Discussion of *The Gender Diversity Gap in Mathematics*

Presenter
Francine Montecinos
Columbia University

Discussant
Tyler Sotomayor
Columbia University

Pre-doctoral Mini Conference
Program for Economic Research
May 29, 2026

COLUMBIA | ECONOMICS
Program for Economic Research

What the paper does

Question

How do mathematics performance and confidence differ across cisgender, transgender, and non-binary students?

Design

Chilean SIMCE panel linking 4th- and 10th-grade records for about 123,000 students; compares students within schools while controlling for baseline math and confidence.

Extensions

Disaggregates six sex-at-birth/gender-identity groups, studies distributional heterogeneity, and examines discrimination, aggression, religious curricula, and school-assignment evidence from centralized admissions.

Where this paper fits

Strand	Key papers	What they establish
Gender gaps in math	Fryer–Levitt (2010); Bharadwaj et al. (2016); Paredes (2014); Alan et al. (2018); Carlana (2019)	Math gaps emerge over schooling and can reflect teacher bias, classroom norms, and upper-tail differences.
Confidence and self-evaluation	Barber–Odean (2001); Niederle–Vesterlund (2007); Coffman (2014); Exley–Kessler (2022)	Beliefs about ability can diverge from performance and matter for later choices.
Identity and measurement	Akerlof–Kranton (2000); NASEM (2022); Brenøe et al. (2022); Coffman et al. (2025)	Gender identity is economically relevant; binary coding can mask meaningful heterogeneity.
LGBTQ+ students and gender minorities	Carpenter (2009); Sansone (2019); Aksoy et al. (2024); Banan et al. (2024)	Gender-minority gaps are documented, but evidence in administrative school settings is still thin.

Positioning

Sits at the intersection of all four strands. Closest benchmark is Aksoy, Exley & Kessler (2024) on U.S. gender-minority confidence gaps in a study/survey setting; this paper brings the question to linked administrative records and four distinct gender-minority categories.

Main takeaway

Gender diversity gaps are multidimensional: achievement gaps are heterogeneous, while confidence gaps are broad, persistent, and closely tied to school experiences.

Achievement

- Gaps differ sharply by identity.
- Quantile patterns matter; averages miss structure.

Confidence

- Gaps are wider for transgender and non-binary students.
- They persist conditional on own math score.

Environment

- Discrimination and aggression explain confidence more than scores.
- Religious-program evidence points to school context.

Implication for the discussion

The key empirical object is the confidence gap conditional on achievement; the main design questions are about interpretation and mechanisms.

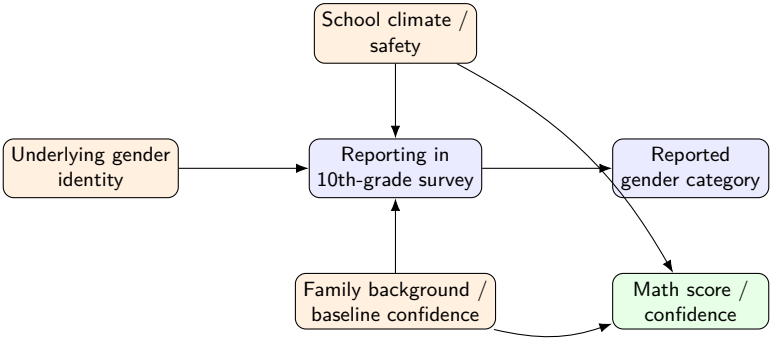
Comment 1: Clarify the estimand

- School FE + 4th-grade math/confidence controls handle baseline composition and pre-existing achievement gaps.
- But gender identity is reported *only in 10th grade* — there is no panel of identity.
- The clean estimand: a within-school conditional gap by *reported 10th-grade gender category*, not a developmental trajectory from 4th grade.
- Reporting itself can vary with within-school safety, peer climate, family support, and confidence.

Suggestion

Use language like "students who report a trans or non-binary identity in 10th grade." For each later-identified group, show 4th-grade math, 4th-grade confidence, and 10th-grade confidence in one figure: does the gap pre-exist disclosure or co-emerge with it?

DAG: reporting can be selected

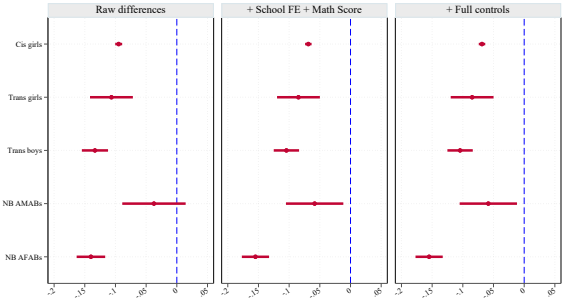


Implication

Estimated gaps are by *reported 10th-grade identity*; school climate can shape both reporting and outcomes.

