

UC Davis

UC Davis Electronic Theses and Dissertations

Title

Essays on Discrimination and Diversity

Permalink

<https://escholarship.org/uc/item/0vv6q8gj>

ISBN

9798189225604

Author

Sun, Christina

Publication Date

2026-06-25

Peer reviewed|Thesis/dissertation

Essays on Discrimination and Diversity

By

CHE (CHRISTINA) SUN
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Economics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Anujit Chakraborty, Chair

Andrés Carvajal

Scott Carrell

Committee in Charge

2026

*To my parents, Haizhen Wang and Xingliang Sun –
you showed me that love makes miracles possible.*

Acknowledgments

Graduate school was perhaps the most challenging period of my life so far — it was marked by much turmoil and struggle, both academic and personal. It was, however, also a period of tremendous growth and discovery, and I owe it to the support, love, and kindness of a great number of people in my life that I am where I am today, at the finish line of my graduate studies. I do not think I possess the words enough to express the depth of my appreciation and gratitude; nevertheless, I shall try.

I owe my deepest thanks to my advisor, Anujit Chakraborty, for introducing me to behavioral and experimental economics when I took his graduate course — an introduction that changed the trajectory of my research and my career. Over the past six years, I am grateful for his tremendous generosity with his time and advice, for his unrelenting support academically and personally, and for his unwavering belief in me even when I doubted myself. He deeply cared about my personal wellbeing, not just my research, and that caring kindness has made all the difference.

I am also deeply grateful to the other members of my committee, Andrés Carvajal and Scott Carrell, for always supporting me with kindness and understanding, and for being generous with their guidance whenever I needed it. Andrés has always been a warm presence in my life at Davis, and Scott always provided incisive and actionable feedback on my research, and continued to support and guide me even after he left UC Davis. I am lucky to have both of them as my mentors.

Several other faculty members shaped this dissertation and my development as a researcher. I am grateful to Mrianne Bitler for being a mentor and friend throughout my time at Davis, for creating a welcoming space for students to find advice and community, and for her unfailing generosity with her time and advice, whether the conversation was academic or personal. I am grateful to Arman Rezaee, for generously taking me into his research group during my job market year when I needed feedback and guidance the most. He showed up when it mattered, and his guidance and support helped get me across the finish line. I am also deeply grateful to Michal Kurlaender, my mentor and honorary advisor at the California Education Lab, for being a warm and maternal presence in my academic life, and for caring about my personal wellbeing and being steadfast in her support. And I

am grateful to Janine Wilson for generously sharing her teaching advice during a formative period.

I have also benefited enormously from the theory and behavioral group at UC Davis. I am grateful to Christoph Schlom, Burkhard Schipper, Giacomo Bonnanno, Kalyani Chaudhury, Tyler Hoppenfeld, Zander Memon, Remy Beauregard, and Tina Danting Zhang, for their thoughtful feedback during the many seminars and presentations, and for their generosity with their time whenever I asked. I am grateful to Brad Barber, for being a mentor and friend, for his generous help preparing me for the job market, and for the steady kindness he has shown me throughout. I am also very fortunate during my job market year to have received helpful feedback from Diana Moreira, Marianne Bitler, James Cloyne, Emile Marin, Paul Feldman, Ellen Anderson, Sebastian Amador, Baiyu Zhou, Matthew Koestner, Tom Crooms, Rita Samudio, and Mitali Mathur. I am especially grateful to Tyler Hoppenfeld, for generously offering advice and support during the job market, and for being a kind and supportive friend. I am also grateful to Kaitlin Aparicio for handling the avalanche of recommendation letter requests during the job market with grace and effectiveness, and for always being a kind and capable help with the administrative side of graduate school.

The California Education Lab has been my home for most of graduate school. The warmth and kindness of the people there gave me a community and a sense of belonging during the years when I needed both, and I am extremely grateful for the friendship and loving embrace I was fortunate to receive. I am deeply grateful to Sherrie Reed, Michal Kurlaender, Paco Martorell, Scott Carrell, Jake Jackson, and Kramer Dykeman, for being mentors who were always ready to give guidance and feedback, and for always being caring, supportive, and kind. I am also grateful to my friends and coworkers at the lab, Teresita, Stephanie, Cassie, Audrey, Robbie, Jaime, Alex, for making the place feel like home.

I also want to express my gratitude to the people working on and supporting the AEA CSQIEP. The first CSQIEP mentoring conference in Berkeley was one of the highlights of my graduate school years, and I am deeply grateful to Kitt Carpenter, Martha Olney, Marianne Bitler, Patrick Button, Mike Martell, Ben Harrell, and the many other people who made it happen. That conference made me believe, for the first time, that I had a place in this field. It was a powerful experience to be in community, and I am grateful to the community they have built for the generous feedback,

mentorship, and friendship I have found there since.

I would probably not have made it through graduate school were it not for my therapists and the many medical professionals who cared for me. I am grateful to Jasleen Kaur, Hali Rowen, Mark Corey, and Shelley Chavoor, for doing the hard work with me and helping me discover who I am and grow into myself. I am grateful to Dr. Clare Rene, who is someone I would call a friend in another life - we share a depth of feeling and intellect, and her care has been profound. I am also grateful to Dr. Cindy Schorzman, for her caring kindness, and for always making time for me when I needed her. To Jeannine Stark - I am grateful to be able to call you a friend.

My time at Davis was made immeasurably better by my friends in the program. I am grateful to Derek Rury for being my mentor during my first year, for inviting me into his home and making me feel welcome in a new and difficult environment. I am fortunate to have become friends with Derek and then coauthors since, and none of this would have happened without his generosity and kindness during my first year. To my friends in the cohort and beyond - Kevin, Stephen, Kalyani, Tyler, Cassie, Zander, Baiyu, Ellen, Sebastian, Seojung, and Reem - thank you for sharing my sorrow and joy, for always being by my side, and for holding me up when I lacked the strength.

Outside of graduate school, I have been carried by a remarkable group of friends. I am grateful to Kennina Ip for being a close friend and confidant since college and a steady presence in my life through every chapter. Your friendship and love was a light during some of the darkest moments in my life. I am deeply grateful to Bret Stevens and Lindsey Hall, two of my best friends, for their steadfast support during the most difficult year of my personal life, the many meals we ate together, the quiet moments of company we shared, and the loving kindness you showed me. To Bret, for the many roads and trails we rode together, and to Lindsey, for giving me a shoulder to cry on. You two are some of the best friends I have ever had, and I would not be where I am today without you. To Jess Uzarski, for sharing her wisdom and kindness - you got me through some of the toughest stretches. To Kairo, Camille, Noa, Jessica, Alyssa, and Noah - thank you for giving me a community.

To Noah Thoron, my partner, my muse, and my best friend: I cannot say in this short space how much you mean to me. I am forever grateful to you for staying by my side through thick and thin, through happiness and sorrow, for standing behind me through every difficult period, for taking

care of me when I was sick, for holding me through darkness, for always being a loving presence by my side, and for believing in me even when I could not believe in myself. The day I met you was the luckiest day of my life. I love you so much, and I could not have done this without you.

To my parents: this dissertation is dedicated to you, but I want to thank you here too. I am grateful to you for everything you have given me — your love, your care, your investment in my education, the chance to come to America to pursue a better life. Our path has not always been an easy one, and there have been chapters that asked a great deal of all of us. When all else failed, your deep and unwavering love carried us through. I owe you a debt of gratitude I cannot repay.

Finally, to everyone else who helped me along the way whose names I have not been able to include here - you know who you are, and thank you. This dissertation belongs in part to all of you.

Abstract

“But it is not those differences between us that are separating us. It is rather our refusal to recognize those differences, and to examine the distortions which result from our misnaming them and their effects upon human behavior and expectation.”

- Audre Lorde, *Sister Outsider* (1984)

Diversity is a hallmark of humanity. Yet wherever diversity is present, so is discrimination, with widespread consequences across markets, institutions, and societies. It shapes the opportunities that individuals and groups can access, and leads to sustained inequality against some of the most vulnerable groups. These inequalities are sustained by a number of mechanisms: the beliefs employers form about applicants, the peers students grow up alongside, and the school climates that shape student aspirations. This dissertation study three mechanisms that contribute to inequality: information acquisition and formation of biased beliefs; long-run peer effects of a diverse student body; and the high school experiences shaping LGBTQ+ students' transition to higher education. Together, they shed light on the forces that sustain economic inequality, and where policy might intervene.

In the first chapter of my dissertation, I ask whether learning from accurate information entrench bias rather than correct it. This chapter shows that endogenous information acquisition sustains biased beliefs despite accurate signals. Through a combination of theory and experiment, I show that even when accurate information is available, individuals who can decide how much to learn and when to stop end up with systematically biased evidence that reinforces their existing belief bias. In contrast, when people acquire similar amounts of information but without the ability to control when to stop, their initial belief bias is significantly reduced after learning. The results identify a novel channel of persistent belief bias - the endogeneity of information acquisition - and point to exogenously structured learning as a mechanism for reducing belief-based discrimination and informing the design of information interventions.

In the second chapter, coauthored with Briana Ballis and Derek Rury, we study the long-run effects of exposure to immigrant peers on native-born students. Using a novel longitudinal dataset linking Texas public school records to higher education and labor market outcomes, we exploit within-family variation across siblings to address endogenous sorting. We find very small long-run impacts: immigrant peer exposure has no meaningful effect on college completion, employment, or earnings, and modest negative effects on college enrollment. These enrollment effects do not persist once we account for within-family sorting using an instrumental variables strategy.

In the third chapter, coauthored with Alexandria Hurtt and Michal Kurlaender, we draw on the results of a large-scale survey ($n = 9,230$) of California high school seniors to examine LGBTQ+ students' high school experiences and college expectations, including their concerns about college prior to matriculation. Findings reveal there is a statistically significant difference between the high school experiences of cisgender students and transgender and gender expansive students, who, on average, report less positive experiences overall and substantially higher rates of identity-based bullying. While transgender and gender expansive students maintained higher levels of worry about discrimination compared to cisgender students, binary-identifying transgender students were more likely to report concerns about discrimination compared to nonbinary and other gender diverse students. Despite heightened concerns about discrimination, transgender and gender expansive students maintained elevated educational aspirations, including higher rates of planned four-year enrollment and advanced degree pursuit compared to cisgender peers. Trans and gender expansive students were also more likely to intend to study humanities and the arts. These results carry implications for California education policy, including the need for targeted anti-bullying interventions, transition supports that build on student resilience, and continued collection and disaggregation of data by gender identity and sexual orientation.

