Acknowledgment of Funding

- ASPES method development is funded by an Institute of Education Sciences Early Career Methods Grant.
- The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305D150016.
- The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.
Endogenous Subgroup Analysis Using ASPES

Part 1: Introduction
Laura Peck
Society for Research on Educational Effectiveness
Washington, DC | March 2017
Agenda

- **Introductions:** to each other, to the course material
- **What this workshop is about:** Mediators and mediation analysis, *with experimental data*
- **Overview of five methods:** Structural Equation Modeling (SEM), Instrumental Variables (IV), Principal Stratification, Propensity Score Matching (PSM), Analysis of Symmetrically-Predicted Endogenous Subgroups (ASPES)
- **Comparison of methods:** research questions, estimation process, assumptions, interpretation
- **Detailed instruction in one method:** ASPES
- **Illustrative example**
**Agenda**

- **Introductions:** to each other, to the course material
- **What this workshop is about:** Mediators and mediation analysis, *with experimental data*
- **Overview of five methods:** Structural Equation Modeling (SEM), Instrumental Variables (IV), Principal Stratification, Propensity Score Matching (PSM), Analysis of Symmetrically-Predicted Endogenous Subgroups (ASPES)
- **Comparison of methods:** research questions, estimation process, assumptions, interpretation
- **Detailed instruction in one method:** ASPES
- **Illustrative example**
Agenda

- **Introductions:** to each other, to the course material
- **What this workshop is about:** Mediators and mediation analysis, *with experimental data*
- **Overview of five methods:** Structural Equation Modeling (SEM), Instrumental Variables (IV), Principal Stratification, Propensity Score Matching (PSM), Analysis of Symmetrically-Predicted Endogenous Subgroups (ASPES)
- **Comparison of methods:** research questions, estimation process, assumptions, interpretation
- **Detailed instruction in one method:** ASPES
- **Illustrative example**
Agenda

- **Introductions:** to each other, to the course material
- **What this workshop is about:** Mediators and mediation analysis, *with experimental data*
- **Overview of five methods:** Structural Equation Modeling (SEM), Instrumental Variables (IV), Principal Stratification, Propensity Score Matching (PSM), Analysis of Symmetrically-Predicted Endogenous Subgroups (ASPES)
- **Comparison of methods:** research questions, estimation process, assumptions, interpretation
- **Detailed instruction in one method:** ASPES
- **Illustrative example**
Introductions

- To instructors:
  - Eleanor Harvill
  - Shawn Moulton
  - Laura Peck

- To each other:
  - Name, affiliation
What this workshop is about

- Mediator = intermediate variable:
  - Program-related: element of program, such as “peer support groups”
  - Person-related: milestone achieved, such as “earned HS degree/GED”
What this workshop is about

Mediator = intermediate variable:
- Program-related: element of program, such as “peer support groups”
- Person-related: milestone achieved, such as “earned HS degree/GED”

Indirect effect (of T on Y, given M) = a*b

Direct effect (of T on Y) = c

Proportion of effect that is indirect
What this workshop is about

- **Mediator** = intermediate variable:
  - Program-related: element of program, such as “peer support groups”
  - Person-related: milestone achieved, such as “earned HS degree/GED”

- **Indirect effect (of T on Y, given M) = a*b**
- **Direct effect (of T on Y) = c**
- Proportion of effect that is indirect

Two kinds of Qs to answer
What this workshop is about

- Structural Equation Modeling
- Instrumental Variables
- Principal Stratification
- Propensity Score Matching
- Analysis of Symmetrically-predicted Endogenous Subgroups
What this workshop is about

- Structural Equation Modeling
- Instrumental Variables
- Principal Stratification
- Propensity Score Matching
- Analysis of Symmetrically-predicted Endogenous Subgroups

Diagram:
- Treatment (T) → Mediator (M) → Outcome (Y)
- a → Mediator (M)
- Mediator (M) → Outcome (Y)
- b → Outcome (Y)
- Treatment (T) → Outcome (Y)
- c
What this workshop is about

- Structural Equation Modeling
- Instrumental Variables
- Principal Stratification
- Propensity Score Matching
- Analysis of Symmetrically-predicted Endogenous Subgroups

Diagram:

- Treatment (T) → Mediator (M) → Outcome (Y)
- a, b, c represent different parameters or effects in the model.
Endogenous Subgroups Conceptually

When exposed to treatment...
- used program feature Z (or not)
- experienced high dosage of intervention
- followed treatment path W-X-Y
- behaved a particular way

If exposed to treatment, *would have*...
- used program feature Z (or not)
- experienced high dosage of intervention
- followed treatment path W-X-Y
- behaved a particular way
Likewise...

