

Effects of Kindergarten Retention on Children's Social-Emotional Development:
An Application of Propensity Score Method to Multivariate Multi-Level Data

Guanglei Hong

Bing Yu

Ontario Institute for Studies in Education of the University of Toronto

Appendix A

Propensity Score Estimation

Here we provide a brief guide for analyzing a two-level hierarchical logistic regression model as specified in Equation (1). In HLM6.0 Window's interface, after specifying the treatment indicator as the outcome variable in the model specification window, open the "Basic Model Specifications" drop-down list and select "Bernoulli (0 or 1)" for the binary outcome. We grand mean center all the predictors of retention at both level 1 and level 2. Choose "Level-1 Residual File" in the "Basic Model Specifications" dialogue window. A variable in the level-1 residual file labeled "fitval" (short-hand for the fitted value for each level-1 unit) will be the estimated logit of propensity score.

Appendix B

Propensity Score Stratification

Let Z denote the treatment indicator. Let $logit_q$ denote a child's estimated logit of propensity score. We use the minimum value of $logit_q$ of the retained group as a cut-off point for identifying children at risk of kindergarten retention. A promoted child whose estimated logit of propensity is below this cut-off value is considered to be not at risk of retention. We then use the 20th, 40th, 60th, and 80th percentiles of $logit_q$ to divide the rest of the sample:

```
RANK  
VARIABLES=  $logit\_q$  (A) /RANK /NTILES (5) /PRINT=YES  
/TIES=MEAN .
```

This procedure will save in the data file a categorical indicator for propensity strata labeled $Nlogit_q$. If the distributions of $logit_q$ are significantly different between the retained group and the promoted group within a stratum, we will subdivide the stratum till we find balance between the two treatment groups in all the strata.

In checking within-stratum balance between the retained group and the promoted group on all the observed pretreatment covariates, we analyze a general linear model for a continuous measure and a logistic or multinomial logistic regression model for a categorical measure. To be brief, suppose that we end up with having five strata as represented by $Nlogit_q$. The SPSS syntax for checking a continuous pretreatment covariate X is:

```
UNIANOVA  
   $X$  BY  $Z$   $Nlogit\_q$   
  /METHOD = SSTYPE(3)  
  /INTERCEPT = INCLUDE  
  /CRITERIA = ALPHA(.05)  
  /DESIGN =  $Z$   $Nlogit\_q$   $Z * Nlogit\_q$  .
```

Let *stratum1*, *stratum2*, *stratum3*, and *stratum4* represent dummy indicators for four of the five propensity strata. Let *Z_stratum1*, *Z_stratum2*, *Z_stratum3*, and *Z_stratum4* represent the respective interactions between *Z* and each of the four stratum indicators.

The SPSS syntax for checking a binary pretreatment covariate is:

```
LOGISTIC REGRESSION VARIABLES X
/METHOD = ENTER Z stratum1 stratum2 stratum3 stratum4
                Z_stratum1 Z_stratum2 Z_stratum3 Z_stratum4
/CONTRAST (Z)=Indicator
/CONTRAST (stratum1)=Indicator /CONTRAST (stratum2)=Indicator
/CONTRAST (stratum3)=Indicator /CONTRAST (stratum4)=Indicator
/CONTRAST (Z_stratum1)=Indicator /CONTRAST (Z_stratum2)=Indicator
/CONTRAST (Z_stratum3)=Indicator /CONTRAST (Z_stratum4)=Indicator
/CRITERIA = PIN(.05) POUT(.10) ITERATE(20) CUT(.5) .
```

Appendix C

Pretreatment True Score Estimation

In preparing the level-1 file for analyzing equation (2), we first stack the two pretreatment measures under the same variable name. We then create two dummy indicators, one for each of these pretreatment measures. After using the reliability information to compute the error variance for each pretreatment measure, we create a new level-1 variable, *errvar*, to represent the error variance. When we specify our model in HLM, we select the “Estimation Settings” option from the “Other Settings” menu, and set *errvar* as “known variance” under the “Weighting” option.

In estimating the pretreatment true scores for teacher rating and parent rating of child interpersonal relationships, we grand mean center the treatment indicator and all the propensity strata indicators at level 2. In the level-2 residual file, a variable labeled “ecind1” corresponding to π_{X1ij} is the estimated true score of pretreatment teacher rating, and another variable labeled “ecind2” corresponding to π_{X2ij} is the estimated true score of pretreatment parent rating.