Contents

1	Seeing is Believing? How Learning Modes Shape Belief Bias and Discrimination	1
1.1	Introduction	1
1.2	Experimental Design	8
1.2.1	Worker Survey	11
1.2.2	Employer Experiment	11
1.2.3	Incentives	17
1.2.4	Implementation	18
1.3	Theoretical Framework	19
1.3.1	Setup	20
1.3.2	Belief Updating	20
1.3.3	The Rational Employer	21
1.3.4	Cognitive Consistency	25
1.4	Hypotheses and Empirical Strategy	33
1.4.1	Prior Beliefs	33
1.4.2	Posterior Beliefs	34
1.4.3	Discriminatory Behavior	36
1.5	Results	38
1.5.1	Sample Characteristics	38
1.5.2	Employer Prior Belief Bias	38
1.5.3	Employer Posterior Belief Bias	42
1.5.4	Discriminatory Behavior	49
1.6	Mechanisms	53

1.6.1	Ruling Out Bias-Confidence Correlation: Evidence from the Matched Sample Treatment	54
1.6.2	Biased Information from Realization-Based Stopping	58
1.6.3	Discrimination: Biased Beliefs or Preferences?	62
1.6.4	Belief Updating Dynamics	65
1.7	Robustness	68
1.7.1	Distributional Balance of Priors across Treatment Arms	68
1.7.2	Self-Reported Mean Belief: Additional Evidence from Alternative Belief Measure	70
1.8	Conclusion	72
1.9	Sample Composition and Balance	75
1.10	Wage Decomposition Panels	80
1.11	Robustness Appendix Tables	85
1.11.1	Survey Interface Screenshots	87
1.12	Theoretical Framework Proofs	91
1.12.1	Proof of Proposition 1 (Commitment Optimum)	91
1.12.2	Proof of Proposition 2 (Voluntary $\equiv$ Commitment)	93
1.12.3	Auxiliary: Posterior-mean evolution as a martingale	94
1.12.4	Auxiliary: Unconditional variance identity $\text{Var}_0(\mu_t) = \sigma_0^2 - \sigma_t^2$	94
1.12.5	Proof of Proposition 3 (Commitment optimum under cognitive consistency) .	96
1.12.6	Proof of Proposition 4 (Voluntary vs. Commitment under cognitive consistency)	97
1.12.7	Proof of Proposition 5 (Commitment vs. Sequential: expected sample size) . .	99
1.12.8	Proof of Proposition 6 (Gap Closure)	100
1.12.9	Proof of Proposition 7 (Voluntary bias toward prior: selection channel)	100
1.12.10	Proof of Corollary 1 (Gap-closure ranking under cognitive consistency)	102
2	Immigrant Peers and the Long-Run Outcomes of US-Born Students	104
2.1	Introduction	104
2.2	Data and Setting	109
2.2.1	Immigration in Texas	109
2.2.2	Data Sources and Sample Construction	109

2.2.3	Sample Construction	112
2.2.4	Summary Statistics	113
2.3	Empirical Strategy	117
2.4	Results	122
2.4.1	Heterogeneous Impacts	125
2.5	Robustness	135
2.5.1	Assessing Sensitivity to Within Family Sorting	135
2.5.2	Immigrant Status or Other Characteristics?	141
2.5.3	Alternative Sample and Immigrant Definitions	142
2.6	Conclusion	142
2.7	Appendix Tables and Figures	145
2.8	Sibling Linkage: Details and External Validity	154
2.8.1	Algorithmic Matching Procedure	154
2.8.2	Incomplete Linkage and Sample Composition	154
2.8.3	External Validity Considerations	155
2.9	Mathematics of Transition Matrix IV	159
3	LGBTQ+ Students' High School Experiences and Academic Plans	161
3.1	Introduction	161
3.1.1	Defining Gender Identity and Sexual Orientation	162
3.2	Theoretical Background and Prior Literature	163
3.3	Data and Methods	166
3.3.1	Survey Instrument and Sample Characteristics	166
3.4	Data Analysis	170
3.5	Findings	172
3.5.1	Bullying in High School	175
3.5.2	Bullying as a Potential Mediator of High School Experiences	178
3.5.3	Concerns About College	180
3.6	Bullying as a Potential Mediator of College Worries	188
3.7	Postsecondary Educational Plans	189

3.7.1	Qualitative Findings	196
3.8	Limitations	197
3.9	Discussion and Conclusion	198
3.10	Additional Demographics	202
3.11	High School Experience Items and Construct	204
3.12	Summary Statistics of Individual High School Experience Items	205
3.13	Summary Statistics of Individual College Concern Items	207
3.14	Additional Demographics	209
3.15	High School Experience by Sexual Orientation Robustness	210
3.16	College Worries by Sexual Orientation Robustness	211
3.17	Predicted Probabilities of Degree Aspirations	214
3.18	Bullying as Mediator for High School Experience	217
3.19	Bullying as Mediator for College Worries	219
Bibliography		219

Seeing is Believing? How Learning Modes Shape Belief Bias and Discrimination

1.1. Introduction

We often form beliefs about others under limited and costly information, through sequential encounters, interviews, and observations. In labor markets, for example, employers decide how much to learn about candidates and when to stop, acquiring information through resumes, interviews, or referrals. Over time, they also form beliefs about different demographic groups based on repeated exposure to applicants from those groups. Such beliefs are an important driver of labor market discrimination (e.g. Arrow, 1973; Phelps, 1972), alongside preference-based mechanisms (G. S. Becker, 1971). Recent research highlights the role of *biased beliefs* in reinforcing discrimination (Bohren et al., 2020). Yet despite growing awareness and policy efforts, belief bias remains persistent, raising a central question: why do biased beliefs endure even when information is abundant?

In this paper, I propose and experimentally test a novel channel in the persistence of biased beliefs: the endogeneity of the information acquisition process. Through a combination of theory and experiment, I show that even when accurate information is available, individuals who can decide how much to learn and when to stop end up with systematically biased evidence that reinforces their existing belief bias. In contrast, when people acquire similar amounts of information but without the ability to control when to stop, their initial belief bias is significantly reduced after learning. This difference arises because endogenous control over information search induces *realization-based stopping*: agents are more likely to stop sampling when the realized signal sequence has produced a posterior consistent with their prior, so the evidence they end up with is systematically more aligned with that prior than the evidence they would have observed under exogenous stopping. Unlike previous accounts that focus on the interpretation or recall of a fixed information set, my approach

highlights how distorted beliefs can emerge upstream from the way people direct their own search for information. The mechanism rests on cognitive-consistency preferences over belief revision: an agent who experiences a small loss each time the running posterior moves away from the prior will, when given the option, stop sampling exactly where the running posterior has stayed close to it.

A robust literature in psychology and economics has documented that the mode through which people learn about risky prospects leads to divergent choice patterns. When people learn through sequential sampling rather than explicit description, their decisions systematically diverge, a phenomenon known as the “Description-Experience Gap” (Hertwig & Erev, 2009). Recently, Oprea and Vieider, 2024 proposed a unifying account rooted in cognitive noise: since noisy coding interacts differently with described vs. experienced probabilities, the two learning modes produce different degrees of Bayesian compression of posterior beliefs towards priors. A natural implication is that the structure of information acquisition may have broader implications for economic behavior beyond risky choice, since beliefs are a key driver in economic decision-making. This paper investigates one such dimension. Within experience-based learning, I ask whether endogenous control over the sampling process, the ability to decide when to stop, generates additional belief distortions relative to exogenous experience. Where the noisy cognition literature points to the noisy encoding of observed information, I identify a complementary, upstream channel: when information acquisition is endogenous, realization-based stopping skews the signals that are observed in the first place, producing distortion towards priors. Both channels generate similar distortion of posterior towards the priors, but operate at different stages of the belief formation process and call for different interventions.

Guided by a theoretical framework of sequential information search by Bayesian agents who derive utility from cognitive consistency, I design and implement a controlled laboratory experiment that isolates the causal impact of different information acquisition modes on belief formation and discriminatory behavior. I set up the experiment using a simulated hiring environment, where participants in the role of “employers” evaluate “workers” from Asian and Hispanic ethnic groups and make wage offers based on perceived productivity. This setting allows me to directly elicit employer beliefs and measure discrimination using their wage offers. To operationalize productivity, I use a 12-question math test taken by participants in the role of “workers”. Using a math task induces

natural belief bias while offering an objective, quantifiable measure of worker productivity. Critically, I construct the Asian and Hispanic worker groups to have identical score distributions, providing a clear “ground truth” benchmark. Any systematic differences in employer beliefs or wage offers between groups thus reflect biased beliefs or preferences, rather than any underlying differences in worker performance. I employ a fully between-subject design, where each employer is randomly assigned to evaluate only one worker ethnicity assignment, thereby reducing potential confounds.

Subjects are randomly assigned to one of three information treatments: **Voluntary**, where employers engage in fully endogenous information search and can decide whether to stop or continue after observing each draw; **Exogenous Mean**, where endogenous control over the search process is removed, and employers must draw a fixed number of random workers equal to the average number drawn in Voluntary¹; and **Exogenous 20**, where information search is also exogenous, but with a fixed sample size of 20 workers. A **Baseline** arm provides a no-information benchmark in which employers report prior beliefs and proceed directly to wage offers. In all three information treatments, employers sequentially draw random workers from their assigned worker group and observe each worker’s individual productivity in the form of score bins (1–3, 4–6, 7–9, or 10–12). This design mirrors real-world settings where information is incomplete and open to interpretation. Comparing employer beliefs in Voluntary against Exogenous Mean isolates the role of endogenous control while holding the average information content the same; Exogenous 20 doubles the amount of information acquired in Exogenous Mean, testing the effect of a larger and more representative sample.

To measure belief bias, I elicit incentivized employer beliefs about the productivity distribution of their worker group before and after information acquisition. In the final stage of the experiment, employers make a series of incentivized wage offers to ten randomly drawn individual workers. Because worker groups are statistically identical, any differences in wage offers reflect either biased beliefs or discriminatory preferences. Full details of the experimental design are provided in [Section 1.2](#).

The findings from my experiment show that the mode of information acquisition substantially affects both belief bias and discrimination. I report four results in [Section 1.5](#). First, employers hold

¹10 draws for employers with Hispanic workers, 8 for those with Asian workers

significantly biased prior beliefs against Hispanic workers. Across all Treatment Arms, employers assigned the Asian worker group expect average productivity that is 1.682 points higher than employers assigned the Hispanic worker group ($p < 0.01$), about one standard deviation of the prior belief distribution. Second, in the Baseline condition, where employers receive no information beyond their priors before submitting wage offers, this biased belief directly translates into wage discrimination: Hispanic workers receive on average about 0.60 experimental tokens less than Asian workers with similar observable profiles, roughly 27% of a standard deviation of the wage distribution. Third, the prior bias persists even when employers can endogenously acquire additional information in the Voluntary condition. Employers revise their beliefs toward greater accuracy for both groups after sampling, yet a residual Asian-Hispanic posterior gap of 0.537 ($p < 0.01$) survives, about a third of the prior gap and roughly half a standard deviation of the posterior distribution. Fourth, the residual posterior gap shrinks substantially when endogenous control over the search process is removed: it is 0.259 in Exogenous Mean (a reduction of about 51% relative to Voluntary, achieved at the same average sample size) and 0.107 in Exogenous 20 (a reduction of about 80%). These belief patterns carry through to behavior: Voluntary employers continue to offer Asian workers a wage premium of about 0.30 tokens, whereas in Exogenous Mean and Exogenous 20 the wage premium is no longer statistically distinguishable from zero.

My findings show that removing endogeneity of information acquisition, even while holding the amount of information constant, can correct distorted beliefs and meaningfully reduce discriminatory behavior. [Section 1.6](#) then asks *why* endogenous information acquisition produces systematically biased beliefs, with a rule-out-then-reveal structure.

The most natural alternative explanation for the Voluntary residual gap is a potential correlation between prior bias and prior confidence: employers with priors farther from the truth might also hold those priors with more confidence (smaller variance), thus stopping after fewer draws, and end up with a posterior that inherits the prior bias. Under this account, the bias would live in the link between an employer’s prior bias and prior confidence, not in selective stopping based on signal realizations. I test this alternative directly with the **Matched Sample** treatment, which randomly assigns the exact signal sequence drawn by a Voluntary counterpart with the same worker ethnicity assignment to a new employer, so that the receiving employer’s own prior characteristics

are statistically independent of the inherited sequence, and the within-employer pairing between prior bias, prior confidence, and stopping behavior that the alternative account relies on cannot operate. If bias-confidence correlation drives the sustained belief bias in Voluntary, the Matched Sample posterior gap should track the gap in Exogenous Mean rather than the gap in Voluntary. The data go the other way: the Matched Sample residual posterior gap is 0.708, comparable in point estimate to Voluntary’s 0.538 and well above Exogenous Mean’s 0.259. The bias is preserved through the Voluntary-to-Matched-Sample transfer even though random assignment destroys the within-employer pairing the alternative account relies on. The bias must therefore live in the realized signal sequence itself.

With the alternative ruled out, the next step is to identify what selects the biased signal sequences. Endogenous control over information search has two features: *realization-based stopping*, the ability to stop search conditional on the signal realizations observed so far, and *endogenous sample size selection*, the ability to choose how many signals to draw in total. To separate these two features, I introduce the **Commitment** treatment, which removes realization-based stopping but preserves sample-size endogeneity: employers commit to a sample size before seeing any signals and then sequentially draw exactly that many workers. The data align with the cognitive-consistency framework on three dimensions: (i) a reduced residual posterior gap in the Commitment arm, suggesting realization-based stopping is resulting in biased signal sequences; (ii) within Voluntary, employers’ stopping hazard rises with the absolute distance between the running posterior and the prior, the cognitive-consistency signature for realization-based stopping; and (iii) the monotonic posterior gap closure pattern of Voluntary > Commitment > Exogenous Mean > Exogenous 20, both for the regression-based posterior gap and for the nonparametric energy distance between Asian-arm and Hispanic-arm posterior belief distributions, which is congruent with the prediction of Corollary 1 in the theoretical framework. These results support the interpretation that biased information sets, shaped by realization-based stopping, are a key driver of belief bias. This reinforces the importance of addressing biased sampling processes as a mechanism for reducing belief-based discrimination.