In the absence of the treatment...
• dropped out of school (or not)
• experienced long-term unemployment
• had less favorable LM outcomes
• behaved a particular way

If not exposed to treatment, would have...
• dropped out of school (or not)
• experienced long-term unemployment
• had less favorable LM outcomes
• behaved a particular way
… the “Story is in the Subgroups”

- Exogenous
  - Uni-dimensional (e.g., women, low-education, prior arrest)
  - Multi-dimensional (e.g., disadvantaged, “at risk”)

- Endogenous
  - Uni-dimensional (e.g., took up offer, experienced intervention delivered with “fidelity”)
  - Multi-dimensional (e.g., experienced some dosage, participated in this package of services)

- Mediation
  - as programmatic factor
  - as personal characteristics
Baron & Kenny (1986)

Addresses both kinds of Qs (direct, indirect effects)

Notation:
- Binary treatment: T=1 (Treatment), T=0 (Control)
- M=Mediator, Y=Outcome, X=Baseline characteristics

Direct and indirect effects are estimated using:

\[ M = \alpha + aT + X\pi + e_1 \]
\[ Y = \beta + bM + cT + X\varphi + e_2 \]

Estimated indirect effect: \( \hat{\delta} = \hat{a} * \hat{b} \)

Estimated direct effect: \( \hat{\gamma} = \hat{c} \)
SEM Assumptions

- Stable Unit Treatment Value Assumption (SUTVA)
- Treatment assignment is random
- Linearity
- No interaction between $M$ and $T$
- Ignorability of observed mediator status: Conditional on $X$, $M$ is not correlated with the error, $e_2$
SEM Assumptions

- Stable Unit Treatment Value Assumption (SUTVA)
- Treatment assignment is random
- Linearity
- No interaction between $M$ and $T$
- Ignorability of observed mediator status: Conditional on $X$, $M$ is not correlated with the error, $e_2$
Instrumental Variables Basics (w/experiment)

- Angrist, Imbens, and Rubin (1996)
- Addresses only the indirect effect question:
  - What is the effect of take up?
- Use exogenous variation in mediator created by treatment to estimate effect of mediator on outcome (the indirect effect)
Instrumental Variables Basics (w/experiment)

- Angrist, Imbens, and Rubin (1996)
- Addresses only the indirect effect question:
  - What is the effect of take up?
- Use exogenous variation in mediator created by treatment to estimate effect of mediator on outcome

![Diagram of Treatment Group and Control Group with 'no shows' and 'took up offer' categories]
Using 2SLS, fit the first stage model:

\[ M = \alpha + aT + X\pi + e_1 \]

Predict \( \hat{M} \) (which is free of unobserved \( W \) and measurement error)

Use the predicted mediator, \( \hat{M} \), in the second stage:

\[ Y = \beta + b\hat{M} + X\varphi + e_2 \]
IV Assumptions

- SUTVA
- Treatment assignment is random
- Linearity
- Treatment effect on the mediator is non-zero
  - Also known as instrument effectiveness
- No direct effect (i.e., $M$ is the only mediator)
  - Also known as the “exclusion restriction”
IV Assumptions

- SUTVA
- Treatment assignment is random
- Linearity
- Treatment effect on the mediator is non-zero
  - Also known as instrument effectiveness
- No direct effect (i.e., $M$ is the only mediator)
  - Also known as the “exclusion restriction”
Principal Stratification Basics

- A generalized case of IV (Frangakis & Rubin, 2002) and ASPES (Bein, 2013)
- Provides a framework for organizing subgroup impacts
- Addresses (indirectly) both kinds of Qs (direct, indirect effects)
- Partition sample into strata based on potential values of mediator and use strata-specific effects to make inferences about a, b, and c
- In practice, it can use varied analytic procedures
Principal Stratification Notation

- Binary mediator; M=high or M=low
- $M_T$: Potential mediator status under treatment
- $M_C$: Potential mediator status under control
- Sample is in one of four groups based on $M_T$ & $M_C$:
  - Always-High (A): $M_T$=high and $M_C$=high
  - Treatment only-High (TO): $M_T$=high and $M_C$=low
  - Control only-High (CO): $M_T$=low and $M_C$=high
  - Never-High (N): $M_T$=low and $M_C$=low
Principal Stratification Notation

- Binary mediator; M=high or M=low
- $M_T$: Potential mediator status under treatment
- $M_C$: Potential mediator status under control
- Sample is in one of four groups based on $M_T$ & $M_C$:
  - Always-High (A): $M_T=$high and $M_C=$high (always takers)
  - Treatment only-High (TO): $M_T=$high and $M_C=$low (compliers)
  - Control only-High (CO): $M_T=$low and $M_C=$high (defiers)
  - Never-High (N): $M_T=$low and $M_C=$low (never takers)
Principal Stratification (cont.)

Principal Strata

<table>
<thead>
<tr>
<th>Control</th>
<th>Treatment</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>M=High</td>
<td>M=High</td>
<td>A</td>
<td>CO</td>
</tr>
<tr>
<td>M=Low</td>
<td>TO</td>
<td></td>
<td>N</td>
</tr>
</tbody>
</table>

Mediation Analysis and PS

- By definition, there are no indirect effects on A and N
- Effects on TO and CO reflect direct and indirect effects
- Estimation challenge: Stratum membership is not observable in both experimental states
PS Assumptions

- SUTVA
- Observed Mediator Status under T or C = Potential Mediator Status under that condition.
- Treatment assignment is random.
- Principal Ignorability: Principal stratum membership is fully explained by pretreatment attributes $X$
PS Assumptions

- SUTVA
- Observed Mediator Status under T or C = Potential Mediator Status under that condition
- Treatment assignment is random
- Principal Ignorability: Principal stratum membership is fully explained by pretreatment attributes $X$
PS-based Estimation

- Page (2012) uses a Bayesian approach
- Stuart and Jo (2012) use propensity score matching
- Unlu et al. (2013) use double propensity scoring
The Analysis of Symmetrically-Predicted Endogenous Subgroups (ASPES) method provides a framework for creating experimentally valid subgroups defined by some post random assignment event or path (Peck, 2003, 2013)

