta_1
\end{pmatrix} +
\begin{pmatrix}
e_{10} \\
e_{11}
\end{pmatrix} , \quad (6)
\]

with \( \delta_1 = \delta_0 + \delta \). \( \delta \) represents the intervention effect. The structural model is

\[
\eta_1 = \tau + \gamma + \beta \eta_0, \quad (7)
\]

whose parameters are defined above.

Similarly, the two measurement models of covariates \( X \) are

\[
\begin{pmatrix}
X_0 \\
X_1
\end{pmatrix} =
\begin{pmatrix}
v_0 \\
v_1
\end{pmatrix} +
\begin{pmatrix}
g_0 & 0 \\
0 & g_1
\end{pmatrix}
\begin{pmatrix}
\xi_0 \\
\xi_1
\end{pmatrix} +
\begin{pmatrix}
e_{20} \\
e_{21}
\end{pmatrix} . \quad (8)
\]

The relationship between \( \xi_1 \) and \( \xi_0 \) is captured by a structural model, \( \xi_1 = \alpha + \pi \xi_0 \). When \( \alpha=0 \) and \( \pi=1 \), the covariates are invariant across two time points.

Further, the relationship of the latent variables of \( X \) and \( Y \) is revealed in the structural model:

\[
\begin{pmatrix}
\eta_0 \\
\eta_1
\end{pmatrix} =
\begin{pmatrix}
A_0 & A_1 \\
B_0 & B_1
\end{pmatrix}
\begin{pmatrix}
\xi_0 \\
\xi_1
\end{pmatrix} +
\begin{pmatrix}
U_0 \\
U_1
\end{pmatrix} . \quad (9)
\]

--- Table 2 is about here ---

\(^1\)For the purpose of simplicity, the superscripts, the group indices, are dropped. However, Table 2 clearly displays each group in a separate row. Adding subscripts may be redundant.
The extended SFGD in SEM framework by nature is a two-level factor analysis model (Muthén, 1994) because of the hierarchically structured data collected from schooling system. Its covariance can be decomposed into within-cluster (denoted as $W$) and between-cluster (detonated as $B$) components (Muthén, 1994). Appendix C has the the variance-covariance of the extended SFGD’s Experimental Group 1. The other three variance-covariance matrices and detailed procedures that derive the matrices will be available in the final paper.

**Mathematical Definition of HEoG**

In the extend-SFGD, data at Time 0 are not collected from Group 2, Experimental and Control. Time-0-SEM is

\[
\begin{align*}
Y_0 &= \delta_0 + \lambda_0 \eta_0 + \epsilon_{10} \\
X_0 &= \nu_0 + g_0 \xi_0 + \epsilon_{20} \\
\eta_0 &= A_0 + B_0 \xi_0 + U_0
\end{align*}
\]

This model is implicit and intangible. Group 2 produces an unbiased intervention effect estimate in the counterfactual sense because one needs to assume that Time-0-SEM implicitly holds equivalently in Group 1 and Group 2. This assumption assures that intervention effect estimate derived from Group 2 is unbiased and equivalent to that derived from Group 1. Detailed proof will be available in the final paper. This defines the SEM-version of the HEoG assumption of the SCD (see Figure 2). That is, **Time-0-SEM holds equally at two 7th grades in Year$_{i-1}$ and Year$_i$**.

**Discussion and Implications**

Mathematically defining the HEoG assumption based on SEM framework provides a way of manipulating SEM model parameters to generate non-comparable cohorts to examine how ad hoc procedures such as matching (Cochran & Rubin, 1973) improves cohort comparability to assure HEoG of the SCD.

**Use Matching To Assure HEoG Assumption.** Using the SCD along with the HEoG assumption involves data sets at three time-cohort knots, $C_2T0$, $C_1T1$ and $C_2T1$. $CiTj$ indicates the knot of cohort i at time j. This correspondingly results in two types of matching: $C2T0 - C1T1Matching$ and $C2T1 - C1T1Matching$. However, the former is only applicable in the case of simulation study. If the covariates are hypothetically unchanged across two time points, two types of matching will be equivalent.