This paper contributes to the growing literature on discrimination driven by biased beliefs (Bohren et al., 2019, 2020; Campos-Mercade & Mengel Friederike, 2022; Coffman et al., 2019, 2024; Eyting, 2022) by identifying a novel source of persistent belief bias: the mode of information

acquisition. Classic models in economics distinguish between taste-based discrimination driven by preferences (G. M. Becker et al., 1964) and statistical discrimination driven by accurate group-level beliefs (Arrow, 1973; Phelps, 1972). Recent work shows that inaccurate beliefs often underlie discriminatory behavior (Bohren et al., 2020), and understanding why such beliefs persist despite access to information remains a central challenge. Closely related, Coffman et al. (2024) study endogenous information acquisition in a two-arm gender-discrimination evaluation experiment and document that contrast-based stereotyping amplifies group-based disparate evaluation. This paper offers a new explanation: beliefs remain biased not because individuals lack access to corrective evidence, but because they endogenously shape the evidence they encounter through realization-based stopping. I show experimentally that even when individuals receive similar amounts of information about group performance, their beliefs and hiring decisions differ sharply depending on whether that information was acquired voluntarily or exogenously. The Asian-Hispanic belief gap persists under voluntary experience, but is eliminated when the same information is provided through an exogenously assigned sample. This demonstrates that the structure of information acquisition, beyond the amount of information, shapes belief formation and downstream decisions. My results also help explain why biased beliefs can persist even in environments rich with information: because information acquisition in everyday life is fully endogenous, individuals tend to stop prematurely when evidence aligns with their prior beliefs. As a result, even when the information source is unbiased, the endogenous nature of learning produces skewed belief updates that reinforce existing views.

This paper also connects to the growing literature on noisy cognition and internal processing constraints in economics, which shows that many behavioral anomalies can arise as a Bayesian response to imprecise internal representations of numerical information (Frydman & Jin, 2023; Oprea & Vieider, 2024). Foundational work by Woodford (2012) proposes that key features of prospect theory emerge from efficient internal coding under capacity constraints, and Khaw et al. (2021) show that a single imprecision parameter generates small-stakes risk aversion and the inverse-S pattern of probability weighting through Bayesian compression of noisy signals toward the prior. Closely related, Bohren et al. (2024) extend this framework to evaluate the effects of attention and memory constraints on

belief formation.² The noisy-cognition literature locates the source of belief compression in the noisy encoding of observed information; this paper identifies a complementary upstream channel located in the choice of which signals to observe in the first place.

I also contribute to a broader literature on biased belief updating documents how non-Bayesian updating patterns shape decisions and discriminatory behavior (Coutts, 2019; Eyting, 2022; Mobius et al., 2014; Zimmermann, 2020). Prior research shows that individuals often interpret information in self-serving ways, overweighting confirming evidence and underweighting contradictory signals (Coutts, 2019; Mobius et al., 2014; Zimmermann, 2020). I extend this literature by showing that biases in belief formation can arise even when individuals interpret signals in a fully Bayesian manner, through a complementary acquisition-side channel: endogenous selection of which signals to observe can sustain belief bias upstream of the interpretation stage where these updating biases operate.

Finally, this paper makes a novel contribution to the literature on discrimination by designing and testing an intervention that eliminates belief-based discrimination in this setting through structural changes in the learning process. While a large body of work has tested information-based interventions aimed at reducing prejudice and discrimination (Alesina et al., 2023, 2024; Bohren et al., 2020; Haaland & Roth, 2019, 2020; Kuziemko et al., 2015; Pedersen & Nielsen, 2024; Wozniak & Macneill, 2018), results are often mixed. I provide evidence that bias persists not just because of insufficient information, but because individuals control how and when to acquire it. In my experiment, when participants sequentially sample signals with full autonomy (Voluntary), belief distortions and discriminatory wage offers emerge. However, structuring the learning process through exogenous assignment of a representative sample completely eliminates both belief bias and wage discrimination, even when the information content is held constant. This identifies an intervention strategy that has been underused in the literature on belief-based discrimination: rather than simply increasing exposure to information, design learning environments that constrain realization-based stopping and motivated sampling.

The rest of the paper is structured as follows. [Section 1.2](#) describes the experimental design. [Section 1.3](#) develops a theoretical model of belief formation under sequential information search

²Their sequential treatment is sampling without replacement until all states are observed, with no optional stopping.

to provide a conceptual underpinning for belief updating under different information acquisition modes. [Section 1.4](#) formulates hypotheses and discusses empirical strategies to test them. [Section 1.5](#) presents my main empirical results on prior bias, posterior bias under Voluntary and the exogenous benchmarks, and Baseline wage discrimination. [Section 1.6](#) uses additional experimental treatments to identify the underlying mechanism, ruling out bias-confidence correlation via the Matched Sample treatment, identifying realization-based stopping as the primary channel via the Commitment treatment and the within-Voluntary hazard model, and decomposing the wage gap into belief-driven and preference-driven components. [Section 1.7](#) reports robustness checks on the distributional balance of priors across Treatment Arms and on the agreement between the estimated mean belief from the PMF and the self-reported mean belief. [Section 1.8](#) concludes and discusses potential policy implications.

1.2. Experimental Design

I design my experiment to answer three research questions. First, what are the impacts of different information acquisition modes - endogenous search and exogenous search - on employer beliefs about worker productivity? Second, how do people search for and interpret information when they have full control over information seeking, and how does it affect belief updating? And third, do these modes of learning meaningfully affect discriminatory behavior against minority workers, and to what degree?

In order to answer these questions, I set up a controlled laboratory experiment to simulate a real-world labor market. I recruit two separate groups of subjects for my experiment, “workers” and “employers”. The “workers” are an ethnicity-balanced sample of Asian and Hispanic subjects who take an incentivized math test with 12 questions, and receive a score equal to the number of questions they answered correctly. I also collect demographic information from the workers to construct worker “resumes” used in the employer hiring task. In order to avoid cross-group comparisons and reduce social desirability bias, I randomly match “employers” with one worker group (Asian or Hispanic). Employers are informed about the calculation of the worker test scores and shown example questions to provide a basis for prior belief formation. I further randomize each employer into one of four information conditions and one control condition. Therefore, there are 2 orthogonal treatment axes

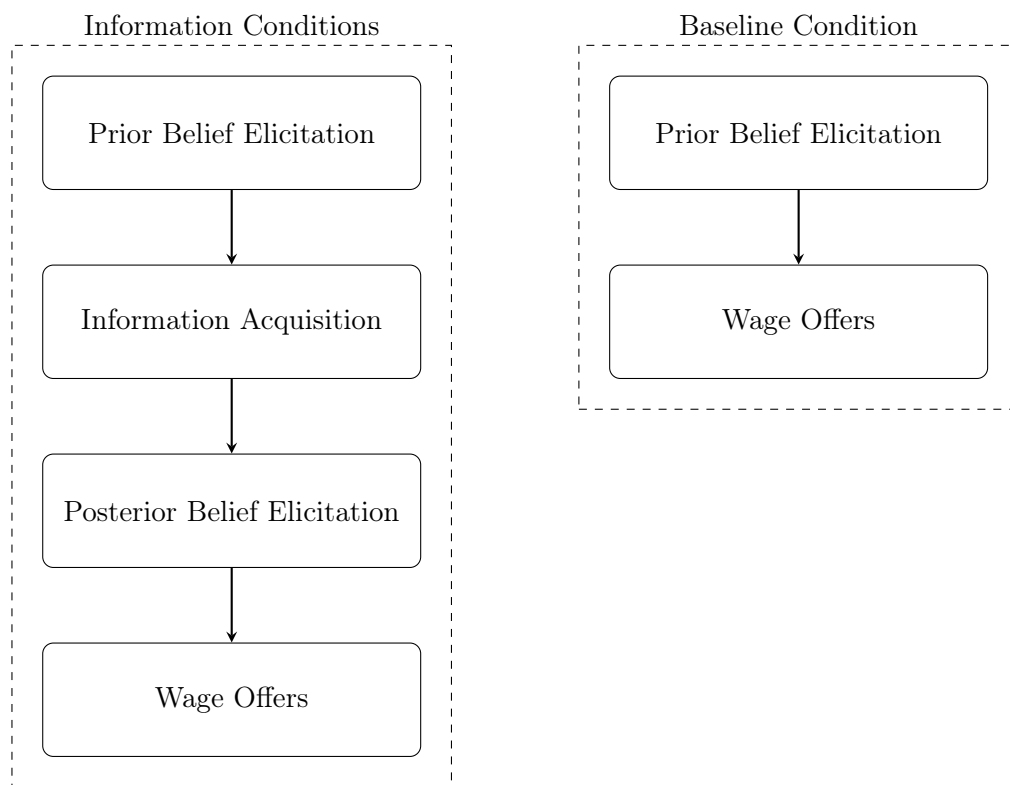
(ethnicity and information mode), with 2 and 5 treatment arms in each axis, respectively. This constitutes a 2×5 full factorial between-subject design. Employers in the information conditions go through four task stages:

- a. **Prior Belief Elicitation:** I elicit employer beliefs about the average score as well as the full distribution of scores in their assigned worker group. This first stage is identical for all employers, which ensures prior beliefs are equally distributed across treatments.
- b. **Information Acquisition:** I randomize each employer into one of three information treatments: Voluntary, Exogenous Mean, and Exogenous 20. In all conditions, employers learn about the score distribution within their assigned worker group, but I experimentally manipulate the mode of information acquisition. In Voluntary, employers sequentially sample individual workers by clicking a button, with each draw revealing which of the four score bins (1–3, 4–6, 7–9, 10–12) this worker falls into. Information in Exogenous is acquired through the same button-click interface, but with an exogenously assigned, fixed number of draws.
- c. **Posterior Belief Elicitation:** After employer learn about their worker group, I elicit updated group-level beliefs about the worker score distribution.
- d. **Wage Offers:** Finally, employers sequentially make wage offers to a series of 6 randomly selected workers from their worker group. I construct a worker profile for each worker and elicit the employers' Willingness to Pay (WTP) for each worker in an incentive compatible way using the BDM value elicitation mechanism (G. M. Becker et al., 1964).

Employers randomized into the Baseline control condition do not receive any information about their worker group and therefore do not report updated beliefs. They only go through two task stages: Prior Belief Elicitation and Wage Offers. This provides a no-information benchmark to study the baseline levels of belief bias and wage discrimination. This flow of tasks is summarized graphically in [Figure 1.1](#).

With the core structure of the experiment in place, I now turn to technical details of the experimental design. In what follows, I describe the worker tasks and the employer experiment, and how each treatment condition in the employer experiment is implemented. I also detail the design of the employer task, how I measure employer beliefs and discrimination, and the incentive structure. I

FIGURE 1.1. Order of tasks in the employer experiment



conclude this section with the implementation procedures.

1.2.1. Worker Survey

The worker survey serves two primary purposes: to obtain an objective measure of productivity and to construct the worker groups used in the employer experiment. I recruit participants who self-identify as Asian or Hispanic to take on the role of “workers”. Each worker completes an incentivized 12-question math test under a six-minute time limit. The questions are adapted from SAT Math questions and workers earn a \$0.2 bonus per correct answer, in addition to a \$2 participation payment. After the time expires, the test is automatically submitted, the workers are shown their score and corresponding bonus payment. Finally, they complete a demographics questionnaire in which I collect their age, gender, race, ethnicity, education, and favorite color. Select demographic characteristics are later used to construct worker resumes shown to the employers.