Requires an experimentally designed evaluation and baseline data
Kinds of Questions

- Discrete endogenous subgroups
  - Potential effects on “no-shows”
  - Treatment dosage or quality (low, medium, high)
  - Treatment components, pathways
  - Control group fall-back experience

- Continuous endogenous indicators
  - Treatment dosage or quality (along a continuum)
  - Continuous mediating factors
  - Control group fall-back experience
Classes of Endogenous Groups

- (1) Potential effects on “no-shows”
- **Examples**
  - **NYCAP**: non-takers still made changes to try and take advantage of new policy structure
  - **MTO**: those who did not lease up still got counseling services and **tried**
Classes of Endogenous Groups

- (2) Treatment dosage or quality

**Examples**

- **BSF:**
  - *what impact does full participation have?* (discrete)
  - *what impact does the number of hours have?* (continuous)

- **HSIS:** *what generates greater impacts…*
  - *two years, rather than one?*
  - *being in a better quality center?* (discrete or continuous)
Classes of Endogenous Groups

- (3) Multi-faceted treatment components/pathways

Examples
- NEWWS: *what impact does [sanction] have?*
- HPOG: *what is it about intervention that drives impacts?*
(4) Subsets of the control group conditions that make particular fall-back choices when denied access to the intervention

**Examples**

- Career Academies: those who dropped out of school
- HSIS: those who stay at home with parent(s)
- JTPA: those with better/worse labor market outcomes
Comparison of Methods for Mediation Analysis

- See Comparison of Methods for Mediation Analysis Handout

- Methods differ in terms of:
  - Research Question Addressed
  - Estimation Process
  - Key Assumptions
  - Interpretation
  - Data Requirements
Break 1

- Up next: Ellie on ASPES Instruction
Endogenous Subgroup Analysis Using ASPES

Part 2: ASPES Instruction
Eleanor Harvill
Society for Research on Educational Effectiveness
Washington, DC | March 2017
In 2004, the U.S. Department of Education’s Institute of Education Sciences contracted with Mathematica Policy Research to conduct the Comprehensive Teacher Induction (CTI) Study.

CTI Study Design: 418 elementary schools in 17 urban districts were assigned by lottery to either:
- a treatment group whose beginning teachers were offered comprehensive teacher induction or
- a control group whose beginning teachers received the district’s “business as usual” induction services

See Impacts of Comprehensive Teacher Induction, Glazerman et al. (2010)
In this section, we will introduce the mechanics of the method using the CTI study as a concrete example.

We are interested in how the intensity of mentorship affects the impact of CTI.

We operationalize the intensity of mentorship in two ways:
- Number of classroom observations by a mentor teacher (continuous measure)
- Indicator for number of observations at or above the median (binary measure)

This section presents methods for analyzing both mediators.

The following section walks through the results of such an analysis.
When exposed to treatment...
• used program feature Z (or not)
• experienced high dosage of intervention
• followed treatment path W-X-Y
• behaved a particular way

If exposed to treatment, *would have*...
• used program feature Z (or not)
• experienced high dosage of intervention
• followed treatment path W-X-Y
• behaved a particular way
A Primer on How To

Continuous Mediator

- Step 1: Predict values of the mediator
  - Use baseline (exogenous) characteristics to predict the value of the mediator
  - Employ an approach to avoid overfitting

Binary Mediator
A Primer on How To

**Continuous Mediator**
- Step 1: Predict values of the mediator
  - Use baseline (exogenous) characteristics to predict the value of the mediator
  - Employ an approach to avoid overfitting
- Step 2: estimate the relationship between the predicted continuous mediator and impact

**Binary Mediator**
A Primer on How To

Continuous Mediator

- Step 1: Predict values of the mediator
  - Use baseline (exogenous) characteristics to predict the value of the mediator
  - Employ an approach to avoid overfitting
- Step 2: estimate the relationship between the predicted continuous mediator and impact

Binary Mediator

- Step 1: Predict values of the mediator and construct predicted subgroups
  - Use baseline (exogenous) characteristics to predict the value of the mediator
  - Employ an approach to avoid overfitting
A Primer on How To

Continuous Mediator
- Step 1: Predict values of the mediator
  - Use baseline (exogenous) characteristics to predict the value of the mediator
  - Employ an approach to avoid overfitting
- Step 2: estimate the relationship between the predicted continuous mediator and impact

Binary Mediator
- Step 1: Predict values of the mediator and construct predicted subgroups
  - Use baseline (exogenous) characteristics to predict the value of the mediator
  - Employ an approach to avoid overfitting
- Step 2: Estimate impacts on predicted subgroups
A Primer on How To

Continuous Mediator
- Step 1: Predict values of the mediator
  - Use baseline (exogenous) characteristics to predict the value of the mediator
  - Employ an approach to avoid overfitting
- Step 2: estimate the relationship between the predicted continuous mediator and impact

Binary Mediator
- Step 1: Predict values of the mediator and construct predicted subgroups
  - Use baseline (exogenous) characteristics to predict the value of the mediator
  - Employ an approach to avoid overfitting
- Step 2: Estimate impacts on predicted subgroups
- Step 3: Convert estimated impacts for predicted subgroups to represent actual subgroups
Step 1: Predict Mentorship

- Predict values of the mediator by:
  - Estimating a model that relates mentorship to baseline characteristics in the treatment group
  - Using these estimates to predict mentorship for both the treatment and control group

- Key points:
  - Predicted subgroups are defined based on exogenous baseline characteristics
  - In expectation, random assignment insures that the predicted values of the mediator is independent of treatment status
Step 1: Predict values of the mediator (Issue: Overfitting)

- What is overfitting?
  - If one uses the entire treatment group to estimate mentorship (as offered by CTI), the model will do a better job of predicting mentorship in treatment group than it does in the control group.

