

THRIVE SHINY APP: WALKTHROUGH

Last Updated: March 11, 2025

SETUP

- Pull the repo [here](https://github.com/michaelwcheung/Conditional-Bootstrap.git), or follow these steps:
 - git clone <https://github.com/michaelwcheung/Conditional-Bootstrap.git>
 - cd Conditional-Bootstrap
 - sh run.sh
- Ensure you have R installed on your machine, and Rscript is on PATH.

The following slides will walk through the four tabs/pages on the THRIVE Shiny app.

I. LANDING PAGE

LANDING PAGE

- Once you launch the Shiny App, you should see the following screen:



Transparent High-dimensional Reproducible Inference through Validation and Exploration (THRIVE)

THRIVE is a method to identify effect modifiers by estimating conditional average treatment effects (CATEs) on resamples of pseudopopulations with different quantile levels of potential effect modifiers.

How to use the Shiny App

- In the 'Upload Data' tab, upload your dataset and select variables for analysis.
- In the 'Set Hyper Parameters' tab, set your parameters for the resampling process.
- In the 'Results' tab, view and download plots and tables about your effect modifiers by quantiles.

Please submit any feedback or questions to [this form](#). Thank you!

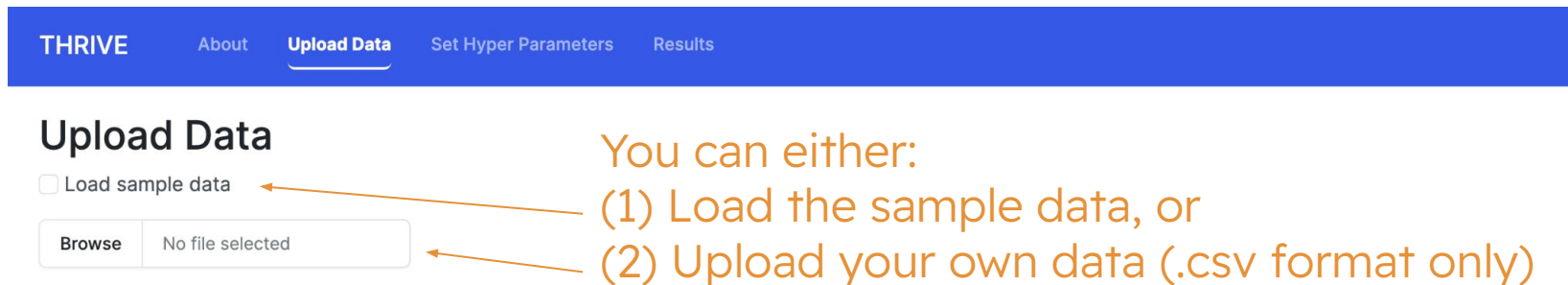


To proceed, click this button!

II. UPLOAD DATA PAGE

UPLOAD DATA PAGE—UPLOAD DATA

- Next, you should see the following screen:



The screenshot shows the 'THRIVE' application interface. The top navigation bar is blue and contains the following links: 'About', 'Upload Data' (which is underlined), 'Set Hyper Parameters', and 'Results'. Below the navigation bar, the page title 'Upload Data' is displayed. Under this title, there is a checkbox labeled 'Load sample data'. Below the checkbox is a file upload button labeled 'Browse' with the text 'No file selected' next to it. Two orange arrows point from the text 'You can either: (1) Load the sample data, or (2) Upload your own data (.csv format only)' to the 'Load sample data' checkbox and the 'Browse' button respectively.

THRIVE About Upload Data Set Hyper Parameters Results

Upload Data

☐ Load sample data

Browse No file selected

You can either:
(1) Load the sample data, or
(2) Upload your own data (.csv format only)

- The “sample data” is a subset of size 10,000 from a case study interested in the effect of drought on stunting among children in Sub-Saharan Africa from the Demographic and Health Survey (DHS) data set.

UPLOAD DATA PAGE—UPLOAD DATA

- Next, you should see a snapshot of your chosen dataset:

THRIVE [About](#) [Upload Data](#) [Set Hyper Parameters](#) [Results](#)

Upload Data

☒ Load sample data

Browse No file selected

Show entries Search:

STUNTED	SEX_MALE	AGE_UNDER2	SMALL_BIRTHSIZE	NEVER_BREASTFED	EDUCATION_NONE	MASS_MEDIA	SINGLE_MOTHER	AGRI_OCCUPATION	RURAL_
0	0	1	0	0	1	0	0		1
1	0	1	1	0	1	1	0		1
1	0	1	0	0	0	1	0		1
1	1	1	1	0	1	0	0		1

Showing 1 to 10 of 10,000 entries

[Previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [...](#) [1000](#) [Next](#)

[Next: Variable Selection](#)

To proceed, click this button!