**Parameter Manipulation in Generating Data for Matching.** Conceptually, the goal is to simulate the SCD by generating Cohort 1’s Time 1 data that are non-comparable with Cohort 2 at Time 0 based on the two-level SEM discussed above. Propensity score matching (Rosenbaum & Rubin, 1983) can be used to ‘assure’ a conditional group equivalence and examine how matching can reduce the “simulated selection bias” to improve schooling effect estimation. The number of possibly simulated parameter combinations can be extremely large. In order to simplify the simulation, appropriate assumption is adapted. For example, the measurement invariance between Cohort 1 at Time 0 and Cohort 2 at Time 1, implies that factorial loadings are invariant and the residuals are identically distributed (Cheung and Rensvold, 2002).
Identify Non-comparability of Hierarchically Structured Settings. The simulated hierarchical structure of the data can capture the multilevel nature of the SCD in educational settings. Three situations can be considered: 1) Non-comparability occurs only on level-1 covariates and level two covariates are identical, e.g., the two 7th grades are located in a school and have been taught by same teachers; 2) Level-2 covariates are non-comparable, e.g., in the clusters randomized design, clusters such as classes are the intervention units; 3) Non-comparability occurs at both levels, e.g., schools are sampling units and students are intervention units. The three situations represent different sources of bias, which need different matching approaches such as individual level matching, cluster level matching, and dual matching.

Conclusions and Limitations

Synthetic Cohort Design as a Special Case of Solomon Four Groups Design. The structural equation modeling (SEM) framework indicates that the SCD is a quasi-longitudinal design, which is equivalent the Group 2 of the SFGD. The SCD requires the historical equivalent groups (HEoG) assumption, in order to assure the comparability of the two 7th grades across Time 0 and Time 1.

SCD Estimates Only Schooling Effect $\delta_T$. The SCD practically allows the estimation of only $\delta_T$, the schooling effect due to one year of schooling. $\delta_C$ is not estimable because “control”, without one year of schooling at 8th grade in Year$_i$, is not applicable in educational settings. Thus, the SCD cannot estimate the true gain, which is the difference between $\delta_T$ and $\delta_C$. The estimator of $\delta_T$ in the SCD, denoted as $\delta_{C2T1-C1T1}$.

HEoG Assures the Unbiasedness of $\delta_{C2T1-C1T1}$. The accuracy of using SCD is determined by how comparable the two 7th grades are across two adjacent years. If the two 7th grades are comparable across two time points is assured by the HEoG assumption, which works in a counterfactual sense. This can be mathematically written as $(HEoG|\text{counterfactual}) \Rightarrow \delta_{C2T1-C1T1} = \delta_T$. If randomization is applicable in the SCD, it can assure the HEoG assumption. Mathematically, it is $(HEoG|\text{randomization}) \Rightarrow \delta_{C2T1-C1T1} = \delta_T$.

Matching Assures HEoG Assumption. In educational settings, randomization is not applicable and it cannot assure the HEoG in the SCD. Theoretically, matching assures HEoG in the SCD. That is, $(HEoG|\text{matching}) \Rightarrow \delta_{C2T1-C1T1} = \delta_T$. This asserts that under the matching condition, HEoG assumption holds and assures that SCD approximates a longitudinal study in terms of estimating the effect of one year schooling. However, matching only reduce bias in practice, but it cannot completely remove bias.
Appendix A:

References


Figure 1. The Solomon Four Group Design: “R” represents group randomization, “T” treatment, and “O” assessment. Besides randomization, matching was proposed as the other approach to create comparable groups in Solomon (1949).
**Figure 2.** Longitudinal vs. Quasi-Longitudinal Comparison.
Table 1: Solomon Four Groups Design in Structural Equation Modeling Framework