- This introduces an imbalance into the analysis of predicted subgroups, which biases estimates.

- How to avoid overfitting?
  - Use a cross-validation approach so that all prediction is out-of-sample
  - Cross-validation allows you to do out-of-sample prediction for all sample members with no loss of sample
Step 1: Predict values of the mediator (Issue: Overfitting)

- **Split Sample Approach**
  - Divide your treatment group in two: a prediction sample and an analysis sample
  - Estimate the prediction model on the treatment group prediction sample
  - Predict values of the mediator for the treatment group analysis sample and the control group
Step 1: Predict values of the mediator (Issue: Overfitting)

- Split Sample Approach
  - Divide your treatment group in two: a prediction sample and an analysis sample
  - Estimate the prediction model on the treatment group prediction sample
  - Predict values of the mediator for the treatment group analysis sample and the control group

- Downside: loss of sample for analysis
Step 1: Predict values of the mediator (Issue: Overfitting)

- Split Sample Approach
  - Divide your treatment group in two: a prediction sample and an analysis sample
  - Estimate the prediction model on the treatment group prediction sample
  - Predict values of the mediator for the treatment group analysis sample and the control group
- Downside: loss of sample for analysis
- Solution: What if you did another out of sample prediction?
Step 1: Predict values of the mediator (Issue: Overfitting)

- **Split Sample Approach**
  - Divide your treatment group in two: a prediction sample and an analysis sample
  - Estimate the prediction model on the treatment group prediction sample
  - Predict values of the mediator for the treatment group analysis sample and the control group

- **Downside:** loss of sample for analysis

- **Solution:** What if you did another out of sample prediction?
  - Estimate prediction model on treatment group analysis sample
  - Predict mediator values for treatment group prediction sample
Step 1: Predict values of the mediator (Solution: Cross-Validation)

Steps in cross-validation:

1. Randomly partition your sample (both T and C) into 10 groups of equal size
Step 1: Predict values of the mediator
(Solution: Cross-Validation)

Steps in cross-validation:

1. Randomly partition your sample (both T and C) into 10 groups of equal size
2. Obtain predictions for group 1 by:
Step 1: Predict values of the mediator (Solution: Cross-Validation)

Steps in cross-validation:

1. Randomly partition your sample (both T and C) into 10 groups of equal size

2. Obtain predictions for group 1 by:
   - Estimating the prediction model on treatment individuals in groups 2-10
Step 1: Predict values of the mediator (Solution: Cross-Validation)

Steps in cross-validation:

1. Randomly partition your sample (both T and C) into 10 groups of equal size

2. Obtain predictions for group 1 by:
   • Estimating the prediction model on treatment individuals in groups 2-10
   • Predicting dosage for both treatment and control individuals in group 1
Step 1: Predict values of the mediator
(Solution: Cross-Validation)

Steps in cross-validation:

1. Randomly partition your sample (both T and C) into 10 groups of equal size

2. Obtain predictions for group 1 by:
   - Estimating the prediction model on treatment individuals in groups 2-10
   - Predicting dosage for both treatment and control individuals in group 1

3. Obtain predictions for group 2 by:
   - Estimating the prediction model on treatment individuals in groups 1 and 3-10
   - Predicting dosage for both treatment and control individuals in group 2

4. ...

Step 1: Predict values of the mediator
(Solution: Cross-Validation)
Step 1: Predict values of the mediator

Continuous Mediator
- Example: Number of classroom observations by a mentor teacher
- Use a cross validation approach to construct predicted number of classroom observations by a mentor teacher

Binary Mediator
- Example: Indicator for number of observations at or above the median
- Use a cross validation approach to construct predicted number of classroom observations by a mentor teacher
- Create an indicator for predicted number of observations at or above the median
- (Alternatively, you could discretize first)
Continuous ASPES: Stage 2: Impact Model

- Estimate the Relationship between the Predicted Mediator and Effect Size:

\[ Y = \beta_0 + \beta_1 \hat{M}^P + \beta_2 T + \beta_3 T \hat{M}^P + \varepsilon_2 \]

- \( Y \) is the outcome being examined;
- \( \hat{M}^P \) is the predicted value of the mediator generated from Stage 1;
- \( T \) indicates whether the member was assigned to the treatment or control group; and
- \( \varepsilon_2 \) is an error term that captures all other factors that influence the outcome.

- The impact of being assigned to the treatment group is given by the following equation:

\[ \frac{\partial Y}{\partial T} = \beta_2 + \beta_3 \hat{M}^P \]
Continuous ASPES: Key Assumption

- We require that:
  - The baseline covariates that predict $M^p$ have no direct or indirect effect on the impact $\Delta$ apart from their indirect effect on $\Delta$ through $M^A$

- If this assumption holds, the coefficient of the predicted mediator reflects the increase in impact associated with a unit increase in the actual mediator
Continuous ASPES: Key Assumption

This assumption may be violated if the baseline characteristics $X$ used to predict the mediator $M1^P$ influence the impact $\Delta$ through channels other than the actual number of observations $M1^A$.