UPLOAD DATA PAGE—VARIABLE SELECTION

- Next, you should see four dropdown boxes and a slider to set your outcome, treatment, covariates, and candidate variables.

[Next: Variable Selection](#)

Variable Selection

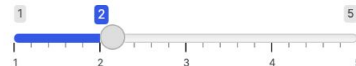
Select outcome variable:

Select covariates:

Select treatment variable:

Select candidate variables:

Select the size of the cross-classifications you would like to test:



UPLOAD DATA PAGE—VARIABLE SELECTION

- THRIVE currently only supports binary treatment & outcome variables, and binary or categorical covariates & candidate variables.
 - For the sample dataset (DHS), the outcome is “stunted”, the treatment/exposure is “drought”, and the covariates are all 10 remaining variables. For the sake of computation, I have chosen three candidate variables: “education_none”, “mass_media”, and “rural_residence”.

Variable Selection

Select outcome variable:

✓
stunted ▼
stunted is a binary variable.

Select covariates:

sex_male, age_under2, small_birthsize ▲

Select treatment variable:

✓
drought ▼
drought is a binary variable.

Select candidate variables:

education_none, mass_media, rural_re ▲

Select the size of the cross-classifications you would like to test:

1 2 3 4 5
1 2 3 4 5

Next: Variable Verification

To proceed, click this button!

UPLOAD DATA PAGE—VARIABLE VERIFICATION

- Next, you should see a table of each column in your dataset with the following information:
 - Their assigned role: “treatment” or “outcome” or “covariate” and/or “candidate”
 - Their data type: “binary” or “categorical” or “continuous”
 - Their missingness

Variable Verification

Show 10 entries

Search:

VARIABLE	ROLE	DATA TYPE	PERCENT NA
stunted	outcome	binary	0
sex_male	covariate	binary	0
age_under2	covariate	binary	0
small_birthsize	covariate	binary	0

Showing 1 to 10 of 12 entries

Previous 1 2 Next

[Next: Confirm Potential Effect Modifier Prevalence](#)

← To proceed, click this button!

UPLOAD DATA PAGE—PREVALENCE

- Next, you should see a table of the prevalence of potential ems.
 - If a potential em's prevalence < 5%, a warning will appear: prevalence close to 0 will lead to biased resampling results (positivity assumption).
 - For the DHS dataset, I have chosen three binary candidate variables with a cross-classification size of 2. So, there are $3 + 2^2 \cdot {}_3C_2 = 15$ potential effect modifiers.

Potential Effect Modifier Prevalence in Data

Show entries

Search:

POTENTIAL EFFECT MODIFIER (CROSS-CLASSIFICATION)	PROPORTION IN DATA (N=10000)	IN-COUNT	OUT-COUNT
education_none	50.66%	5066	4934
mass_media	44.53%	4453	5547
rural_residence	72.29%	7229	2771
education_none_1:mass_media_1	17%	1700	8300

Showing 1 to 10 of 15 entries

Previous **1** 2 Next

[Next: Set Hyper Parameters](#)

To proceed, click this button!

III. SET HYPER PARAMETERS

SET HYPER PARAMETER PAGE

- Next, you should see a few selection items; the first are the quantile levels for the potential effect modifiers at which to resample.
 - You may also enter a custom quantile; please ensure the format follows those provided: 0, ..., 100 (in increasing order).

THRIVE[About](#)[Upload Data](#)[Set Hyper Parameters](#)[Results](#)

Set Hyper Parameters

Select quantile:

☒ Median: 0, 50, 100

☐ Tertile: 0, 33.33, 66.66, 100

☐ Quartile: 0, 25, 50, 75, 100

☐ Quintile: 0, 20, 40, 60, 80, 100

☐ Decile: 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100

Or enter a custom quantile (please ensure 0 and 100 are included!):

Add custom quantile

SET HYPER PARAMETER PAGE

- The remaining selection items are as required in a traditional effect modification analysis.

Select the type of estimand:

- ☒ Risk difference
☐ Risk ratio
☐ Odds ratio

Select the heterogeneity test:

- ☒ Cochran's Q
☐ Bootstrap hypothesis difference of means
☐ Confidence interval overlap
☐ Pre-bootstrap Cochran's Q

Select the estimation method for causal effect:

- ☒ Outcome regression (adjusted)
☐ IP weighting
☐ Doubly robust (variance estimates may be biased)

Number of bootstrap samples:

100

👉 Start Resampling!

👉 Stop Resampling

Note: the doubly robust estimator (AIPTW) uses a closed-form formula for the variance estimate rather than a bootstrap estimate, which may be biased if not under certain conditions ([Lunceford and Davidian, 2004](#)).

