Online Appendix to Accompany

"Tackling Biased Beliefs to Foster Women's Entry into Economics: Experimental Evidence from Mexico"

Stanislao Maldonado
Tecnológico de Monterrey

Edgar Castro Méndez
Tecnológico de Monterrey

Elvira Naranjo
Tecnológico de Monterrey

Joana Chapa
Universidad Autónoma de Nuevo León

This version: October 31, 2026

Not for Publication

Contents

A Formal Framework and Proofs **A-5**

A.1 Environment and Baseline Decision A-5

A.2 Treatment-Induced Updating A-6

A.3 Motivated Processing and Messenger Relevance A-7

A.4 Applications as a Threshold Outcome A-8

A.5 Heterogeneity by Prior Beliefs and Gender A-9

A.6 Direct Exposure, Spillovers, and Persistence A-10

A.7 Mapping the Model to the Empirical Analysis A-11

B Outcomes and Index Construction **A-13**

B.1 Primary Outcome A-13

B.2 Survey Domains and Wave-Specific Measurement A-13

B.3 Standardization and Index Weights A-14

B.4 Missing Responses and Multiple Testing A-15

C Interpretation and Statistical Power **A-16**

C.1 What the short-term wave can and cannot show A-16

C.2 What the realized design can detect A-17

References **A-18**

D Survey Instruments **A-19**

D.1 Baseline Survey A-19

D.2 Immediate Post-Intervention Survey A-23

D.3 Follow-Up Survey A-28

D.4 Complementary Online Endline Survey (August 2026) A-33

Online Appendix Figures **A-36**

Online Appendix Tables **A-44**

List of Figures

A1 Representative Materials from the Implemented Role-Model Presentation A-36

A2 Screenshots from the Tec Intervention Videos A-37

A3 Screenshots from the UANL Intervention Videos A-38

A4 Screenshots from the Interviews with Women Economists A-39

A5 Two-Stage Assignment and Direct-versus-Spillover Contrasts A-40

A6 Locations of Study Schools A-41

A7 Common-Item Persistence in the Balanced Panel A-42

A8 Leave-One-School-Out Estimates A-43

List of Tables

A1 Exposure, Recall, and Implementation Process A-44

A2 Randomization, Analytic Samples, and Realized Power A-45

A3 Baseline Survey Completion Relative to Classroom Enrollment A-46

A4 Baseline Gender Gaps A-47

A5 Scale Items, Non-Response, and Reliability: Baseline A-48

A6 Scale Items, Non-Response, and Reliability: Immediate Post-Intervention A-49

A7 Scale Items, Non-Response, and Reliability: Follow-Up A-50

A8 Item-Level Effects: Beliefs about Economics A-51

A9 Item-Level Effects: Identity and Self-Concept A-52

A10 Item-Level Effects: Intentions and Aspirations A-53

A11 Item-Level Effects: Gender Stereotypes A-54

A12 Persistence for Items Common across Waves A-55

A13 Applications by Institution: Women in the Final Semester A-56

A14 Applications by Institution: All Final-Semester Students A-57

A15 Administrative Linkage and Self-Report Validation A-58

A16 Sources of the Combined Application Measure A-59

A17 Applications to Economics by Gender A-60

A18 Survey Outcomes: Direct and Spillover Effects, Immediate Post-Intervention A-61

A19 Survey Outcomes: Direct and Spillover Effects, Follow-Up A-62

A20 Applications to Economics: Direct and Spillover Effects A-63

A21 Heterogeneity by Baseline Beliefs about Economics A-64

A22 Heterogeneity by Baseline Interest in Economics A-65

A23 Heterogeneity by Baseline Mathematics Confidence A-66

A24 Heterogeneity by Student Gender A-67

A25 Differential Attrition A-68

A26 Attrition Corrections: Inverse-Probability Weighting, Follow-Up A-69

A27 Attrition Corrections: Inverse-Probability Weighting, Immediate Post-Intervention . .A-70

A28 Attrition Corrections: Lee Bounds A-71

A29 Full Sample versus Balanced Panel A-72
A30 Main Estimates without Covariate Adjustment A-73
A31 Alternative ANCOVA Specifications: Immediate Post-Intervention Outcomes A-74
A32 Alternative ANCOVA Specifications: Follow-up Outcomes A-75
A33 Alternative ANCOVA Specifications: Administrative Outcomes A-76
A34 Alternative Index Constructions: Immediate Post-Intervention Outcomes A-77
A35 Alternative Index Constructions: Follow-up Outcomes A-78
A36 Alternative Definitions of an Economics Application: Women in the Final Semester . .A-79
A37 Alternative Definitions of an Economics Application: All Final-Semester Students . . .A-81
A38 Design-Based and Small-Cluster Inference A-83
A39 Placebo Outcomes and Subject-Specific Benchmarks A-88

A Formal Framework and Proofs

This appendix develops the decision problem summarized in Section 2 of the paper. The model has two purposes. First, it separates the three margins targeted by the role-model intervention: beliefs about succeeding in economics, the perceived value of the field, and the perceived cost of pursuing it. Second, it shows why treatment effects on survey beliefs, applications, gender heterogeneity, persistence, and within-school spillovers need not move together. The framework is deliberately parsimonious and should be read as an organizing device rather than as a structural model of major choice.

A.1 Environment and Baseline Decision

Student i decides whether to apply to an economics program. Let $d_i^r \in \{0,1\}$ denote this decision in state $r \in \{0,1\}$, where $r = 0$ is the counterfactual without the role-model signal and $r = 1$ is the state after receiving it. The student has underlying analytical ability a_i, which she observes imperfectly. Given information set I_i^r, she holds the posterior mean

$$b_i^r = E[a_i \mid I_i^r]. \tag{A.1}$$

Let $p(b_i^r)$ denote the perceived probability of succeeding in economics, $V_i^r > 0$ the perceived value of studying economics, and C_i^r the perceived cost of pursuing it. The value term includes expected earnings, career opportunities, and the social value the student assigns to the work. The cost term includes application and study costs as well as anticipated failure, discrimination, social disapproval, and lack of belonging. The perceived net return is

$$M_i^r = p(b_i^r)V_i^r - C_i^r, \tag{A.2}$$

and the student applies whenever

$$d_i^r = \mathbf{1}\{M_i^r \geq 0\}. \tag{A.3}$$

Assumption A.1 (Regularity)**.** The function $p : \mathbb{R} \to [0,1]$ is continuously differentiable and satisfies $p'(b) > 0$ for every b. Perceived value is positive in both states: $V_i^r > 0$ for $r \in \{0,1\}$.

Assumption A.1 imposes only that students who believe they are better prepared assign themselves a higher probability of succeeding. It does not require women and men to differ in objective ability. A gender gap in applications can arise from differences in b_i^0, V_i^0, or C_i^0, even when the distribution of a_i is identical.

A.2 Treatment-Induced Updating

The role-model session provides a signal σ_i about economics and about the experience of women in the field. We write the induced changes as

$$b_i^1 = b_i^0 + \delta_{bi}, \qquad V_i^1 = V_i^0 + \delta_{Vi}, \qquad C_i^1 = C_i^0 + \delta_{Ci}. \tag{A.4}$$

The intervention is designed to produce $\delta_{bi} \geq 0$, $\delta_{Vi} \geq 0$, and $\delta_{Ci} \leq 0$: a student becomes more confident about succeeding, assigns greater value to economics, or perceives lower barriers to entering it. These inequalities are hypotheses about the direction of updating, not assumptions imposed on the empirical estimates. A student with favorable priors may learn little, an identity-incongruent signal may receive little weight, and an informative session may even lead some students to revise beliefs downward.

The change in the perceived net return follows directly from Equations (A.2) and (A.4):

$$\begin{aligned} \Delta M_i &\equiv M_i^1 - M_i^0 \\ &= \left[p(b_i^1) - p(b_i^0)\right] V_i^0 + p(b_i^1)\delta_{Vi} - \delta_{Ci}. \end{aligned} \tag{A.5}$$

The first term is the ability-belief channel, the second is the perceived-value channel, and the third is the perceived-cost channel. The decomposition is exact and includes the interaction between belief and value updating through $p(b_i^1)\delta_{Vi}$.

Proposition A.1 (Direction of the perceived-return effect)**.** *Under Assumption A.1, the perceived net return is increasing in b_i^r and V_i^r and decreasing in C_i^r. If $\delta_{bi} \geq 0$, $\delta_{Vi} \geq 0$, and $\delta_{Ci} \leq 0$, then $\Delta M_i \geq 0$. The inequality is strict whenever at least one shift is strict.*

Proof. From Equation (A.2),

$$\frac{\partial M_i^r}{\partial b_i^r} = p'(b_i^r)V_i^r > 0, \qquad \frac{\partial M_i^r}{\partial V_i^r} = p(b_i^r) \geq 0, \qquad \frac{\partial M_i^r}{\partial C_i^r} = -1.$$

Under the stated sign restrictions, every term on the right-hand side of Equation (A.5) is weakly positive. Strict monotonicity of p and positivity of V_i^0 imply that a strict belief shift produces a strict increase. A strict value shift produces a strict increase whenever $p(b_i^1) > 0$, and a strict cost reduction always does so. Hence $\Delta M_i \geq 0$, with strict inequality under any non-degenerate strict shift. □

Proposition A.1 states the directional prediction generated by favorable updating. Without the sign restrictions, the net effect is ambiguous. This qualification matters empirically because the treatment is a bundle and the experiment does not separately manipulate the three components of Equation (A.5).

A.3 Motivated Processing and Messenger Relevance

To make the motivated-belief channel explicit, suppose the role-model signal contains an assessment s_i relevant to the student's expected success. Let m_i denote the perceived relevance or identity congruence of the messenger, and let $\omega_i(m_i) \in [0, 1]$ denote the weight placed on the signal. A convenient updating rule is

$$b_i^1 = (1 - \omega_i(m_i))b_i^0 + \omega_i(m_i)s_i. \tag{A.6}$$

In a standard Bayesian interpretation, ω_i summarizes relative signal precision. In a motivated-belief interpretation, it may also depend on whether accepting the signal preserves self-image or appears compatible with the student's identity (Bénabou and Tirole, 2002, 2016; Golman, Hagmann and Loewenstein, 2017).

Proposition A.2 (Messenger relevance and belief updating)**.** *Suppose $\omega_i'(m_i) > 0$. If $s_i > b_i^0$, then both b_i^1 and the belief component of M_i^1 increase with messenger relevance m_i. If $s_i < b_i^0$, the inequalities reverse.*

Proof. Differentiating Equation (A.6) yields

$$\frac{\partial b_i^1}{\partial m_i} = \omega_i'(m_i)(s_i - b_i^0).$$

The derivative is positive when $s_i > b_i^0$ and negative when $s_i < b_i^0$. Holding V_i^1 and C_i^1 fixed, the corresponding derivative of perceived net returns is

$$\frac{\partial M_i^1}{\partial m_i} = p'(b_i^1)V_i^1 \frac{\partial b_i^1}{\partial m_i},$$

which has the same sign by Assumption A.1. □

This result clarifies both the appeal and the limitation of a female role model. When the message is favorable, a more relevant messenger can strengthen updating. However, the experiment jointly varies the message, the presenter's profession and seniority, and the salience of a successful woman in economics. It therefore cannot separately estimate $\omega_i'(m_i)$ or attribute the reduced-form effect to messenger identity.

A.4 Applications as a Threshold Outcome

The survey outcomes measure components of M_i^r, whereas the administrative outcome records the discrete decision in Equation (A.3). This distinction generates a direct implication for a rare application outcome.

Proposition A.3 (Marginal applicants)**.** *Suppose $\Delta M_i \geq 0$. Then treatment changes student i's application decision if and only if*

$$-\Delta M_i \leq M_i^0 < 0. \tag{A.7}$$

Consequently,

$$E[d_i^1 - d_i^0] = \Pr\left(-\Delta M_i \leq M_i^0 < 0\right). \tag{A.8}$$

If $\Delta M_i = \Delta > 0$ is constant and M_i^0 has continuous distribution function F_M, the effect equals $F_M(0) - F_M(-\Delta)$.

Proof. When $\Delta M_i \geq 0$, treatment cannot move a student from applying to not applying. A student with $M_i^0 \geq 0$ applies in both states. A student with $M_i^0 < -\Delta M_i$ remains below the threshold

in both states. The decision changes only when $M_i^0 < 0$ but $M_i^1 = M_i^0 + \Delta M_i \geq 0$, which is equivalent to Equation (A.7). Taking expectations gives Equation (A.8). With a constant shift, the probability of the marginal set is $\Pr(-\Delta \leq M_i^0 < 0) = F_M(0) - F_M(-\Delta)$. □

Corollary A.1 (Beliefs and applications need not move together)**.** *A positive average change in beliefs can coexist with a small or statistically undetectable change in applications when few students lie near the application threshold. If F_M has density f_M that is continuous at zero, then for a small common shift Δ,*

$$F_M(0) - F_M(-\Delta) \approx f_M(0)\Delta.$$

Proof. The approximation follows from a first-order expansion of $F_M(-\Delta)$ around zero. The result is small when the density of baseline perceived returns near the application threshold is small, even if the underlying belief shift is positive. □

The corollary is especially relevant in this setting because applying to economics is rare. It also shows why a null application estimate is not an equivalence result: a wide range of perceived-return effects may be consistent with a small realized number of marginal applicants.

A.5 Heterogeneity by Prior Beliefs and Gender

Students with less favorable priors may have more scope to update, but this conclusion requires structure. The following result states one sufficient condition.

Proposition A.4 (Updating from lower priors)**.** *Suppose p is weakly concave. Consider two students, L and H, with $b_L^0 \leq b_H^0$, the same V^0, and common shifts $\delta_b \geq 0$, δ_V, and δ_C. Then the belief component of Equation* (A.5) *is weakly larger for student L than for student H. The inequality is strict if p is strictly concave, $b_L^0 < b_H^0$, and $\delta_b > 0$.*

Proof. The belief component for a student with prior b is

$$B(b) = [p(b + \delta_b) - p(b)]\, V^0.$$

Because p is concave, its increments over an interval of fixed length δ_b are weakly decreasing in the starting point. Hence $p(b_L^0 + \delta_b) - p(b_L^0) \geq p(b_H^0 + \delta_b) - p(b_H^0)$. Multiplying by $V^0 > 0$ preserves

the inequality. Strict concavity, different priors, and a positive shift imply a strict inequality. □

Proposition A.4 rationalizes the pre-specified heterogeneity test by baseline confidence. It does not by itself predict a larger application effect for women. Such a conclusion would additionally require women to have less favorable priors, comparable shifts in the other components of perceived returns, and sufficient mass near the application threshold. In general, for gender $g \in \{F, M\}$, Proposition A.3 implies

$$ATE_g = \Pr\left(-\Delta M_i \leq M_i^0 < 0 \mid g\right). \tag{A.9}$$

Gender differences can therefore reflect both differences in updating and differences in the distribution of baseline perceived returns around zero. This is why the empirical analysis relies on direct treatment-by-gender interaction tests rather than on comparisons of separate significance levels.

A.6 Direct Exposure, Spillovers, and Persistence

The two-stage design allows the role-model signal to reach students with different intensity. Let $\lambda \in [0, 1]$ denote the fraction of each treatment-induced shift that reaches a student. Classrooms assigned to the role-model session receive the full signal ($\lambda = 1$). Control classrooms receive the active placebo without economics or gender content ($\lambda = 0$). Spillover classrooms also receive the active placebo, but may receive an intermediate signal if information diffuses through classmates or teachers in treated schools. Define

$$M_i(\lambda) = p(b_i^0 + \lambda\delta_{bi})(V_i^0 + \lambda\delta_{Vi}) - (C_i^0 + \lambda\delta_{Ci}). \tag{A.10}$$

Proposition A.5 (Attenuated exposure)**.** *Under Assumption A.1, suppose* $\delta_{bi} \geq 0$, $\delta_{Vi} \geq 0$, *and* $\delta_{Ci} \leq 0$. *Then* $M_i(\lambda)$ *is weakly increasing in* λ. *For any spillover intensity* $\lambda_S \in [0, 1]$,

$$M_i(1) \geq M_i(\lambda_S) \geq M_i(0). \tag{A.11}$$

Proof. Differentiating Equation (A.10) gives

$$\frac{\partial M_i(\lambda)}{\partial \lambda} = p'(b_i^0 + \lambda\delta_{bi})\delta_{bi}(V_i^0 + \lambda\delta_{Vi}) + p(b_i^0 + \lambda\delta_{bi})\delta_{Vi} - \delta_{Ci}.$$

Every term is weakly positive under the stated conditions. Hence $M_i(\lambda)$ is weakly increasing, and the ordering in Equation (A.11) follows from $1 \geq \lambda_S \geq 0$. □

The proposition maps directly into the experimental contrasts. The comparison between direct-treatment and control classrooms captures $M_i(1) - M_i(0)$, the spillover contrast captures $M_i(\lambda_S) - M_i(0)$, and the comparison between direct-treatment and spillover classrooms in treated schools captures $M_i(1) - M_i(\lambda_S)$. The model does not impose $\lambda_S > 0$. When students do not transmit the message, $\lambda_S = 0$ and the spillover contrast is null; with complete diffusion, $\lambda_S = 1$ and direct and spillover classrooms receive the same effective signal.

The same representation can describe persistence. Let $\rho(\tau) \in [0, 1]$ denote the fraction of the initial shift retained τ periods after the session. Under the sign conditions of Proposition A.5, perceived net returns at horizon τ are $M_i(\rho(\tau))$. If a common retention rate declines over time, the treatment-induced shift also weakly declines. However, the survey domains may decay at different rates, and the available design does not require a common ρ. The sign and composition of the follow-up effect therefore remain empirical questions.

A.7 Mapping the Model to the Empirical Analysis

The model yields four practical implications. First, the beliefs-about-economics index primarily captures updating about V_i and about the information used to assess the field. Confidence and perceived fit capture components of b_i and C_i, while stereotype and barrier measures capture components of C_i. These mappings are not one-to-one, which is why the paper analyzes pre-specified outcome families rather than assigning each survey item to a single structural parameter. Second, the administrative application outcome records the threshold decision in Equation (A.3); Proposition A.3 explains why it can be less responsive than the underlying survey measures. Third, heterogeneity by gender or baseline beliefs requires direct interaction tests because updating and

threshold location can differ simultaneously. Finally, the two-stage assignment separately identifies direct exposure and within-school diffusion, corresponding to the contrasts in Proposition A.5.

Taken together, the formal results support a restrained interpretation of the experiment. The model predicts favorable effects only under explicit conditions on belief, value, and cost updating. It also permits immediate belief effects without persistent changes or applications, and it does not guarantee larger effects for women. These are empirical margins rather than assumptions built into the framework.

B Outcomes and Index Construction

This section documents the outcome definitions and index-construction rules summarized in the main text. The survey domains correspond to the channels in the framework developed in Section A.

B.1 Primary Outcome

We measure applications to an undergraduate economics program during the 2026 admissions cycle (April–June 2026) using university records and a combined-source indicator. The administrative indicator equals one for UANL Economics or Tec LEC and zero when neither registry identifies an economics application. The combined-source indicator also counts economics choices reported in the August survey, including choices at other universities and cases not matched in the study institutions' records. Both measures retain the full eligible sample; a combined-source zero denotes no detected event, including cases without a survey response or administrative match. The gateway indicator combines UANL Economics with Tec ECSG, changing program coverage rather than adding another data source. The primary population consists of women observed at baseline in the final year of their four- or six-semester program; all-student results pool the same eligible programs. The PAP specifies applications measured through administrative records. The combined-source measure incorporates additional information on students' educational choices, including destinations outside UANL and Tec. Applications, acceptance, and enrollment remain distinct events in the 2026 school-level intention-to-treat analysis. We do not use the 2025 cohort as a pre-intervention comparison because its consent and sample-selection rules are not comparable.

B.2 Survey Domains and Wave-Specific Measurement

Secondary outcomes capture the belief- and identity-based channels through which role models may affect applications. Following best practice in survey-based evaluations of role-model interventions (e.g., Breda et al., 2023), conceptually related items are combined into standardized indices within four pre-specified domains:

A. **Intentions and aspirations:** likelihood of applying to economics, preferred field after high school, perceived family support, and perceived affordability, summarized in an *Intentions and Aspirations Index*.

B. **Beliefs about economics (content and social value):** understanding of what economists do, perceived social relevance of economics, and the belief that economics is mainly about money or finance (reverse-coded), summarized in an *Economics Beliefs Index*.

C. **Identity alignment and self-concept:** perceived fit of economics with one's self-concept, identification with economists, nervousness about studying economics (reverse-coded), self-assessed ability in economics and mathematics, and comparative ability relative to same- and other-gender peers, summarized in an *Identity and Self-Concept Index*.

D. **Gender stereotypes and perceived barriers:** the belief that men are naturally better at economics than women and perceived discrimination in economics-related careers are summarized in a *Gender Stereotypes and Barriers Index*. The perceived main barrier to studying economics and the ability to name a female economist (recall proxy) are reported separately.

The four pre-specified domains are measured with different numbers of items across waves. The immediate instrument supports a three-item Economics Beliefs index and a two-item Intentions index, but only a single perceived-fit item for Identity and Self-Concept and a single stereotype-endorsement item for Gender Stereotypes. We identify these last two outcomes as single-item measures rather than multi-item indices. The longer follow-up instrument supports all four domain summaries. Tables A5–A7 report the constituent items, non-response rates, the listwise-complete sample size, mean inter-item correlations, and Cronbach's α for each domain and wave.