Assumes no direct or indirect effect of $X$ on $\Delta$.
To interpret the coefficient $\beta_3$ as the causal increase in student achievement expected from each additional teacher observation, we require:

- A given mediator-value-defined subpopulation would experience the same impact as an alternative mediator-value-defined subpopulation if they were coerced to receive the corresponding alternative value of the mediator.

- This assumption may be violated if study members attributes (e.g., motivation, ability, etc.) vary significantly across subpopulations since these differences in attributes may drive differential subpopulation effects.
Continuous Mediator

- Step 1: Predict values of the mediator
  - Use baseline (exogenous) characteristics to predict the value of the mediator
  - Employ an approach to avoid overfitting
- Step 2: estimate the relationship between the predicted continuous mediator and impact

Binary Mediator

- Step 1: Predict values of the mediator and construct predicted subgroups
  - Use baseline (exogenous) characteristics to predict the value of the mediator
  - Employ an approach to avoid overfitting
- Step 2: Estimate impacts on predicted subgroups
- Step 3: Convert estimated impacts for predicted subgroups to represent actual subgroups
Consider two groups, A & B:
- \( I_A = \bar{Y}_{TA} - \bar{Y}_{CA} \) and \( I_B = \bar{Y}_{TB} - \bar{Y}_{CB} \)

Or, estimate:
- \( y_i = \alpha + \delta T_i + \beta X_i + e_i \)
- \( y \) is the outcome;
- \( \alpha \) is the intercept (interpreted as the control mean outcome);
- \( T \) is the treatment indicator (treatment = 1; control = 0);
- \( \delta \) is the impact of the treatment (on subgroup of interest);
- \( X \) is a vector of baseline characteristics;
- \( \beta \) are the coefficients on the baseline characteristics;
- \( e \) is the residual; and
- the subscript \( i \) indexes individuals.
In the two group case:

- $I_A = w_A A_A + (1 - w_A) B_A$
- $I_B = w_B B_B + (1 - w_B) A_B$

where

- $I$ is the impact on predicted Subgroup members;
- $A$ is the impact on actual Subgroup $A$;
- $B$ is the impact on actual Subgroup $B$;
- $w$ is the proportion of predicted Subgroup members who are actually in the Subgroup; and
- the subscripts $A$ & $B$ denote Subgroup membership.
Binary ASPES: Step 3: Convert from Predicted to Actual

In the two group case:

- \( I_A = w_A A_A + (1 - w_A) B_A \)
- \( I_B = w_B B_B + (1 - w_B) A_B \)

where

- \( I \) is the impact on predicted Subgroup members;
- \( A \) is the impact on actual Subgroup A;
- \( B \) is the impact on actual Subgroup B;
- \( w \) is the proportion of predicted Subgroup members who are actually in the Subgroup; and
- the subscripts A & B denote Subgroup membership.
Binary ASPES: Step 3: Convert from Predicted to Actual

- In the two group case:
  - \( I_A = w_A A_A + (1 - w_A) B_A \)
  - \( I_B = w_B B_B + (1 - w_B) A_B \)

where

- \( I \) is the impact on predicted Subgroup members;
- \( A \) is the impact on actual Subgroup A;
- \( B \) is the impact on actual Subgroup B;
- \( w \) is the proportion of predicted Subgroup members who are actually in the Subgroup; and
- the subscripts A & B denote Subgroup membership.
In the two group case:

- $I_A = w_A A_A + (1 - w_A) B_A$
- $I_B = w_B B_B + (1 - w_B) A_B$

where

- $I$ is the impact on predicted Subgroup members;
- $A$ is the impact on actual Subgroup A;
- $B$ is the impact on actual Subgroup B;
- $w$ is the proportion of predicted Subgroup members who are actually in the Subgroup; and
- the subscripts A & B denote Subgroup membership.
Binary ASPES: Step 3: Convert from Predicted to Actual

- In the two group case:
  - \( I_A = w_A A_A + (1 - w_A) B_A \)
  - \( I_B = w_B B_B + (1 - w_B) A_B \)

where
- \( I \) is the impact on predicted Subgroup members;
- \( A \) is the impact on actual Subgroup A;
- \( B \) is the impact on actual Subgroup B;
- \( w \) is the proportion of predicted Subgroup members who are actually in the Subgroup; and
- the subscripts \( A \) & \( B \) denote Subgroup membership.
Binary ASPES: Step 3: Conversion Assumptions

- With the following assumptions...
  - $A_A = A_B$
  - $B_A = B_B$

- ... we can rearrange the equations to solve for the unknowns as a function of the knowns:

\[
A_A = \frac{(I_A)(w_B) - (1 - w_A)(I_B)}{w_B + w_A - 1}
\]

\[
B_B = \frac{(I_B)(w_A) - (1 - w_B)(I_A)}{w_B + w_A - 1}
\]
Break 2

- Up next: Shawn on ASPES in Practice, with CTI Illustration
Endogenous Subgroup Analysis Using ASPES

Part 3: ASPES in Practice
Shawn Moulton
*Society for Research on Educational Effectiveness*
Washington, DC | March 2017
ASPES Method in Practice: Outline

- Design requirements
- ASPES example using data from the Comprehensive Teacher Induction Study (Glazerman et al., 2010)
- Introduction to SPI-Path User Guide
Design requirements