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experimental</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention Effect: Slope $\beta$</td>
<td>$Y_{0}^{E_1} = \delta_0 + \Lambda_0 \eta_0 + \varepsilon_0$</td>
<td>$Y_{0}^{E_2} = \delta_0 + \Lambda_0 \eta_0 + \varepsilon_0$</td>
</tr>
<tr>
<td>Maturation Effect: Intercept $\gamma$</td>
<td>$Y_{1}^{E_1} = (\delta_0 + \delta) + \Lambda_1 \eta_1 + \varepsilon_1$</td>
<td>$Y_{1}^{E_2} = (\delta_0 + \delta) + \Lambda_1 \eta_1 + \varepsilon_1$</td>
</tr>
<tr>
<td>Pre-test Effect: $\tau$</td>
<td>$\eta_1 = \tau + \gamma + \beta \eta_0$</td>
<td>$\eta_1 = \tau + \beta \eta_0$</td>
</tr>
<tr>
<td><strong>Control</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maturation Effect: $\gamma$</td>
<td>$Y_{0}^{C_2} = \delta_0 + \Lambda_0 \eta_0 + \varepsilon_0$</td>
<td>$Y_{0}^{C_2} = \delta_0 + \Lambda_0 \eta_0 + \varepsilon_0$</td>
</tr>
<tr>
<td>Pre-test Effect: $\tau$</td>
<td>$Y_{1}^{C_2} = \delta_0 + \Lambda_1 \eta_1 + \varepsilon_1$</td>
<td>$Y_{1}^{C_2} = \delta_0 + \Lambda_1 \eta_1 + \varepsilon_1$</td>
</tr>
<tr>
<td></td>
<td>$\eta_1 = \tau + \gamma + \eta_0$</td>
<td>$\eta_1 = \tau + \eta_0$</td>
</tr>
</tbody>
</table>
Table 2: SEMs of the Extended Solomon Four Group Design and Covariance Matrices

<table>
<thead>
<tr>
<th>Extended Solomon Four Group Design</th>
<th>SEMs and Constraints</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Experimental group 1:</strong></td>
<td></td>
<td>Appendix C</td>
</tr>
<tr>
<td>Time 0 : Pre-test</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Treatment = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₀ Y₀</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 1 : Post-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X₁ Y₁</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constraints on Experimental group 1’s Model:</td>
<td>Zero treatment effect: $\delta_1 = \delta_0 + \delta$ and $\delta=0$ ;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unity acceleration effect: Slope $\beta=1$.</td>
<td></td>
</tr>
<tr>
<td><strong>Control group 1:</strong></td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Time 0 : Pre-test</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Treatment = 0</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>X₀ Y₀</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Time 1 : Post-test</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Treatment = 0</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>X₁ Y₁</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Constraints on Experimental group 1’s Model:</td>
<td>Zero treatment effect: $\delta_1 = \delta_0 + \delta$ and $\delta=0$ ;</td>
<td>Final Paper</td>
</tr>
<tr>
<td></td>
<td>No pre-test effect: $\tau=0$.</td>
<td>Final Paper</td>
</tr>
<tr>
<td><strong>Experimental group 2:</strong></td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Time 0 : Pre-test</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Treatment = 1</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>N/A</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Time 1 : Post-test</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Treatment = 1</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>X₁ Y₁</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Constraints on Experimental group 1’s Model:</td>
<td>Zero treatment effect: $\delta_1 = \delta_0 + \delta$ and $\delta=0$ ;</td>
<td>Final Paper</td>
</tr>
<tr>
<td></td>
<td>No pre-test effect: $\tau=0$.</td>
<td>Final Paper</td>
</tr>
<tr>
<td></td>
<td>Unity acceleration effect: Slope $\beta=1$.</td>
<td>Final Paper</td>
</tr>
<tr>
<td><strong>Control group 2:</strong></td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Time 1 : Pre-test</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Treatment = 0</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>N/A</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Time 1 : Post-test</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Treatment = 0</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>X₁ Y₁</td>
<td></td>
<td>Final Paper</td>
</tr>
<tr>
<td>Constraints on Experimental group 1’s Model:</td>
<td>Zero treatment effect: $\delta_1 = \delta_0 + \delta$ and $\delta=0$ ;</td>
<td>Final Paper</td>
</tr>
<tr>
<td></td>
<td>No pre-test effect: $\tau=0$.</td>
<td>Final Paper</td>
</tr>
<tr>
<td></td>
<td>Unity acceleration effect: Slope $\beta=1$.</td>
<td>Final Paper</td>
</tr>
</tbody>
</table>
Appendix C: Variance-Covariance Matrix of SFGD Experimental Group 1: Pre-test = 0, Post-test = 1

\[ Cov(Y, X) = \begin{bmatrix}
\text{Var}(Y_0, Y) & \text{Cov}(Y_0, X) \\
\text{Cov}(Y_0, X) & \text{Var}(X_0, X)
\end{bmatrix} = 
\begin{bmatrix}
\lambda_0 & 0 & 0 & 0 \\
0 & \lambda_0 & 0 & 0 \\
0 & 0 & \lambda_0 & 0 \\
0 & 0 & 0 & \lambda_0
\end{bmatrix}
\]