

Using Predictive Models to Assign Treatment Groups for NTPS 2017-18 Teacher Incentives Experiment

Allison Zotti

Center for Optimization and Data Science (CODS)

US Census Bureau

National Teacher and Principal Survey (NTPS)

- National cross sectional survey of public and private schools
- Sponsor: National Center for Education Statistics (NCES)
- Multi-level survey
 - School and principal questionnaires (school level)
 - Teacher listing form - TLF (school level)
 - Teacher questionnaires (teacher level)
- 8-9 month data collection (August – May/June)
- Survey cycle is every two or three years
- 2017-18 is the second data collection cycle for public schools and first data collection cycle for private schools

Teacher Incentives Experiment

- Motivation: Increase overall teachers response rates
- Challenges:
 - The teacher response rate is a two-stage response rate
 - Teachers are sampled in waves
- Plan: Incentivize on multiple levels
 - Teacher
 - Principal and/or School Coordinator
 - Targeted contingency incentives

Experimental Design

- Phase 1 – Teacher sample waves 1-12
 - Two treatment groups:
 - Teacher Incentive (\$5 for priority and non-priority)
 - No Teacher Incentive
- Phase 2 – Teacher Sample waves 13-21
 - Four treatment groups:
 - Teacher Incentive and School Coordinator Incentive (\$5 for non-priority, \$10 for priority)
 - Teacher Incentive and no School Coordinator Incentive
 - School Coordinator Incentive and no Teacher Incentive
 - No Teacher Incentive and no School Coordinator Incentive

Goals of Modeling/Analysis

- Goal #1: Evenly distribute schools that are predicted to return the TLF across the eight experimental groups
 - Schools are assigned to the eight experimental groups before the start of data collection
 - Attempt to parse out the effect of the incentive vs. the schools that are more likely to respond randomly being assigned to receive incentives
- Goal #2: Evenly distribute schools that are predicted to return the TLF **early** across the eight experimental groups
 - Attempt to obtain a similar number of schools in each group to increase the strength of the analyses of the second phase of the experiment

Models

- Logistic regression:
 - TLF Response Model
 - Binary response variable: returned the TLF vs. did not return the TLF
 - TLF Early vs. Late Response Model
 - Binary response variable: returned the TLF early vs. did not return the TLF early
- Time-to-event Model
 - Predicts the number of days of data collection before a school will return the TLF

Data Available for Creating Models

- 2015-16 NTPS Frame and Sample Data
 - Includes geography and school characteristics
- 2015-16 NTPS response status data
 - Includes TLF response status and TLF response date

Potential Model Covariates

Charter Status	Lowest grade offered
Grade level (High, Middle, Combined, Elementary)	Highest grade offered
Locale Code (City, Suburbs, Town, Rural)	Number of Full-time Teachers
Poverty Status – 2 categories	Shared Time Indicator
Certainty school flag	Number of American Indian Students
Poverty Status – 4 categories	Number of Asian Students
Adjusted School Type	Number of Hispanic Students
All students – Total	Number of Pacific Islander Students
Categorical recode of MEMBER variable	Number of Black Students
Priority Flag	Number of White Students
Special District Flag	Number of Multi-race Students
Census Regional Office	Percentage of American Indian Students
State code	Percentage of Asian Students
School Type	Percentage of Hispanic Students
School Status	Percentage of Pacific Islander Students
NCES urban-centric locale code	Percentage of Black Students
Magnet School Flag	Percentage of White Students
Total Students in Free/Reduced-Price Lunch Program	Percentage of Multi-race Students
School matched to vendor list	Title I eligibility Flag

Model Building

- Stepwise logistic regression procedures were used to select significant characteristics for the two response propensity models
- Significant characteristics were included and eliminated using the $\alpha = .1$ significance level
- Correlations between characteristics were analyzed for potential multicollinearity between predictors
- The time-to-event model uses Cox proportional hazard modeling to predict the number of days until the school would return the TLF

Final Model Covariates

TLF Response Model	TLF Early Response Model	Time-to-TLF-Response
Charter Status	Charter Status	Charter Status
Locale Code	Enrollment	Locale Code
Priority Flag	Status	Priority Flag
Special District Flag	State	Special District Flag
Region	Percent Black Students	State
Title I Indicator	Number of Hispanic Students	Status
Enrollment	Vendor Flag	Lowest Grade
Percent Hispanic students		Highest Grade
Vendor Flag		Vendor Flag
		Full-time teachers
		Percent Black Students