- ASPES uses data from an experimental evaluation
- Data must Include:
  - Outcome of interest
  - An indicator for treatment/control status
  - Measure of the mediator of interest
  - Baseline data that can be used to model the endogenous subgroups of interest
- Sufficient Sample Size:
  - For Predicted Subgroups: A sample size of at least 560 is needed to detect an effect size of 0.30 or larger
  - For Actual Subgroups: A sample size of at least 3,380 is needed to detect an effect size of 0.30 or larger (assuming correct placement rates of 65 percent)
ASPES Method in Practice: Outline

- Design requirements
- *ASPES example using data from the Comprehensive Teacher Induction Study (Glazerman et al., 2010)*
- Introduction to SPI-Path User Guide
Comprehensive Teacher Induction (CTI) Study

- In 2004, the U.S. Department of Education’s Institute of Education Sciences contracted with Mathematica Policy Research to conduct the Comprehensive Teacher Induction (CTI) Study.

- CTI Study Design: 418 elementary schools in 17 urban districts were assigned by lottery to either:
  1. a treatment group whose beginning teachers were offered comprehensive teacher induction or
  2. a control group whose beginning teachers received the district’s “business as usual” induction services

- See *Impacts of Comprehensive Teacher Induction*, Glazerman et al. (2010)
For teachers who received two years of comprehensive induction:

- There was no impact on student achievement in the first two years
- In the third year, there was a positive and statistically significant impact on student math and reading achievement (Glazerman et al., 2010)
One key component of teacher induction programs is mentorship, or personal guidance from experienced teachers.

Mentorship activities include:
- Observing instruction or providing a demonstration lessons;
- Reviewing lesson plans, instructional materials, or student work; or
- Delivering constructive feedback (Glazerman et al., 2010).

**Research Question:** What role did mentorship play in improving student achievement outcomes?
## Research Questions and Methods

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Method Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the impact of CTI on students taught by teachers who are predicted to receive a high [low] dosage of mentorship?</td>
<td>Discrete Version of the ASPES Method (predicted subgroup impacts)</td>
</tr>
<tr>
<td>What is the impact of CTI on students taught by teachers who receive a high [low] dosage of mentorship?</td>
<td>Discrete Version of the ASPES Method (actual subgroup impacts)</td>
</tr>
<tr>
<td>How does mentorship for beginning teachers influence the impact of CTI on student outcomes?</td>
<td>Continuous Version of the ASPES Method</td>
</tr>
</tbody>
</table>
ASPES Stage 1: Predict Mentorship Receipt

- The first stage of the ASPES analysis involves employing a strategy that ensures the symmetric prediction of the mediator of interest for the treatment and control groups using baseline covariates.
CTI Application: Measures

- **Mediator of interest:** We constructed a continuously-defined proxy for mentorship defined as follows:
  - *The Average Number of Times Teacher was Observed Teaching by Mentor in Past Three Months (Averaged over Fall Year 1, Spring Year 1, Fall Year 2, and Spring Year 2)*

- **Baseline characteristics:** Teacher background data used for prediction (e.g., teacher professional backgrounds, current teaching assignments, and demographic characteristics)
Which baseline characteristics to include in the prediction model?

- Strategy 1: The “kitchen sink” approach to covariate selection to achieve “best” prediction
- Strategy 2: Use empirical approach to select covariates that are strong predictors of mediator
- Strategy 3: Include baseline covariates that strongly predict mediator, but otherwise bear little relationship to impact magnitude
  - Seeking “instrumental variables as predictors that affect impact through mediator but not by other means” (Bell and Peck, 2013)
The mediator, which is the outcome of interest in the prediction model, is defined at the teacher-level.

Issue: relatively few degrees of freedom at the prediction stage (220 teacher-level observations).

Implication: must be selective in choosing which teacher-level covariates to include as covariates at the prediction stage.

Solution: use a backward selection procedure to strategically select the set of covariates included in the prediction model.
<table>
<thead>
<tr>
<th>Baseline Covariates Selected for Inclusion in the Prediction Model</th>
<th>Baseline Covariates Selected for Inclusion in the Prediction Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Demographic characteristics</td>
<td>Teacher Demographic characteristics</td>
</tr>
<tr>
<td>Age</td>
<td>Age</td>
</tr>
<tr>
<td>Age-squared</td>
<td>Age-squared</td>
</tr>
<tr>
<td>Male teacher</td>
<td>Male teacher</td>
</tr>
<tr>
<td>Teacher is Hispanic or Latino</td>
<td>Teacher is Hispanic or Latino</td>
</tr>
<tr>
<td>Teacher is black</td>
<td>Teacher is black</td>
</tr>
<tr>
<td>Teacher is white</td>
<td>Teacher is white</td>
</tr>
<tr>
<td>Married</td>
<td>Married</td>
</tr>
<tr>
<td>Any children living in the home</td>
<td>Any children living in the home</td>
</tr>
<tr>
<td>Number of children under 18 years in the home</td>
<td>Number of children under 18 years in the home</td>
</tr>
<tr>
<td>Teacher Professional Background Characteristics</td>
<td>Teacher Professional Background Characteristics</td>
</tr>
<tr>
<td>Has Master’s or Doctoral degree</td>
<td>Has Master’s or Doctoral degree</td>
</tr>
<tr>
<td>Earned a Bachelor’s degree from a highly selected college</td>
<td>Earned a Bachelor’s degree from a highly selected college</td>
</tr>
<tr>
<td>Earned a degree with education-related major or minor</td>
<td>Earned a degree with education-related major or minor</td>
</tr>
<tr>
<td>Entered profession through traditional route</td>
<td>Entered profession through traditional route</td>
</tr>
<tr>
<td>Career changer</td>
<td>Career changer</td>
</tr>
<tr>
<td>Late hire during the school year</td>
<td>Late hire during the school year</td>
</tr>
<tr>
<td>First year teacher</td>
<td>First year teacher</td>
</tr>
<tr>
<td>Currently pursuing state certification</td>
<td>Currently pursuing state certification</td>
</tr>
</tbody>
</table>