Using a math test allows me to induce natural belief bias at baseline, since math is a task domain in which Asian workers are stereotyped as higher performing relative to Hispanic workers. A score out of 12 allows me to elicit the entire subjective belief distribution of employers without imposing excessive effort and time costs, which allows identification of the source of discrimination. After collecting data from the worker survey, I construct the two worker groups: an Asian worker group and a Hispanic worker group, each consisting of 100 workers. Importantly, I construct the two groups to have identical test score distributions. This ensures that the two worker groups are in fact equally productive, providing a clear objective benchmark to measure updated beliefs against. By holding actual performance constant across groups, I can rule out accurate statistical discrimination, and distinguish the source of any observed discrimination between taste-based discrimination and discrimination driven by biased beliefs.

1.2.2. Employer Experiment

I recruit a separate group of participants for the employer experiment. Upon entering the experiment, participants are informed that they will be playing the role of employers and make evaluation and hiring decisions about some workers. I inform the employer that they have been matched with a worker group - Asian or Hispanic - consisting of 100 Prolific workers based in the United States.

This between-subject design ensures that employers never engage in explicit cross-group comparison, reducing the impact of other behavioral factors that can alter behavior, such as social desirability bias. Specifically, employer know that each worker completed a timed SAT-level math test with 12 questions, and that each worker received a score equal to the number of correct answers. I also inform them that no worker received a score of 0. This gives a common support of integers from 1 to 12 for the employer belief distribution. The employers can view sample math questions from the worker survey which aids in their initial belief formation.

Prior Belief Elicitation

After employers learn about the experiment setup, I elicit their initial beliefs about the score distribution in their worker group using two complementary measures: the employer’s full subjective belief distribution over worker scores in the group (the probability mass function), and the employer’s self-reported belief about average worker performance. The first distributional measure pins down employer beliefs and allows me to derive the estimated mean belief from the PMF (the primary belief measure throughout the paper); the second elicits a single point belief directly, providing an intuitive measure that is easy for employers to understand, enabling additional robustness checks against the PMF-derived measure (see Section 1.7). Both elicitations are incentivized.

To elicit the full belief distribution, I ask employers to report how many workers they think have each possible score in the group. This involves 12 questions: they report how many workers they think received a score of 1, how many workers received a score of 2, etc. Their answers for these 12 questions must add up to 100, reflecting the total number of workers in the group. To elicit the self-reported mean belief, I use a text box where they can enter a positive number less than or equal to 12 (with up to two decimal points), indicating the average worker score they expect in the group. [Figure 1.12](#) and [Figure 1.13](#) display the survey interface for belief elicitation, which is identical for both prior and posterior beliefs.

Belief elicitations are incentivized to ensure truthful reporting. At the end of the experiment, one question is randomly selected to be paid a bonus. If a belief question is selected, the employer receives a \$1 bonus ³ if their answer is within ± 2 of the correct answer. To ensure participants

³Bonus is calculated in experimental tokens with an exchange rate of 10 tokens to \$1 USD. The bonus for belief

understood the structure of these incentives, the experiment included detailed written instructions, annotated screenshots of the elicitation interfaces, and a mandatory comprehension quiz with four questions. Participants could only proceed if they answered all questions correctly within two attempts. [Section 1.2.3](#) provides details on the incentive structure in this experiment.

Information Acquisition

After reporting their prior beliefs about group-level productivity, each employer is randomly assigned to one of three information conditions: Voluntary, Exogenous Mean, Exogenous 20, and a no-information benchmark, the Baseline condition. In the information treatments, all employers receive data about the distribution of worker scores within their assigned group, but the mode of information delivery varies systematically across conditions. This variation allows me to isolate the effects of different information environments on belief updating and hiring behavior. I describe each condition in detail below.

Voluntary. In the Voluntary condition, I study the impact of experience-based learning when employers can endogenously seek out information about workers. Employers in this treatment are presented with a button they can click to sequentially sample individual workers from their assigned group, without replacement. Each click triggers the computer to randomly select a worker and display that worker’s score range (1–3, 4–6, 7–9, and 10–12), rather than the exact score. Data on worker productivity distribution in all information conditions are delivered in the same binned format. This choice serves two purposes. First, it introduces a degree of ambiguity into the information, giving employers wiggle room in forming their beliefs. Rather than seeing precise averages or raw scores, they must mentally interpolate and reason about what the data implies—closer to how information is processed in real-world decision-making. Second, binning mimics how data is often summarized in practice. Outside of lab settings, individuals rarely encounter complete information about underlying distributions; instead, they work with aggregated, approximate data.

To encourage reflection and preserve the sequential structure of experience-based information acquisition, I implement a 3-second cool-down between clicks. During this period, the draw button is temporarily disabled, mimicking the temporal cost of acquiring real-world information and

elicitation is 10 tokens.

discouraging rapid, thoughtless clicking to accumulate the sample and viewing them simultaneously. In addition, each new worker’s score range is appended below the previous draws, creating a cumulative list on screen. This design feature minimizes memory constraints and ensures that employers have access to the full history of their observations throughout the task. [Figure 1.14](#) shows a screenshot of the button interface, which is identical through all three experience conditions.

Employers can draw as many workers as they like (including 0 and up to 100), allowing me to observe not only how they update their beliefs in response to experience, but also how they choose to seek out information when the process is completely endogenous. This condition captures the dynamics of endogenous information acquisition, where individuals may selectively seek out certain outcomes—potentially reinforcing prior beliefs or biases.

Exogenous Mean. In the Exogenous Mean condition, employers sequentially sample workers without replacement from their assigned group by clicking a button. The interface is identical to the Voluntary treatment arm: each click reveals a randomly selected worker’s score bin (1–3, 4–6, 7–9, or 10–12), and the new observation is appended below the previous ones to create a visible sampling history. Unlike in the Voluntary condition, however, employers in this treatment are required to draw exactly the average number of workers drawn by the employers with the same worker group in the Voluntary condition. That is, employers in the Exogenous Mean-Asian treatment draw the mean sample size from Voluntary-Asian employers (8 workers), and Exogenous Mean-Hispanic employers drawn the mean sample size from Voluntary-Hispanic employers (10 workers). This eliminates the endogeneity of information acquisition by fixing the sample size, making it comparable to the voluntary condition while holding equal the information format and average information content. All other aspects of the task—including the interface, timing constraints, and binned feedback format—are identical to the Voluntary treatment.

Exogenous 20 (20 Draws). The Exogenous 20 condition follows the same structure as the Exogenous Mean treatment, with one key difference: employers are required to sample 20 workers, a larger and more representative sample. Sampling is again done without replacement, and the information is displayed using the same button-click interface and binned score format as in all other experience conditions. Each new draw is appended below the previous ones, allowing employers

to view their full sampling history and eliminating memory constraints. All other aspects of the design remain identical to the Voluntary and Forced Benchmark conditions. This treatment allows me to examine whether reducing sampling bias by providing a larger sample of exogenously assigned experiential data leads to more accurate belief updating and less discriminatory behavior.

Baseline. The Baseline condition serves as a control group in which employers receive no information about the productivity distribution of the worker group they are assigned to evaluate. After reporting their prior beliefs, they proceed directly to the wage offer tasks without any opportunity to learn about the workers. This condition allows me to measure the extent of initial belief bias and discriminatory behavior in the absence of any information provision.

Posterior Belief Elicitation

After the information acquisition phase, I elicit posterior beliefs from employers in all information treatments. The belief elicitation mirrors the prior belief stage exactly, using the same two complementary measure and the same interface: the PMF over scores 1 through 12 (12 questions summing to 100, yielding the estimated mean posterior belief from the PMF) and the text-box entry for the self-reported mean posterior belief (a number from 0 to 12 with up to two decimal places).

To minimize memory constraints, employers have full access to the information they received during the learning phase while reporting their posterior beliefs. The complete history of sampled draws is displayed in order, allowing employers to refer back to the observed data. The Baseline condition skips this step, as employers receive no new information prior to making hiring decisions.

Wage Offers

After employers report their updated group-level beliefs, they enter the wage offer stage, where I elicit their willingness to pay for individual workers. Each employer sees 10 randomly selected workers from their matched group of 100, presented sequentially. I present a short “resume” for each worker, which includes an avatar, anonymized nickname, country of residence (United States), gender, age (18–45 or 46+), and favorite color. Ethnicity is conveyed subtly through the avatar and nickname to minimize social desirability bias. Each worker group includes 90 ethnically representative names (e.g., Mateo for Hispanic workers, Ming for Asian workers) and 10 white-sounding names (e.g., Mark),

reflecting the real-world practice of ethnic minorities sometimes adopting Anglicized names. [Figure 1.15](#) shows example profiles for a Hispanic and an Asian worker, respectively. These profiles also allow me to control for observable worker characteristics in the analysis of wage offers.

To elicit willingness to pay in an incentive compatible manner, I implement the BDM elicitation mechanism (G. M. Becker et al., 1964), which is incentive compatible for arbitrary risk preferences. For each worker, employers are endowed with 12 experimental tokens as their hiring budget, and they are asked to make an integer wage offer between 0 and 12. After they make the wage offer, the computer randomly generates an asking wage for the worker, also an integer between 0 and 12. If the asking wage is less than or equal to the wage offer, the employer hires the worker at the asking wage, and earns $(12 - \text{asking wage} + \text{actual worker score})$ tokens. If the asking wage is higher than the wage offer, the employer does not hire the worker and keep the 12 tokens. Importantly, employers are informed that their decisions do not affect the compensation of the workers they evaluate. This removes the influence of interpersonal preferences or prosocial considerations from the decision environment, and ensures that wage offers reflect only beliefs about worker productivity and preference-driven bias. [Figure 1.16](#) displays the interface for the wage offer.

I implement several design features to ensure comprehension and data quality. First, the instructions for the BDM mechanism is framed in an intuitive, easily understandable manner, using the comparison between asking wage and wage offer to determine hiring outcomes. Second, employers receive detailed instructions and must correctly calculate their payment in 2 hypothetical scenarios in the comprehension quiz in order to proceed with the experiment. Because worker groups have identical productivity distributions, any systematic difference in wage offers between the Asian and Hispanic groups reflects discriminatory behavior, rather than differences in actual performance.

After completing the wage offer stage, employers answer a brief demographics questionnaire. I include an open-ended question—“What do you think this experiment is trying to study?”—to assess potential social desirability bias and the degree to which participants infer the study’s purpose. I utilize supervised machine learning to code this question and classify respondents by whether they correctly inferred the true study purpose. As a robustness check, I exclude participants who were able to infer the true purpose of the experiment and test whether results are robust to this sample

restriction.

1.2.3. Incentives

The experiment was carefully designed to incentivize both accurate beliefs and realistic hiring behavior. At the end of the study, one decision was randomly selected for bonus payment to ensure that participants took each decision seriously. This Random Problem Selection (RPS) incentive scheme has been shown to be incentive compatible for experiments with multiple tasks (Azrieli et al., 2018).

If a belief elicitation question was selected, participants received 10 experimental tokens for bonus (the experimental tokens have a conversion rate of 10 tokens = \$1 USD) if their reported belief about the average score or distribution was within a ± 2 margin of the correct value. This scoring rule is incentive-compatible and encourages truthful reporting of beliefs under risk neutrality, following standard practice in belief elicitation.

If a wage offer decision was selected, the bonus depended on the realized hiring outcome and the true productivity score of the matched worker, implemented using the Becker–DeGroot–Marschak (BDM) mechanism. For each worker, employers are endowed with 12 experimental tokens as their hiring budget, and they are asked to make an integer wage offer between 0 and 12. After they make the wage offer, the computer randomly generates an asking wage for the worker, also an integer between 0 and 12. If the asking wage is less than or equal to the wage offer, the employer hires the worker at the asking wage, and earns $(12 - \text{asking wage} + \text{actual worker score})$ tokens. If the asking wage is higher than the wage offer, the employer does not hire the worker and keeps the 12 tokens (i.e. the possible amounts of bonus an employer can earn in a given wage offer range from \$0.1 to \$2.3).

To ensure participants understood the structure of these incentives, the experiment included detailed written instructions, annotated screenshots of the elicitation interfaces, and a mandatory comprehension quiz with four questions. Participants could only proceed if they answered all questions correctly within two attempts. To further promote careful reading, participants received a \$1 bonus for answering all questions correctly on the first try. Approximately 66% of participants earned this bonus, indicating a high level of engagement and comprehension. This supports the

validity of the responses and high quality of the experimental data.

