Causal detailed decompositions

Philipp Ketz, Blaise Melly and Miguel Salema

Paris School of Economics and University of Bern

2025 Swiss Stata Conference

Motivation

- Decomposition methods are widely used in social sciences:
  - Wage gaps, test score gaps, employment gaps, etc.
- Aggregate decompositions:
  - Existing methods (e.g. Oaxaca-Blinder, distributional methods) have a clear causal interpretation under ignorability and common support.
- Beyond aggregate decompositions:
  - Researchers often want to understand the contribution of each variable or group of variables to the explained component.
- Problem:
  - Detailed decompositions are treated as accounting exercises, without a clear causal interpretation.

What we do

- Show that naive detailed Oaxaca–Blinder decompositions are not causally interpretable.
- Provide a detailed decomposition with a transparent causal meaning.
- Key ingredients:
  - A triangular (recursive) structure for the covariates.
  - A sequential independence assumption.
- Develop a general nonparametric, distributional framework that applies to linear, nonlinear and quantile decompositions.
- Show that the method can be implemented with standard commands for aggregate decompositions.

Running example: black–white test score gap

- Outcome: standardized test scores of children.
- Groups: black ($G = b$) and white ($G = w$).
- Key covariates:
  - Parental income I
  - School quality S
- Typical questions:
  - How much of the test score gap is explained by income and school quality?
  - Within the explained component, how much is due to income and how much to schooling?

Aggregate Oaxaca–Blinder decomposition

- Two groups $G \in \{0, 1\}$, outcome Y, covariates X.
- Separate linear models:

$$E[Y \mid X, G = g] = X'\beta_g.$$

- Aggregate Oaxaca–Blinder decomposition:

$$\bar{Y}_1 - \bar{Y}_0 = \underbrace{\left(\bar{X}_1'\hat{\beta}_1 - \bar{X}_1'\hat{\beta}_0\right)}_{\text{unexplained / structure}} + \underbrace{\left(\bar{X}_1'\hat{\beta}_0 - \bar{X}_0'\hat{\beta}_0\right)}_{\text{explained / endowments}}.$$

- Under linearity and ignorability:
  - The unexplained component is a causal group effect for group 1. It is called the average treatment effect on the treated when G is a treatment.
  - The explained component is the causal effect of changing covariate distributions from group 0 to group 1.

Naive detailed decomposition for characteristics

- Suppose $X = (I, S)$. A common practice is:

$$(\bar{X}_1 - \bar{X}_0)'\hat{\beta}_0 = (\bar{S}_1 - \bar{S}_0)\hat{\beta}_0^S + (\bar{I}_1 - \bar{I}_0)\hat{\beta}_0^I.$$

- Interpreted as contribution of S and I to the explained component.
- Problems:
  - No clear counterfactual experiment: changing S without adjusting I is not well defined when they are causally linked.
  - If S is a causal consequence of I, part of the effect of I is mechanically allocated to S.
- Conclusion:
  - Naive detailed Oaxaca–Blinder is an accounting device.

Illustrative linear example

- Consider the following system in each group $g \in \{0,1\}$:

$$S_g = \gamma_g^C + \gamma_g^I I_g + U_g^S,$$

$$I_g = \delta_g^C + \delta_g^S S_g + U_g^I.$$

- Outcome equation in group g:

$$Y_g = \beta_g^C + \beta_g^S S_g + \beta_g^I I_g + U_g^Y.$$

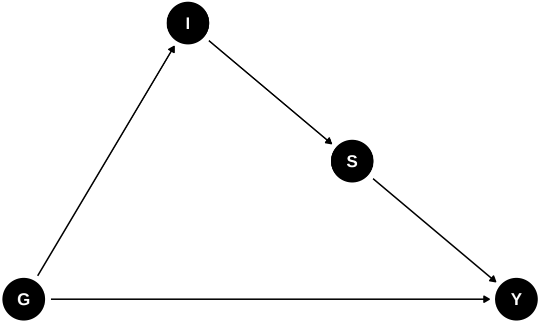
- Even in this simple system:

$$\left(\bar{S}_1 - \bar{S}_0\right) \hat{\beta}_0^S \underset{p}{\rightarrow} = \left(\frac{\gamma_1^S + \gamma_1^I \delta_1^S}{1 - \gamma_1^I \delta_1^S} - \frac{\gamma_0^S + \gamma_0^I \delta_0^S}{1 - \gamma_0^I \delta_0^S} \right) \beta_0^S,$$

$$\left(\bar{I}_1 - \bar{I}_0\right) \hat{\beta}_0^I \underset{p}{\rightarrow} = \left(\frac{\delta_1^C + \delta_1^S \gamma_1^C}{1 - \delta_1^S \gamma_1^I} - \frac{\delta_0^C + \delta_0^S \gamma_0^C}{1 - \delta_0^S \gamma_0^I} \right) \beta_0^I.$$

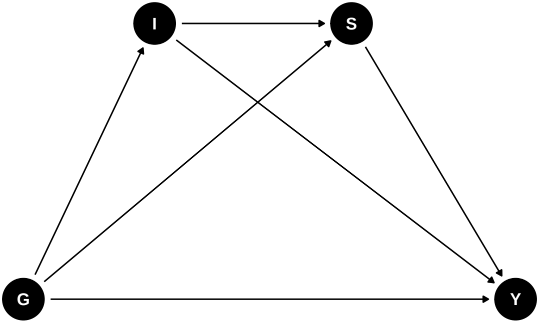
No causal interpretation

- Unless I and S are uncorrelated, both naive detailed decomposition terms depend on the whole system.
- Even in this example, the detailed decomposition will attribute a part of the difference to S:



Identification by triangularity

- Without further assumptions the causal detailed decomposition is not identified.
- In this paper, we impose a causal ordering of the covariates (triangular structure):



Linear model with triangularity

- With this model, the causal effect of exogenously changing I is

$$\left(\delta_1^C - \delta_0^C\right)\left(\beta_0^W + \gamma_0^W \beta_0^V\right)$$

and the effect of changing S is

$$\left(\gamma_1^C - \gamma_0^C + \left(\gamma_1^W - \gamma_0^W\right)\delta_1^C\right)\beta_0^V.$$

- Note that the sum of these two components is equal to the total explained difference.
- This decomposition can be implemented by incorporating the regressors sequentially: first only I, then I and S.
- This sequential procedure is often used in practice without justification.

General nonparametric framework

- Groups: $G \in \mathcal{G} = \{b, w\}$ (black and white).
- Covariates:
  - Parental income I.
  - Schooling S.
- Potential outcomes:

$$Y(g, i, s) \text{ is the outcome if } G = g, I = i, S = s.$$

- Potential covariates:
  - $I(g, s)$ is potential income if $G = g$ and $S = s$.
  - $S(g, i)$ is potential schooling if $G = g$ and $I = i$.

Triangularity and building blocks

$$I(g, s) = I(g) \quad \forall g \in \mathcal{G}.$$

- Parental income is determined before schooling.
- By the observation rule:

$$I = I(G), \quad S = S(G, I(G)), \quad Y = Y(G, I(G), S(G, I(G))).$$

- Building block counterfactuals:

$$Y(g, I(g'), S(g'', I(g')))$$

  where $g, g', g'' \in \mathcal{G}$.

- These objects allow us to decompose the gap and attribute parts of it to each covariate.

Sequential independence assumption

For any g, g', g'', i, s, we assume:

1. $Y(g, i, s), S(g'', i), I(g') \perp G$.
2. $Y(g, i, s), S(g'', i) \perp I \mid G$.
3. $Y(g, i, s) \perp S \mid G, I$.

- Interpretation:
  - Potential income is independent of group, once we fix the group in the potential notation.
  - Conditional on group and income, potential schooling does not depend on the realized value of income beyond that.
  - Conditional on group and covariates, potential outcomes are independent of the realized covariates.
- This is a sequential version of unconfoundedness along the causal chain $G \to I \to S \to Y$.