These diagnostics limit a latent-scale interpretation of some summaries. The baseline and follow-up alphas are -0.406 and -0.442 for Economics Beliefs, and -0.104 and -0.174 for Gender Stereotypes; the corresponding two stereotype items are weakly negatively correlated. In contrast, Identity and Self-Concept has alphas of 0.798 and 0.813. We retain the four domains specified in the pre-analysis plan, but interpret the low-reliability indices as pre-specified weighted summaries of distinct responses, not as validated unidimensional scales. The binary recall item and the multiple-response barrier item are reported separately and do not enter the stereotype index.

B.3 Standardization and Index Weights

Within each domain, items are oriented so that higher values reflect more pro-equality attitudes or greater alignment with the target construct. Each item's values at all waves are standardized

using its mean and standard deviation among eligible final-year control-group students at baseline, pooling fourth and sixth semesters. We then construct the inverse-covariance-weighted (ICW) index of Anderson (2008): the weights are proportional to the inverse of the covariance matrix of the standardized baseline items among those controls, multiplied by a vector of ones, and are normalized to sum to one. For Economics Beliefs, the weights are 0.221 for reported understanding of economists' work, 0.331 for perceived social value, and 0.448 for rejection of the money-and-finance stereotype. The same baseline-derived weights are used at later waves. Each wave uses only items administered both at baseline and in that wave, with weights renormalized over that subset; items without a corresponding baseline control measurement are excluded from ICW construction. The matched baseline and post-treatment indices are finally expressed in baseline control-group standard deviations.

B.4 Missing Responses and Multiple Testing

For each domain and wave, item non-response is calculated among students in the analysis sample who answered at least one item in that domain. If an item's non-response rate is strictly below 10%, a missing response is replaced by that student's mean of the other observed, standardized items in the domain. Otherwise, the index uses the observed items with their ICW weights renormalized; it remains missing if none is observed. This below-10% rule comes from the pre-analysis plan, rather than from Anderson's index construction. The robustness analysis includes a sensitivity specification controlling for the count of missing baseline domain items. For families with multiple correlated endpoints, we report the domain summary and key constituent items. Benjamini–Hochberg (Benjamini and Hochberg, 1995) q-values adjust the four domain tests separately within each wave and estimand; the administrative application measure is not adjusted.

C Interpretation and Statistical Power

C.1 What the short-term wave can and cannot show

Because the baseline and the short-term instrument were administered in the same classroom session, minutes apart, the short-term contrast should be read as an *immediate reaction* to the intervention rather than as evidence about settled beliefs. Students answered in the same room within minutes of the session, sometimes with the presenter still present, and several items use language close to that of the presentation.

The active placebo mitigates this concern. Students in spillover and control classrooms also received a session, heard a personal narrative, watched videos, and completed the same questionnaire minutes afterward. This shared format reduces the concern that the contrast captures a generic response to outside attention. A demand effect may nevertheless remain *specific to the economics content*— students inferring that this particular visitor wanted them to say warmer things about economics. So the concern is narrower than a general acquiescence story, but it is not eliminated. The gain in precision from conditioning on a baseline measured the same day is real (McKenzie, 2012), but it does not address this concern; if anything, the short recall interval sharpens it. In comparison, Breda et al. (2023) field their student survey between one and six months after the visits and present it as unconnected to the program, precisely to avoid this problem.

We respond in three ways. First, we examine the follow-up survey and administrative applications to assess whether the immediate responses persisted or translated into choices. Second, we report placebo items on fields unrelated to economics (Table A39); uniform positive movement across all items, including placebos, would be consistent with generalized reporting effects, although neither a uniform nor a selective response uniquely identifies demand. Third, we examine whether the pattern of effects across the belief domains is differentiated—a broad, undifferentiated shift would be less specific to the intervention content, whereas the framework of Section A predicts movement concentrated in the domains the session actually targets.

A separate question is whether the baseline measurement is a valid conditioning variable at all. It is pre-treatment in time, being completed before the intervention begins. Role models did not participate in this first part of the session: they normally entered only after the baseline survey, and, when already present for logistical reasons, remained silent and made no comment until students had

completed it. The field protocol likewise instructed role models not to intervene during either survey. This sequencing rules out baseline responses to the role model's narrative or to direct interaction with her, although students could still anticipate that a classroom visit would follow. Because assignment preceded consent, differences among baseline respondents cannot establish whether the randomized populations were balanced, and they may combine such anticipation with selective participation. We therefore distinguish preassignment balance from respondent composition and examine unadjusted comparisons alongside ANCOVA, with fixed-item balanced-panel changes used to assess persistence.

C.2 What the realized design can detect

Table A2 reconciles the analysis populations and translates the realized assignment-aligned standard errors into approximate minimum detectable effects (MDE $\approx 2.8\times$ SE), using 80 percent power and a two-sided 5 percent test. The benchmarks are 0.165 baseline control-group standard deviations for immediate beliefs and 0.144 at follow-up, compared with the pre-analysis plan's 0.18 benchmark. These are descriptive normal-approximation calculations; few clusters, incomplete response, and the choice of inference procedure affect their interpretation. The positive follow-up estimate of 0.169 is supported by conventional, wild-bootstrap, and randomization tests, while selection and measurement timing remain separate limitations.

For the combined application measure, the minimum detectable effect is 0.89 p.p. for women and 1.08 p.p. for all final-semester students. The administrative measure includes all eligible baseline students and codes an absence from both registries as zero. Among 2,009 final-semester women, the registries record eight economics applications, with a control rate of 0.196 percent. The school-ITT estimate is 0.404 percentage points (standard error 0.271), giving an approximate MDE of 0.76 percentage points. Its conventional 95 percent confidence interval, $[-0.153, 0.961]$ percentage points, includes both zero and substantively important increases relative to the control rate. We therefore emphasize absolute effects, event counts, and assignment-aligned inference. An imprecise primary estimate does not establish that the intervention leaves applications unchanged, nor does a large proportional point estimate establish a precise response.

References

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association*, 103(484): 1481–1495.

Bénabou, Roland, and Jean Tirole. 2002. "Self-Confidence and Personal Motivation." *Quarterly Journal of Economics*, 117(3): 871–915.

Bénabou, Roland, and Jean Tirole. 2016. "Mindful Economics: The Production, Consumption, and Value of Beliefs." *Journal of Economic Perspectives*, 30(3): 141–164.

Benjamini, Yoav, and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1): 289–300.

Breda, Thomas, Julien Grenet, Marion Monnet, and Clémentine Van Effenterre. 2023. "How Effective Are Female Role Models in Steering Girls Towards STEM? Evidence from French High Schools." *The Economic Journal*, 133(653): 1773–1809.

Golman, Russell, David Hagmann, and George Loewenstein. 2017. "Information Avoidance." *Journal of Economic Literature*, 55(1): 96–135.

McKenzie, David. 2012. "Beyond Baseline and Follow-Up: The Case for More T in Experiments." *Journal of Development Economics*, 99(2): 210–221.

D Survey Instruments

The following sections reproduce English translations of the study questionnaires in administration order: baseline, immediate post-intervention, follow-up, and the complementary online endline (August 2026). The first three use archived printable forms; the fourth uses an archived transcript of the online questionnaire. Question identifiers, response options, scale anchors, instructions, and stated routing are retained. All surveys were administered in Spanish. The printable forms do not reproduce every digital screen or mode-specific wording; their blank adaptive monetary amounts remain blank here. Gray annotations in the August instrument identify recorded fields, response codes, and display conditions separately from participant-facing text.

D.1 Baseline Survey

Student Survey. Protocol: P-ECSG-202511-01.

Unique Participant Code

To protect your identity while allowing us to follow up on your responses, please create your unique code by completing the following three fields.

1. First 3 letters of your paternal surname (in uppercase): _____.
2. Last 4 digits of your mobile phone number: _____.
3. Month of birth:

01 – January	04 – April	07 – July	10 – October
02 – February	05 – May	08 – August	11 – November
03 – March	06 – June	09 – September	12 – December

Section A: General and Academic Information

A1. What is your gender?

Response options: Woman; Man; Prefer not to say; Other.

A2. How old are you?

_____ (age in completed years).

A3. What grade/year are you in?

Response options: Fourth semester; Sixth semester.

A3b. Do you work for pay during the school week?

Response options: Yes; No.

If you answered "Yes": Approximately how many hours per week do you work?

_____ (0–80 hours; leave blank if you do not work).

A4a. Highest level of education completed by your father or male guardian:

Response options: Primary school; Lower secondary school; Upper secondary school; Technical education; University; Postgraduate education; Don't know.

A4b. Highest level of education completed by your mother or female guardian:

Response options: Primary school; Lower secondary school; Upper secondary school; Technical education; University; Postgraduate education; Don't know.

A5. Is your father's, mother's, or guardian's occupation related to economics, business, public policy, or finance?

Response options: Yes; No.

A6. Do you know an economist in your social circle, other than your father or mother?

Response options: Yes; No.

A7. What was your grade average last semester?

_____ (0–100; example: 85.5).

A8. Which of the following do you have at home?

Check all that apply.

Response options: Your own computer; Internet at home; More than 25 books at home; Your own room.

Section B: Interests and Self-Perceptions

B9. How much do you like the following subjects?

0 = Not at all; 10 = Very much.

Economics

0 1 2 3 4 5 6 7 8 9 10

Mathematics

0 1 2 3 4 5 6 7 8 9 10

History / Social sciences

0 1 2 3 4 5 6 7 8 9 10

Spanish

0 1 2 3 4 5 6 7 8 9 10

English

0 1 2 3 4 5 6 7 8 9 10

B10. How would you rate your level of ability?

0 = Very low; 10 = Very high.

Economics

0 1 2 3 4 5 6 7 8 9 10

Mathematics

0 1 2 3 4 5 6 7 8 9 10

Compared with other students, how would you rate your level in economics? 0 = Much lower; 10 = Much higher.

B11. Compared with students of your own gender, your level in economics is:

0 1 2 3 4 5 6 7 8 9 10 Not applicable

B12. Compared with students of the other gender, your level in economics is:

0 1 2 3 4 5 6 7 8 9 10

Section C: Beliefs, Identity, and Stereotypes

Indicate how much you agree with each statement. 0 = Strongly disagree; 10 = Strongly agree.

C13. I understand what economists do in their actual jobs.

0 1 2 3 4 5 6 7 8 9 10 Don't know

C14. Economics helps improve people's lives and reduce social problems.

0 1 2 3 4 5 6 7 8 9 10 Don't know

C15. Economics is mainly about money and finance.

0 1 2 3 4 5 6 7 8 9 10 Don't know

C16. Men are naturally better at economics than women.

0 1 2 3 4 5 6 7 8 9 10 Don't know

C17. Women face discrimination in economics-related careers.

0 1 2 3 4 5 6 7 8 9 10 Don't know

C18. Most students of my gender think that economics is a suitable field for women.

0 1 2 3 4 5 6 7 8 9 10 Don't know

C19. Thinking about studying economics makes me nervous or insecure.

0 1 2 3 4 5 6 7 8 9 10

C20. Studying economics fits who I am.

0 1 2 3 4 5 6 7 8 9 10

C21. When I think of an economist, I imagine someone like me.

0 1 2 3 4 5 6 7 8 9 10

0 = No prestige; 10 = A great deal of prestige.

C22. How much prestige do you think studying economics brings?

0 1 2 3 4 5 6 7 8 9 10 Don't know

0 = Not at all; 10 = Completely.

C23. If an economist gives you advice, how much would you trust their word?

0 1 2 3 4 5 6 7 8 9 10 Don't know

Section D: Career Aspirations and Intentions

D24. Which field interests you most after high school?

Response options: Economics / Finance; Business / Management; Engineering; Health; Education; Law; Other (please specify): _____.

0 = Not at all likely; 10 = Very likely.

D25. How likely are you to apply to an economics program?

0 1 2 3 4 5 6 7 8 9 10

0 = Not at all; 10 = Very much.

D26. How much would your family or friends support you if you decided to study economics?

0 1 2 3 4 5 6 7 8 9 10 Don't know

0 = Definitely not; 10 = Definitely yes.

D27. Could you afford to study economics if you wanted to?

0 1 2 3 4 5 6 7 8 9 10 Don't know

Section E: Optional Laboratory Items

Imagine that all the students in your class took a test in mathematics and economics.

E28. If you took a mathematics/economics test with all the students in your class, what percentile do you think you would be in?

_____ (0 = the lowest score in the class; 100 = the highest score in the class).

E29. What is the probability, from 0 to 100, that your score would be above the class median?

_____ (0 = impossible; 100 = completely certain).

End of the survey. Thank you for participating.

D.2 Immediate Post-Intervention Survey

Student Survey. Protocol: P-ECSG-202511-01.

Unique Participant Code

To protect your identity while allowing us to follow up on your responses, please create your unique code by completing the following three fields.

1. First 3 letters of your paternal surname (in uppercase): _____.
2. Last 4 digits of your mobile phone number: _____.
3. Month of birth:

01 – January	04 – April	07 – July	10 – October
02 – February	05 – May	08 – August	11 – November
03 – March	06 – June	09 – September	12 – December

Email Address for Compensation

To provide the compensation for your participation, we need a contact email address. Compensation will be sent approximately one month after data collection ends. Your email address will be stored separately from your responses and deleted once compensation has been provided.

1. Email address: _____.
2. Email address of your guardian or responsible adult: _____.

Section S: Exposure and Immediate Perceptions

S1. Did you attend the session in which a female economist visited your class?

Response options: Yes; No; Not applicable.

0 = Strongly disagree; 10 = Strongly agree; N/A.

S2. I can recall the speaker's name or affiliation.

0 1 2 3 4 5 6 7 8 9 10 Not applicable

S3. I enjoyed the session.

0 1 2 3 4 5 6 7 8 9 10 Not applicable

S4. The visit increased my interest in economics.

0 1 2 3 4 5 6 7 8 9 10 Not applicable

S5. The visit increased my confidence in my quantitative ability.

0 1 2 3 4 5 6 7 8 9 10 Not applicable

S6. After the visit, I can imagine myself as an economist.

0 1 2 3 4 5 6 7 8 9 10 Not applicable

S7. My perception of what economists do changed.

0 1 2 3 4 5 6 7 8 9 10 Not applicable

Section T: Beliefs and Stereotypes

0 = Strongly disagree; 10 = Strongly agree.

T8. I understand what economists do in actual jobs.

0 1 2 3 4 5 6 7 8 9 10

0 = Strongly disagree; 10 = Strongly agree; Don't know.

T9. Economics helps improve people's lives and reduce social problems.

0 1 2 3 4 5 6 7 8 9 10 Don't know

T10. Economics is mainly about money and finance.

0 1 2 3 4 5 6 7 8 9 10 Don't know

T11. Men are naturally better at economics than women.

0 1 2 3 4 5 6 7 8 9 10 Don't know

0 = Strongly disagree; 10 = Strongly agree.

T12. Studying economics fits who I am.

0 1 2 3 4 5 6 7 8 9 10

Section U: Intentions and Social Support

0 = Not at all likely; 10 = Very likely.

U13. How likely are you to apply to an economics program?

0 1 2 3 4 5 6 7 8 9 10

0 = Not at all; 10 = Very much; Don't know.

U14. How much would your family or friends support you if you decided to study economics?

0 1 2 3 4 5 6 7 8 9 10 Don't know

Section V: Qualitative GPS Preferences

For the following questions, indicate your level of agreement on a scale from 0 to 10.

0 = Not at all willing; 10 = Very willing.

V15. How willing are you to take risks in general?

0 1 2 3 4 5 6 7 8 9 10

V16. How willing are you to give up something beneficial today to obtain more in the future?

0 1 2 3 4 5 6 7 8 9 10

0 = Strongly disagree; 10 = Strongly agree.

V17. When someone treats me very unfairly, I am willing to punish them even if it costs me something.

0 1 2 3 4 5 6 7 8 9 10

V18. When someone treats others very unfairly, I am willing to punish them even if it costs me something.

0 1 2 3 4 5 6 7 8 9 10

0 = Not at all willing; 10 = Very willing.

V19. How willing are you to donate to a good cause without expecting anything in return?

0 1 2 3 4 5 6 7 8 9 10

0 = Strongly disagree; 10 = Strongly agree.

V20. When someone does me a favor, I am willing to return it.

0 1 2 3 4 5 6 7 8 9 10

V21. If I am treated unfairly, I will take revenge at the first opportunity, even if it costs me something.

0 1 2 3 4 5 6 7 8 9 10

V22. I generally assume that people have the best intentions.

0 1 2 3 4 5 6 7 8 9 10

V23. How would you rate your level of ability in mathematics?

0 = Very low; 10 = Very high.

0 1 2 3 4 5 6 7 8 9 10

V24. I tend to put things off.

0 1 2 3 4 5 6 7 8 9 10

Section W: GPS Risk Decisions

In each question, choose between receiving a guaranteed amount or taking part in a lottery in which you have a 50% chance of winning 450 pesos and a 50% chance of winning nothing.

W25-1. Which do you prefer?

Response options: Lottery (50% chance of 450 pesos); Receive _____ pesos for certain.

W25-2. Which do you prefer?

Response options: Lottery (50% chance of 450 pesos); Receive _____ pesos for certain.

W25-3. Which do you prefer?

Response options: Lottery (50% chance of 450 pesos); Receive _____ pesos for certain.

W25-4. Which do you prefer?

Response options: Lottery (50% chance of 450 pesos); Receive _____ pesos for certain.

W25-5. Which do you prefer?

Response options: Lottery (50% chance of 450 pesos); Receive _____ pesos for certain.

Note: The guaranteed amounts vary according to your previous answers (staircase method).

Section X: GPS Trust and Altruism

X26. Imagine that you receive 60 pesos and can give part of it to a stranger. How much would you give?

Response options: 0; 5; 10; 15; 20; 25; 30; 35; 40; 45; 50; 55; 60 pesos.

X27. Imagine that you won 1,600 pesos in a lottery. How much would you donate to a good cause?

_____ pesos (from 0 to 1,600).

Section Y: GPS Patience Decisions

In each question, choose between receiving 200 pesos today or receiving a larger amount in 12 months.

Y28-1. Which do you prefer?

Response options: 200 pesos today; _____ pesos in 12 months.

Y28-2. Which do you prefer?

Response options: 200 pesos today; _____ pesos in 12 months.

Y28-3. Which do you prefer?

Response options: 200 pesos today; _____ pesos in 12 months.

Y28-4. Which do you prefer?

Response options: 200 pesos today; _____ pesos in 12 months.

Y28-5. Which do you prefer?

Response options: 200 pesos today; _____ pesos in 12 months.

Note: The future amounts vary according to your previous answers (staircase method).

Undergraduate Program Registration

Have you already registered for the admission examination for an undergraduate degree or program?

Response options: Yes; No.

If you answered "Yes," continue with the following questions. If you answered "No," go to the next section.

Which degree program(s) did you register for? (Select all that apply.)

Response options: Economics / Finance; Business / Management; Engineering; Health; Education; Law; Other: _____.

How certain are you about that choice?

0 = Not at all certain; 10 = Completely certain.

0 1 2 3 4 5 6 7 8 9 10

Demand Check

What do you think is the purpose of this survey?

_____

_____

_____

_____

Responsible researchers: Edgar Castro Méndez (ecastrom@tec.mx), Stanislao Maldonado (stanislao.maldonado@tec.mx). Research Ethics Committee of Tecnológico de Monterrey. Protocol: P-ECSG-202511-01.

End of the survey. Thank you for participating.

D.3 Follow-Up Survey

Student Survey. Protocol: P-ECSG-202511-01.

Unique Participant Code

To protect your identity while allowing us to link your responses to earlier surveys, please create your unique code by completing the following three fields. This code does not allow you to be personally identified.

1. First 3 letters of your paternal surname (in uppercase): _____.
2. Last 4 digits of your mobile phone number: _____.
3. Month of birth:

01 – January	04 – April	07 – July	10 – October
02 – February	05 – May	08 – August	11 – November
03 – March	06 – June	09 – September	12 – December

Section L: Exposure, Recall, and Diffusion

The following questions refer to activities, talks, or conversations about economics that you may have had at school in recent weeks. Please answer as honestly as possible.

L1. In recent weeks, did you participate in a talk or activity in which a female economist spoke about her experience, economics, or careers in economics?

Response options: Yes; No; Not sure; Not applicable.

L2. Do you remember the female economist who visited your class, if your class received that visit?

Response options: Yes; No; Partly; Not applicable.

0 = Strongly disagree; 10 = Strongly agree.

L3. I can describe at least one topic covered in that talk or activity.

0 1 2 3 4 5 6 7 8 9 10 Not applicable

L4. Since that talk or activity, I have discussed economics with family members, friends, or classmates.

0 1 2 3 4 5 6 7 8 9 10 Not applicable

L5. In recent weeks, a classmate, teacher, family member, or friend has talked to me about studying economics or about careers in economics.

0 1 2 3 4 5 6 7 8 9 10

Section M: Interests and Self-Perceptions

M1. How much do you like the following subjects?

0 = Not at all; 10 = Very much.

Economics

0 1 2 3 4 5 6 7 8 9 10

Mathematics

0 1 2 3 4 5 6 7 8 9 10

History / Social sciences

0 1 2 3 4 5 6 7 8 9 10

Spanish

0 1 2 3 4 5 6 7 8 9 10

English

0 1 2 3 4 5 6 7 8 9 10

M2. How would you rate your level of ability?

0 = Very low; 10 = Very high.