1.2.4. Implementation

The experiment was approved by the UC Davis Institutional Review Board and pre-registered on AsPredicted (#213,758). The Exogenous Mean condition, which uses the average number of draws from the Voluntary condition as a design parameter, was conducted later and pre-registered separately as a follow-up study (#237,319).

All surveys were programmed in Qualtrics using custom JavaScript to implement dynamic interfaces. I recruited participants for both the worker survey and employer experiment via the online platform Prolific, which has been shown to produce data quality comparable to traditional in-person lab experiments and superior to other online platforms such as Amazon Mechanical Turk (Gupta et al., 2021; Peer et al., 2022). Prolific enforces strict privacy protections and prohibits the collection of personally identifiable information, helping reduce potential experimenter demand and social desirability bias—especially important in studies of sensitive topics like discrimination. Additionally, recent evidence suggests that demand effects in online experiments are relatively small, and even when participants infer the purpose of a study, their behavior is largely unaffected (De Quidt et al., 2018; Mummolo & Peterson, 2019).

All participants were at least 18 years old and residents of the United States. For the worker survey, I used Prolific’s screening tools to recruit participants who self-identify as either Asian or Hispanic. Workers received a \$2 base participation payment and earned a bonus of \$0.20 for each correct answer on the math test. Employers were paid \$2 for completing the experiment, with an additional bonus ranging from \$0 to \$3.4. The average time taken to complete the survey for employers was 15 minutes, with an average total compensation of around \$3.72. resulting in an hourly compensation of approximately \$15. The final sample includes approximately 400 workers and 2,400 employers, generating rich data on belief formation, sampling behavior, and wage offers.

1.3. Theoretical Framework

This section presents a model of belief updating under sequential information acquisition in which an employer weighs instrumental utility, from forming accurate beliefs about an unknown worker score, against belief utility, a psychological benefit from cognitive consistency with her prior. Exogenously assigned and endogenously acquired information draw signals from the same data-generating process; the modes differ in the employer's control over when sampling stops. Endogenous control opens a channel for distorted inference whenever the employer places positive weight on belief utility.

I formalize three modes of learning that vary in endogenous control:

- a. **Ex-ante commitment to a sample size (Commitment).** The employer chooses a sample size τ in advance, then is required to draw exactly τ signals.
- b. **Sequential search with optional stopping (Voluntary).** The employer observes signals one at a time and chooses at each step whether to continue.
- c. **Exogenously assigned sample size (Exogenous).** The employer receives a fixed, exogenously assigned number of signals and observes the realizations sequentially.

Voluntary and Commitment are observationally equivalent for a fully Bayesian employer, but they generate different predictions when belief utility is present.

I begin with a baseline model of a fully rational employer, then extend the model to incorporate belief utility in the spirit of cognitive-consistency preferences (Yariv 2001). Under cognitive consistency, search decisions depend on the realized signal path, producing earlier stopping times, skewed information sets, and stickier posteriors. The framework explains how two employers exposed to the same average amount of information can arrive at systematically different posteriors depending on whether they chose that information themselves, and generates qualitative comparative statics across the three regimes.

The model adopts a stylized normal-normal Bayesian framework for tractability. Its content is the qualitative comparative statics derived in Section 1.3.4: the gap-closure ranking across regimes, the Voluntary-vs-Commitment wedge under cognitive consistency, and the bias-toward-prior in

Voluntary. These results depend on the structure of the cognitive-consistency mechanism rather than on the specific Gaussian functional form. Quantitative parameter values $(\sigma_0^2, \sigma_d^2, c)$ are illustrative.

1.3.1. Setup

Let $\theta \in \mathbb{R}$ denote the true average score of a worker group. The employer does not observe θ but knows that it is drawn from a known prior distribution. Time is discrete and indexed by $n = 0, 1, 2, \dots, T$, where T is a finite horizon. I use n as both a time index and a sample-size counter.

In each period the employer may draw one signal. Each draw incurs a constant cost $c \geq 0$. After observing the signal history $D_n = \{s_1, \dots, s_n\}$, the employer updates her beliefs by Bayes' rule. For tractability I specialize to the conjugate normal-normal pair:

$$\theta \sim \mathcal{N}(\mu_0, \sigma_0^2), \quad (1.1)$$

$$s_i \mid \theta \sim \mathcal{N}(\theta, \sigma_d^2), \quad i = 1, 2, \dots, n, \quad (1.2)$$

where μ_0 is the prior mean, $\sigma_0^2 > 0$ is the prior variance, and $\sigma_d^2 > 0$ is the signal-noise variance. The cross-group analysis in Section 1.3.4 below considers two groups indexed by $r \in \{H, L\}$ that share σ_0^2 and σ_d^2 but may have distinct prior means μ_0^H and μ_0^L .

At any stopping time n the employer reports a point estimate $\hat{\theta} \in \mathbb{R}$ and earns instantaneous instrumental utility

$$u(\hat{\theta}, \theta) = -(\hat{\theta} - \theta)^2, \quad (1.3)$$

with total instrumental payoff $u(\hat{\theta}, \theta) - cn$. The quadratic-loss specification rewards the employer for accurate predictions of θ , while also improving tractability. Sampling costs accumulate linearly with no time discount factor.

1.3.2. Belief Updating

After observing D_n , the employer's posterior is normal: $\theta \mid D_n \sim \mathcal{N}(\mu_n, \sigma_n^2)$ with

$$\sigma_n^2 = \frac{\sigma_0^2 \sigma_d^2}{n\sigma_0^2 + \sigma_d^2} = \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma_d^2} \right)^{-1}, \quad (1.4)$$

$$\mu_n = \omega_n \bar{s}_n + (1 - \omega_n) \mu_0, \quad (1.5)$$

where $\bar{s}_n = \frac{1}{n} \sum_{i=1}^n s_i$ and the shrinkage weight on the sample mean is $\omega_n = n\sigma_0^2 / (n\sigma_0^2 + \sigma_d^2) \in [0, 1]$. The posterior mean is a convex combination of the prior mean and the sample mean; the weight ω_n is increasing in n and in σ_0^2 . As $n \rightarrow \infty$, $\omega_n \rightarrow 1$.

Two structural facts about the posterior drive the analysis below. First, the posterior variance σ_n^2 is a deterministic function of n alone; it does not depend on signal realizations. This property is the source of the equivalence in Proposition 2 below. Second, the posterior mean is a martingale:

$$\mathbb{E}[\mu_{n+1} \mid \mu_n] = \mu_n, \quad (1.6)$$

with conditional variance $\text{Var}(\mu_{n+1} \mid \mu_n) = \omega_{n+1}^2 (\sigma_{n+1}^2 + \sigma_d^2)$ (Lemma 1 in Appendix 1.12).

Under quadratic loss, the loss-minimizing prediction at any stopping time is the posterior mean $\hat{\theta}^* = \mu_n$, and the resulting expected loss equals the posterior variance σ_n^2 . Because σ_n^2 is non-stochastic given n , the ex ante expected loss from stopping at sample size n also equals σ_n^2 . Subtracting cumulative sampling costs, the employer's **stopping value** is

$$V^{\text{stop}}(n) = -\sigma_n^2 - cn. \quad (1.7)$$

1.3.3. The Rational Employer

I first analyze the decision problem of a fully rational employer, who derives only instrumental utility from beliefs and therefore cares only about belief accuracy. I begin with the simpler Commitment problem because its objective admits a closed-form sample size; I then turn to the Sequential (Voluntary) problem and show that its solution coincides with the Commitment solution under rationality.

Ex-Ante Commitment

In the Commitment regime, the employer chooses the sample size n at the outset, before observing any signal realizations, and is then required to draw exactly n signals. Because the stopping value

$-\sigma_n^2$ depends on n alone and the cumulative cost is cn , the problem reduces to a deterministic optimization:

$$n^{\text{commit}} = \underset{n \in \{0,1,\dots,T\}}{\operatorname{argmax}} V^{\text{stop}}(n) = \underset{n}{\operatorname{argmax}} \{-\sigma_n^2 - cn\}.$$

Treating n as continuous yields a closed-form expression for the optimum, the tractability gain of the quadratic-loss specification.

PROPOSITION 1 (Optimal Sample Size). *The unique interior maximizer of $V^{\text{stop}}(n)$ in (1.7) is*

$$n^* = \frac{\sigma_d}{\sigma_0^2} \left(\frac{\sigma_0^2}{\sqrt{c}} - \sigma_d \right) = \frac{\sigma_d}{\sqrt{c}} - \frac{\sigma_d^2}{\sigma_0^2}, \quad (1.8)$$

provided $\sigma_0^2/\sqrt{c} > \sigma_d$ (otherwise the corner $n^* = 0$ binds). The minimized expected loss at n^* is

$$V^{\text{stop}}(n^*) = -2\sigma_d\sqrt{c} + \frac{c\sigma_d^2}{\sigma_0^2}. \quad (1.9)$$

The comparative statics are

$$\frac{\partial n^*}{\partial \sigma_0^2} > 0, \quad \frac{\partial n^*}{\partial c} < 0,$$

and $\partial n^*/\partial \sigma_d^2$ has a hump shape: positive when $\sigma_d < \sigma_0^2/(2\sqrt{c})$ and negative otherwise.

PROOF. See Appendix 1.12. □

Although the actual sample size is integer-valued, the continuous relaxation locates the discrete optimum to within ± 1 : since $V^{\text{stop}}(n)$ is strictly concave in n , it is unimodal, so the integer maximizer is one of $\lfloor n^* \rfloor$ or $\lceil n^* \rceil$. The closed-form n^* in (1.8) is therefore an exact characterization of the discrete optimum up to integer rounding, not an asymptotic approximation. The comparative statics have natural interpretations. More diffuse priors raise the marginal value of information at every step, so the employer samples longer. Higher marginal cost reduces the optimal sample size monotonically. Signal noise has competing effects: noisier signals reduce per-draw variance reduction, but also lower the precision available at any sample size, so the employer needs more draws to reach a given precision target.

Figure 1.2 plots $V^{\text{stop}}(n)$ at an illustrative parameter setting, showing the strict concavity of

the stopping value in n and the location of the closed-form optimum.

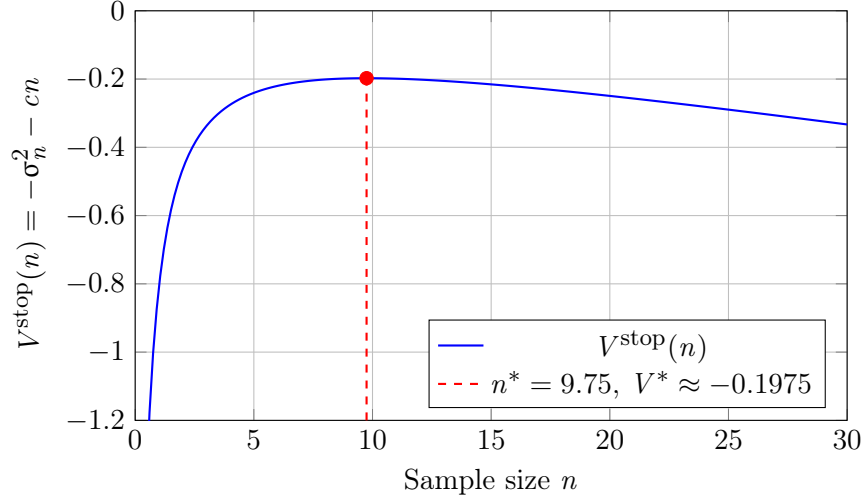


FIGURE 1.2. Commitment value $V^{\text{stop}}(n) = -\sigma_n^2 - cn$ at illustrative parameters ($\sigma_0^2 = 4$, $\sigma_d^2 = 1$, $c = 0.01$). The unique interior maximum is at $n^* = 9.75$ (Proposition 1), with optimized value $V^{\text{stop}}(n^*) \approx -0.1975$. The function is strictly concave in n in the interior.

The Sequential Problem (Voluntary)

Under endogenous sequential search, the employer faces a discrete-time optimal-stopping problem. In each period she draws a signal, updates her posterior, and decides whether to stop or continue. If she stops at τ , she reports μ_τ and earns instantaneous instrumental utility $-(\mu_\tau - \theta)^2$. If she continues, she draws another signal, updates again, and faces the same decision next period. The horizon T is finite, and there is no time discounting.