Final Model Covariates

TLF Response Model	TLF Early Response Model	Time-to-TLF-Response
Charter Status	Charter Status	Charter Status
Locale Code	Enrollment	Locale Code
Priority Flag	Status	Priority Flag
Special District Flag	State	Special District Flag
Region	Percent Black Students	State
Title I Indicator	Number of Hispanic Students	Status
Enrollment	Vendor Flag	Lowest Grade
Percent Hispanic students		Highest Grade
Vendor Flag		Vendor Flag
		Full-time teachers
		Percent Black Students

Final Model Covariates

TLF Response Model

Charter Status

Locale Code

Priority Flag

Special District Flag

Region

Title I Indicator

Enrollment

Percent Hispanic students

Vendor Flag

TLF Early Response Model

Charter Status

Enrollment

Status

State

Percent Black Students

Number of Hispanic Students

Vendor Flag

Time-to-TLF-Response

Charter Status

Locale Code

Priority Flag

Special District Flag

State

Status

Lowest Grade

Highest Grade

Vendor Flag

Full-time teachers

Percent Black Students

Determining the Appropriate Number of Sort Variables

Each additional sort variable further refines the sorting strata.

Having less than eight schools within one sorting strata would result in a school with particular combination of characteristics going into only a few of the experimental groups.

Variable	Number of Levels	Cumulative Sorting Strata
Vendor Flag	2	2
Priority Flag	2	4
Region	4	16
Locale Code	4	64
Charter Status	2	128
Special District Flag	2	256

Determining the Appropriate Number of Sort Variables

Potential sort order:

- Vendor Flag
- Priority Flag
- Region
- Locale Code
- Charter Status
- Special District Flag

Variables Included	Total Sorting Strata	Empty Sorting Strata	Sorting Strata with less than Eight Schools (and at least one School)	Sorting Strata with at least Eight Schools
First Six	256	98	73	85
First Five	128	26	28	74
First Four	64	6	9	49
First Three	16	0	0	16

Determining the Appropriate Number of Sort Variables

Potential sort order:

- Vendor Flag
- Priority Flag
- Region
- Locale Code
- Charter Status
- ~~Special District Flag~~

Variables Included	Total Sorting Strata	Empty Sorting Strata	Sorting Strata with less than Eight Schools (and at least one School)	Sorting Strata with at least Eight Schools
First Six	256	98	73	85
First Five	128	26	28	74
First Four	64	6	9	49
First Three	16	0	0	16

Determining the Appropriate Number of Sort Variables

For each potential sort, the distribution of schools predicted to return the TLF early vs late within each experimental group was compared:

	16 Sorting Strata		64 Sorting Strata		128 Sorting Strata	
Experimental Group	Early	Late	Early	Late	Early	Late
1	811	227	810	228	811	227
2	809	229	813	225	805	233
3	809	228	810	227	814	223
4	808	229	807	230	811	226
5	812	226	807	231	811	227
6	802	236	806	232	805	233
7	809	228	809	228	807	230
8	811	226	809	228	807	230
Largest Difference	10	10	7	5	9	10

Determining the Appropriate Number of Sort Variables

For each potential sort, the distribution of schools predicted to return the TLF early vs late within each experimental group was compared:

	16 Sorting Strata		64 Sorting Strata		128 Sorting Strata	
Experimental Group	Early	Late	Early	Late	Early	Late
1	811	227	810	228	811	227
2	809	229	813	225	805	233
3	809	228	810	227	814	223
4	808	229	807	230	811	226
5	812	226	807	231	811	227
6	802	236	806	232	805	233
7	809	228	809	228	807	230
8	811	226	809	228	807	230
Largest Difference	10	10	7	5	9	10

Final Sort Order

- Vendor Flag
- Priority Flag
- Region
- Locale Code
- TLF Response Propensity (Public Schools only)
- Random number (to break ties)

TLF Response Rates

Treatment Group	Public Schools
1	84.04%
2	84.37%
3	84.82%
4	83.40%
5	85.73%
6	84.85%
7	83.45%
8	85.36%
All	84.52%

* No two experimental group's TLF response rates tested significantly different at the .05 significance level for either the public schools nor the private schools

Public Teacher Sample Sizes

Treatment Group	Phase One	Phase Two
1	3,820	3,150
2	3,820	3,470
3	3,690	3,530
4	3,600	3,510
5	4,010	3,330
6	3,850	3,530
7	3,740	3,470
8	3,740	3,520

Summary

- Response propensity modeling was used to establish a way to evenly distribute schools (and therefore, teachers) into the incentive treatment groups
- The final TLF response rates were all comparatively the same at the end of data collection
- This resulted in an even distribution of teachers across the eight experimental treatment groups
- Experiment Results will be presented at AAPOR Conference in Toronto, CA on May 18th, 2019

Thank you!

Contact:

Allison Zotti

allison.zotti@census.gov