Economics

0 1 2 3 4 5 6 7 8 9 10

Mathematics

0 1 2 3 4 5 6 7 8 9 10

Compared with other students, how would you rate your level in economics? 0 = Much lower; 10 = Much higher.

M3. Compared with students of your own gender, your level in economics is:

0 1 2 3 4 5 6 7 8 9 10 Not applicable

M4. Compared with students of the other gender, your level in economics is:

0 1 2 3 4 5 6 7 8 9 10

Section N: Beliefs, Identity, and Stereotypes

Indicate how much you agree with each statement. 0 = Strongly disagree; 10 = Strongly agree.

N1. I understand what economists do in their actual jobs.

0 1 2 3 4 5 6 7 8 9 10 Don't know

N2. Economics helps improve people's lives and reduce social problems.

0 1 2 3 4 5 6 7 8 9 10 Don't know

N3. Economics is mainly about money and finance.

0 1 2 3 4 5 6 7 8 9 10 Don't know

N4. Men are naturally better at economics than women.

0 1 2 3 4 5 6 7 8 9 10 Don't know

N5. Women face discrimination in economics-related careers.

0 1 2 3 4 5 6 7 8 9 10 Don't know

N6. Most students of my gender think that economics is a suitable field for women.

0 1 2 3 4 5 6 7 8 9 10 Don't know

N7. Thinking about studying economics makes me nervous or insecure.

0 1 2 3 4 5 6 7 8 9 10

N8. Studying economics fits who I am.

0 1 2 3 4 5 6 7 8 9 10

N9. When I think of an economist, I imagine someone like me.

0 1 2 3 4 5 6 7 8 9 10

N10. Can you name a female economist?

Response options: Yes; No.

If your answer is yes, write her name: _____.

0 = No prestige; 10 = A great deal of prestige.

N11. How much prestige do you think studying economics brings?

0 1 2 3 4 5 6 7 8 9 10 Don't know

0 = Not at all; 10 = Completely.

N12. If an economist gives you advice, how much would you trust their word?

0 1 2 3 4 5 6 7 8 9 10 Don't know

N13. To succeed in studying economics, it is necessary to have very strong mathematical skills.

0 1 2 3 4 5 6 7 8 9 10 Don't know

Section O: Aspirations, Decisions, and Feasibility

The following questions refer to your recent academic decisions and plans. Please answer honestly.

O1. Did you apply for admission to an economics program this semester?

Response options: Yes, I have already submitted a complete application; I started an application but have not completed it yet; No; Don't know / Don't remember.

O2. Did you apply for admission to any other degree program?

Response options: Yes; No; Don't know / Don't remember.

O3. Which degree program(s) did you apply to? (Select all that apply.)

Response options: Economics / Finance; Business / Management; Engineering; Health; Education; Law; Other (please specify): _____.

O4. If you selected Economics / Finance, which of the following best describes your application or main interest?

Response options: Economics; Finance; Both; Don't know; Not applicable.

0 = Not at all likely; 10 = Very likely.

O5. How likely are you to apply to an economics program?

0 1 2 3 4 5 6 7 8 9 10

0 = Not at all; 10 = Very much.

O6. How much would your family or friends support you if you decided to study economics?

0 1 2 3 4 5 6 7 8 9 10 Don't know

0 = Definitely not; 10 = Definitely yes.

O7. Could you afford to study economics if you wanted to?

0 1 2 3 4 5 6 7 8 9 10 Don't know

Section P: Barriers and Demand for Information

P1. What are the main obstacles that could prevent you from studying economics? (Select up to two options.)

Response options: Cost; Distance; Family expectations; Low confidence in mathematics; I do not know much about what economists do; I cannot imagine myself pursuing a career in economics; I prefer another field; Other (please specify): _____.

P2. What information would you like to receive? (Select all that apply.)

Response options: Brochures about economics programs; Admission dates; Financial aid / Scholarships; Talks about careers in economics; Testimonials from students or female economics graduates; Information about employment opportunities in economics; None.

P3. Would you like to receive reminders or updates about economics programs?

Response options: Yes; No.

Section Q: Optional Laboratory Items

Imagine that all the students in your class took a test in mathematics and economics.

Q1. If you took a mathematics/economics test with all the students in your class, what percentile do you think you would be in?

_____ (0 = the lowest score in the class; 100 = the highest score in the class).

Q2. What is the probability, from 0 to 100, that your score would be above the class median?

_____ (0 = impossible; 100 = completely certain).

Section R: Brief Quantitative Reasoning Exercises

Solve the following exercises without using a calculator. Select only one option for each question. These questions do not affect your school grades; their only purpose is to learn how students solve some brief exercises.

R1. If a product costs $200 and has a 25% discount, what is the final price?

Response options: $50; $150; $175; $250; Don't know.

R2. If 3 notebooks cost $48, how much do 5 notebooks cost at the same unit price?

Response options: $60; $72; $80; $96; Don't know.

R3. If $2x + 6 = 20$, what is the value of x?

Response options: 5; 7; 10; 13; Don't know.

R4. A pizza is divided into 8 equal slices and 3 are eaten. What fraction of the pizza remains?

Response options: 3/8; 5/8; 1/2; 3/5; Don't know.

R5. A class has 40 students. If 60% are women, how many women are in the class?

Response options: 16; 20; 24; 28; Don't know.

R6. The mean of three grades is 80. Two grades are 70 and 90. What is the third grade?

Response options: 70; 80; 90; 100; Don't know.

End of the survey. Thank you for participating.

D.4 Complementary Online Endline Survey (August 2026)

Student Survey. Protocol: P-ECSG-202511-01.

Raffle Results Page

Research-team note (not shown to participants). This is an English translation of the archived transcript of the online questionnaire completed when participants checked the tablet-raffle results on the `rifa-resultados-rm` page in August 2026. Gray text identifies recorded variables, response codes, and display conditions. Questionnaire responses are stored in the `respuestas` table; Section C (the laboratory invitation) is stored separately in `lab_consent`.

Section A: Identify Yourself Using Your Code

This is the same anonymous code you created at the start of the survey. It consists of just three pieces of information.

The code is constructed from the three fields and stored in `codigo_unico`. If it is not on the study list, the page does not allow the participant to proceed.

1. First 3 letters of your paternal surname: _____. *Example: García → GAR.*
2. Last 4 digits of your mobile phone number: _____.
3. Month of birth:

January	April	July	October
February	May	August	November
March	June	September	December

Section B: Questions

Before showing you your result, could you help us by answering 2 questions?

It takes 30 seconds, and this is the last thing we will ask of you. Your responses are confidential and will be analyzed in aggregate.

No question is mandatory. At the bottom of the screen, the button "I would rather see my result now" skips the questions (`omitio` = true).

1. Will you be studying next semester (August–December 2026)? [q_estudia]

- Yes, at a university [1]
- Yes, but not at a university [2]
- I will not be studying [3]
- I do not know yet [4]

2. What will you study, and where?

Hidden if the answer to question 1 is "I will not be studying."

Degree program: _____ [q_carrera]

Free text with suggestions (another answer may be entered): Economics; Finance; Management / Business; Accounting; Engineering; Health sciences / Medicine; Education; Law; Social sciences / Humanities; I have not decided yet.

University or institution: [q_universidad]

- UANL — Universidad Autónoma de Nuevo León [UANL]
- Tec de Monterrey (ITESM) [ITESM]
- UDEM — Universidad de Monterrey [UDEM]
- U-ERRE — Universidad Regiomontana [UERRE]
- Universidad TecMilenio [TECMILENIO]
- UMM — Universidad Metropolitana de Monterrey [UMM]
- UVM — Universidad del Valle de México [UVM]
- Instituto Tecnológico de Nuevo León [ITNL]
- State polytechnic / technological university [UPN]
- Teacher-training college (Escuela Normal) [NORMAL]
- Another institution in Nuevo León [OTRA_NL]
- An institution in another state [OTRO_ESTADO]
- An institution in another country [OTRO_PAIS]
- I have not decided yet [NO_DECIDO]

If "Another institution in Nuevo León," "another state," or "another country" is selected:

Which one? *(Name of the institution)*: _____. [q_uni_otra]

If a named institution is selected (neither "Another…" nor "I have not decided yet"):

Faculty or campus *(if you already know)*: _____. [q_facultad]

Submit and see my result I would rather see my result now

Section C: An Invitation

This screen appears after the raffle result and is reached by pressing "Continue."

Behavioral and Experimental Economics Laboratory

Tec de Monterrey is opening a new laboratory for behavioral and experimental economics studies. People who participate in its studies (in person, online, or in the field) receive **monetary compensation**.

Would you like to participate in the laboratory's studies and receive invitations? `[consiente]`

- Yes, I would like to participate and receive invitations `[1]`
- No, not at this time `[0]`

Only if the answer is "Yes":

Where should we send you the invitations? *(At least one)*

1. Email address: _____ `[email]`
2. Mobile phone number: _____. *(10 digits)* `[telefono]`

Study approved by the Research Ethics Committee. Protocol P-ECSG-202511-01.



(a) What is economics?

(b) Is economics for women?

(c) Women economists as role models

(d) Closing messages and aspirations

Figure A1: Representative Materials from the Implemented Role-Model Presentation

Notes: The panels reproduce representative Spanish-language slides from the role-model presentation delivered in UANL classrooms. They illustrate the implemented content rather than a retrospective summary: the sequence defines economics and the work of economists, addresses gendered beliefs about aptitude and professional fit, presents women economists as concrete role models, and provides information about studying economics. Presenters followed a common training protocol but could adapt examples and transitions to classroom discussion. Parallel materials with institution-specific application information were used in Tec classrooms. The reproduced slides are a subset of the complete presentation materials and should be read together with the implementation description in the main text.

A. First video: What is economics?



B. Second video: Is economics a field for women?



Figure A2: Screenshots from the Tec Intervention Videos

Notes: Each panel reproduces six selected frames from the archived video prepared for the Tecnológico de Monterrey high-school network (PrepaTec). Frames follow the order of the video, from left to right and then from the upper to the lower row. Panel A introduces economics through everyday decisions and its applications to education, employment, the environment, and social problems. Panel B addresses gendered beliefs about studying economics and illustrates women's participation in the profession. The implementation protocol places these videos alongside the visiting role model's personal reflections; a third video of interviews with women economists is shown in Figure A4. The original Spanish-language images and text are preserved, while panel headings summarize the content in English. The complete videos are available online: Tec Video 1 and Tec Video 2.

A. First video: What is economics?



B. Second video: Is economics a field for women?

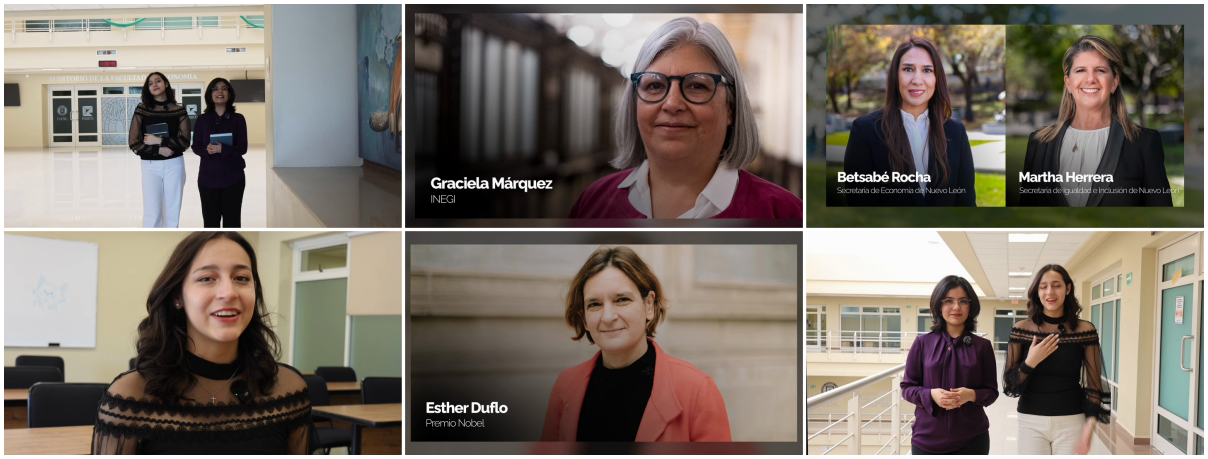


Figure A3: Screenshots from the UANL Intervention Videos

Notes: Each panel reproduces six selected frames from the archived video prepared for the Universidad Autónoma de Nuevo León (UANL) high-school network. Frames follow the order of the video, from left to right and then from the upper to the lower row. Panel A introduces economics through everyday choices and applications to inequality, urban life, and social justice. Panel B addresses whether economics is a suitable field for women and presents examples of women in professional and academic roles. These institution-specific videos occupy the same positions in the treatment protocol as the Tec videos in Figure A2; the third video is reproduced in Figure A4. The original Spanish-language images and text are retained, while panel headings summarize the content in English. The complete videos are available online: UANL Video 1 and UANL Video 2.

Third video: Career choices and professional experiences

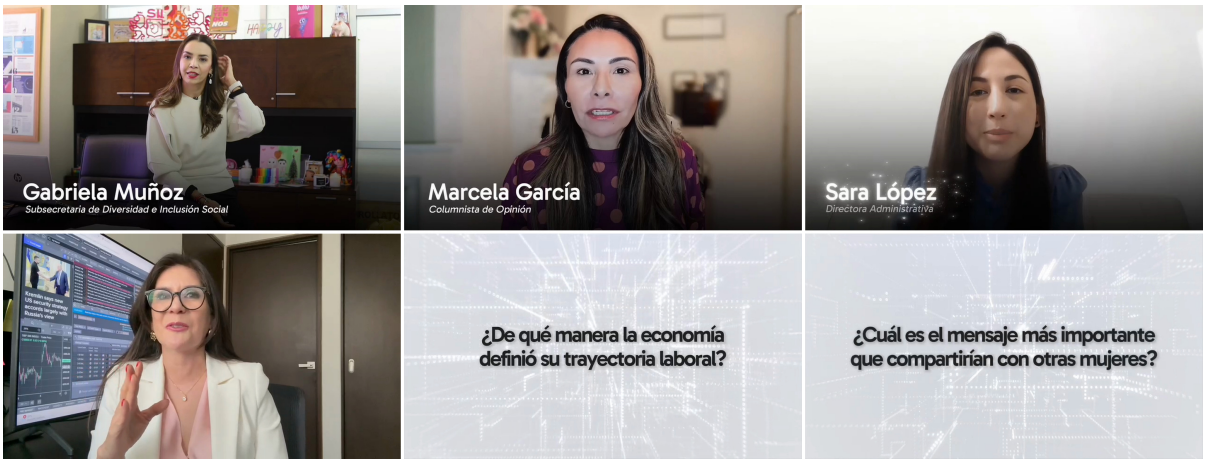


Figure A4: Screenshots from the Interviews with Women Economists

Notes: The figure reproduces six selected frames from *Economistas expertas*, the third video specified in the role-model implementation manual and structured script. This video complements the institution-specific videos in Figures A2 and A3 with interviews about educational choices, professional trajectories, and advice to other women. The protocol places it after the visiting economist's second personal reflection and before the motivational closing. Frames appear in chronological order, from left to right and then from the upper to the lower row. The two reproduced question cards ask how economics shaped the interviewees' professional trajectories and what message they would share with other women. Original Spanish-language images, names, and text are preserved. The complete video is available online: Economistas expertas.

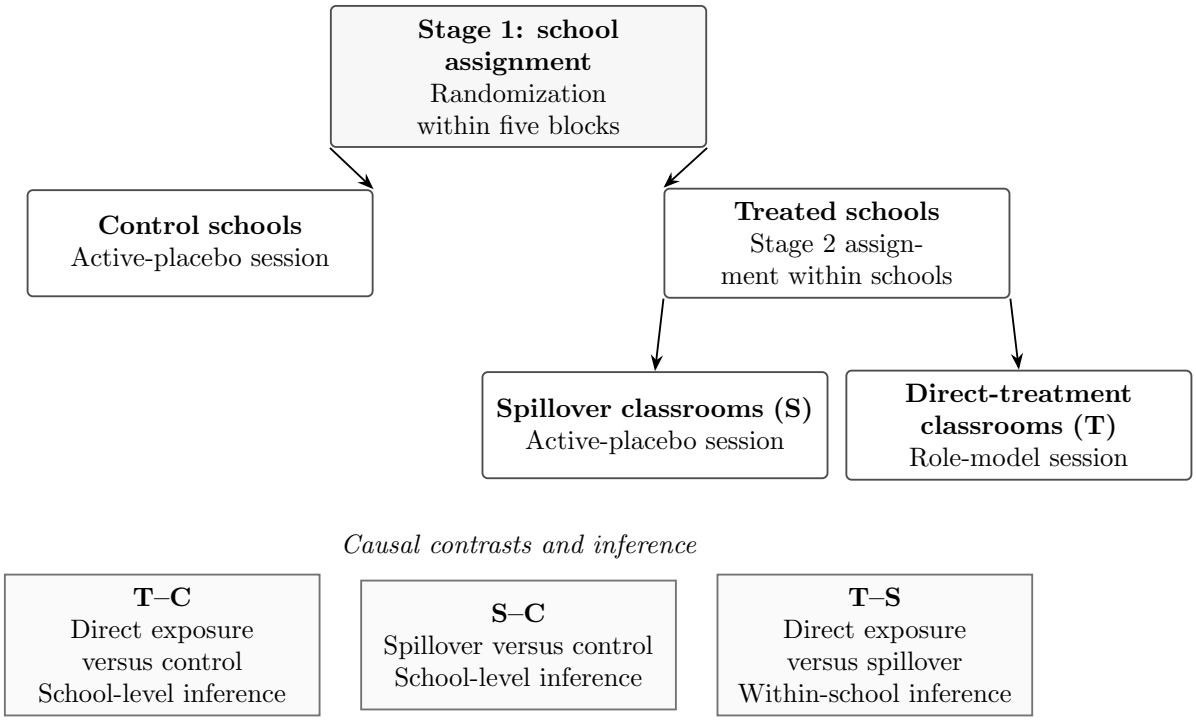
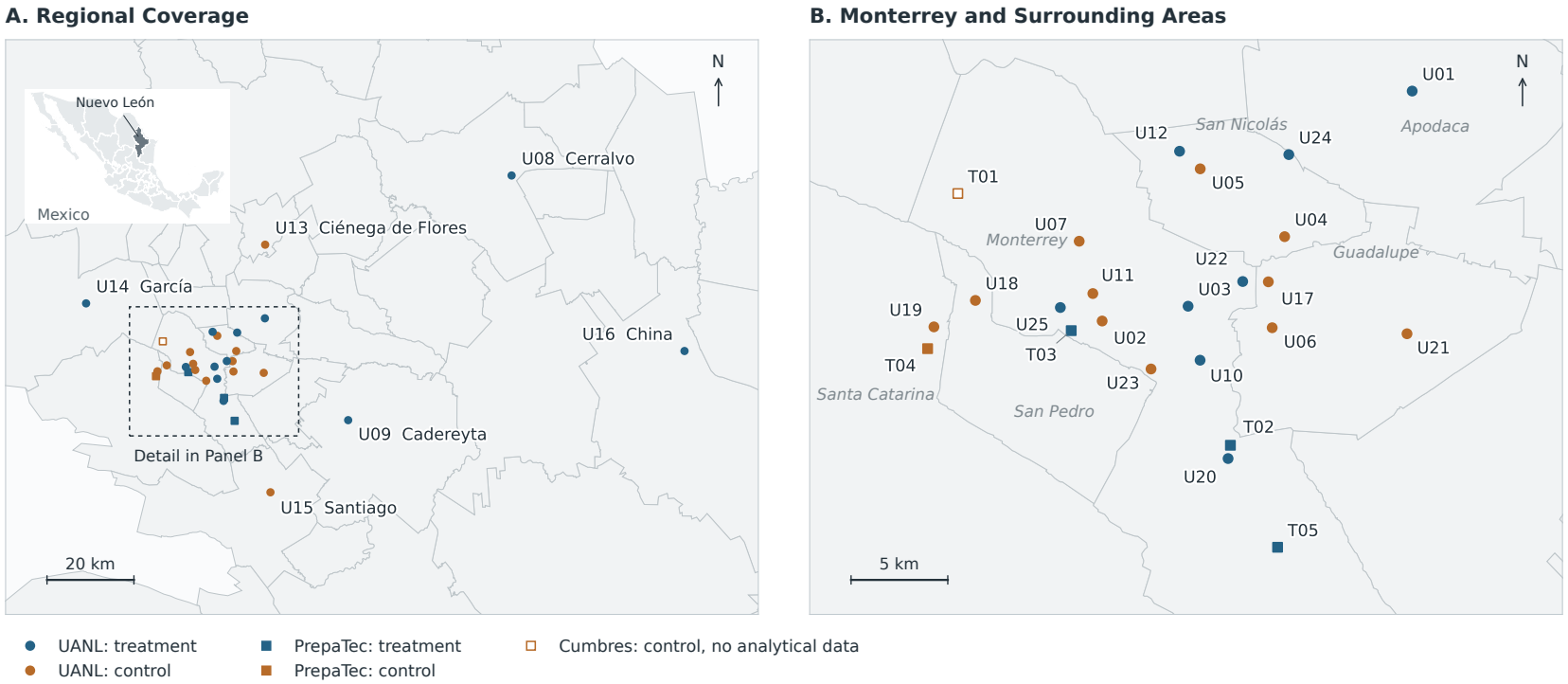


Figure A5: Two-Stage Assignment and Direct-versus-Spillover Contrasts

Notes: The figure summarizes the experiment's two assignment stages and the resulting estimands. First, school assignment units were randomized to treatment or control within five blocks. Second, within treated schools, eligible classrooms were assigned either to the role-model session or to the active placebo, forming the direct-treatment and spillover groups, respectively. Classrooms in control schools also received the active placebo. Treated-school assignment compares all classrooms in treated schools with classrooms in control schools. Direct treatment versus control and spillover versus control are cross-school contrasts estimated with the original randomization-block fixed effects and school-clustered standard errors. Direct treatment versus spillover is identified within treated schools and is estimated with school fixed effects and classroom-clustered standard errors. Arrows represent assignment and comparison structure; they do not imply that spillovers occurred.