The Bellman equation in period n is

$$V_n(\mu) = \max \left\{ \underbrace{-\sigma_n^2}_{\text{Stop}}, \underbrace{\mathbb{E}[V_{n+1}(\mu_{n+1}) \mid \mu_n = \mu] - c}_{\text{Continue}} \right\}, \quad (1.10)$$

with terminal condition $V_T(\mu) = -\sigma_T^2$. The stopping value $-\sigma_n^2$ depends on n alone, not on the realized μ_n . By backward induction (Appendix 1.12), the value function V_n inherits this property: it is constant in μ_n at every n . The Bellman therefore collapses to a deterministic recursion

$$V_n = \max \{ -\sigma_n^2, V_{n+1} - c \}, \quad n = T-1, \dots, 0, \quad V_T = -\sigma_T^2,$$

and the optimal stopping rule reduces to a deterministic threshold on n . Equivalently, the employer stops at the first n at which the marginal value of an additional signal, $\Delta_n := \sigma_n^2 - \sigma_{n+1}^2$, falls below the marginal cost c . This marginal rule is the first-order condition for the maximization in Section 1.3.3 and characterizes the same n^* . The collapse of the value function is what produces the rational equivalence between Sequential and Commitment.

PROPOSITION 2 (Voluntary $\equiv$ Commitment Under Rationality). *For the rational employer with quadratic loss, the optimal Voluntary stopping time and the optimal Commitment sample size coincide:*

$$\tau_{rat}^{vol} = n_{rat}^{commit} = n^*.$$

The induced posterior distributions are identical: $\mu_{n^}^{vol}$ and $\mu_{n^*}^{commit}$ have the same distribution at every prior (μ_0, σ_0^2) .*

PROOF. See Appendix 1.12. □

Because V_n is constant in μ_n , the option to re-optimize after each signal has no value: deterministic and signal-dependent stopping rules give the same expected payoff, and Sequential and Commitment policies therefore coincide for a fully Bayesian employer whose only objective is prediction accuracy. Section 1.3.4 shows that this equivalence breaks once the employer also derives utility from belief stability.

The marginal rule has its own visual intuition. Figure 1.3 plots the per-draw expected variance reduction against the constant marginal cost. The two curves cross at the optimum n^* : above the crossing, an additional draw is worth its cost; below, it is not. This threshold view is the local counterpart to the global maximization in Figure 1.2.

Exogenously Assigned Sample Size

When the sample size is exogenously assigned, there is no active stopping decision. The rational employer simply observes the realized signals, updates her beliefs by Bayes' rule, and reports the posterior mean μ_n^{exog} . Conditional on a given exogenous sample size n , the posterior is identical in

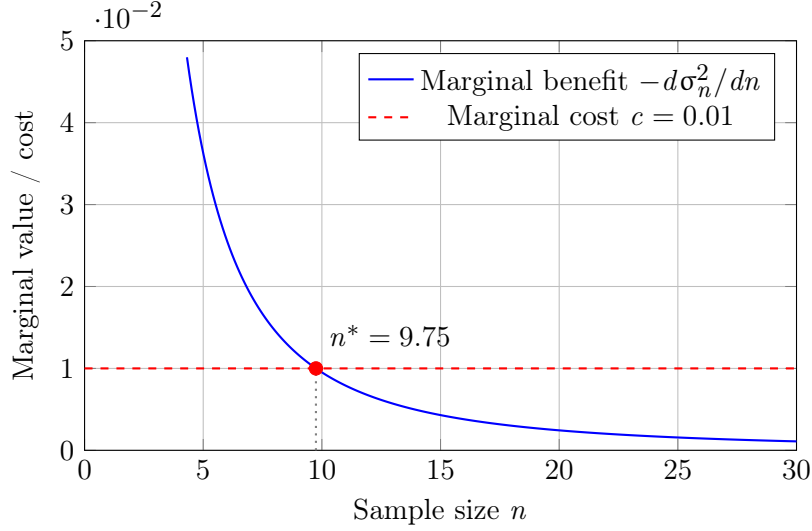


FIGURE 1.3. Marginal stopping rule. The marginal value of an additional signal, $-d\sigma_n^2/dn = \sigma_0^4\sigma_d^2/(n\sigma_0^2 + \sigma_d^2)^2$ (blue), is strictly decreasing in n : information has diminishing returns. The marginal cost c (red dashed) is constant. The optimum n^* is the unique crossover: continue while the next draw’s variance reduction exceeds c , stop when it falls below. At illustrative parameters $(\sigma_0^2, \sigma_d^2, c) = (4, 1, 0.01)$, the crossover is at $n^* = 9.75$, the same value as in Figure 1.2.

distribution to the posterior at sample size n in either Voluntary or Commitment. In particular, in the special case $n^{\text{exog}} = \mathbb{E}[\tau^{\text{vol}}]$ (the “ExogenousMean” regime), the rational benchmark gives $\mu_n^{\text{exog}} \stackrel{d}{=} \mu_n^{\text{vol}} \stackrel{d}{=} \mu_n^{\text{commit}}$ at every $n = n^*$. The three learning regimes are observationally equivalent under the rational-quadratic-loss benchmark.

1.3.4. Cognitive Consistency

The rational benchmark predicts equivalence across the three learning regimes. This rests on a strong assumption: that the employer cares only about the accuracy of her belief and assigns no value to its content. A long tradition in behavioral economics relaxes this by allowing decision-makers to derive utility directly from the beliefs they hold; in particular, agents may prefer beliefs close to a prior position and incur a psychological cost from updating away. I extend the rational model along these lines by introducing belief utility in the form of cognitive-consistency disutility from updating away from the prior. The employer’s per-period flow payoff becomes

$$U_t = \underbrace{u(\mu_t, \theta)}_{\text{instrumental utility}} + \underbrace{\gamma v(\mu_0, \mu_t)}_{\text{belief utility}}, \quad (1.11)$$

with $v(\mu_0, \mu_t) = -(\mu_t - \mu_0)^2$, a **cognitive-consistency** preference following the Yariv (2001) tradition, parameterized by $\gamma \geq 0$. When $\gamma = 0$, the employer is fully rational and only cares about belief accuracy; as $\gamma \rightarrow \infty$, she cares only about staying close to her prior. The functional form $-(\mu_t - \mu_0)^2$ represents a cost for belief deviation from the prior μ_0 , capturing the employer's preference for beliefs close to her starting position. Under cognitive consistency, the employer's decision problem becomes path-dependent: utility depends not only on the variance of her belief but also on its location relative to the prior.

Ex-Ante Commitment Under Cognitive Consistency

In the Commitment regime under cognitive consistency, the employer chooses $\tau \leq T$ at $t = 0$ to maximize the expected total payoff

$$\tilde{V}^{\text{commit}}(\tau) = \mathbb{E}_0 \left[u(\mu_\tau, \theta) + \sum_{t=1}^{\tau} \gamma v(\mu_0, \mu_t) - c\tau \right].$$

The expected instrumental payoff is $-\sigma_\tau^2$. The cumulative expected belief disutility simplifies in two steps. First, by the martingale property (1.6), the expected posterior mean at any future date equals the prior: $\mathbb{E}_0[\mu_t] = \mu_0$. The squared deviation from the prior is therefore the squared deviation from the unconditional mean of μ_t , so

$$\mathbb{E}_0[(\mu_t - \mu_0)^2] = \mathbb{E}_0[(\mu_t - \mathbb{E}_0[\mu_t])^2] = \text{Var}_0(\mu_t),$$

the unconditional variance of μ_t at date 0. Second, standard normal-normal algebra (Appendix 1.12) gives a closed-form expression for this variance as the difference between the prior variance and the posterior variance at t :

$$\text{Var}_0(\mu_t) = \sigma_0^2 - \sigma_t^2. \quad (1.12)$$

Substituting both expressions into the Commitment objective gives

$$\tilde{V}^{\text{commit}}(\tau) = -\sigma_\tau^2 - \gamma \sum_{t=1}^{\tau} (\sigma_0^2 - \sigma_t^2) - c\tau. \quad (1.13)$$

The Commitment objective is a deterministic function of τ alone, with a closed-form optimum.

PROPOSITION 3 (Commitment Optimum Under Cognitive Consistency). *Let $\tilde{n}_\gamma^*$ denote the maximizer of $\tilde{V}^{commit}(\tau)$ in (1.13), indexed by the belief-utility weight $\gamma \geq 0$. At $\gamma = 0$ the belief cost vanishes and $\tilde{n}_0^* = n^*$, the rational Commitment optimum from Proposition 1. For $\gamma > 0$, the marginal benefit of an additional signal is*

$$\underbrace{(\sigma_\tau^2 - \sigma_{\tau+1}^2)}_{\text{instrumental gain}} - \underbrace{\gamma(\sigma_0^2 - \sigma_{\tau+1}^2)}_{\text{cumulative belief cost}}, \quad (1.14)$$

and the employer continues while this exceeds c . The belief-cost term is non-negative and weakly increasing in τ , so $\tilde{n}_\gamma^ \leq n^*$.*

PROOF. See Appendix 1.12. □

The Commitment employer internalizes the full cumulative belief cost ex ante, so she rationally undersamples relative to the rational Commitment benchmark.

The Sequential Problem (Voluntary) Under Cognitive Consistency

The Voluntary employer, unlike her Commitment counterpart, can defer the belief-utility cost on histories where realized μ_t stays close to μ_0 . This flexibility breaks the rational equivalence between Voluntary and Commitment policies. Her Bellman equation under cognitive consistency becomes

$$\tilde{V}_t(\mu) = \max \left\{ \underbrace{-\sigma_t^2}_{\text{Stop}}, \underbrace{\mathbb{E} \left[\tilde{V}_{t+1}(\mu_{t+1}) - \gamma(\mu_{t+1} - \mu_0)^2 \mid \mu_t = \mu \right] - c}_{\text{Continue}} \right\}, \quad (1.15)$$

with terminal condition $\tilde{V}_T(\mu) = -\sigma_T^2$. The stopping value is unchanged from the rational case, but the continuation value now includes the expected belief disutility $-\gamma \mathbb{E}[(\mu_{t+1} - \mu_0)^2 \mid \mu_t]$. This continuation value depends on μ through the belief-cost term, which breaks the rational equivalence:

PROPOSITION 4 (Voluntary $\neq$ Commitment Under Cognitive Consistency). *Under cognitive consistency with $\gamma > 0$, the value function $\tilde{V}_t(\mu)$ is non-trivially a function of μ at every $t < T$. Consequently, the optimal Voluntary stopping rule $\tilde{\tau}_\gamma^{vol}$ depends on the realized posterior mean μ_t and is not a deterministic function of t . There exist parameter tuples $(\sigma_0^2, \sigma_d^2, c, \gamma, T)$ and posterior-history pairs (D_t, D'_t) at the same t such that the optimal rule prescribes stop at D_t and continue at*

D'_t .

PROOF. See Appendix 1.12. □

The mechanism is transparent in the continuation value. Conditioning on μ_t , applying the martingale identity (1.6) and the Law of Total Variance gives

$$\mathbb{E}\left[(\mu_{t+1} - \mu_0)^2 \mid \mu_t\right] = \text{Var}(\mu_{t+1} \mid \mu_t) + (\mu_t - \mu_0)^2. \quad (1.16)$$

The first term, $\text{Var}(\mu_{t+1} \mid \mu_t) = \omega_{t+1}^2(\sigma_{t+1}^2 + \sigma_d^2)$, depends on t alone. The second term, the location term $(\mu_t - \mu_0)^2$, depends non-trivially on the realized μ_t . The continuation value at t is therefore

$$\tilde{C}_t(\mu_t) = \mathbb{E}\left[\tilde{V}_{t+1}(\mu_{t+1}) \mid \mu_t\right] - \gamma \text{Var}(\mu_{t+1} \mid \mu_t) - \gamma (\mu_t - \mu_0)^2 - c, \quad (1.17)$$

which depends on μ_t explicitly through the $-\gamma(\mu_t - \mu_0)^2$ term.