Mediator: Average Number of Times Teacher was Observed Teaching By Mentor in Past Three Months
Baseline Covariates Selected for Inclusion in the Prediction Model (Continued)

<table>
<thead>
<tr>
<th>Teacher Professional Background Characteristics (Cont.)</th>
<th>Baseline Covariates Selected for Inclusion in the Prediction Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any student teaching</td>
<td></td>
</tr>
<tr>
<td>Number of weeks spent student teaching</td>
<td></td>
</tr>
<tr>
<td>Current school year salary</td>
<td>✓</td>
</tr>
<tr>
<td>Any outstanding student loans</td>
<td></td>
</tr>
<tr>
<td>Amount of student loans</td>
<td></td>
</tr>
<tr>
<td>Member of a teacher’s union or professional association</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Teacher College Entrance Exams</strong></td>
<td></td>
</tr>
<tr>
<td>SAT combined score (or ACT equivalent)</td>
<td>✓</td>
</tr>
<tr>
<td>SAT math score</td>
<td></td>
</tr>
<tr>
<td><strong>Teaching Assignments</strong></td>
<td></td>
</tr>
<tr>
<td>Responsible for reading outcomes</td>
<td></td>
</tr>
<tr>
<td>Responsible for math outcomes</td>
<td></td>
</tr>
<tr>
<td>Grade level</td>
<td>✓</td>
</tr>
<tr>
<td>Teaching in preferred grade and subject</td>
<td></td>
</tr>
<tr>
<td><strong>School Characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>Type of school: K-5, K-6 or K-8</td>
<td>✓</td>
</tr>
<tr>
<td>District</td>
<td>✓</td>
</tr>
<tr>
<td>District X Grade</td>
<td>✓</td>
</tr>
</tbody>
</table>

Mediator: Average Number of Times Teacher was Observed Teaching By Mentor in Past Three Months
This assumption may be violated if the baseline characteristics $X$ (e.g., teacher salary, SAT scores) used to predict the mediator $M_1^P$ influence the impact $\Delta$ through channels other than the actual number of observations $M_1^A$.
Performance of the Prediction Model

Relationship Between Actual and Predicted Mentorship

The Average Number of Times Teacher was Observed Teaching By Mentor in Past Three Months

(1)

**Predicted Mediator**

0.976***

(0.025)

**T-Statistic**

38.68

**Number of Teachers**

220

**Number of Schools**

90

**Number of Districts**

10

**R-Squared**

0.871

Sample limited to teachers in the treatment group. Standard errors clustered at the school level. *** p<0.01. Reported sample sizes are rounded to the nearest 10 to minimize disclosure risk.
Gauging Prediction Success

- We regressed the Actual Mediator on the Predicted Mediator for observations in the treatment group.
- The T-statistic of 38.7 indicates a strong relationship between the actual and predicted values of the mediator.
- The regression coefficient of 0.98 indicates that increasing the predicted mediator by one unit is associated with a 0.98 increase in the actual mediator, representing a near one-to-one correspondence.
## Research Questions and Methods

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Method Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the impact of CTI on students taught by teachers who are predicted to receive a high [low] dosage of mentorship?</td>
<td>Discrete Version of the ASPES Method (predicted subgroup impacts)</td>
</tr>
</tbody>
</table>
For discrete ASPES method: Define Subgroups of Interest

- Create “interesting” subgroups
- Ensure sufficient sample sizes in each predicted subgroup
- In the CTI Study application:
  - Predicted high dosage subgroup includes teachers in the treatment and control groups who are predicted to receive at least the median dose of mentorship
  - Predicted low dosage subgroup includes teachers predicted to receive less than the median dose of mentorship
- Percent of treatment group members predicted to be in their true subgroup (correct placement rate): 73 percent
Stage 2: Estimate Impacts on Predicted Subgroups

### Treatment
- Predicted Low Dosage
- Predicted High Dosage

### Control
- Predicted Low Dosage
- Predicted High Dosage
## Impacts on Predicted Subgroups

<table>
<thead>
<tr>
<th></th>
<th>Math Achievement</th>
<th>Reading Achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTI Study Impact for pooled sample</td>
<td>0.20***</td>
<td>0.11**</td>
</tr>
<tr>
<td>Predicted High Dosage Subgroup</td>
<td>0.360***</td>
<td>0.241***</td>
</tr>
<tr>
<td></td>
<td>(0.096)</td>
<td>(0.078)</td>
</tr>
<tr>
<td>Predicted Low Dosage Subgroup</td>
<td>-0.020</td>
<td>-0.138</td>
</tr>
<tr>
<td></td>
<td>(0.092)</td>
<td>(0.112)</td>
</tr>
</tbody>
</table>

Notes: *p<0.10, ** p<0.05, *** p<0.01. CTI Study Impact from Glazerman et al. (2010).
The CTI intervention had a large positive impact on math and reading achievement for students taught by teachers most likely to receive a high dosage of mentorship.