School Key

U01	High School 1 · Apodaca	**U09**	High School 12 · Cadereyta	**U17**	High School 22 · Linda Vista	**U25**	Medical Technical High School
U02	High School 2	**U10**	High School 15 · Florida	**U18**	High School 23 · San Pedro	**T01**	Cumbres
U03	High School 3	**U11**	High School 15 · Madero	**U19**	High School 23 · Santa Catarina	**T02**	Eugenio Garza Lagüera
U04	High School 7 · Oriente	**U12**	High School 16	**U20**	CIDEB · Mederos	**T03**	Eugenio Garza Sada
U05	High School 7 · Las Puentes	**U13**	High School 17 · Ciénega de Flores	**U21**	EIAO · Guadalupe	**T04**	Santa Catarina
U06	High School 8	**U14**	High School 19 · García	**U22**	EIAO · Monterrey I	**T05**	Valle Alto
U07	High School 9	**U15**	High School 20 · Santiago	**U23**	EIAO · Monterrey II		
U08	High School 11 · Cerralvo	**U16**	High School 21 · China	**U24**	EIAO · San Nicolás		

Figure A6: Locations of Study Schools

Notes: The map shows the locations of the 30 physical campuses in the original assignment frame in Nuevo León, Mexico. Panel A presents regional coverage; Panel B enlarges Monterrey and surrounding areas. The campuses belong to 29 school assignment units, of which 15 were assigned to treatment and 14 to control. Blue markers denote treatment and orange markers control; circles identify UANL and squares PrepaTec. Treatment schools include both directly treated and spillover classrooms. The two Prepa 23 campuses, San Pedro (U18) and Santa Catarina (U19), share one control assignment. Cumbres (T01), shown as an open orange square, was assigned to control but contributes no analytical data; the analytical sample therefore covers 28 assignment units spanning 29 campuses. UANL is the Autonomous University of Nuevo León and PrepaTec is the Tecnológico de Monterrey high-school network. EIAO denotes UANL's Álvaro Obregón industrial and technical high school; CIDEB denotes UANL's Center for Research and Development in Bilingual Education. The U/T codes identify networks and are specific to this map. Locations are institutional map points, not field GPS measurements. Both panels use WGS 84 / UTM zone 14N; scale bars are in kilometers. *Sources:* Project assignment records, UANL and PrepaTec institutional maps, and INEGI municipal and state boundaries, accessed October 25, 2026.

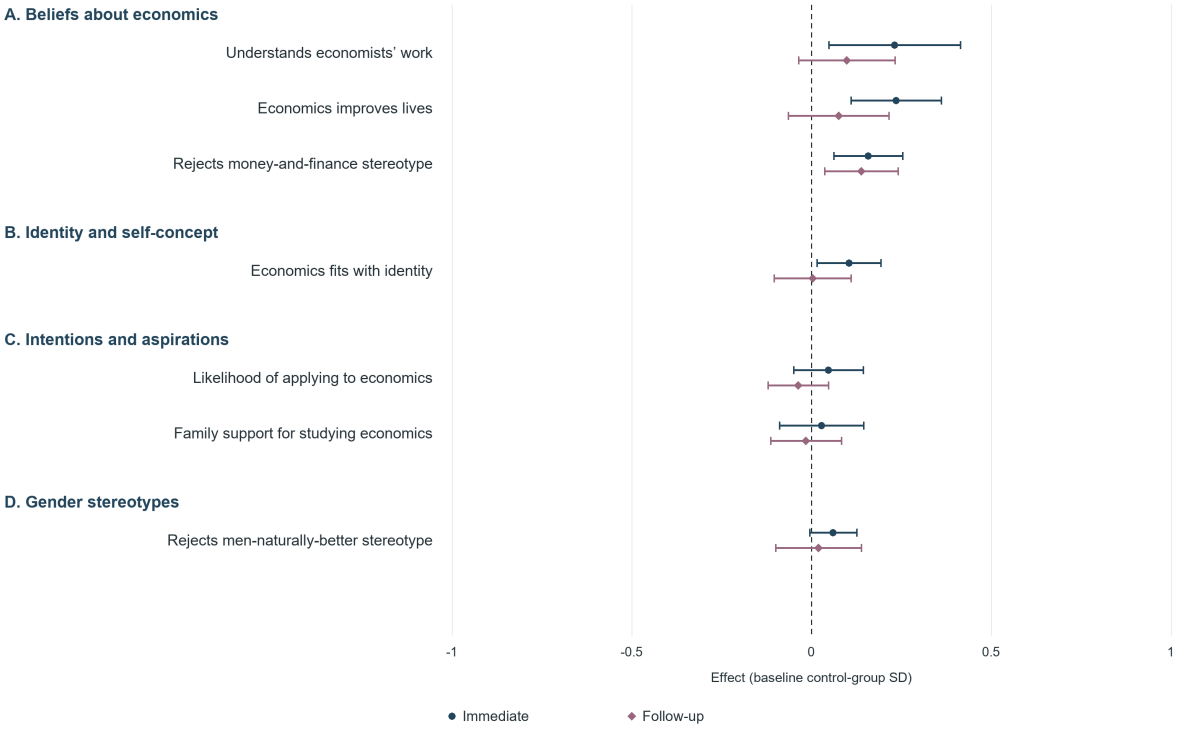


Figure A7: Common-Item Persistence in the Balanced Panel

Notes: The sample is a balanced panel of students observed in both the immediate and follow-up surveys, and the figure includes only items worded and coded identically at the two waves. Within each outcome family, separate markers report the treated-school assignment effect immediately after the classroom visit and at follow-up; horizontal bars are 95 percent confidence intervals and the dashed vertical line denotes zero. Each estimate comes from a wave-specific ANCOVA with the original randomization-block fixed effects and school-clustered standard errors. Items are standardized using baseline control-group moments and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. The common -1 to 1 scale facilitates comparison across items and waves; formal follow-up-minus-immediate tests are reported in the corresponding appendix table.

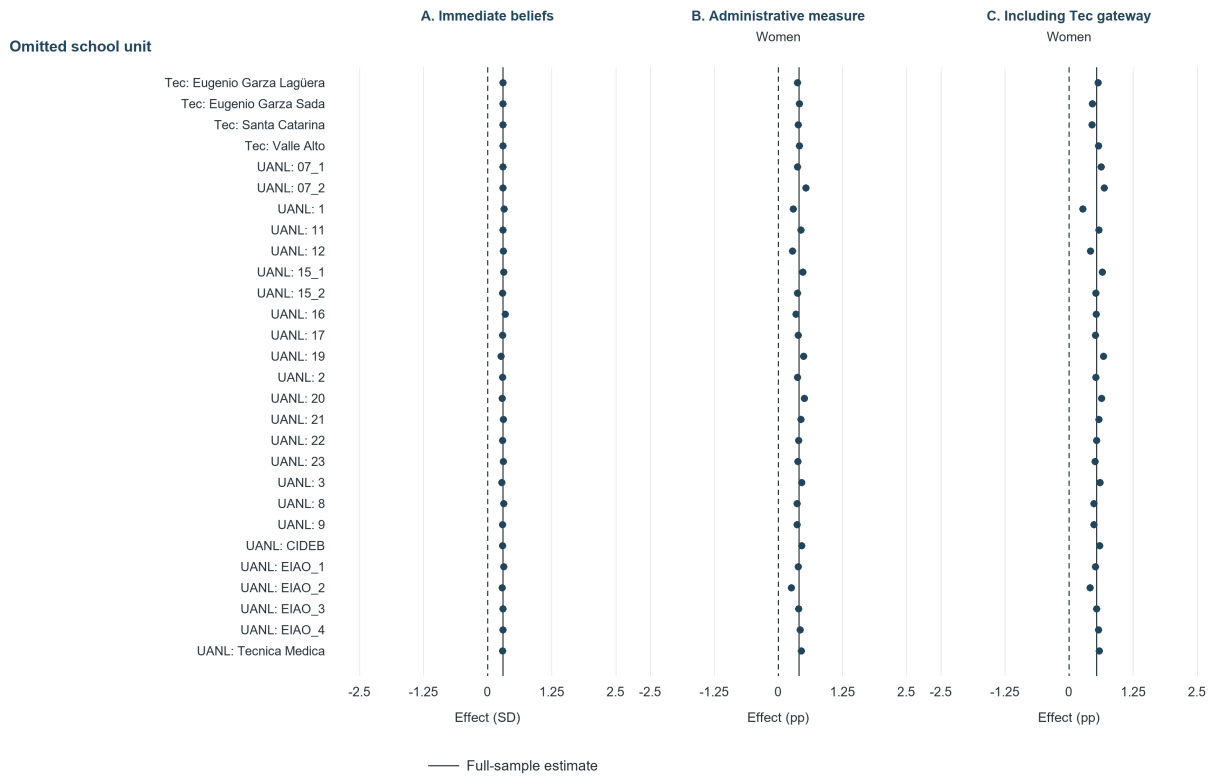


Figure A8: Leave-One-School-Out Estimates

Notes: Each marker re-estimates the treated-school assignment coefficient after omitting the named school assignment unit. The gray dashed vertical line denotes zero and the solid black line the corresponding full-sample estimate. The panels report the immediate beliefs-about-economics index for final-semester students and, for women in the final semester, the administrative and gateway application outcomes. Survey effects are measured in baseline control-group standard deviations; application effects are measured in percentage points. Every re-estimation retains the original randomization-block fixed effects and clusters standard errors at the remaining school units, although the figure displays point estimates only. School units appear in the same order in all panels, each on a common -2.5 to 2.5 scale. The exercise diagnoses whether a result is driven by one school; it is not a family of independent hypothesis tests and should not be used for selective inference.

Table A1: Exposure, Recall, and Implementation Process

Outcome	Assignment contrast (1)	Effect (2)	p-value [q-value] (3)	Control mean (4)	Spillover mean (5)	Direct-treatment mean (6)
Panel A. Outcome family: Follow-up exposure and recall						
Reports a female-economist activity	Treated-school assignment	-1.899 (2.084)	0.370 [0.559]	90.252	88.008	85.577
Discussed economics with family, friends, or peers	Treated-school assignment	0.054 (0.046)	0.248 [0.495]	3.564	3.870	3.986
Can name a female economist	Treated-school assignment	7.032** (3.057)	0.029 [0.078]	17.188	20.200	26.306
Someone discussed studying economics with student	Treated-school assignment	-0.025 (0.060)	0.679 [0.679]	3.547	3.488	3.694
Any or partial recall of visiting economist	Treated-school assignment	34.117** (12.392)	0.010 [0.042]	43.702	56.286	90.385
Exact recall of visiting economist	Treated-school assignment	33.917*** (12.078)	0.009 [0.042]	40.854	51.838	88.846
Can describe at least one topic from activity	Treated-school assignment	-0.047 (0.057)	0.419 [0.559]	6.114	6.197	6.090
Requests economics-program updates	Treated-school assignment	-1.425 (2.620)	0.591 [0.675]	27.848	25.051	27.000
Panel B. Outcome family: Exposure-coding robustness						
Reports activity, among yes/no responses	Treated-school assignment	-0.194 (1.039)	0.853 [0.853]	94.931	94.008	92.133
Panel C. Outcome family: Immediate implementation process						
Reports attending the female-economist session	Treated-school assignment	33.540*** (8.667)	$<$ 0.001 [0.004]	43.759	62.129	89.860
Visit increased quantitative confidence	Treated-school assignment	-0.110* (0.060)	0.079 [0.184]	6.857	6.521	6.479
Enjoyed the session	Treated-school assignment	-0.056 (0.062)	0.369 [0.516]	8.640	8.565	8.365
After visit can imagine being an economist	Treated-school assignment	0.029 (0.077)	0.710 [0.710]	4.511	4.193	4.880
Visit increased interest in economics	Treated-school assignment	0.078 (0.073)	0.299 [0.516]	5.220	5.042	5.748
Visit changed perception of economists' work	Treated-school assignment	0.242*** (0.075)	0.003 [0.012]	6.002	5.917	7.238
Can recall speaker name or affiliation	Treated-school assignment	0.030 (0.059)	0.617 [0.710]	7.072	7.075	7.389

Notes: Panel A reports follow-up exposure and recall measures; subsequent panels report implementation and information-seeking outcomes. Treated-school assignment compares all classrooms in treated schools with classrooms in control schools. Effect is the coefficient from a regression with the original randomization-block fixed effects and school-clustered standard errors. Binary outcomes and their arm means are multiplied by 100 and expressed in percentage points; continuous measures are reported in their displayed scale or in control-group standard deviations when a standardized measure is defined. Control, spillover, and direct-treatment means are unadjusted and show the three experimental arms separately. The bracketed q-value controls the false-discovery rate within the displayed process-outcome family. Recall is a diagnostic of reported exposure and should not be treated as a manipulation check free of reporting error. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A2: Randomization, Analytic Samples, and Realized Power

Panel A. Analytic samples

Sample	Sample size (1)	Role in the analysis (2)
Full student sample	4,632	Linkage across surveys and university records
Baseline respondents	3,763	Baseline characteristics and survey participation
Baseline and immediate survey respondents	3,571	Immediate effects on survey outcomes
Final-semester baseline respondents	3,763	Pre-specified survey outcomes
Women in the final-semester baseline sample	2,009	Applications to economics programs
Follow-up survey respondents	2,767	Effects on survey outcomes at follow-up
August online survey respondents	1,001	Self-reported university and field of study
Records affected by the measurement problem	120	Excluded from estimation

Panel B. Realized minimum detectable effects

Outcome	Survey wave or sample (1)	S.E. (2)	MDE (3)	Units (4)
Beliefs about economics	Follow-up	0.052	0.144	SD
Identity and self-concept	Follow-up	0.046	0.129	SD
Intentions and aspirations	Follow-up	0.039	0.110	SD
Gender stereotypes	Follow-up	0.057	0.159	SD
Beliefs about economics	Immediate	0.059	0.165	SD
Identity and self-concept	Immediate	0.025	0.071	SD
Intentions and aspirations	Immediate	0.065	0.182	SD
Gender stereotypes	Immediate	0.037	0.103	SD
Combined application measure	Women, final year	0.318	0.89	p.p.
Administrative applications	Women, final year	0.271	0.76	p.p.
Application to UANL Economics	Women, final year	0.271	0.76	p.p.
Application to Tec LEC	Women, final year	0.036	0.10	p.p.
Economics outside both registries	Women, final year	0.036	0.10	p.p.
Combined application measure	All students, final year	0.387	1.08	p.p.
Administrative applications	All students, final year	0.325	0.91	p.p.
Application to UANL Economics	All students, final year	0.254	0.71	p.p.
Application to Tec LEC	All students, final year	0.164	0.46	p.p.
Economics outside both registries	All students, final year	0.058	0.16	p.p.

Notes: The pre-specified samples include baseline respondents in the final semester of four- and six-semester programs. The survey samples overlap. We exclude the original records affected by the measurement problem and retain valid records from replacement classrooms. In Panel B, each minimum detectable effect (MDE) equals approximately 2.8 times the design-based cluster-robust standard error (S.E.), corresponding to 80 percent power for a two-sided test at the 5 percent level. Application benchmarks use the same block-adjusted school-assignment regressions as the main estimates. Survey outcomes are expressed in standard deviations (SD) relative to the baseline control group and applications in percentage points (p.p.). These calculations can overstate power when events are rare and the number of clusters is small.

Table A3: Baseline Survey Completion Relative to Classroom Enrollment

Assignment arm	Classrooms	All students	Women
	(1)	(2)	(3)
Control	75	1,905/2,843 (67.0%)	1,020/1,559 (65.4%)
Spillover	41	939/1,375 (68.3%)	518/774 (66.9%)
Direct treatment	43	919/1,417 (64.9%)	471/731 (64.4%)
Total	159	3,763/5,635 (66.8%)	2,009/3,064 (65.6%)

Notes: Each classroom cell reports completed baseline surveys divided by rostered enrollment, followed by the corresponding percentage. Total and women's enrollment come from the assignment roster; respondent gender comes from the baseline survey. Because the roster does not report men's enrollment separately, no male-specific completion rate is shown. Sixty-five respondents without a binary gender classification enter the all-student numerator but not the women's numerator. The table includes every classroom with a baseline response. The confirmed transcription error for classroom 108 is corrected from 409 to 40 rostered students. Five identifiers reused for randomly selected replacement classrooms are counted once in the denominator. The exact individual crosswalk excludes 120 original wrong-instrument records while retaining valid replacements. In classroom 21, 23 respondents exceed the roster count of 21. Ratios can reflect absence, refusal, non-completion, roster error, or timing differences and do not measure treatment receipt.

Table A4: Baseline Gender Gaps

Outcome	Men (1)	Women (2)	Women - men (3)	p-value [q-value] (4)
Panel A. Outcome family: Beliefs about economics				
Rejects money-and-finance stereotype	-0.075	0.049	0.127***	$<$ 0.001 [$<$ 0.001]
			(0.028)	
Economics improves lives	0.071	-0.071	-0.143***	$<$ 0.001 [$<$ 0.001]
			(0.036)	
Understands what economists do	0.148	-0.169	-0.320***	$<$ 0.001 [$<$ 0.001]
			(0.035)	
Panel B. Outcome family: Identity and self-concept				
Economics fits with identity	0.185	-0.267	-0.451***	$<$ 0.001 [$<$ 0.001]
			(0.027)	
Panel C. Outcome family: index				
Beliefs about economics index	0.031	-0.059	-0.090**	0.015 [0.016]
			(0.035)	
Identity and self-concept index	0.347	-0.245	-0.593***	$<$ 0.001 [$<$ 0.001]
			(0.030)	
Intentions and aspirations index	0.105	-0.069	-0.172***	$<$ 0.001 [$<$ 0.001]
			(0.028)	
Rejects gender stereotypes index	-0.305	0.262	0.570***	$<$ 0.001 [$<$ 0.001]
			(0.039)	
Panel D. Outcome family: Intentions and aspirations				
Likelihood of applying to economics	0.168	-0.186	-0.353***	$<$ 0.001 [$<$ 0.001]
			(0.033)	
Family support for studying economics	-0.037	0.019	0.058*	0.061 [0.061]
			(0.030)	
Panel E. Outcome family: orientation				
Self-assessed economics ability	0.249	-0.200	-0.449***	$<$ 0.001 [$<$ 0.001]
			(0.044)	
Self-assessed mathematics ability	0.293	-0.103	-0.395***	$<$ 0.001 [$<$ 0.001]
			(0.038)	
Interest in economics	0.279	-0.256	-0.533***	$<$ 0.001 [$<$ 0.001]
			(0.041)	
Panel F. Outcome family: Gender stereotypes				
Rejects men-naturally-better stereotype	-0.311	0.259	0.571***	$<$ 0.001 [$<$ 0.001]
			(0.040)	

Notes: The sample consists of final-semester baseline respondents in control schools. Men and Women columns report unadjusted group means. Women–men is estimated in a regression with the original randomization-block fixed effects and school-clustered standard errors. Outcomes are standardized using baseline control-group moments and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. The bracketed q-value controls the false-discovery rate across the displayed baseline outcome family. The estimates are descriptive gender differences and are not interpreted causally. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A5: Scale Items, Non-Response, and Reliability: Baseline

Item	Respondents	Missing	Non-resp. rate	Listwise sample size	Mean inter-item r	Cronbach's α
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Intentions and aspirations						
Preferred field after high school is economics or finance	3,751	61	0.016			
Likelihood of applying to economics	3,751	2	0.001			
Perceived family support for studying economics	3,751	378	0.101			
Perceived affordability of studying economics	3,751	687	0.183			
Scale summary				2,801	0.194	0.484
Panel B. Beliefs about economics						
Understands what economists do	3,637	734	0.202			
Perceived social relevance of economics	3,637	231	0.064			
Economics is mainly about money or finance (reverse-coded)	3,637	253	0.070			
Scale summary				2,641	-0.102	-0.406
Panel C. Identity and self-concept						
Economics fits self-concept	3,758	11	0.003			
Identification with economists	3,758	10	0.003			
Nervousness about studying economics (reverse-coded)	3,758	7	0.002			
Self-assessed ability in economics	3,758	2	0.001			
Self-assessed ability in mathematics	3,758	1	0.000			
Ability relative to same-gender peers	3,758	402	0.107			
Ability relative to other-gender peers	3,758	423	0.113			
Scale summary				3,296	0.362	0.798
Panel D. Gender stereotypes						
Men are naturally better at economics (reverse-coded)	3,196	231	0.072			
Perceived discrimination in economics-related careers	3,196	687	0.215			
Scale summary				2,278	-0.049	-0.104

Notes: Each panel describes one confirmatory measurement domain at the baseline wave for final-year students. The item list reproduces the exact variables entering the paper-facing measure. Respondents in the instrument sample answered at least one item in the panel; item non-response rates use that sample as the denominator. The listwise-complete sample answered every displayed item. The mean inter-item correlation averages the off-diagonal entries of the listwise correlation matrix, and Cronbach's alpha is computed from the listwise covariance matrix after standardizing items with baseline control-group moments. Reliability statistics are omitted for single-item measures. These diagnostics describe measurement consistency and do not enter the treatment-effect specification.