Economically, the cognitive-consistency employer finds it costly to update away from her prior, so she stops earlier when realized signals push μ_t far from μ_0 and continues when μ_t remains close to μ_0 . Selection on realizations enters the Voluntary employer's stopping decision, which truncates the unconditional distribution of her terminal posterior toward the prior. Because she defers the belief cost on prior-aligned histories, the Voluntary employer also samples weakly more in expectation than the Commitment employer:

PROPOSITION 5 (Expected Sample-Size Ranking). *Under cognitive consistency with $\gamma > 0$, the Commitment employer's deterministic sample size is weakly less than the Voluntary employer's expected stopping time:*

$$\tilde{n}_\gamma^* \leq \mathbb{E}_0\left[\tilde{\tau}_\gamma^{vol}\right].$$

PROOF. See Appendix 1.12. □

Exogenously Assigned Sample Size Under Cognitive Consistency

When the sample size is exogenously imposed, the employer has no discretion over stopping. She still incurs belief disutility $-\gamma(\mu_t - \mu_0)^2$ at each t , but the disutility does not feed back into a stopping decision. Her posterior at sample size n^{exog} is therefore identical in distribution to the rational-Bayesian posterior at the same sample size: μ_n^{exog} has unconditional mean and variance given by the standard Bayesian formulas (1.5)–(1.4) with $n = n^{\text{exog}}$. There is no realization-conditional selection on the stopping decision under Exogenous, even when $\gamma > 0$.

Qualitative Implications

The three regimes' interaction with cognitive consistency produces an ordering of expected belief bias and a corresponding ordering of expected sample sizes. I measure belief bias two ways: at the cross-group level by the residual gap Δ_n/Δ_0 (how much of the prior gap between groups remains after sampling), and at the within-group level by the squared distance $\mathbb{E}_0[(\mu_n^r - \mu_0^r)^2]$ (how close a single group's posterior remains to its own prior).

Suppose the employer faces two worker groups indexed by $r \in \{H, L\}$, with priors $\mathcal{N}(\mu_0^r, \sigma_0^2)$ and common signal noise σ_d^2 , where H is the higher-prior group ($\mu_0^H > \mu_0^L$). Both groups have the same prior variance σ_0^2 : the employer is equally uncertain about each group's mean productivity, even though she expects the means to differ. Let $\Delta_0 \equiv \mu_0^H - \mu_0^L$ denote the prior gap and $\Delta_n \equiv \mu_n^H - \mu_n^L$ denote the posterior gap at sample size n . When n is independent of the realized signal path, as in Commitment, Exogenous, and rational Voluntary, the unconditional gap-closure formula

$$\frac{\mathbb{E}_0[\Delta_n]}{\Delta_0} = 1 - \omega_n = \frac{\sigma_d^2}{n\sigma_0^2 + \sigma_d^2} \quad (1.18)$$

follows from (1.5) (Proposition 6 below).

PROPOSITION 6 (Unconditional gap-closure rate). *Suppose n is determined independently of the signal realizations. Then the gap shrinkage factor between groups with the same prior variance σ_0^2 is given by (1.18). The shrinkage factor is strictly decreasing in n and in σ_0^2 , and strictly increasing in σ_d^2 .*

PROOF. See Appendix 1.12. □

The unconditional rate is the right benchmark for Commitment, Exogenous, and rational Voluntary, where the stopping rule is realization-independent. In the cognitive-consistency Voluntary regime, n is itself a function of realizations, and the conditional expectation of Δ_n exceeds this baseline through selection on realizations, a deviation formalized in Proposition 7 and Corollary 1 below.

Figure 1.4 plots the gap-closure rate Δ_n/Δ_0 for three values of prior variance σ_0^2 . More diffuse priors close the gap faster per signal, because ω_n places more weight on the sample mean.

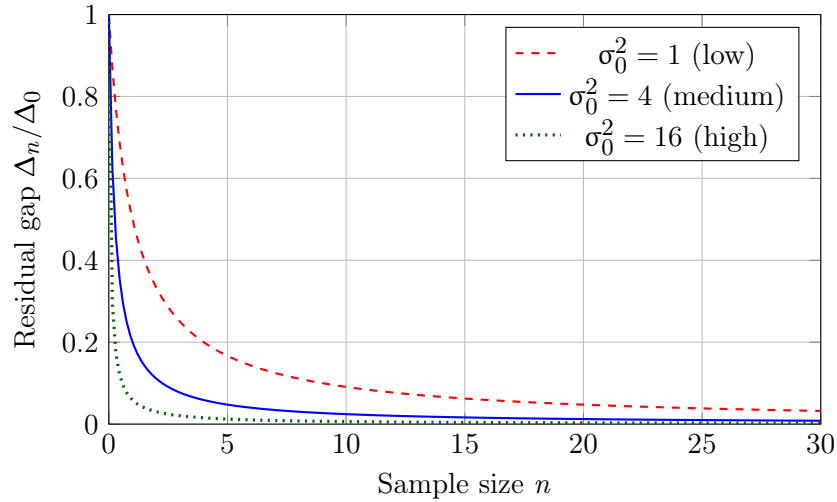


FIGURE 1.4. Gap-closure rate $\Delta_n/\Delta_0 = \sigma_d^2/(n\sigma_0^2 + \sigma_d^2)$ as a function of n , for three values of prior variance σ_0^2 (with $\sigma_d^2 = 1$). The function is strictly decreasing in both n and σ_0^2 . Higher σ_0^2 closes the gap faster per signal. At the illustrative setting $\sigma_0^2 = 4$, the gap closes to $\approx 2.5\%$ of its prior value at $n = n^* = 9.75$.

The gap-shrinkage formula (1.18) implies that, for a fixed sample size n , the residual gap is independent of the regime: any difference across regimes must come either from differences in expected n or from selection on realizations within a regime. Both channels are active under cognitive consistency. The first prediction isolates the selection channel: the Voluntary employer's stopping rule truncates the distribution of her terminal posterior toward her prior, relative to the rational Bayesian benchmark at the same realized sample size. To distinguish the cognitive-consistency Voluntary employer's posterior from the unselected rational benchmark at the same sample size, let μ_n^{cons} denote the posterior mean at sample size n under the cognitive-consistency Voluntary employer's joint distribution (signal histories conditioned on her stopping rule), and let μ_n^{rat} denote

the posterior mean at sample size n under the rational benchmark's joint distribution (signal histories drawn unconditionally, with n treated as exogenous). Both posteriors apply the Bayesian update formula (1.5)–(1.4); the superscript marks the joint distribution under which the expectation is taken. Under the model's stated prior, the rational benchmark satisfies $\mathbb{E}_0[(\mu_n^{\text{rat}} - \mu_0)^2] = \sigma_0^2 - \sigma_n^2$ (the unconditional variance identity from the appendix).

PROPOSITION 7 (Voluntary Bias Toward Prior). *Under cognitive consistency with $\gamma > 0$,*

$$\mathbb{E}_0[(\mu_{\tilde{\tau}_\gamma^{\text{vol}}}^{\text{cons}} - \mu_0)^2] \leq \mathbb{E}_0[(\mu_{\tilde{\tau}_\gamma^{\text{vol}}}^{\text{rat}} - \mu_0)^2] = \mathbb{E}_0[\sigma_0^2 - \sigma_{\tilde{\tau}_\gamma^{\text{vol}}}^2],$$

where $\mathbb{E}_0$ denotes the prior expectation, taken over the joint distribution of θ and signal realizations (with $\tilde{\tau}_\gamma^{\text{vol}}$ a function of the signal sequence under the Voluntary stopping rule under cognitive consistency, μ_n^{cons} the cognitive-consistency Voluntary employer's posterior mean at sample size n , and μ_n^{rat} the unselected Bayesian posterior mean at sample size n). The middle term is the rational Bayesian's expected squared deviation evaluated at the same stopping distribution; the right-hand equality follows from the unconditional variance identity $\text{Var}_0(\mu_n^{\text{rat}}) = \sigma_0^2 - \sigma_n^2$. Equality holds throughout at $\gamma = 0$.

PROOF. See Appendix 1.12. □

The Voluntary stopping rule selects against histories with large $|\mu_t - \mu_0|$ (the location-term cost in $\tilde{C}_t$ pushes the employer to stop on such histories), so under the model's assumption $\theta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ the realized distribution of $\mu_{\tilde{\tau}_\gamma^{\text{vol}}}^{\text{cons}}$ has smaller expected squared deviation from μ_0 than the unselected Bayesian benchmark at the same stopping time. This is the selection channel of cognitive consistency in the Voluntary regime; the corresponding sample-size channel is captured separately by Proposition 5. Combining the two channels across regimes gives the gap-closure ordering stated in Corollary 1 below.

COROLLARY 1 (Gap-Closure Ranking Across Regimes). *Under cognitive consistency with $\gamma > 0$ and*

equal prior variance across the two groups,

$$\mathbb{E}_0 \left[\frac{\Delta_n^{vol}}{\Delta_0} \right] \geq \mathbb{E}_0 \left[\frac{\Delta_n^{commit}}{\Delta_0} \right] \geq \mathbb{E}_0 \left[\frac{\Delta_n^{ExogMean}}{\Delta_0} \right],$$

where $\mathbb{E}_0$ denotes the prior expectation, taken over the joint distribution of (θ_H, θ_L) drawn from their respective group priors $\theta_r \sim \mathcal{N}(\mu_0^r, \sigma_0^2)$ and the resulting signal realizations. The first inequality is strict when $\gamma > 0$; the second is strict when in addition $n^{ExogMean} > \tilde{n}_\gamma^*$, which holds under Proposition 5 when the realization-conditional location-cost $\gamma(\mu_t - \mu_0)^2$ has nonzero variance. Equivalently in closure form, Voluntary closes the gap least, Commitment is intermediate, and ExogenousMean closes the gap most. As $\gamma \rightarrow 0$ the inequalities collapse to equality and all three regimes converge to the rational gap-closure rate $1 - \omega_{n^*}$.

PROOF. See Appendix 1.12. □

Two qualitative implications distinguish cognitive consistency from the rational benchmark. First, the across-regime gap-closure ordering in Corollary 1 reflects three forces: the Voluntary selection wedge (Proposition 7), Commitment's ex-ante undersampling (Proposition 3), and Voluntary's expected oversampling relative to Commitment (Proposition 5). The constraint $n^{ExogMean} = \mathbb{E}_0[\tau^{vol}]$ then fixes the ExogenousMean sample size at the Voluntary mean, leaving ExogenousMean with the largest sample size and the most gap closure. Second, within Voluntary, the employer's stopping probability rises with the prior-distance $|\mu_t - \mu_0|$, since her continuation value falls in $(\mu_t - \mu_0)^2$. The rational benchmark exhibits neither pattern.

Scope.. The ranking in Corollary 1 follows from the prior assumption $\theta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ stated in equation (1.1). The corollary applies under correctly specified priors, that is, when θ is actually drawn from the same distribution the employer uses as her prior. The unconditional variance identity $\text{Var}_0(\mu_n) = \sigma_0^2 - \sigma_n^2$ used in the proof of Proposition 7 is a consequence of this assumption. Relaxing it alters the geometry of the selection mechanism, and the corollary's strict ordering is not characterized in those regimes.

1.4. Hypotheses and Empirical Strategy

In this section, I formulate hypotheses about three categories of employer behavior: prior beliefs, posterior beliefs across the Voluntary treatment, the Exogenous benchmark, and the high-information condition, and Baseline wage discrimination. I also discuss the empirical strategies I use to test these hypotheses. Hypotheses and empirical strategies for the Match and Commitment treatment arms are introduced in [Section 1.6](#) alongside the corresponding empirical exhibits.

Throughout the paper, the primary measure of employer beliefs is the estimated mean belief from the elicited probability mass function over worker scores; the self-reported mean belief is a companion measure, and the agreement between the two is reported in [Section 1.7.2](#). Belief bias refers to the gap between employers assigned the Asian worker group and employers assigned the Hispanic worker group in their beliefs about the same underlying worker pool. The hypotheses below test this gap in priors and in posteriors. Individual deviation from the truth is a related but distinct concept; the paper’s bias claims are gap-based unless explicitly indicated otherwise. Distributional balance of priors across Treatment Arms is reported in [Section 1.7.1](#).