Comment 2: Put confidence at the center

- The paper's cleanest new fact is about confidence, not only test scores.
- Confidence gaps are large, stable across specifications, and persist after conditioning on own 10th-grade math score.
- This connects the paper to self-assessment, STEM investment, course-taking, and later educational choices.



Presenter evidence: confidence gaps relative to cis boys

Suggestion

Make confidence the headline outcome: show score, GPA, and confidence gaps side by side, and bring the ordered-confidence results into the main text.

Comment 3: The mechanism decomposition

Headline finding

Discrimination and aggression explain up to 30–40% of confidence gaps for trans men and NB-AMAB, operating via differences in *exposure* rather than *returns*.

Three concerns

- *Common-method bias.* Discrimination, aggression, and confidence all from the same student, same survey, same moment.
- *Reverse causality.* Low confidence → social withdrawal → perceived exclusion.
- *Endogenous interpretation.* What counts as discrimination is itself filtered through the respondent.

Suggestion

Build a leave-one-out peer-level discrimination rate at the school × cohort level as a more exogenous exposure proxy.

Or

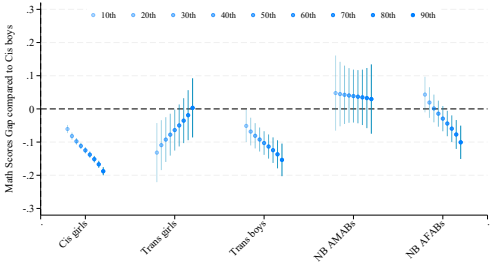
Lean on visible-discrimination measures (looks / expression) — the least interpretively filtered category.

Comment 4: A structural asymmetry worth confronting

The pattern

- AFAB groups (cis girls, trans men, NB-AFAB) widen at upper quantiles.
- Trans women (AMAB) concentrate in the lower tail.
- NB-AMAB is statistically flat.

At least two distinct processes.



Presenter quantile evidence: score gaps relative to cis boys

Descriptive characterization + a puzzle

Within-natal-sex comparisons (cis girls → trans men → NB-AFAB) can *characterize* this pattern descriptively — but can't isolate a causal effect of gender-nonconformity, since identity disclosure is itself selected (Comment 1). **The puzzle:** trans men show *larger* math gaps than cis girls — the opposite of naive identity-prescription. Reconciliation likely runs through selection-into-disclosure and minority-stress channels, not pure identity payoffs.

Comment 5: Use preferences to sharpen school context

Problem

Religious-program attendance is selected. Families who rank or attend these schools may differ in values, information, constraints, and safety concerns.

What it helps identify

Among students with similar observables and similar revealed school preferences, are gender gaps still different in religious-program schools?

Control function

SAE rank lists reveal relative demand for religious versus non-religious options. The control function summarizes that latent demand and adds it to the outcome model.

Why plausible

Chile's centralized ranking mechanism is plausibly strategy-proof: conditional on feasible options and information, families have weak incentives to rank schools in preference order.

Takeaway

This is a strong addition: it moves the religious-program section from raw heterogeneity toward selection-adjusted evidence of school-context moderation.

Where to take this next

Caveat: causal inference here is genuinely fraught

Gender itself isn't a manipulable treatment. Major Chilean LGBTQ+ laws (Ley 21.120 gender identity 2019; Ley 21.400 marriage 2022) rolled out nationwide and simultaneously — thin DID variation. Comuna-level rollouts (MOVILH *Sello de Diversidad*, *Oficinas de Diversidad*, local ordenanzas) exist but are politically selected by mayoral coalition. **The descriptive contribution here is the right step at this stage of the field.**

Most tractable: downstream outcomes

SIMCE links to PSU/PAES, university enrollment, and STEM major. Does *residual* confidence (orthogonal to score) predict STEM avoidance? Doesn't require new identification — turns this into a paper about life trajectories.

Where causal traction may eventually come

Future SIMCE cohorts spanning Ley 21.120; hand-coded comuna policy adoption with comuna FE + parallel-trends defense, ideally with mayoral-election timing as an IV; or cross-country comparisons where rollout sequencing differs. Teacher–student gender match (already in SIMCE) is a smaller separate exercise.

Closing

Bottom line

This is an important measurement contribution to the economics of education: linked SIMCE records make gender-diversity gaps in math achievement, confidence, and school context visible at national scale.

- The six-category gender measure is the central strength.
- The confidence result is the clearest new fact: gaps persist even conditional on achievement.
- The SAE control function is a valuable addition because it takes selection into school environments seriously.
- A broader takeaway is the need for better data on underrepresented LGBTQ+ populations: this paper shows how much becomes visible when measurement improves.