Table A6: Scale Items, Non-Response, and Reliability: Immediate Post-Intervention

Item	Respondents	Missing	Non-resp. rate	Listwise sample size	Mean inter-item r	Cronbach's α
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Intentions and aspirations						
Likelihood of applying to economics	3,552	0	0.000			
Perceived family support for studying economics	3,552	318	0.090			
Scale summary				3,234	0.278	0.429
Panel B. Beliefs about economics						
Understands what economists do	3,421	336	0.098			
Perceived social relevance of economics	3,421	186	0.054			
Economics is mainly about money or finance (reverse-coded)	3,421	223	0.065			
Scale summary				2,839	-0.094	-0.397
Panel C. Identity and self-concept						
Economics fits self-concept	3,554	0	0.000			
Single-item measure				3,554	–	–
Panel D. Gender stereotypes						
Men are naturally better at economics (reverse-coded)	2,873	0	0.000			
Single-item measure				2,873	–	–

Notes: Each panel describes one confirmatory measurement domain at the immediate post-intervention wave for final-year students. The item list reproduces the exact variables entering the paper-facing measure. Respondents in the instrument sample answered at least one item in the panel; item non-response rates use that sample as the denominator. The listwise-complete sample answered every displayed item. The mean inter-item correlation averages the off-diagonal entries of the listwise correlation matrix, and Cronbach's alpha is computed from the listwise covariance matrix after standardizing items with baseline control-group moments. Reliability statistics are omitted for single-item measures. These diagnostics describe measurement consistency and do not enter the treatment-effect specification.

Table A7: Scale Items, Non-Response, and Reliability: Follow-Up

Item	Respondents	Missing	Non-resp. rate	Listwise sample size	Mean inter-item r	Cronbach's α
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Intentions and aspirations						
Likelihood of applying to economics	1,896	0	0.000			
Perceived family support for studying economics	1,896	94	0.050			
Perceived affordability of studying economics	1,896	191	0.101			
Scale summary				1,662	0.307	0.573
Panel B. Beliefs about economics						
Understands what economists do	1,904	31	0.016			
Perceived social relevance of economics	1,904	49	0.026			
Economics is mainly about money or finance (reverse-coded)	1,904	70	0.037			
Scale summary				1,779	-0.121	-0.442
Panel C. Identity and self-concept						
Economics fits self-concept	1,934	13	0.007			
Identification with economists	1,934	13	0.007			
Nervousness about studying economics (reverse-coded)	1,934	13	0.007			
Self-assessed ability in economics	1,934	0	0.000			
Self-assessed ability in mathematics	1,934	0	0.000			
Ability relative to same-gender peers	1,934	62	0.032			
Ability relative to other-gender peers	1,934	5	0.003			
Scale summary				1,864	0.369	0.813
Panel D. Gender stereotypes						
Men are naturally better at economics (reverse-coded)	1,774	53	0.030			
Perceived discrimination in economics-related careers	1,774	193	0.109			
Scale summary				1,528	-0.080	-0.174

Notes: Each panel describes one confirmatory measurement domain at the follow-up wave for final-year students. The item list reproduces the exact variables entering the paper-facing measure. Respondents in the instrument sample answered at least one item in the panel; item non-response rates use that sample as the denominator. The listwise-complete sample answered every displayed item. The mean inter-item correlation averages the off-diagonal entries of the listwise correlation matrix, and Cronbach's alpha is computed from the listwise covariance matrix after standardizing items with baseline control-group moments. Reliability statistics are omitted for single-item measures. These diagnostics describe measurement consistency and do not enter the treatment-effect specification.

Table A8: Item-Level Effects: Beliefs about Economics

Survey item	Effect (1)	p-value [q-value] (2)	Control mean (3)
Panel A. Survey wave: Follow-up			
Rejects money-and-finance stereotype	0.116**	0.026 [0.077]	0.284
	(0.049)		
Economics improves lives	0.065	0.352 [0.352]	-0.422
	(0.069)		
Understands what economists do	0.098	0.145 [0.218]	-0.023
	(0.066)		
Panel B. Survey wave: Immediate			
Rejects money-and-finance stereotype	0.225***	$<$ 0.001 [$<$ 0.001]	0.139
	(0.050)		
Economics improves lives	0.089	0.231 [0.231]	-0.131
	(0.072)		
Understands what economists do	0.178**	0.029 [0.044]	0.251
	(0.077)		

Notes: The table disaggregates the beliefs about economics index into its component items. Each effect is the treated-school assignment coefficient from the same ANCOVA used for the corresponding index: the specification includes the original randomization-block fixed effects and the item's baseline value when available, and it uses school-clustered standard errors. Items are coded so that positive values indicate more economics-oriented responses or less stereotyped views. Continuous items are standardized using baseline control-group moments. The control mean is computed for the relevant control-group analytic sample. The bracketed q-value controls the false-discovery rate within the displayed item family and survey wave. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A9: Item-Level Effects: Identity and Self-Concept

Survey item	Effect (1)	p-value [q-value] (2)	Control mean (3)
Panel A. Survey wave: Follow-up			
Self-assessed economics ability	0.007	0.895 [0.905]	-0.296
	(0.054)		
Self-assessed mathematics ability	0.018	0.756 [0.905]	-0.105
	(0.058)		
Feels comfortable considering economics	0.052	0.244 [0.905]	0.015
	(0.044)		
Ability relative to other-gender peers	0.022	0.722 [0.905]	-0.308
	(0.060)		
Ability relative to same-gender peers	0.024	0.664 [0.905]	-0.329
	(0.054)		
Economics fits with identity	-0.006	0.905 [0.905]	-0.003
	(0.054)		
Can imagine an economist like me	-0.036	0.388 [0.905]	0.053
	(0.041)		
Panel B. Survey wave: Immediate			
Economics fits with identity	0.097***	$<$ 0.001 [$<$ 0.001]	-0.015
	(0.025)		

Notes: The table disaggregates the identity and self-concept index into its component items. Each effect is the treated-school assignment coefficient from the same ANCOVA used for the corresponding index: the specification includes the original randomization-block fixed effects and the item's baseline value when available, and it uses school-clustered standard errors. Items are coded so that positive values indicate more economics-oriented responses or less stereotyped views. Continuous items are standardized using baseline control-group moments. The control mean is computed for the relevant control-group analytic sample. The bracketed q-value controls the false-discovery rate within the displayed item family and survey wave. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A10: Item-Level Effects: Intentions and Aspirations

Survey item	Effect (1)	p-value [q-value] (2)	Control mean (3)
Panel A. Survey wave: Follow-up			
Could afford studying economics	0.011	0.816 [0.816]	-0.168
	(0.048)		
Likelihood of applying to economics	-0.042	0.345 [0.816]	-0.094
	(0.044)		
Family support for studying economics	0.024	0.630 [0.816]	-0.311
	(0.049)		
Panel B. Survey wave: Immediate			
Likelihood of applying to economics	0.027	0.487 [0.487]	-0.038
	(0.039)		
Family support for studying economics	-0.060	0.413 [0.487]	-0.140
	(0.072)		

Notes: The table disaggregates the intentions and aspirations index into its component items. Each effect is the treated-school assignment coefficient from the same ANCOVA used for the corresponding index: the specification includes the original randomization-block fixed effects and the item's baseline value when available, and it uses school-clustered standard errors. Items are coded so that positive values indicate more economics-oriented responses or less stereotyped views. Continuous items are standardized using baseline control-group moments. The control mean is computed for the relevant control-group analytic sample. The bracketed q-value controls the false-discovery rate within the displayed item family and survey wave. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A11: Item-Level Effects: Gender Stereotypes

Survey item	Effect (1)	p-value [q-value] (2)	Control mean (3)
Panel A. Survey wave: Follow-up			
Recognizes discrimination in economics	0.047	0.506 [0.506]	-0.061
	(0.069)		
Rejects men-naturally-better stereotype	0.040	0.485 [0.506]	-0.069
	(0.057)		
Panel B. Survey wave: Immediate			
Rejects men-naturally-better stereotype	0.052	0.170 [0.170]	0.010
	(0.037)		

Notes: The table disaggregates the gender stereotypes and barriers index into its component items. Each effect is the treated-school assignment coefficient from the same ANCOVA used for the corresponding index: the specification includes the original randomization-block fixed effects and the item's baseline value when available, and it uses school-clustered standard errors. Items are coded so that positive values indicate more economics-oriented responses or less stereotyped views. Continuous items are standardized using baseline control-group moments. The control mean is computed for the relevant control-group analytic sample. The bracketed q-value controls the false-discovery rate within the displayed item family and survey wave. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A12: Persistence for Items Common across Waves

Item	Immediate (1)	Follow-up (2)	Follow-up - immediate (3)	p-value change (4)	q-value change (5)
Understands what economists do	0.231**	0.099	-0.133*	0.099	0.174
	(0.089)	(0.065)	(0.078)		
Economics improves lives	0.236***	0.076	-0.160*	0.079	0.174
	(0.061)	(0.068)	(0.088)		
Rejects money-and-finance stereotype	0.158***	0.139***	-0.019	0.686	0.686
	(0.047)	(0.050)	(0.046)		
Economics fits with identity	0.104**	0.003	-0.101**	0.048	0.174
	(0.043)	(0.052)	(0.049)		
Likelihood of applying to economics	0.047	-0.037	-0.084*	0.095	0.174
	(0.047)	(0.041)	(0.048)		
Family support for studying economics	0.028	-0.015	-0.043	0.519	0.662
	(0.057)	(0.048)	(0.066)		
Rejects men-naturally-better stereotype	0.061*	0.020	-0.041	0.567	0.662
	(0.032)	(0.058)	(0.070)		

Notes: The sample is a balanced panel of students observed in both the immediate and follow-up surveys, and the rows are restricted to items worded and coded identically at both waves. Immediate and Follow-up are treated-school assignment effects from wave-specific ANCOVA models with randomization-block fixed effects and school-clustered standard errors. All items are standardized using baseline control-group moments and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. Follow-up–immediate is the formal difference between the two effects for the common respondent sample. The change q-value controls the false-discovery rate across the displayed inter-wave tests. This design holds item composition and respondents fixed, but the change test remains a comparison of two estimated treatment effects rather than an individual-level growth regression. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A13: Applications by Institution: Women in the Final Semester

Specification	Contrast	Effect	p-value	Control mean
	(1)	(2)	(3)	(4)
Panel A. Outcome: UANL economics application				
PAP specification	Spillover vs. control	0.374	0.302	0.196
		(0.355)		
PAP specification	Treated-school assignment	0.404	0.148	0.196
		(0.271)		
PAP specification	Direct treatment vs. control	0.440	0.160	0.196
		(0.304)		
PAP specification	Direct treatment vs. spillover	0.101	0.797	0.579
		(0.385)		
Covariate-adjusted	Treated-school assignment	0.432	0.128	0.196
		(0.275)		
Panel B. Outcome: Tec economics (LEC) application				
PAP specification	Spillover vs. control	-0.060	0.426	0.000
		(0.075)		
PAP specification	Treated-school assignment	0.031	0.401	0.000
		(0.036)		
PAP specification	Direct treatment vs. control	0.138	0.313	0.000
		(0.134)		
PAP specification	Direct treatment vs. spillover	0.210	0.341	0.000
		(0.213)		
Covariate-adjusted	Treated-school assignment	0.024	0.475	0.000
		(0.033)		
Panel C. Outcome: Tec ECSG gateway application				
PAP specification	Spillover vs. control	-0.091	0.693	0.196
		(0.227)		
PAP specification	Treated-school assignment	0.169	0.420	0.196
		(0.206)		
PAP specification	Direct treatment vs. control	0.473	0.146	0.196
		(0.316)		
PAP specification	Direct treatment vs. spillover	0.357	0.421	0.772
		(0.430)		
Covariate-adjusted	Treated-school assignment	0.134	0.475	0.196
		(0.186)		

Notes: The sample consists of women in the final semester. Each panel defines a different institution-specific or combined application outcome. UANL economics and Tec LEC are the components of the administrative measure; Tec ECSG is the broader gateway containing the economics program. Application outcomes are binary indicators multiplied by 100; effects, standard errors, and control means are therefore reported in percentage points. The treated-school assignment row compares all classrooms in treated schools with control-school classrooms. Direct treatment versus control and spillover versus control are cross-school contrasts with randomization-block fixed effects and school-clustered standard errors. Direct treatment versus spillover is identified within treated schools with school fixed effects and classroom-clustered standard errors. The control mean refers to the omitted comparison group for the displayed contrast. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A14: Applications by Institution: All Final-Semester Students

Specification	Contrast	Effect	p-value	Control mean
	(1)	(2)	(3)	(4)
Panel D. Outcome: UANL economics application				
PAP specification	Spillover vs. control	0.217	0.506	0.315
		(0.322)		
PAP specification	Treated-school assignment	0.659**	0.015	0.315
		(0.254)		
PAP specification	Direct treatment vs. control	1.133**	0.013	0.315
		(0.428)		
PAP specification	Direct treatment vs. spillover	0.891	0.148	0.532
		(0.582)		
Covariate-adjusted	Treated-school assignment	0.741***	0.006	0.315
		(0.251)		
Panel E. Outcome: Tec economics (LEC) application				
PAP specification	Spillover vs. control	-0.230	0.212	0.157
		(0.180)		
PAP specification	Treated-school assignment	-0.193	0.250	0.157
		(0.164)		
PAP specification	Direct treatment vs. control	-0.154	0.460	0.157
		(0.206)		
PAP specification	Direct treatment vs. spillover	0.073	0.729	0.106
		(0.207)		
Covariate-adjusted	Treated-school assignment	-0.166	0.280	0.157
		(0.150)		
Panel F. Outcome: Tec ECSG gateway application				
PAP specification	Spillover vs. control	-0.434	0.126	0.367
		(0.275)		
PAP specification	Treated-school assignment	-0.284	0.319	0.367
		(0.280)		
PAP specification	Direct treatment vs. control	-0.124	0.695	0.367
		(0.313)		
PAP specification	Direct treatment vs. spillover	0.252	0.200	0.639
		(0.187)		
Covariate-adjusted	Treated-school assignment	-0.249	0.362	0.367
		(0.268)		

Notes: The sample consists of all final-semester students. Each panel defines a different institution-specific or combined application outcome. UANL economics and Tec LEC are the components of the administrative measure; Tec ECSG is the broader gateway containing the economics program. Application outcomes are binary indicators multiplied by 100; effects, standard errors, and control means are therefore reported in percentage points. The treated-school assignment row compares all classrooms in treated schools with control-school classrooms. Direct treatment versus control and spillover versus control are cross-school contrasts with randomization-block fixed effects and school-clustered standard errors. Direct treatment versus spillover is identified within treated schools with school fixed effects and classroom-clustered standard errors. The control mean refers to the omitted comparison group for the displayed contrast. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A15: Administrative Linkage and Self-Report Validation

Tec program covered by university records	All final-semester students		Women in the final semester	
	Social Sciences and Government	Economics degree	Social Sciences and Government	Economics degree
	(1)	(2)	(3)	(4)
Panel A. Agreement between university records and the follow-up survey				
Share with a recorded application	0.010	0.007	0.007	0.004
Share reporting a completed application	0.034	0.034	0.028	0.028
Share classified alike by both sources	0.968	0.971	0.971	0.974
Agreement beyond chance (Cohen's kappa)	0.274	0.295	0.153	0.171
Sensitivity	0.632	0.857	0.429	0.750
Specificity	0.972	0.972	0.975	0.975
Positive predictive value	0.185	0.185	0.100	0.100
Negative predictive value	0.996	0.999	0.996	0.999
Panel B. Effect of assignment to a treatment school (percentage points)				
University records	0.695	0.494	0.926	0.613
	(0.372)	(0.293)	(0.541)	(0.310)
p-value	0.073	0.103	0.099	0.059
Follow-up survey	-1.085	-1.085	-0.980	-0.980
	(0.570)	(0.570)	(0.401)	(0.401)
p-value	0.068	0.068	0.021	0.021
Panel C. Respondents and applications				
Sample size	1,900	1,900	1,070	1,070
Students with a recorded application	19	14	7	4
Students reporting a completed application	65	65	30	30

Notes: All four columns count applications to study economics at UANL. At Tec, columns (1) and (3) count applications to the School of Social Sciences and Government, which offers several majors. Columns (2) and (4) count applications to the economics degree program. The sample includes students who answered the follow-up question about applying to economics that semester. The follow-up question covers applications to economics at any institution, whereas administrative records cover UANL and Tec only. Differences between sources therefore need not indicate reporting or linkage errors. We code a survey application as one if the student reported submitting a complete application and zero for all other nonblank answers. Sensitivity and specificity measure the share of students with and without a recorded application, respectively, whose survey answer agrees with the record. Positive and negative predictive values measure agreement with the records among students coded one and zero in the survey, respectively. Panel B includes randomization-block fixed effects and reports school-clustered standard errors in parentheses. We assess significance using two-sided p-values below 0.10.

Table A16: Sources of the Combined Application Measure

Sample	Arm	Sample size	Endline linked	Admin positive	Survey positive	Admin only	Survey only	Both	Combined
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
All final-semester students	ALL	3,763	968	28	18	19	9	9	37
All final-semester students	Control	1,905	435	9	7	5	3	4	12
All final-semester students	Spillover	939	289	6	4	5	3	1	9
All final-semester students	Direct treatment	919	244	13	7	9	3	4	16
Women in the final semester	ALL	2,009	527	8	5	5	2	3	10
Women in the final semester	Control	1,020	227	2	2	1	1	1	3
Women in the final semester	Spillover	518	173	3	0	3	0	0	3
Women in the final semester	Direct treatment	471	127	3	3	1	1	2	4

Notes: Rows partition the final-semester samples by experimental arm. Sample size is the full linked administrative sample, whereas Endline linked is the subset with a August-survey self-report. Admin positive identifies an economics application in the UANL or Tec registry. Survey positive identifies an economics-degree report in the August survey, including reports at UANL, Tec, or other institutions. Admin only, Survey only, and Both separate the overlap among students observed in the relevant sources; Combined is their union. The combined measure records detected events in the full eligible sample, without restricting it to respondents or controlling for response. Its survey component nevertheless depends on August-survey coverage. The administrative records and survey reports provide complementary coverage of the same educational choices.