Identification: key result

Proposition

Under triangularity, sequential independence and common support,

$$F_{Y(g,I(g'),S(g'',I(g')))}(y) = F_{Y\langle g,g',g''\rangle}(y)$$

where $F_{Y\langle g,g',g''\rangle}(y)$

$$\int\int F_Y(y \mid G = g, I = i, S = s)\, dF_S(s \mid G = g'', I = i)\, dF_I(i \mid G = g').$$

- Right-hand side is expressed entirely in terms of observable conditional distributions.
- This gives an identified representation of the cross-world counterfactuals $Y(g, I(g'), S(g'', I(g')))$.

From building blocks to decompositions

- Total gap:

$$F_{Y\langle b,b,b\rangle}(y) - F_{Y\langle w,w,w\rangle}(y).$$

- Decompose into:
  - Outcome structure effect (change g holding covariates at w).
  - Endowment effect (change covariates from w to b holding structure fixed).
- Further decompose the endowment effect:
  - Effect of changing income:

$$F_{Y\langle b,b,b\rangle}(y) - F_{Y\langle b,w,b\rangle}(y).$$

  - Effect of changing schooling:

$$F_{Y\langle b,w,b\rangle}(y) - F_{Y\langle b,w,w\rangle}(y).$$

- Each term has a clear causal interpretation as a change in one structural equation.

Estimation and inference

1. Plug-in estimation
   - Parametric or nonparametric estimation of conditional distributions (e.g. quantile regression, distribution regression).
   - Or reweighting methods.
2. Sequential aggregate decompositions using an increasing set of regressors.
   - In Stata, the detailed decomposition can be implemented by repeatedly calling the command `cdeco` from the `counterfactual` package.

- Inference:
  - Functional central limit theorems.
  - Bootstrap; all steps of the procedure must be bootstrapped jointly.

Relation to mediation analysis

- Closely related to the mediation literature:
  - Daniel et al. (2015), Zhou (2022), among others.
- They study causal pathways $G \to I \to S \to Y$ and identify direct and indirect effects.
- Our angle:
  - Provide a link between mediation analysis and the decomposition literature.
  - Use similar potential outcome structures and assumptions.
  - Provide results for the whole distributions, rather than on average mediation effects.

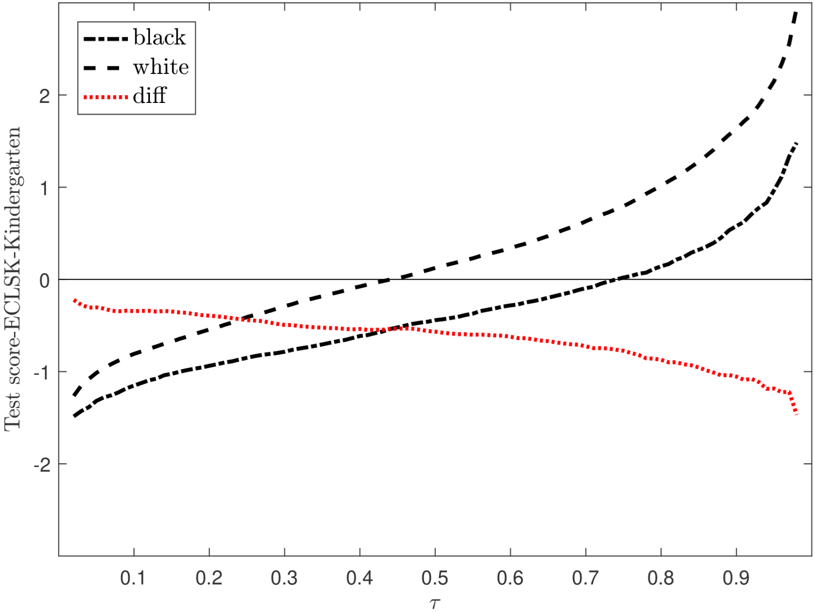
Application: black–white test score gap

- Standard dataset used by Fryer and Levitt (2004): early Childhood Longitudinal Study kindergarten cohort (ECLSK) - 1998/99.
- Outcome: Item Response Theory (IRT) test scores in math and reading.
- Decomposition:
  - Aggregate gap in the distribution of test scores.
  - Outcome structure vs characteristics.
  - Within characteristics: four groups of variables.
- Implementation: quantile regression.