No effect on students taught be teachers who are predicted to receive comparatively little mentorship.

Impacts on the predicted high dosage subgroup are larger in magnitude than CTI Impacts using the full sample.
Predicted subgroup impacts:
- Are asymptotically unbiased (Harvill, Peck & Bell, 2013)
- Not everyone in the predicted high dosage subgroup actually received a high dosage
- Provide estimate of CTI’s impact on students taught by teachers who are most likely to receive a high (or low) dosage of mentorship

Researchers may be more interested in impacts on actual subgroups (e.g., those who actually received a high dosage of mentorship)
### Research Questions and Methods

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Method Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the impact of CTI on students taught by teachers who receive a high [low] dosage of mentorship?</td>
<td>Discrete Version of the ASPES Method (actual subgroup impacts)</td>
</tr>
</tbody>
</table>
Stage 3: Convert from Predicted to Actual Impacts

### Treatment

- **Actual**
  - Low Dosage
  - High Dosage

### Control

- **Actual**
  - Low Dosage
  - High Dosage
## Comparison to Impacts on Actual High Dosage Subgroup

<table>
<thead>
<tr>
<th></th>
<th>Math Achievement</th>
<th>Reading Achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CTI Study Impact for pooled sample</strong></td>
<td>0.20***</td>
<td>0.11**</td>
</tr>
<tr>
<td><strong>Predicted High Dosage Subgroup</strong></td>
<td>0.360*** (0.096)</td>
<td>0.241*** (0.078)</td>
</tr>
<tr>
<td><strong>Actual High Dosage Subgroup</strong></td>
<td>0.691*** (0.197)</td>
<td>0.504*** (0.153)</td>
</tr>
</tbody>
</table>

Notes: *p<0.10, **p<0.05, ***p<0.01.
CTI Study Impact from Glazerman et al. (2010).
Comparison to Impacts on Actual High Dosage Subgroup: Summary of Findings

- Impacts on actual high dosage subgroup is larger in magnitude than CTI Study impacts and predicted subgroup impacts.
- Standard errors on actual high dosage subgroup impacts are large, limiting our ability to reject more modest effects.
- Unbelievably large impacts on actual subgroups may indicate that ASPES conversation assumptions are not satisfied.
Predicted Subgroup Impacts Vs. Actual Subgroup Impacts

- Often greater interest in impacts on actual subgroups
- Requires additional assumptions to be considered asymptotically unbiased
- Requires larger sample sizes
Research Questions and Methods

<table>
<thead>
<tr>
<th>Research Question</th>
<th>Method Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>How does mentorship for beginning teachers influence the impact of CTI on student outcomes?</td>
<td>Continuous Version of the ASPES Method</td>
</tr>
</tbody>
</table>
Stage 2: Impact Model

- Estimate the Relationship between the Predicted Mediator and Effect Size:
  \[
  Y = \beta_0 + \beta_1 \hat{M}^P + \beta_2 T + \beta_3 T \hat{M}^P + \varepsilon_2
  \]
  - \(Y\) is the outcome being examined;
  - \(\hat{M}^P\) is the predicted value of the mediator generated from Stage 1;
  - \(T\) indicates whether the member was assigned to the treatment or control group; and
  - \(\varepsilon_2\) is an error term that captures all other factors that influence the outcome.

- The impact of being assigned to the treatment group is given by the following equation:
  \[
  \frac{\partial Y}{\partial T} = \beta_2 + \beta_3 \hat{M}^P
  \]
The effect of mentorship on the impact of CTI

<table>
<thead>
<tr>
<th>The Average Number of Times Teacher was Observed Teaching By Mentor in Past Three Months</th>
<th>Math Achievement</th>
<th>Reading Achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.266**</td>
<td>0.278***</td>
</tr>
<tr>
<td></td>
<td>(0.109)</td>
<td>(0.088)</td>
</tr>
</tbody>
</table>

Notes: *p<0.10, ** p<0.05, *** p<0.01.
The effect of mentorship on the impact of CTI: Summary of Findings

- A one unit increase in the *Average Number of Times Teacher was Observed Teaching By Mentor in Past Three Months* is akin to jumping from the ~25th percentile to the 75th percentile in terms of the amount of mentorship received.

- This increase in mentorship is associated with a large increase in the impact of CTI on both math and reading achievement:
  - Boosts the CTI Impact on math by 27 percent of a standard deviation
  - Boosts the CTI Impact on reading by 28 percent of a standard deviation
ASPES Method in Practice: Outline

- Design requirements
- ASPES example using data from the Comprehensive Teacher Induction Study (Glazerman et al., 2010)
- *Introduction to SPI-Path User Guide*
Social Policy Impact Pathfinder (SPI-Path) User Guide

- Describes the ASPES method in detail
- Includes SAS and Stata code for conducting analysis
- Provides sample table shells and interpretation assistance
- Practical considerations and examples from literature throughout
- **Download User Guide:**

- **Stay in touch:**
  - [Eleanor_Harvill@abtassoc.com](mailto:Eleanor_Harvill@abtassoc.com) (301) 347-5638
  - [Shawn_Moulton@abtassoc.com](mailto:Shawn_Moulton@abtassoc.com) (617) 520-2459
  - [Laura_Peck@abtassoc.com](mailto:Laura_Peck@abtassoc.com) (301) 347-5537