Table A17: Applications to Economics by Gender

Assignment contrast	Men (1)	Women (2)	Women - men (3)	p-value [wild q-value] (4)
Panel A. Outcome family: Applications; Outcome: Applied to UANL economics or Tec ECSG gateway				
Spillover vs. control	-0.720	0.329	1.049*	0.075 [0.197]
	(0.630)	(0.484)	(0.567)	
Treated-school assignment	0.050	0.538	0.488	0.614 [0.693]
	(0.783)	(0.470)	(0.957)	
Direct treatment vs. control	0.793	0.793	-0.001	1.000 [0.999]
	(1.194)	(0.516)	(1.461)	
Direct treatment vs. spillover	1.509	0.367	-1.142	0.399 [0.460]
	(1.010)	(0.744)	(1.348)	
Panel B. Outcome family: Applications; Outcome: Applied to economics at UANL or Tec				
Spillover vs. control	-0.184	0.223	0.407	0.396 [0.615]
	(0.523)	(0.398)	(0.472)	
Treated-school assignment	0.624	0.246	-0.379	0.612 [0.693]
	(0.611)	(0.354)	(0.739)	
Direct treatment vs. control	1.401	0.290	-1.111	0.363 [0.806]
	(1.025)	(0.410)	(1.201)	
Direct treatment vs. spillover	1.540	0.052	-1.487	0.150 [0.460]
	(0.938)	(0.530)	(1.023)	

Notes: The table reports pooled treatment-by-gender models for final-semester students. Men and Women are subgroup-specific linear combinations; Women–men is the interaction coefficient. The administrative measure identifies an application to UANL Economics or Tec LEC; broad gateway additionally counts Tec ECSG. Treated-school assignment compares all students in treated schools with controls, using randomization-block fixed effects and school-clustered standard errors. Direct treatment versus control compares direct-treatment classrooms with classrooms in control schools; spillover versus control compares spillover classrooms in treated schools with classrooms in control schools; and direct treatment versus spillover compares direct-treatment and spillover classrooms within treated schools. The two cross-school contrasts include randomization-block fixed effects and use school-clustered standard errors. The within-school contrast includes school fixed effects and uses classroom-clustered standard errors. Application outcomes are binary indicators multiplied by 100; effects, standard errors, and control means are therefore reported in percentage points. The conventional p-value tests the gender interaction. The bracketed wild-bootstrap q-value retains the Benjamini–Hochberg adjustment across combined, administrative, and gateway application definitions, separately for each assignment contrast. The combined definition remains in that family even though it is not displayed here. Sparse events imply wide confidence intervals. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A18: Survey Outcomes: Direct and Spillover Effects, Immediate Post-Intervention

Contrast	Effect (SD) (1)	p-value (2)	q-value (3)	Control mean (4)
Beliefs about economics				
Spillover vs. control	0.037	0.495	0.608	0.093
	(0.054)			
Treated-school assignment	0.301***	$<$ 0.001	$<$ 0.001	0.093
	(0.059)			
Direct treatment vs. control	0.579***	$<$ 0.001	$<$ 0.001	0.093
	(0.076)			
Direct treatment vs. spillover	0.540***	$<$ 0.001	$<$ 0.001	0.113
	(0.062)			
Identity and self-concept				
Spillover vs. control	0.015	0.608	0.608	-0.015
	(0.029)			
Treated-school assignment	0.097***	$<$ 0.001	0.001	-0.015
	(0.025)			
Direct treatment vs. control	0.185***	$<$ 0.001	$<$ 0.001	-0.015
	(0.047)			
Direct treatment vs. spillover	0.176***	$<$ 0.001	$<$ 0.001	-0.049
	(0.044)			
Intentions and aspirations				
Spillover vs. control	-0.048	0.368	0.608	-0.131
	(0.052)			
Treated-school assignment	-0.038	0.567	0.567	-0.131
	(0.065)			
Direct treatment vs. control	-0.027	0.779	0.779	-0.131
	(0.096)			
Direct treatment vs. spillover	0.026	0.673	0.762	-0.211
	(0.062)			
Gender stereotypes				
Spillover vs. control	0.049	0.250	0.608	-0.009
	(0.041)			
Treated-school assignment	0.052	0.170	0.226	-0.009
	(0.037)			
Direct treatment vs. control	0.055	0.263	0.351	-0.009
	(0.049)			
Direct treatment vs. spillover	0.015	0.762	0.762	0.022
	(0.048)			

Notes: The table reports immediate post-intervention survey outcomes under the experiment's two-stage assignment. Direct treatment versus control compares direct-treatment classrooms with classrooms in control schools; spillover versus control compares spillover classrooms in treated schools with classrooms in control schools; and direct treatment versus spillover compares direct-treatment and spillover classrooms within treated schools. The two cross-school contrasts include randomization-block fixed effects and use school-clustered standard errors. The within-school contrast includes school fixed effects and uses classroom-clustered standard errors. Survey-index effects and control means are expressed in baseline control-group standard deviations and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. The control mean corresponds to the omitted group for each contrast. The bracketed q-value controls the false-discovery rate within the displayed outcome-by-contrast family. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A19: Survey Outcomes: Direct and Spillover Effects, Follow-Up

Contrast	Effect (SD) (1)	p-value (2)	q-value (3)	Control mean (4)
Beliefs about economics				
Spillover vs. control	0.050	0.415	0.831	-0.081
	(0.061)			
Treated-school assignment	0.169***	0.003	0.012	-0.081
	(0.052)			
Direct treatment vs. control	0.305***	$<$ 0.001	$<$ 0.001	-0.081
	(0.061)			
Direct treatment vs. spillover	0.301***	$<$ 0.001	$<$ 0.001	-0.044
	(0.049)			
Identity and self-concept				
Spillover vs. control	0.059	0.201	0.805	-0.114
	(0.045)			
Treated-school assignment	0.035	0.451	0.601	-0.114
	(0.046)			
Direct treatment vs. control	0.008	0.882	0.882	-0.114
	(0.053)			
Direct treatment vs. spillover	-0.029	0.466	0.466	-0.060
	(0.039)			
Intentions and aspirations				
Spillover vs. control	-0.004	0.927	0.927	-0.334
	(0.047)			
Treated-school assignment	0.020	0.618	0.618	-0.334
	(0.039)			
Direct treatment vs. control	0.047	0.354	0.473	-0.334
	(0.050)			
Direct treatment vs. spillover	0.067	0.233	0.310	-0.358
	(0.056)			
Gender stereotypes				
Spillover vs. control	0.018	0.755	0.927	-0.175
	(0.058)			
Treated-school assignment	0.057	0.324	0.601	-0.175
	(0.057)			
Direct treatment vs. control	0.102	0.123	0.247	-0.175
	(0.064)			
Direct treatment vs. spillover	0.098*	0.060	0.121	-0.116
	(0.051)			

Notes: The table reports follow-up survey outcomes under the experiment's two-stage assignment. Direct treatment versus control compares direct-treatment classrooms with classrooms in control schools; spillover versus control compares spillover classrooms in treated schools with classrooms in control schools; and direct treatment versus spillover compares direct-treatment and spillover classrooms within treated schools. The two cross-school contrasts include randomization-block fixed effects and use school-clustered standard errors. The within-school contrast includes school fixed effects and uses classroom-clustered standard errors. Survey-index effects and control means are expressed in baseline control-group standard deviations and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. The control mean corresponds to the omitted group for each contrast. The bracketed q-value controls the false-discovery rate within the displayed outcome-by-contrast family. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A20: Applications to Economics: Direct and Spillover Effects

Specification	Contrast	Effect	p-value	Control mean
	(1)	(2)	(3)	(4)
Panel A. Sample: Women in the final semester; Outcome: Combined application measure				
PAP specification	Spillover vs. control	0.218	0.584	0.294
		(0.394)		
PAP specification	Direct treatment vs. control	0.467	0.205	0.294
		(0.360)		
PAP specification	Direct treatment vs. spillover	0.224	0.584	0.579
		(0.400)		
Panel B. Sample: Women in the final semester; Outcome: Administrative applications				
PAP specification	Spillover vs. control	0.374	0.302	0.196
		(0.355)		
PAP specification	Direct treatment vs. control	0.440	0.160	0.196
		(0.304)		
PAP specification	Direct treatment vs. spillover	0.101	0.797	0.579
		(0.385)		
Panel C. Sample: All final-semester students; Outcome: Combined application measure				
PAP specification	Spillover vs. control	0.115	0.800	0.630
		(0.451)		
PAP specification	Direct treatment vs. control	0.892	0.139	0.630
		(0.584)		
PAP specification	Direct treatment vs. spillover	0.731	0.335	0.958
		(0.732)		
Panel D. Sample: All final-semester students; Outcome: Administrative applications				
PAP specification	Spillover vs. control	0.035	0.929	0.472
		(0.391)		
PAP specification	Direct treatment vs. control	0.817	0.103	0.472
		(0.484)		
PAP specification	Direct treatment vs. spillover	0.755	0.239	0.639
		(0.614)		

Notes: Panels A–B report women and Panels C–D all final-semester students. The administrative measure identifies applications to UANL Economics or Tec LEC. The combined measure uses administrative records and August-survey economics-degree reports, including cases absent from the registries. Both outcomes retain the full eligible sample: zero denotes no detected event. No model conditions on survey response; incomplete coverage limits the mixed-source measure. Direct treatment versus control compares direct-treatment classrooms with control-school classrooms; spillover versus control compares spillover classrooms in treated schools with controls. Both include randomization-block fixed effects and school-clustered standard errors. Direct treatment versus spillover compares classrooms within treated schools, using school fixed effects and classroom-clustered standard errors. Binary outcomes are multiplied by 100: effects and standard errors are in percentage points, and reference means are percentages. Control mean refers to the omitted arm, including spillover for the within-school contrast. The p-value tests the displayed contrast. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A21: Heterogeneity by Baseline Beliefs about Economics

Outcome	Low subgroup (1)	Effect low (2)	High subgroup (3)	Effect high (4)	High - low (5)	p-value [wild q-value] (6)
Panel A. Economic applications						
Economics application: broad gateway	Low beliefs	0.337	High beliefs	0.799	0.462	0.624 [0.648]
		(0.513)		(0.680)	(0.930)	
Administrative applications	Low beliefs	0.200	High beliefs	0.708	0.509	0.306 [0.401]
		(0.157)		(0.517)	(0.487)	
Panel B. Follow-up outcomes						
Beliefs about economics	Low beliefs	0.168**	High beliefs	0.178***	0.011	0.892 [0.977]
		(0.076)		(0.050)	(0.079)	
Identity and self-concept	Low beliefs	0.071	High beliefs	0.000	-0.071	0.297 [0.977]
		(0.058)		(0.056)	(0.067)	
Intentions and aspirations	Low beliefs	-0.015	High beliefs	0.056	0.070	0.433 [0.977]
		(0.053)		(0.063)	(0.088)	
Gender stereotypes	Low beliefs	0.047	High beliefs	0.066	0.019	0.819 [0.977]
		(0.078)		(0.056)	(0.082)	
Panel C. Immediate post-intervention outcomes						
Beliefs about economics	Low beliefs	0.376***	High beliefs	0.227***	-0.149*	0.053 [0.396]
		(0.069)		(0.071)	(0.073)	
Identity and self-concept	Low beliefs	0.087**	High beliefs	0.085*	-0.002	0.977 [0.978]
		(0.037)		(0.045)	(0.064)	
Intentions and aspirations	Low beliefs	-0.058	High beliefs	-0.001	0.057	0.289 [0.978]
		(0.064)		(0.070)	(0.053)	
Gender stereotypes	Low beliefs	0.050	High beliefs	0.058	0.008	0.888 [0.978]
		(0.053)		(0.042)	(0.059)	

Notes: The moderator is the baseline beliefs-about-economics index; low and high labels follow the prespecified split shown in the table. Panel A reports economics applications, Panel B follow-up survey outcomes, and Panel C immediate post-intervention survey outcomes. Each row is estimated from a pooled model containing treated-school assignment, the displayed subgroup indicator, and their interaction. Effect low and Effect high are subgroup-specific linear combinations; High–low is the interaction coefficient. Models include randomization-block fixed effects and use school-clustered standard errors. Survey-index effects and control means are expressed in baseline control-group standard deviations and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. Application outcomes are binary indicators multiplied by 100; effects, standard errors, and control means are therefore reported in percentage points. The conventional p-value tests the interaction. The bracketed wild q-value applies a false-discovery-rate adjustment to null-imposed school wild-bootstrap p-values within the displayed moderator family. Because the moderator is measured before assignment, the interaction is not mechanically conditioned on a post-treatment variable. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A22: Heterogeneity by Baseline Interest in Economics

Outcome	Low subgroup (1)	Effect low (2)	High subgroup (3)	Effect high (4)	High - low (5)	p-value [wild q-value] (6)
Panel A. Economic applications						
Economics application: broad gateway	Low interest	-0.266	High interest	1.227*	1.493*	0.076 [0.400]
		(0.363)		(0.655)	(0.809)	
Administrative applications	Low interest	0.012	High interest	0.736	0.724	0.162 [0.400]
		(0.118)		(0.489)	(0.504)	
Panel B. Follow-up outcomes						
Beliefs about economics	Low interest	0.167**	High interest	0.169***	0.002	0.983 [0.982]
		(0.078)		(0.056)	(0.085)	
Identity and self-concept	Low interest	0.128*	High interest	-0.022	-0.150*	0.082 [0.977]
		(0.075)		(0.053)	(0.083)	
Intentions and aspirations	Low interest	0.009	High interest	0.022	0.014	0.877 [0.977]
		(0.060)		(0.055)	(0.087)	
Gender stereotypes	Low interest	0.000	High interest	0.095	0.095	0.356 [0.977]
		(0.082)		(0.072)	(0.101)	
Panel C. Immediate post-intervention outcomes						
Beliefs about economics	Low interest	0.445***	High interest	0.221***	-0.223***	0.008 [0.110]
		(0.080)		(0.064)	(0.078)	
Identity and self-concept	Low interest	0.074*	High interest	0.099**	0.025	0.701 [0.978]
		(0.043)		(0.036)	(0.064)	
Intentions and aspirations	Low interest	-0.060	High interest	-0.028	0.033	0.634 [0.978]
		(0.069)		(0.073)	(0.068)	
Gender stereotypes	Low interest	0.033	High interest	0.063	0.031	0.646 [0.978]
		(0.057)		(0.043)	(0.066)	

Notes: The moderator is baseline interest in economics; low and high labels follow the prespecified split shown in the table. Panel A reports economics applications, Panel B follow-up survey outcomes, and Panel C immediate post-intervention survey outcomes. Each row is estimated from a pooled model containing treated-school assignment, the displayed subgroup indicator, and their interaction. Effect low and Effect high are subgroup-specific linear combinations; High–low is the interaction coefficient. Models include randomization-block fixed effects and use school-clustered standard errors. Survey-index effects and control means are expressed in baseline control-group standard deviations and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. Application outcomes are binary indicators multiplied by 100; effects, standard errors, and control means are therefore reported in percentage points. The conventional p-value tests the interaction. The bracketed wild q-value applies a false-discovery-rate adjustment to null-imposed school wild-bootstrap p-values within the displayed moderator family. Because the moderator is measured before assignment, the interaction is not mechanically conditioned on a post-treatment variable. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A23: Heterogeneity by Baseline Mathematics Confidence

Outcome	Low subgroup (1)	Effect low (2)	High subgroup (3)	Effect high (4)	High - low (5)	p-value [wild q-value] (6)
Panel A. Economic applications						
Economics application: broad gateway	Low math confidence	0.080 (0.394)	High math confidence	1.069 (0.795)	0.988 (0.978)	0.321 [0.401]
Administrative applications	Low math confidence	0.039 (0.098)	High math confidence	0.812 (0.569)	0.773 (0.560)	0.178 [0.400]
Panel B. Follow-up outcomes						
Beliefs about economics	Low math confidence	0.152** (0.065)	High math confidence	0.178*** (0.058)	0.026 (0.070)	0.716 [0.977]
Identity and self-concept	Low math confidence	0.014 (0.055)	High math confidence	0.038 (0.058)	0.024 (0.073)	0.747 [0.977]
Intentions and aspirations	Low math confidence	0.049 (0.062)	High math confidence	-0.019 (0.056)	-0.068 (0.089)	0.453 [0.977]
Gender stereotypes	Low math confidence	0.062 (0.090)	High math confidence	0.043 (0.076)	-0.019 (0.124)	0.878 [0.977]
Panel C. Immediate post-intervention outcomes						
Beliefs about economics	Low math confidence	0.288*** (0.064)	High math confidence	0.309*** (0.074)	0.021 (0.076)	0.781 [0.978]
Identity and self-concept	Low math confidence	0.118** (0.043)	High math confidence	0.073** (0.027)	-0.045 (0.054)	0.412 [0.978]
Intentions and aspirations	Low math confidence	-0.042 (0.082)	High math confidence	-0.039 (0.059)	0.003 (0.065)	0.961 [0.978]
Gender stereotypes	Low math confidence	0.053 (0.047)	High math confidence	0.051 (0.048)	-0.003 (0.060)	0.965 [0.978]

Notes: The moderator is baseline self-assessed mathematics ability; low and high labels follow the prespecified split shown in the table. Panel A reports economics applications, Panel B follow-up survey outcomes, and Panel C immediate post-intervention survey outcomes. Each row is estimated from a pooled model containing treated-school assignment, the displayed subgroup indicator, and their interaction. Effect low and Effect high are subgroup-specific linear combinations; High–low is the interaction coefficient. Models include randomization-block fixed effects and use school-clustered standard errors. Survey-index effects and control means are expressed in baseline control-group standard deviations and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. Application outcomes are binary indicators multiplied by 100; effects, standard errors, and control means are therefore reported in percentage points. The conventional p-value tests the interaction. The bracketed wild q-value applies a false-discovery-rate adjustment to null-imposed school wild-bootstrap p-values within the displayed moderator family. Because the moderator is measured before assignment, the interaction is not mechanically conditioned on a post-treatment variable. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A24: Heterogeneity by Student Gender

Outcome	Low subgroup (1)	Effect low (2)	High subgroup (3)	Effect high (4)	High - low (5)	p-value [wild q-value] (6)
Panel A. Economic applications						
Economics application: broad gateway	Male	0.050	Female	0.538	0.488	0.614 [0.696]
		(0.783)		(0.470)	(0.957)	
Administrative applications	Male	0.624	Female	0.246	-0.379	0.612 [0.696]
		(0.611)		(0.354)	(0.739)	
Panel B. Follow-up outcomes						
Beliefs about economics	Male	0.127*	Female	0.194***	0.067	0.362 [0.500]
		(0.074)		(0.053)	(0.072)	
Identity and self-concept	Male	-0.024	Female	0.089	0.113*	0.078 [0.369]
		(0.052)		(0.057)	(0.062)	
Intentions and aspirations	Male	0.034	Female	0.026	-0.008	0.931 [0.932]
		(0.074)		(0.048)	(0.094)	
Gender stereotypes	Male	0.113	Female	0.016	-0.097	0.285 [0.500]
		(0.084)		(0.064)	(0.089)	
Panel C. Immediate post-intervention outcomes						
Beliefs about economics	Male	0.268***	Female	0.342***	0.074	0.289 [0.477]
		(0.071)		(0.067)	(0.069)	
Identity and self-concept	Male	0.017	Female	0.157***	0.140**	0.011 [0.070]
		(0.042)		(0.030)	(0.051)	
Intentions and aspirations	Male	-0.028	Female	-0.035	-0.007	0.913 [0.911]
		(0.087)		(0.060)	(0.066)	
Gender stereotypes	Male	0.097*	Female	0.018	-0.080	0.335 [0.477]
		(0.057)		(0.051)	(0.081)	

Notes: The moderator is student gender, with men as the low/reference subgroup and women as the high subgroup. Panel A reports economics applications, Panel B follow-up survey outcomes, and Panel C immediate post-intervention survey outcomes. Each row is estimated from a pooled model containing treated-school assignment, the displayed subgroup indicator, and their interaction. Effect low and Effect high are subgroup-specific linear combinations; High–low is the interaction coefficient. Models include randomization-block fixed effects and use school-clustered standard errors. Survey-index effects and control means are expressed in baseline control-group standard deviations and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. Application outcomes are binary indicators multiplied by 100; effects, standard errors, and control means are therefore reported in percentage points. The conventional p-value tests the interaction. The bracketed wild q-value applies a false-discovery-rate adjustment to null-imposed school wild-bootstrap p-values within the displayed moderator family. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.

Table A25: Differential Attrition

Outcome/wave	Estimand	Difference	p-value	Control mean	Sample size	Schools
	(1)	(2)	(3)	(4)	(5)	(6)
Observed in the immediate survey	Treated-school assignment	0.006 (0.013)	0.634	0.950	3,763	28
Observed in the immediate survey	Direct treatment vs. control	0.013 (0.015)	0.377	0.950	3,763	28
Observed in the immediate survey	Spillover vs. control	-0.001 (0.016)	0.956	0.950	3,763	28
Observed in the immediate survey	Direct treatment vs. spillover	0.016 (0.019)	0.410		1,858	15
Observed in the follow-up survey	Treated-school assignment	0.137* (0.070)	0.061	0.483	3,763	28
Observed in the follow-up survey	Direct treatment vs. control	0.153* (0.081)	0.070	0.483	3,763	28
Observed in the follow-up survey	Spillover vs. control	0.122* (0.069)	0.086	0.483	3,763	28
Observed in the follow-up survey	Direct treatment vs. spillover	0.016 (0.055)	0.770		1,858	15
Observed in the online endline	Treated-school assignment	0.031 (0.025)	0.229	0.228	3,763	28
Observed in the online endline	Direct treatment vs. control	0.006 (0.027)	0.810	0.228	3,763	28
Observed in the online endline	Spillover vs. control	0.053* (0.028)	0.068	0.228	3,763	28
Observed in the online endline	Direct treatment vs. spillover	-0.047** (0.022)	0.048		1,858	15

Notes: The analysis population consists of final-year students observed at baseline. Each row uses an indicator for being observed in the named survey or analytic sample as the dependent variable. Treated-school assignment, direct treatment versus control, and spillover versus control include the original randomization-block fixed effects and use school-clustered standard errors. Direct treatment versus spillover is estimated within treated schools with school fixed effects and classroom-clustered standard errors. Difference is the assignment effect on observation probability; control mean is the observation rate in the omitted comparison group. Estimates are reported in probability units, so 0.043 corresponds to 4.3 percentage points. The table diagnoses selection into observed samples and motivates the weighting and bounding exercises; it does not by itself correct outcome estimates. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A26: Attrition Corrections: Inverse-Probability Weighting, Follow-Up

Method	Estimate	p-value	Control mean
	(1)	(2)	(3)
Panel A. Beliefs about economics			
Complete cases	0.169***	0.003	-0.078
	(0.052)		
Stabilized inverse-probability weights	0.141**	0.010	-0.083
	(0.051)		
IPW, weights trimmed at the 99th percentile	0.141**	0.010	-0.084
	(0.051)		
Expanded ANCOVA	0.143***	0.009	-0.081
	(0.051)		
Panel B. Identity and self-concept			
Complete cases	0.035	0.451	-0.116
	(0.046)		
Stabilized inverse-probability weights	0.027	0.564	-0.089
	(0.046)		
IPW, weights trimmed at the 99th percentile	0.027	0.562	-0.092
	(0.045)		
Expanded ANCOVA	0.027	0.554	-0.114
	(0.045)		
Panel C. Intentions and aspirations			
Complete cases	0.020	0.618	-0.333
	(0.039)		
Stabilized inverse-probability weights	0.017	0.688	-0.347
	(0.041)		
IPW, weights trimmed at the 99th percentile	0.015	0.720	-0.342
	(0.041)		
Expanded ANCOVA	0.011	0.781	-0.334
	(0.038)		
Panel D. Gender stereotypes			
Complete cases	0.057	0.324	-0.120
	(0.057)		
Stabilized inverse-probability weights	0.079	0.188	-0.239
	(0.059)		
IPW, weights trimmed at the 99th percentile	0.078	0.191	-0.237
	(0.058)		
Expanded ANCOVA	0.063	0.264	-0.175
	(0.055)		