The **resampling process** is computationally intense, depending on the number of bootstraps and number of potential effect modifiers. A progress bar will appear in the bottom right of the screen, but please use the Shiny app for demonstration or testing purposes only; the function call will have shorter runtime than the Shiny app in general on the same dataset and parameters.

If everything looks good, **start resampling!**

SET HYPER PARAMETER PAGE

- Once the resampling is complete, the following button will appear to take you to the results page.
 - You may also choose to stop the resampling process—this will delete all resampling progress.

Number of bootstrap samples:

👉 Start Resampling!

👉 Stop Resampling

👉 Next: See Results

← To proceed, click this button!

IV. RESULTS PAGE

RESULTS PAGE

- There are three result items, all of which are downloadable.
- The first is the plot of the resampled CATE estimates by quantile for identified (significant) effect modifiers.

THRIVE

About

Upload Data

Set Hyper Parameters

Results

Results

Observed ATE: 0.01266 with 95% CI (-0.01382, 0.03913)

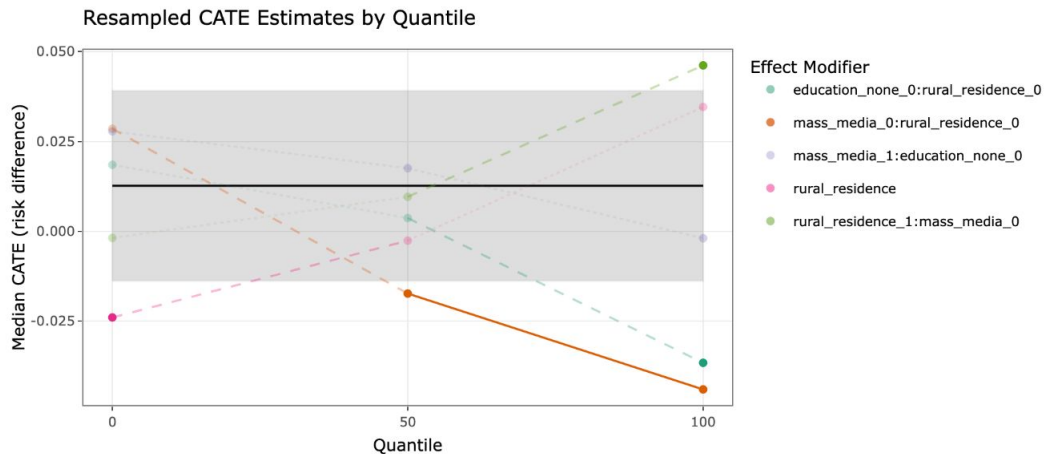
☒ Show ATE on plot

Effect Modifiers:

education_none_0:rural_residence_0, mass_media_0:1

Linetype Key:

- Solid: Significant
- Dashed: Between significance threshold
- Dotted: Insignificant



RESULTS PAGE

- The first table contains the resampled CATEs by quantile and effect modifier for identified effect modifiers (same as those on the plot) ranked by magnitude across quantile and effect modifier.

Resampled CATEs by Quantile and Subgroup

Show 10 ▾ entries

Search:

QUANTILE ▾	EFFECT MODIFIER ▾	MEDIAN CATE ▾	MEAN CATE ▾	LOWER 95% CI ▾	UPPER 95% CI ▾
100	rural_residence_1:mass_media_0	0.046	0.044	0.021	0.068
100	mass_media_0:rural_residence_0	-0.044	-0.047	-0.072	-0.033
100	education_none_0:rural_residence_0	-0.037	-0.036	-0.049	-0.021
100	rural_residence	0.035	0.033	0.011	0.05

Showing 1 to 10 of 15 entries

Previous 1 2 Next

Download

RESULTS PAGE

- The second table contains the linear regression coefficients of the significant effect modifiers (i.e. the slope of the lines in the plot)
 - We use linear regression because the lines in the plot appear to be linear.

Linear Regression Coefficients of CATEs on Quantile by Effect Modifier

Show entries

Search:

EFFECT MODIFIER	COEFFICIENT ESTIMATE	P-VALUE
mass_media_0:rural_residence_0	-0.00073	0.096
rural_residence	0.00059	0.098
education_none_0:rural_residence_0	-0.00055	0.166
rural_residence_1:mass_media_0	0.00048	0.187

Showing 1 to 5 of 5 entries

Previous **1** Next

 Download

REFERENCES

Lunceford JK and Davidian M.
“Stratification and weighting
via the propensity score in
estimation of causal
treatment effects: a
comparative study”. *Statist.
Med.* 2004; 23:2937–2960.
[[Link](#)]