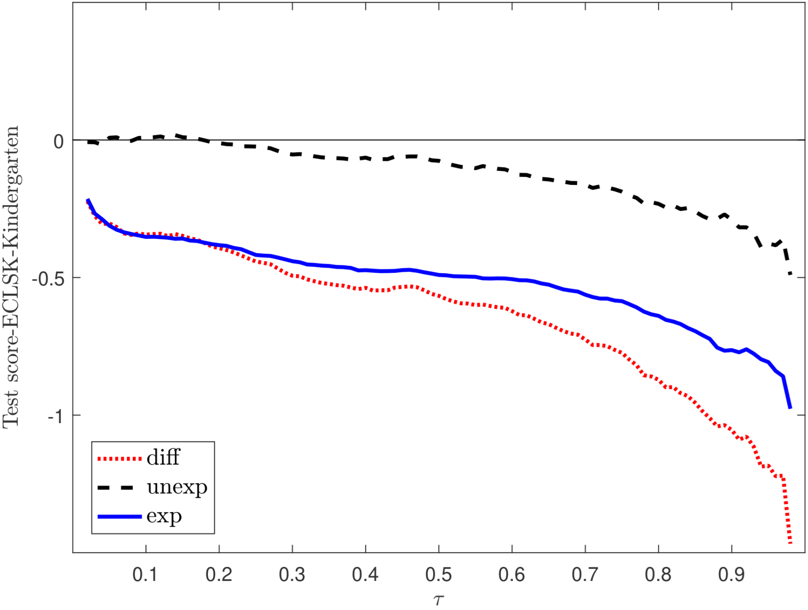
Covariates

1. Socioeconomic variables: SES, WIC (Women, Infant, and Children — Food and Nutrition Service), WIC for mother
2. Variables determined at birth: sex, birth weight, indicators for teenage mother and mother above 30 at first birth
3. Variables measuring the home environment: number of children books (and its square)
4. School quality variables: school fixed effects¡

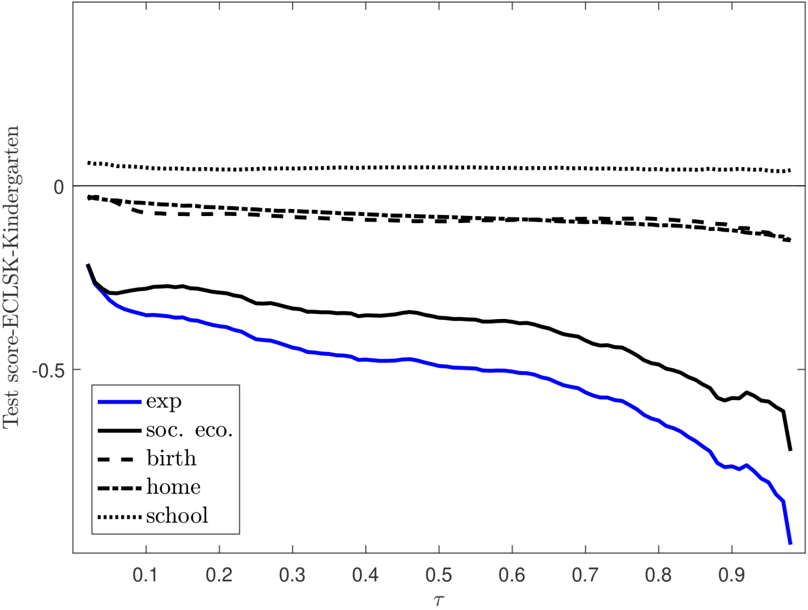
Total gap in the math score



Aggregate decomposition



Detailed decomposition



Conclusion and limitations

- Aggregate decompositions have a clear causal interpretation under standard assumptions.
- Detailed decompositions require additional structure.
- Our approach:
  - Uses a triangular structure and sequential independence to identify causal contributions of individual covariates.
  - Links sequential decompositions used in practice to an explicit structural model.
- Limitations:
  - Requires a credible causal ordering of covariates.
  - Sequential independence is strong and may require rich conditioning sets.