Notes: Panels correspond to the four follow-up survey-index families. Complete cases reproduce the unweighted ANCOVA on observed respondents. Stabilized inverse-probability weighting (IPW) models follow-up availability as a function of assignment, randomization strata, the matching baseline index, and predetermined baseline covariates. The capped specification limits weights at the 99th percentile. Outcome models retain randomization-block fixed effects and school-clustered standard errors. Estimates and control means are in baseline control-group standard deviations. IPW recovers the assignment effect under correct response-propensity specification and selection on observables; it does not address selection on unobserved determinants of the outcome. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A27: Attrition Corrections: Inverse-Probability Weighting, Immediate Post-Intervention

Method	Estimate (1)	p-value (2)	Control mean (3)
Panel A. Beliefs about economics			
Complete cases	0.301***	$<$ 0.001	0.102
	(0.059)		
Stabilized inverse-probability weights	0.297***	$<$ 0.001	0.084
	(0.057)		
IPW, weights trimmed at the 99th percentile	0.298***	$<$ 0.001	0.086
	(0.057)		
Expanded ANCOVA	0.299***	$<$ 0.001	0.093
	(0.058)		
Panel B. Identity and self-concept			
Complete cases	0.097***	$<$ 0.001	-0.015
	(0.025)		
Stabilized inverse-probability weights	0.088***	$<$ 0.001	-0.014
	(0.022)		
IPW, weights trimmed at the 99th percentile	0.090***	$<$ 0.001	-0.013
	(0.024)		
Expanded ANCOVA	0.090***	$<$ 0.001	-0.015
	(0.024)		
Panel C. Intentions and aspirations			
Complete cases	-0.038	0.567	-0.131
	(0.065)		
Stabilized inverse-probability weights	-0.036	0.564	-0.138
	(0.061)		
IPW, weights trimmed at the 99th percentile	-0.035	0.568	-0.138
	(0.061)		
Expanded ANCOVA	-0.034	0.582	-0.131
	(0.061)		
Panel D. Gender stereotypes			
Complete cases	0.052	0.170	0.010
	(0.037)		
Stabilized inverse-probability weights	0.025	0.487	-0.051
	(0.036)		
IPW, weights trimmed at the 99th percentile	0.023	0.518	-0.049
	(0.036)		
Expanded ANCOVA	0.032	0.327	-0.009
	(0.032)		

Notes: Panels correspond to the four immediate post-intervention survey-index families. Complete cases reproduce the unweighted ANCOVA on observed respondents. Stabilized inverse-probability weighting (IPW) models immediate-survey availability as a function of assignment, randomization strata, the matching baseline index, and predetermined baseline covariates. The capped specification limits weights at the 99th percentile. Outcome models retain randomization-block fixed effects and school-clustered standard errors. Estimates and control means are in baseline control-group standard deviations. IPW recovers the assignment effect under correct response-propensity specification and selection on observables; it does not address selection on unobserved determinants of the outcome. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A28: Attrition Corrections: Lee Bounds

Outcome family	Estimated bounds	95% CI: lower bound	95% CI: upper bound	Trim share
	(1)	(2)	(3)	(4)
Panel A. Follow-up				
Beliefs about economics	[-0.075, 0.431]	[-0.408, 0.258]	[0.175, 0.687]	0.144
Identity and self-concept	[-0.165, 0.316]	[-0.519, 0.189]	[0.060, 0.572]	0.144
Intentions and aspirations	[-0.207, 0.364]	[-0.594, 0.179]	[0.006, 0.722]	0.141
Gender stereotypes	[-0.155, 0.357]	[-0.533, 0.224]	[0.022, 0.691]	0.165
Panel B. Immediate post-intervention				
Beliefs about economics	[0.221, 0.351]	[0.084, 0.357]	[0.231, 0.471]	0.021
Identity and self-concept	[0.032, 0.040]	[-0.084, 0.149]	[-0.065, 0.145]	0.002
Intentions and aspirations	[-0.066, -0.051]	[-0.222, 0.090]	[-0.203, 0.101]	0.002
Gender stereotypes	[0.000, 0.211]	[-0.099, 0.099]	[0.037, 0.386]	0.063

Notes: Panel A reports follow-up outcomes and Panel B immediate post-intervention outcomes, all in baseline control-group standard deviations. The estimated bounds are the Lee lower and upper bounds under monotone sample selection: the outcome distribution in the arm with the higher observation rate is trimmed to equalize response rates. The next two columns report pointwise 95% normal-approximation confidence intervals for the lower and upper bounds, respectively. These describe uncertainty in each endpoint; the estimated bounds are not a confidence interval. Trim share is the exact fractional share removed from the relevant tail. Uncertainty uses 999 bootstrap replications that resample the original school assignment units within the original randomization blocks. The monotonicity restriction is not testable from the observed outcomes; the bounds are sensitivity analyses and do not replace the preregistered ANCOVA or the IPW estimates.

Table A29: Full Sample versus Balanced Panel

Outcome	Full: effect	Full: p-value	Full: sample size	Balanced: effect	Balanced: p-value	Balanced: sample size
	(1)	(2)	(3)	(4)	(5)	(6)
Immediate outcomes						
Likelihood of applying to economics	0.027	0.487	3,549	0.047	0.328	1,846
	(0.039)			(0.047)		
Family support for studying economics	-0.060	0.413	3,059	0.028	0.623	1,554
	(0.072)			(0.057)		
Rejects the money-and-finance stereotype	0.225***	< 0.001	2,998	0.158***	0.002	1,536
	(0.050)			(0.047)		
Economics fits the student's identity	0.097***	< 0.001	3,551	0.104**	0.023	1,871
	(0.025)			(0.043)		
Rejects the men-are-naturally-better stereotype	0.052	0.170	2,563	0.061*	0.068	1,290
	(0.037)			(0.032)		
Economics improves people's lives	0.089	0.231	3,035	0.236***	< 0.001	1,577
	(0.072)			(0.061)		
Understands what economists do	0.178**	0.029	2,538	0.231**	0.015	1,331
	(0.077)			(0.089)		
Follow-up outcomes						
Likelihood of applying to economics	-0.042	0.345	1,894	-0.037	0.378	1,846
	(0.044)			(0.041)		
Family support for studying economics	0.024	0.630	1,651	-0.015	0.758	1,554
	(0.049)			(0.048)		
Rejects the money-and-finance stereotype	0.116**	0.026	1,677	0.139***	0.009	1,536
	(0.049)			(0.050)		
Economics fits the student's identity	-0.006	0.905	1,920	0.003	0.948	1,871
	(0.054)			(0.052)		
Rejects the men-are-naturally-better stereotype	0.040	0.485	1,432	0.020	0.737	1,290
	(0.057)			(0.058)		
Economics improves people's lives	0.065	0.352	1,709	0.076	0.278	1,577
	(0.069)			(0.068)		
Understands what economists do	0.098	0.145	1,454	0.099	0.144	1,331
	(0.066)			(0.065)		

Notes: Panels separate the two waves. Full and Balanced compare the full available sample at each wave with a balanced panel requiring the same student to be observed in both the immediate and follow-up surveys. Each estimate uses 28 school clusters and reports the treated-school assignment effect from an ANCOVA with the original randomization-block fixed effects and school-clustered standard errors. Survey-index effects and control means are expressed in baseline control-group standard deviations and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. The full sample maximizes precision but allows respondent composition to change across waves; the balanced panel holds respondents fixed but conditions on observation at both post-treatment waves. Differences between rows can therefore reflect both precision and selection and are not formal tests of persistence unless a change test is reported. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A30: Main Estimates without Covariate Adjustment

Outcome	Effect (1)	p-value (2)	Control mean (3)
Immediate post-intervention outcomes			
Intentions and aspirations	-0.059	0.474	-0.131
	(0.081)		
Beliefs about economics	0.284***	$<$ 0.001	0.093
	(0.058)		
Identity and self-concept	0.037	0.480	-0.015
	(0.052)		
Gender stereotypes	0.046	0.458	-0.009
	(0.061)		
Follow-up outcomes			
Intentions and aspirations	0.028	0.674	-0.334
	(0.067)		
Beliefs about economics	0.177**	0.011	-0.081
	(0.064)		
Identity and self-concept	0.068	0.352	-0.114
	(0.072)		
Gender stereotypes	0.094	0.243	-0.175
	(0.079)		
Administrative outcomes			
Administrative applications	0.411	0.141	0.196
	(0.271)		
Economics application: broad gateway	1.226*	0.063	0.392
	(0.633)		

Notes: Rows report unadjusted treated-school assignment effects. These regressions omit randomization-block fixed effects, the baseline outcome, and all additional precision covariates. They include a constant and use school-clustered standard errors. Survey-index effects and control means are expressed in baseline control-group standard deviations and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. Application outcomes are binary indicators multiplied by 100; effects, standard errors, and control means are therefore reported in percentage points. Control mean is the unadjusted mean in control schools for the corresponding analytic sample. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A31: Alternative ANCOVA Specifications: Immediate Post-Intervention Outcomes

Specification	Effect (1)	p-value (2)	Control mean (3)
Panel A. Outcome: Intentions and aspirations			
Randomization-block fixed effects only	-0.084	0.316	-0.131
	(0.082)		
Baseline-outcome ANCOVA	-0.038	0.567	-0.131
	(0.065)		
Registered covariates	-0.043	0.508	-0.131
	(0.063)		
Baseline ANCOVA with item-nonresponse adjustment	-0.039	0.548	-0.131
	(0.065)		
Expanded precision covariates	-0.037	0.543	-0.131
	(0.060)		
Panel B. Outcome: Beliefs about economics			
Randomization-block fixed effects only	0.269***	$<$ 0.001	0.093
	(0.061)		
Baseline-outcome ANCOVA	0.301***	$<$ 0.001	0.093
	(0.059)		
Registered covariates	0.296***	$<$ 0.001	0.093
	(0.060)		
Baseline ANCOVA with item-nonresponse adjustment	0.302***	$<$ 0.001	0.093
	(0.059)		
Expanded precision covariates	0.301***	$<$ 0.001	0.093
	(0.060)		
Panel C. Outcome: Identity and self-concept			
Randomization-block fixed effects only	0.012	0.802	-0.015
	(0.048)		
Baseline-outcome ANCOVA	0.097***	$<$ 0.001	-0.015
	(0.025)		
Registered covariates	0.091***	$<$ 0.001	-0.015
	(0.023)		
Baseline ANCOVA with item-nonresponse adjustment	0.097***	$<$ 0.001	-0.015
	(0.025)		
Expanded precision covariates	0.082***	$<$ 0.001	-0.015
	(0.021)		
Panel D. Outcome: Gender stereotypes			
Randomization-block fixed effects only	0.059	0.348	-0.009
	(0.062)		
Baseline-outcome ANCOVA	0.052	0.170	-0.009
	(0.037)		
Registered covariates	0.049	0.182	-0.009
	(0.036)		
Baseline ANCOVA with item-nonresponse adjustment	0.052	0.170	-0.009
	(0.037)		
Expanded precision covariates	0.046	0.207	-0.009
	(0.036)		

Notes: Panels report immediate post-intervention survey-index outcomes. All models estimate treated-school assignment effects with randomization-block fixed effects and school-clustered standard errors. Baseline-outcome ANCOVA adds the wave-matched baseline index. Registered covariates add gender, GPA, mathematics self-assessment, and household assets, with mean imputation and missing-value indicators. Item-nonresponse adjustment instead adds the number of unanswered baseline items in that index to baseline ANCOVA, without the registered covariates. The expanded set adds age, baseline application probability, and economics interest to the registered controls, using the same missing-data procedure. Survey-index effects and control means are expressed in baseline control-group standard deviations and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. Control mean is the unadjusted mean in control schools for the corresponding analytic sample. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A32: Alternative ANCOVA Specifications: Follow-up Outcomes

Specification	Effect (1)	p-value (2)	Control mean (3)
Panel E. Outcome: Intentions and aspirations			
Randomization-block fixed effects only	-0.004	0.948	-0.334
	(0.062)		
Baseline-outcome ANCOVA	0.020	0.618	-0.334
	(0.039)		
Registered covariates	0.000	0.991	-0.334
	(0.037)		
Baseline ANCOVA with item-nonresponse adjustment	0.018	0.617	-0.334
	(0.037)		
Expanded precision covariates	0.000	0.993	-0.334
	(0.037)		
Panel F. Outcome: Beliefs about economics			
Randomization-block fixed effects only	0.142**	0.022	-0.081
	(0.058)		
Baseline-outcome ANCOVA	0.169***	0.003	-0.081
	(0.052)		
Registered covariates	0.157***	0.005	-0.081
	(0.051)		
Baseline ANCOVA with item-nonresponse adjustment	0.169***	0.002	-0.081
	(0.050)		
Expanded precision covariates	0.160***	0.003	-0.081
	(0.048)		
Panel G. Outcome: Identity and self-concept			
Randomization-block fixed effects only	0.030	0.705	-0.114
	(0.079)		
Baseline-outcome ANCOVA	0.035	0.451	-0.114
	(0.046)		
Registered covariates	0.024	0.589	-0.114
	(0.045)		
Baseline ANCOVA with item-nonresponse adjustment	0.033	0.487	-0.114
	(0.046)		
Expanded precision covariates	0.025	0.585	-0.114
	(0.046)		
Panel H. Outcome: Gender stereotypes			
Randomization-block fixed effects only	0.078	0.261	-0.175
	(0.068)		
Baseline-outcome ANCOVA	0.057	0.324	-0.175
	(0.057)		
Registered covariates	0.031	0.572	-0.175
	(0.055)		
Baseline ANCOVA with item-nonresponse adjustment	0.057	0.317	-0.175
	(0.056)		
Expanded precision covariates	0.029	0.596	-0.175
	(0.055)		

Notes: Panels report follow-up survey-index outcomes. All models estimate treated-school assignment effects with randomization-block fixed effects and school-clustered standard errors. Baseline-outcome ANCOVA adds the wave-matched baseline index. Registered covariates add gender, GPA, mathematics self-assessment, and household assets, with mean imputation and missing-value indicators. Item-nonresponse adjustment instead adds the number of unanswered baseline items in that index to baseline ANCOVA, without the registered covariates. The expanded set adds age, baseline application probability, and economics interest to the registered controls, using the same missing-data procedure. Survey-index effects and control means are expressed in baseline control-group standard deviations and oriented so that positive values indicate more economics-oriented responses or less stereotyped views. Control mean is the unadjusted mean in control schools for the corresponding analytic sample. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A33: Alternative ANCOVA Specifications: Administrative Outcomes

Specification	Effect (1)	p-value (2)	Control mean (3)
Panel I. Outcome: Administrative applications			
Randomization-block fixed effects only	0.404	0.148	0.196
	(0.271)		
Expanded precision covariates	0.432	0.128	0.196
	(0.275)		
Panel J. Outcome: Economics application: broad gateway			
Randomization-block fixed effects only	0.542	0.151	0.392
	(0.367)		
Expanded precision covariates	0.541	0.141	0.392
	(0.357)		

Notes: Panels report final-semester administrative application outcomes. The table compares a model with only the original randomization-block fixed effects with models adding the preregistered or expanded predetermined precision covariates. Standard errors are clustered at the school assignment unit. Application outcomes are binary indicators multiplied by 100; effects, standard errors, and control means are therefore reported in percentage points. The administrative measure equals one for an application to UANL Economics or Tec LEC; broad gateway additionally counts Tec ECSG. Control mean is the unadjusted mean in control schools for the corresponding analytic sample. Covariate adjustment is used to improve precision and does not redefine the school-level intention-to-treat estimand. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A34: Alternative Index Constructions: Immediate Post-Intervention Outcomes

Index construction	Effect (1)	p-value (2)	Control mean (3)
Panel E. Outcome: Beliefs about economics			
Inverse-covariance weighted index	0.286***	< 0.001	0.150
	(0.051)		
Equal-weight mean index	0.260***	< 0.001	0.173
	(0.063)		
First principal component	0.021	0.788	-0.012
	(0.075)		
Panel F. Outcome: Identity and self-concept			
Inverse-covariance weighted index	0.097***	< 0.001	-0.015
	(0.025)		
Equal-weight mean index	0.097***	< 0.001	-0.015
	(0.025)		
First principal component	0.097***	< 0.001	-0.015
	(0.025)		
Panel G. Outcome: Intentions and aspirations			
Inverse-covariance weighted index	-0.038	0.580	-0.048
	(0.068)		
Equal-weight mean index	-0.027	0.673	-0.031
	(0.064)		
First principal component	-0.028	0.673	-0.106
	(0.066)		
Panel H. Outcome: Gender stereotypes			
Inverse-covariance weighted index	0.052	0.170	0.010
	(0.037)		
Equal-weight mean index	0.052	0.170	0.010
	(0.037)		
First principal component	0.052	0.170	0.010
	(0.037)		

Notes: Panels report immediate post-intervention outcomes on a common complete-item sample within each domain. The inverse-covariance-weighted index is the preregistered measure; the equal-weight mean and first principal component assess sensitivity to item aggregation. Items retain their preregistered orientation. The first principal component maximizes baseline variation and may have loadings of opposite signs. For beliefs, it loads positively on reported understanding and social value and negatively on rejection of the money-and-finance stereotype. It therefore differs in interpretation from the positively weighted ICW and equal-weight summaries. Models include the original randomization-block fixed effects, the corresponding baseline index, and school-clustered standard errors. Effects and control means are standardized within the construction's common sample and should be compared across rows of the same panel, not directly with estimates using a larger respondent sample. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A35: Alternative Index Constructions: Follow-up Outcomes

Index construction	Effect (1)	p-value (2)	Control mean (3)
Panel A. Outcome: Beliefs about economics			
Inverse-covariance weighted index	0.154**	0.016	0.015
	(0.060)		
Equal-weight mean index	0.148**	0.039	-0.038
	(0.068)		
First principal component	0.025	0.709	-0.278
	(0.067)		
Panel B. Outcome: Identity and self-concept			
Inverse-covariance weighted index	0.021	0.616	-0.079
	(0.042)		
Equal-weight mean index	0.015	0.733	-0.175
	(0.045)		
First principal component	0.010	0.830	-0.232
	(0.044)		
Panel C. Outcome: Intentions and aspirations			
Inverse-covariance weighted index	-0.005	0.912	-0.174
	(0.047)		
Equal-weight mean index	-0.014	0.765	-0.146
	(0.047)		
First principal component	-0.008	0.861	-0.244
	(0.047)		
Panel D. Outcome: Gender stereotypes			
Inverse-covariance weighted index	0.058	0.437	-0.143
	(0.073)		
Equal-weight mean index	0.058	0.433	-0.144
	(0.073)		
First principal component	-0.042	0.442	0.050
	(0.054)		

Notes: Panels report follow-up outcomes on a common complete-item sample within each domain. The inverse-covariance-weighted index is the preregistered measure; the equal-weight mean and first principal component assess sensitivity to item aggregation. Items retain their preregistered orientation. The first principal component maximizes baseline variation and may have loadings of opposite signs. For beliefs, it loads positively on reported understanding and social value and negatively on rejection of the money-and-finance stereotype. It therefore differs in interpretation from the positively weighted ICW and equal-weight summaries. Models include the original randomization-block fixed effects, the corresponding baseline index, and school-clustered standard errors. Effects and control means are standardized within the construction's common sample and should be compared across rows of the same panel, not directly with estimates using a larger respondent sample. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A36: Alternative Definitions of an Economics Application: Women in the Final Semester

Specification	Contrast	Effect	*p*-value	Control mean
	(1)	(2)	(3)	(4)
Panel A. Outcome: Combined application measure				
PAP specification	Spillover vs. control	0.218	0.584	0.294
		(0.394)		
PAP specification	Treated-school assignment	0.333	0.304	0.294
		(0.318)		
PAP specification	Direct treatment vs. control	0.467	0.205	0.294
		(0.360)		
PAP specification	Direct treatment vs. spillover	0.224	0.584	0.579
		(0.400)		
Covariate-adjusted	Treated-school assignment	0.340	0.310	0.294
		(0.329)		
Panel B. Outcome: Administrative applications				
PAP specification	Spillover vs. control	0.374	0.302	0.196
		(0.355)		
PAP specification	Treated-school assignment	0.404	0.148	0.196
		(0.271)		
PAP specification	Direct treatment vs. control	0.440	0.160	0.196
		(0.304)		
PAP specification	Direct treatment vs. spillover	0.101	0.797	0.579
		(0.385)		
Covariate-adjusted	Treated-school assignment	0.432	0.128	0.196
		(0.275)		
Panel C. Outcome: Tec economics (LEC) application				
PAP specification	Spillover vs. control	-0.060	0.426	0.000
		(0.075)		
PAP specification	Treated-school assignment	0.031	0.401	0.000
		(0.036)		
PAP specification	Direct treatment vs. control	0.138	0.313	0.000
		(0.134)		
PAP specification	Direct treatment vs. spillover	0.210	0.341	0.000
		(0.213)		
Covariate-adjusted	Treated-school assignment	0.024	0.475	0.000
		(0.033)		

Continued on next page

Table A36 continued

Specification	Contrast	Effect	p-value	Control mean
	(1)	(2)	(3)	(4)
Panel D. Outcome: Economics application: broad gateway				
PAP specification	Spillover vs. control	0.343	0.427	0.392
		(0.426)		
PAP specification	Treated-school assignment	0.542	0.151	0.392
		(0.367)		
PAP specification	Direct treatment vs. control	0.775*	0.062	0.392
		(0.399)		
PAP specification	Direct treatment vs. spillover	0.248	0.579	1.351
		(0.436)		
Covariate-adjusted	Treated-school assignment	0.542	0.145	0.392
		(0.362)		
Panel E. Outcome: Tec ECSG gateway application				
PAP specification	Spillover vs. control	-0.091	0.693	0.196
		(0.227)		
PAP specification	Treated-school assignment	0.169	0.420	0.196
		(0.206)		
PAP specification	Direct treatment vs. control	0.473	0.146	0.196
		(0.316)		
PAP specification	Direct treatment vs. spillover	0.357	0.421	0.772
		(0.430)		
Covariate-adjusted	Treated-school assignment	0.134	0.475	0.196
		(0.186)		

Notes: The sample consists of women in the final semester. Panels compare the combined measure, which uses administrative records and survey reports, with the administrative measure for UANL Economics or Tec LEC, the measure including the Tec ECSG gateway, and the institution-specific components. Application outcomes are binary indicators multiplied by 100; effects, standard errors, and control means are therefore reported in percentage points. Direct treatment versus control compares direct-treatment classrooms with classrooms in control schools; spillover versus control compares spillover classrooms in treated schools with classrooms in control schools; and direct treatment versus spillover compares direct-treatment and spillover classrooms within treated schools. The two cross-school contrasts include randomization-block fixed effects and use school-clustered standard errors. The within-school contrast includes school fixed effects and uses classroom-clustered standard errors. The combined measure uses administrative records and August-survey economics-degree reports, including cases at UANL, Tec, and other institutions not detected in the registries. This outcome is defined for the full eligible sample, with zero indicating no detected event. Models do not condition on survey response, although incomplete coverage limits interpretation of the mixed-source measure. These comparisons assess whether the findings depend on program coverage or the sources used to observe educational choices. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A37: Alternative Definitions of an Economics Application: All Final-Semester Students

Specification	Contrast	Effect	p-value	Control mean
	(1)	(2)	(3)	(4)
Panel F. Outcome: Combined application measure				
PAP specification	Spillover vs. control	0.115	0.800	0.630
		(0.451)		
PAP specification	Treated-school assignment	0.490	0.216	0.630
		(0.387)		
PAP specification	Direct treatment vs. control	0.892	0.139	0.630
		(0.584)		
PAP specification	Direct treatment vs. spillover	0.731	0.335	0.958
		(0.732)		
Covariate-adjusted	Treated-school assignment	0.614	0.109	0.630
		(0.371)		
Panel G. Outcome: Administrative applications				
PAP specification	Spillover vs. control	0.035	0.929	0.472
		(0.391)		
PAP specification	Treated-school assignment	0.412	0.215	0.472
		(0.325)		
PAP specification	Direct treatment vs. control	0.817	0.103	0.472
		(0.484)		
PAP specification	Direct treatment vs. spillover	0.755	0.239	0.639
		(0.614)		
Covariate-adjusted	Treated-school assignment	0.516	0.107	0.472
		(0.309)		
Panel H. Outcome: Tec economics (LEC) application				
PAP specification	Spillover vs. control	-0.230	0.212	0.157
		(0.180)		
PAP specification	Treated-school assignment	-0.193	0.250	0.157
		(0.164)		
PAP specification	Direct treatment vs. control	-0.154	0.460	0.157
		(0.206)		
PAP specification	Direct treatment vs. spillover	0.073	0.729	0.106
		(0.207)		
Covariate-adjusted	Treated-school assignment	-0.166	0.280	0.157
		(0.150)		

Continued on next page

Table A37 continued

Specification	Contrast	Effect	p-value	Control mean
	(1)	(2)	(3)	(4)
Panel I. Outcome: Economics application: broad gateway				
PAP specification	Spillover vs. control	-0.169	0.711	0.682
		(0.451)		
PAP specification	Treated-school assignment	0.321	0.427	0.682
		(0.398)		
PAP specification	Direct treatment vs. control	0.847	0.110	0.682
		(0.512)		
PAP specification	Direct treatment vs. spillover	0.934	0.121	1.171
		(0.566)		
Covariate-adjusted	Treated-school assignment	0.432	0.272	0.682
		(0.385)		
Panel J. Outcome: Tec ECSG gateway application				
PAP specification	Spillover vs. control	-0.434	0.126	0.367
		(0.275)		
PAP specification	Treated-school assignment	-0.284	0.319	0.367
		(0.280)		
PAP specification	Direct treatment vs. control	-0.124	0.695	0.367
		(0.313)		
PAP specification	Direct treatment vs. spillover	0.252	0.200	0.639
		(0.187)		
Covariate-adjusted	Treated-school assignment	-0.249	0.362	0.367
		(0.268)		

Notes: The sample consists of all final-semester students. Panels compare the combined measure, which uses administrative records and survey reports, with the administrative measure for UANL Economics or Tec LEC, the measure including the Tec ECSG gateway, and the institution-specific components. Application outcomes are binary indicators multiplied by 100; effects, standard errors, and control means are therefore reported in percentage points. Direct treatment versus control compares direct-treatment classrooms with classrooms in control schools; spillover versus control compares spillover classrooms in treated schools with classrooms in control schools; and direct treatment versus spillover compares direct-treatment and spillover classrooms within treated schools. The two cross-school contrasts include randomization-block fixed effects and use school-clustered standard errors. The within-school contrast includes school fixed effects and uses classroom-clustered standard errors. The combined measure uses administrative records and August-survey economics-degree reports, including cases at UANL, Tec, and other institutions not detected in the registries. This outcome is defined for the full eligible sample, with zero indicating no detected event. Models do not condition on survey response, although incomplete coverage limits interpretation of the mixed-source measure. These comparisons assess whether the findings depend on program coverage or the sources used to observe educational choices. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed using two-sided, unrounded p-values below 0.10.

Table A38: Design-Based and Small-Cluster Inference

Outcome	Contrast	Estimate	Cluster p-value	Wild p-value	Classroom wild p-value	RI p-value	Wild CI	Sample size	Clusters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Applications: School assignment									
Combined measure	School ITT	0.490	0.216	0.280	0.203	0.218	[-0.428, 1.413]	3,763	28
		(0.387)							
Combined measure	School ITT	0.333	0.304	0.361	0.378	0.371	[-0.430, 1.041]	2,009	28
		(0.318)							
Broad gateway	School ITT	0.321	0.427	0.496	0.361	0.457	[-0.628, 1.185]	3,763	28
		(0.398)							
Broad gateway	School ITT	0.542	0.151	0.219	0.216	0.215	[-0.328, 1.438]	2,009	28
		(0.367)							
Outside both registries	School ITT	0.096	0.110	0.198	0.120	0.201	[-0.024, 0.251]	3,763	28
		(0.058)							
Outside both registries	School ITT	0.031	0.401	0.386	0.408	0.895	[-0.011, 0.095]	2,009	28
		(0.036)							
Administrative measure	School ITT	0.412	0.215	0.280	0.210	0.232	[-0.345, 1.162]	3,763	28
		(0.325)							
Administrative measure	School ITT	0.404	0.148	0.182	0.211	0.211	[-0.213, 0.991]	2,009	28
		(0.271)							
Tec LEC	School ITT	-0.193	0.250	0.414	0.235	0.162	[-0.577, 0.110]	3,763	28
		(0.164)							
Tec LEC	School ITT	0.031	0.401	0.377	0.426	0.895	[-0.011, 0.094]	2,009	28
		(0.036)							
UANL economics	School ITT	0.659**	0.015	0.020	0.022	0.039	[0.089, 1.234]	3,763	28
		(0.254)							
UANL economics	School ITT	0.404	0.148	0.187	0.202	0.211	[-0.222, 1.016]	2,009	28
		(0.271)							
Survey outcomes: School assignment									
Follow-up: Beliefs	School ITT	0.169***	0.003	0.009	0.001	0.008	[0.044, 0.293]	1,850	28
		(0.052)							
Follow-up: Identity	School ITT	0.035	0.451	0.498	0.382	0.394	[-0.074, 0.139]	1,933	28
		(0.046)							
Follow-up: Intentions	School ITT	0.020	0.618	0.651	0.705	0.676	[-0.070, 0.106]	1,894	28
		(0.039)							
Follow-up: Stereotypes	School ITT	0.057	0.324	0.347	0.258	0.350	[-0.068, 0.181]	1,575	28
		(0.057)							
Immediate: Beliefs	School ITT	0.301***	< 0.001	< 0.001	< 0.001	< 0.001	[0.160, 0.451]	3,350	28
		(0.059)							
Immediate: Identity	School ITT	0.097***	< 0.001	0.004	0.007	0.004	[0.038, 0.155]	3,551	28
		(0.025)							

Continued on next page

Table A38 continued

Outcome	Contrast	Estimate	Cluster p-value	Wild p-value	Classroom wild p-value	RI p-value	Wild CI	Sample size	Clusters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Immediate: Intentions	School ITT	-0.038	0.567	0.632	0.515	0.607	[-0.204, 0.120]	3,549	28
		(0.065)							
Immediate: Stereotypes	School ITT	0.052	0.170	0.230	0.142	0.179	[-0.034, 0.141]	2,563	28
		(0.037)							
Applications: Direct and spillover contrasts									
Combined measure	S–C	0.115	0.800	0.814	0.769	0.820	[-0.878, 1.143]	3,763	28
		(0.451)							
Combined measure	T–C	0.892	0.139	0.179	0.169	0.066	[-0.395, 2.218]	3,763	28
		(0.584)							
Combined measure	T–S	0.731	0.277	0.296	0.296	0.348	[-0.583, 2.033]	1,858	84
		(0.668)							
Combined measure	S–C	0.218	0.584	0.606	0.624	0.619	[-0.666, 1.112]	2,009	28
		(0.394)							
Combined measure	T–C	0.467	0.205	0.248	0.315	0.285	[-0.359, 1.287]	2,009	28
		(0.360)							
Combined measure	T–S	0.224	0.643	0.654	0.654	0.692	[-0.720, 1.168]	989	83
		(0.481)							
Broad gateway	S–C	-0.169	0.711	0.748	0.662	0.735	[-1.184, 0.810]	3,763	28
		(0.451)							
Broad gateway	T–C	0.847	0.110	0.144	0.137	0.084	[-0.290, 1.956]	3,763	28
		(0.512)							
Broad gateway	T–S	0.934	0.107	0.113	0.113	0.139	[-0.197, 2.075]	1,858	84
		(0.572)							
Broad gateway	S–C	0.343	0.427	0.459	0.546	0.518	[-0.604, 1.334]	2,009	28
		(0.426)							
Broad gateway	T–C	0.775*	0.062	0.102	0.183	0.143	[-0.152, 1.712]	2,009	28
		(0.399)							
Broad gateway	T–S	0.248	0.733	0.730	0.730	0.741	[-1.206, 1.677]	989	83
		(0.725)							
Outside both registries	S–C	0.049	0.589	0.636	0.656	0.627	[-0.108, 0.227]	3,763	28
		(0.089)							
Outside both registries	T–C	0.148	0.281	0.352	0.370	0.126	[-0.132, 0.465]	3,763	28
		(0.134)							
Outside both registries	T–S	0.079	0.674	0.715	0.715	0.855	[-0.295, 0.445]	1,858	84
		(0.186)							
Outside both registries	S–C	-0.060	0.426	0.529	0.424	0.743	[-0.198, 0.053]	2,009	28
		(0.075)							
Outside both registries	T–C	0.138	0.313	0.447	0.447	0.296	[-0.060, 0.422]	2,009	28
		(0.134)							

Continued on next page

Table A38 continued

Outcome	Contrast	Estimate	Cluster p-value	Wild p-value	Classroom wild p-value	RI p-value	Wild CI	Sample size	Clusters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Outside both registries	T–S	0.210	0.299	0.321	0.321	0.456	[-0.126, 0.559]	989	83
		(0.201)							
Administrative measure	S–C	0.035	0.929	0.935	0.927	0.937	[-0.819, 0.919]	3,763	28
		(0.391)							
Administrative measure	T–C	0.817	0.103	0.130	0.118	0.049	[-0.219, 1.876]	3,763	28
		(0.484)							
Administrative measure	T–S	0.755	0.182	0.193	0.193	0.225	[-0.357, 1.875]	1,858	84
		(0.561)							
Administrative measure	S–C	0.374	0.302	0.351	0.386	0.331	[-0.415, 1.185]	2,009	28
		(0.355)							
Administrative measure	T–C	0.440	0.160	0.209	0.275	0.239	[-0.220, 1.145]	2,009	28
		(0.304)							
Administrative measure	T–S	0.101	0.830	0.836	0.836	0.839	[-0.817, 1.009]	989	83
		(0.468)							
Tec LEC	S–C	-0.230	0.212	0.265	0.146	0.176	[-0.620, 0.107]	3,763	28
		(0.180)							
Tec LEC	T–C	-0.154	0.460	0.560	0.439	0.370	[-0.597, 0.268]	3,763	28
		(0.206)							
Tec LEC	T–S	0.073	0.693	0.721	0.721	0.897	[-0.283, 0.426]	1,858	84
		(0.185)							
Tec LEC	S–C	-0.060	0.426	0.532	0.444	0.743	[-0.201, 0.052]	2,009	28
		(0.075)							
Tec LEC	T–C	0.138	0.313	0.461	0.464	0.296	[-0.060, 0.417]	2,009	28
		(0.134)							
Tec LEC	T–S	0.210	0.299	0.303	0.303	0.456	[-0.120, 0.559]	989	83
		(0.201)							
UANL economics	S–C	0.217	0.506	0.540	0.496	0.595	[-0.480, 0.931]	3,763	28
		(0.322)							
UANL economics	T–C	1.133**	0.013	0.011	0.020	0.003	[0.246, 2.089]	3,763	28
		(0.428)							
UANL economics	T–S	0.891	0.106	0.106	0.106	0.142	[-0.186, 1.960]	1,858	84
		(0.545)							
UANL economics	S–C	0.374	0.302	0.343	0.374	0.331	[-0.418, 1.191]	2,009	28
		(0.355)							
UANL economics	T–C	0.440	0.160	0.210	0.278	0.239	[-0.235, 1.153]	2,009	28
		(0.304)							
UANL economics	T–S	0.101	0.830	0.839	0.839	0.839	[-0.799, 1.008]	989	83
		(0.468)							

Continued on next page

Table A38 continued

Outcome	Contrast	Estimate	Cluster p-value	Wild p-value	Classroom wild p-value	RI p-value	Wild CI	Sample size	Clusters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Survey outcomes: Direct and spillover contrasts									
Follow-up: Beliefs	S–C	0.050	0.415	0.471	0.384	0.544	[-0.090, 0.196]	1,850	28
		(0.061)							
Follow-up: Beliefs	T–C	0.305***	< 0.001	< 0.001	< 0.001	< 0.001	[0.150, 0.449]	1,850	28
		(0.061)							
Follow-up: Beliefs	T–S	0.301***	< 0.001	< 0.001	< 0.001	< 0.001	[0.200, 0.400]	988	81
		(0.049)							
Follow-up: Identity	S–C	0.059	0.201	0.248	0.174	0.221	[-0.051, 0.157]	1,933	28
		(0.045)							
Follow-up: Identity	T–C	0.008	0.882	0.894	0.868	0.861	[-0.114, 0.134]	1,933	28
		(0.053)							
Follow-up: Identity	T–S	-0.029	0.466	0.493	0.493	0.541	[-0.106, 0.049]	1,030	81
		(0.039)							
Follow-up: Intentions	S–C	-0.004	0.927	0.930	0.943	0.941	[-0.110, 0.094]	1,894	28
		(0.047)							
Follow-up: Intentions	T–C	0.047	0.354	0.386	0.441	0.417	[-0.061, 0.155]	1,894	28
		(0.050)							
Follow-up: Intentions	T–S	0.067	0.233	0.253	0.253	0.295	[-0.050, 0.184]	1,007	81
		(0.056)							
Follow-up: Stereotypes	S–C	0.018	0.755	0.768	0.734	0.789	[-0.109, 0.144]	1,575	28
		(0.058)							
Follow-up: Stereotypes	T–C	0.102	0.123	0.145	0.138	0.151	[-0.045, 0.239]	1,575	28
		(0.064)							
Follow-up: Stereotypes	T–S	0.098*	0.060	0.060	0.060	0.108	[-0.005, 0.202]	844	80
		(0.051)							
Immediate: Beliefs	S–C	0.037	0.495	0.513	0.447	0.693	[-0.084, 0.162]	3,350	28
		(0.054)							
Immediate: Beliefs	T–C	0.579***	< 0.001	< 0.001	< 0.001	< 0.001	[0.395, 0.761]	3,350	28
		(0.076)							
Immediate: Beliefs	T–S	0.540***	< 0.001	< 0.001	< 0.001	< 0.001	[0.418, 0.661]	1,668	83
		(0.062)							
Immediate: Identity	S–C	0.015	0.608	0.593	0.675	0.723	[-0.049, 0.075]	3,551	28
		(0.029)							
Immediate: Identity	T–C	0.185***	< 0.001	0.002	< 0.001	< 0.001	[0.072, 0.289]	3,551	28
		(0.047)							
Immediate: Identity	T–S	0.176***	< 0.001	< 0.001	< 0.001	< 0.001	[0.086, 0.263]	1,755	83
		(0.044)							
Immediate: Intentions	S–C	-0.048	0.368	0.381	0.370	0.539	[-0.162, 0.066]	3,549	28
		(0.052)							

Continued on next page

Table A38 continued

Outcome	Contrast	Estimate	Cluster p-value	Wild p-value	Classroom wild p-value	RI p-value	Wild CI	Sample size	Clusters
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Immediate: Intentions	T–C	-0.027	0.779	0.857	0.795	0.720	[-0.283, 0.183]	3,549	28
		(0.096)							
Immediate: Intentions	T–S	0.026	0.673	0.695	0.695	0.728	[-0.096, 0.151]	1,753	83
		(0.062)							
Immediate: Stereotypes	S–C	0.049	0.250	0.294	0.191	0.274	[-0.045, 0.145]	2,563	28
		(0.041)							
Immediate: Stereotypes	T–C	0.055	0.263	0.306	0.292	0.216	[-0.056, 0.163]	2,563	28
		(0.049)							
Immediate: Stereotypes	T–S	0.015	0.762	0.776	0.776	0.789	[-0.084, 0.113]	1,285	83
		(0.048)							

Notes: The table compares three inferential approaches for the same displayed assignment coefficient. Cluster p-value uses the design-aligned cluster-robust standard error shown beneath the estimate. Wild p-value and Wild CI use a null-imposed wild cluster bootstrap; the classroom variant is reported as a finer-cluster sensitivity. RI p-value is obtained by conditional randomization inference: school assignments are permuted within the original blocks for the school ITT, direct-versus-control, and spillover-versus-control contrasts, whereas classroom assignments are permuted within treated schools for the direct-versus-spillover contrast. Survey estimates are in baseline control-group standard deviations and application estimates in percentage points. The methods address small-cluster and design-based uncertainty but do not alter the point estimand. Significance is assessed using two-sided, unrounded p-values below 0.10. School ITT denotes treated-school assignment; T–C, S–C, and T–S denote direct treatment versus control, spillover versus control, and direct treatment versus spillover. The RI design is conditional school-block assignment for School ITT, conditional two-stage assignment for T–C and S–C, and classroom assignment within treated schools for T–S. CI denotes confidence interval. All original outcome, contrast, and sample rows are retained.

Table A39: Placebo Outcomes and Subject-Specific Benchmarks

Outcome	Role (1)	Effect (2)	p-value [q-value] (3)	Control mean (4)
Panel A. Unrelated-subject placebos				
Interest in English	PAP placebo: English interest	-0.012 (0.038)	0.745 [0.745]	-0.086
Interest in history	PAP placebo: history interest	-0.030 (0.037)	0.427 [0.640]	-0.091
Interest in Spanish	PAP placebo: Spanish interest	-0.043 (0.029)	0.152 [0.455]	-0.055
Panel B. Mathematics outcomes				
Self-assessed mathematics ability	Related benchmark: mathematics confidence	0.018 (0.058)	0.756 [0.756]	-0.104
Interest in mathematics	Related benchmark: mathematics interest	-0.012 (0.031)	0.714 [0.756]	-0.041
Panel C. Economics outcomes				
Self-assessed economics ability	Target benchmark: economics confidence	0.007 (0.054)	0.895 [0.895]	-0.295
Interest in economics	Target benchmark: economics interest	0.036 (0.037)	0.337 [0.675]	-0.319

Notes: Panel A contains the preregistered placebo outcomes: interest in English, history, and Spanish. Panel B reports interest and self-assessed ability in mathematics, a related subject; Panel C reports the corresponding economics outcomes targeted by the intervention. Self-assessed ability records perceived competence, not performance on an objective skills test. Outcomes in Panels B and C may respond to the intervention and are not placebo tests. Each row is the treated-school assignment effect from a follow-up ANCOVA including the corresponding baseline outcome and original randomization-block fixed effects, with school-clustered standard errors. The sample includes eligible baseline students in the final year of four- and six-semester programs with the required outcome observations. Outcomes are standardized using baseline control-group moments; positive effects indicate greater interest or self-assessed ability. The bracketed q-values apply the Benjamini–Hochberg adjustment separately within the original families: the three unrelated-subject placebos, the two mathematics outcomes, and the two economics outcomes. Only Panel A serves as a falsification exercise for generalized changes in subject interest. Failure to reject those null hypotheses does not establish zero effects or rule out selective reporting, especially at the immediate survey. Standard errors are reported in parentheses. *, **, and *** denote two-sided conventional p-values below 0.10, 0.05, and 0.01, respectively. Significance is assessed at 10 percent using unrounded p-values; adjusted tests use q-values below 0.10.