1.4.1. Prior Beliefs

In order to study the evolution of belief bias, I must first induce bias in prior beliefs. I accomplish this using the math task coupled with random assignment of Asian vs. Hispanic workers. Math is a task domain in which Asians are stereotyped to perform better, therefore worker ethnicity assignment should induce employers to form differential priors. This leads me to formulate the first hypothesis:

HYPOTHESIS 1 (Bias in Prior Beliefs). *Employers hold higher prior beliefs for the average Asian worker productivity compared to Hispanic workers.*

To test for belief bias in the priors, I compare the subjective beliefs about worker scores of employers assigned to Hispanic versus Asian worker groups before any additional performance information is provided. I estimate the following Ordinary Least Squares (OLS) regression of prior belief on worker ethnicity assignment:

$$Prior_i = \beta_0 + \beta_1 Asian_i + \beta_2 X_i + \epsilon_i \quad (1.19)$$

where $Prior_i$ is the prior belief of employer i on the average performance of the assigned worker group, $Asian_i$ is a binary indicator that equals 1 if employer i is randomly assigned the Asian worker group, and X_i is a vector of employer demographic controls. The coefficient $\beta_1 > 0$ would provide evidence on prior belief bias in favor of Asian workers. I estimate this regression on the pooled dataset of employers across all information conditions; the cell-by-cell version of the same regression with treatment-arm interactions, reported in [Table 1.1](#), also serves as the cell-level balance check on prior distributions across Treatment Arms. Standard errors are heteroskedasticity-robust at the individual employer level. I report the multivariate energy-distance test on the elicited PMFs as a distributional companion to the regression-based test in [Section 1.7.1](#).

1.4.2. Posterior Beliefs

The central question I ask is how seeking information endogenously versus acquiring information exogenously affects the persistence of belief bias in posteriors. Two hypotheses follow from the theoretical framework. First, when employers retain endogenous control over the information search process in the Voluntary treatment, the residual posterior gap between Asian-arm and Hispanic-arm employers does not close to zero. Second, removing endogeneity in information search compresses this residual gap. The Exogenous benchmark (ExogenousMean) holds the average sample size equal to its Voluntary counterpart and isolates the realization-based-stopping channel by neutralizing the sample-size channel; the high-information condition (Exogenous20) provides a fixed sample of 20 random workers and additionally provides more information than the Voluntary mean. Corollary 1 predicts that the residual posterior gap is largest in Voluntary, smaller under ExogenousMean, and smallest under Exogenous20. The full cross-arm ordering, including the Commitment column, is reported in [Section 1.6](#).

HYPOTHESIS 2 (Voluntary Residual Posterior Gap). *Within the Voluntary treatment, the Asian-Hispanic gap in posterior beliefs is positive: self-directed information acquisition does not eliminate the prior belief bias.*

To test [Hypothesis 2](#), I estimate the following cross-sectional regression of posterior beliefs on

the Asian worker group indicator, restricted to employers in the Voluntary treatment:

$$Posterior_i = \beta_0 + \beta_1 Asian_i + \beta_2 X_i + \epsilon_i \quad (1.20)$$

where $Posterior_i$ is employer i 's updated belief about the average score of the assigned worker group, $Asian_i$ is a binary indicator that equals 1 if employer i is randomly assigned the Asian worker group, and X_i is a vector of employer demographic controls. The coefficient $\beta_1 > 0$ would provide evidence that the prior belief bias survives self-directed sampling. I estimate this regression on the Voluntary subsample and report the results in [Table 1.2](#). Standard errors are heteroskedasticity-robust at the individual employer level.

HYPOTHESIS 3 (Exogenous-Benchmark Posterior Gap Compression). *The Asian-Hispanic gap in posterior beliefs is smaller in the ExogenousMean condition than in the Voluntary condition. The gap is the smallest in the Exogenous20 condition.*

To test [Hypothesis 3](#), I compare the differences in the Asian-Hispanic belief gap in updated employer beliefs between information conditions. Specifically, I estimate the following difference-in-differences regression for employer i :

$$\begin{aligned} Posterior_i = & \beta_0 + \beta_1 Asian_i + \rho_1 ExogenousMean_i + \rho_2 Exogenous20_i \\ & + \phi_1 Asian_i \times ExogenousMean_i + \phi_2 Asian_i \times Exogenous20_i \\ & + \beta_2 X_i + \epsilon_i \end{aligned} \quad (1.21)$$

where $Posterior_i$ is the employer's updated belief about the average score of the assigned worker group, $Asian_i$ is a binary indicator that equals 1 if employer i 's assigned worker group is Asian, and $ExogenousMean_i$ and $Exogenous20_i$ are indicators for whether employer i is randomly assigned to the ExogenousMean or Exogenous20 treatment, respectively. The omitted reference group is the Voluntary treatment. The vector X_i contains employer demographic characteristics.

In this regression, coefficient β_1 represents the Asian-Hispanic posterior gap in the Voluntary treatment. Coefficients ρ_1 and ρ_2 measure the effects of the ExogenousMean and Exogenous20 treatments, respectively, on updated beliefs about Hispanic workers, relative to Voluntary. The

interaction coefficients ϕ_1 and ϕ_2 represent difference-in-differences estimates, capturing how the belief gap between Asian and Hispanic workers changes under each alternative treatment compared to Voluntary. The coefficient ϕ_1 quantifies how the belief gap differs in ExogenousMean relative to Voluntary, that is, the treatment effect of removing endogeneity of information search on belief bias while holding average sample size constant; ϕ_2 captures the analogous effect for Exogenous20. [Hypothesis 3](#) predicts $\phi_1 < 0$ and $\phi_2 < \phi_1 < 0$. I report the regression in [Table 1.3](#); the table also includes a Commitment column that I return to in [Section 1.6](#). Standard errors are heteroskedasticity-robust at the individual employer level.

1.4.3. Discriminatory Behavior

With prior and posterior belief hypotheses set, I now ask whether wage offers reflect the prior belief gap in the absence of any information about the worker pool. I examine wage offers in the Baseline condition, where employers receive no productivity information about the assigned worker group beyond what they were told in the instructions. Cross-regime wage gap patterns across Voluntary, ExogenousMean, and Exogenous20, the formal belief-mediation decomposition, and the Commitment differential prior loading are post-hoc mechanism exhibits, all reported in [Section 1.6.3](#).

HYPOTHESIS 4 (Baseline Wage Discrimination). *The wage offers for Hispanic workers are lower than those for Asian workers in the Baseline condition.*

To formally test [Hypothesis 4](#), I estimate the following specification for the Baseline condition:

$$Wage_{i,j} = \beta_0 + \beta_1 Asian_i + \delta W_{i,j} + \gamma X_i + \epsilon_i \quad (1.22)$$

where $Wage_{i,j}$ is the wage offer by employer i to worker j , and $Asian_i$ is a binary indicator for worker ethnicity assignment for employer i . The vector $W_{i,j}$ includes worker characteristics presented in the worker profile, such as age category, gender, favorite color, and X_i is a vector of employer demographic characteristics. The coefficient β_1 captures the average wage gap between Asian and Hispanic workers within the Baseline condition. The coefficient $\beta_1 > 0$ indicates that employers discriminate against Hispanic workers in their wage offers. I report the regression in [Table 1.5](#). Standard errors are clustered at the employer level to account for within-employer correlation across

the multiple wage offers per employer in the long-format wage panel.

Next, I examine treatment differences in the Asian-Hispanic wage gap across information acquisition modes. If different modes of information acquisition lead to systematically different beliefs, then those beliefs should also predict subsequent wage offers. I formulate the following hypothesis based on predictions in [Hypothesis 3](#).

HYPOTHESIS 5 (Wage Discrimination After Learning). *The Asian-Hispanic wage gap in the Exogenous Mean treatment is smaller than the wage gap in Voluntary. The wage gap is the smallest in Exogenous 20.*

To formally test [Hypothesis 5](#), I estimate the following difference-in-differences specification:

$$\begin{aligned}
Wage_{i,j} = & \beta_0 + \beta_1 Asian_i + \rho_1 Voluntar y_i + \rho_2 ExogenousMean_i \\
& + \rho_3 Exogenous20_i + \phi_1 Asian_i \times Voluntar y_i \\
& + \phi_2 Asian_i \times ExogenousMean_i + \phi_3 Asian_i \times Exogenous20_i \\
& + \delta W_{i,j} + \gamma X_i + \epsilon_i
\end{aligned} \tag{1.23}$$

In this regression, $Wage_{i,j}$ is the wage offer by employer i to worker j , and $Asian_i$ is a binary indicator for worker ethnicity assignment for employer i . The dummy variables $Voluntar y_i$, $ExogenousMean_i$, and $Exogenous20_i$ are indicators for information treatment assignment, with Baseline serving as the omitted reference group. The coefficient β_1 reflects the average wage gap between Asian and Hispanic workers in the Baseline condition. The interaction coefficients ϕ_1 through ϕ_3 represent treatment-specific differences in the wage gap relative to the Baseline. For example, ϕ_1 measures the change in the wage gap under Voluntary relative to Baseline, that is, the treatment effect on wage discrimination, and ϕ_3 captures the analogous effect for the Exogenous 20 treatment. To compare wage discrimination across treatments directly, I test for statistical differences between the interaction terms ϕ_1 through ϕ_3 . [Hypothesis 5](#) predicts $\phi_2 < \phi_1 < 0$ and $\phi_3 < \phi_2 < 0$. I report the regression in [Table 1.6](#). Standard errors are clustered at the employer level to account for within-employer correlation across the multiple wage offers per employer in the long-format wage panel.

1.5. Results

In this section, I present results from the employer experiment. [Section 1.5.1](#) describes the analysis sample, and [Section 1.5.2](#) reports results on bias in employer prior beliefs before learning. [Section 1.5.3](#) discusses results on the employers’ posterior belief bias after information search, and [Section 1.5.4](#) reports results on wage discrimination.

1.5.1. Sample Characteristics

The analysis sample consists of 2,906 employers, with Treatment Arm counts ranging from 443 to 510 employers per cell across the six treatment arms. Each employer in the information-acquisition treatments and the Baseline arm submits ten wage offers in the wage-offer stage, yielding a long-format wage panel of roughly 19,670 employer-worker observations after the analysis-sample exclusions documented in [Appendix 1.9](#). [Table 1.15](#) in the appendix reports the full demographic composition of the analysis sample, broken out by employer race, Hispanic indicator, gender, age group, and Treatment Arm.

Employer demographics are balanced across treatment arms within each worker ethnicity assignment. Appendix balance tables ([Tables 1.16](#) and [1.17](#)) report mean demographic characteristics by information treatment, separately for employers assigned the Hispanic and Asian worker groups, and confirm that the random assignment produced comparable employer pools across cells. The sample is predominantly female (60%) and White (76%), with a median age in the 35–44 bracket. Asian and Hispanic employers comprise about 9% and 9% of the sample, respectively, so in-group effects are unlikely to drive the main results reported below.

1.5.2. Employer Prior Belief Bias

I begin by examining employer prior beliefs across different information conditions in each worker ethnicity treatment arm. Importantly, my experimental design standardizes the prior elicitation stage: all employers assigned to the same worker group (Asian or Hispanic) respond to identical belief questions before receiving any information, regardless of their treatment condition. This design ensures that any differences in prior belief distributions across treatments arise solely from

random variation rather than information treatment assignment. If this condition holds, it enables clean comparisons of posterior beliefs across treatments, since differences in posteriors can then be attributed to the effects of information acquisition rather than prior belief differences.

To verify this design feature, I estimate prior beliefs cell-by-cell across the six treatment arms and report the regression in [Table 1.1](#). The Asian worker group main effect is 1.682 ($p < 0.01$), confirming substantial prior bias against Hispanic workers. None of the treatment-arm main effects on the estimated mean belief from the PMF are statistically significant, and none of the Asian $\times$ treatment interactions are significant either. Together, these coefficients indicate that prior belief distributions are statistically indistinguishable across Treatment Arms, supporting the internal validity of the cross-treatment comparisons that follow. I corroborate this regression-based balance check with a nonparametric Kruskal–Wallis (K–W) test of equality of the prior distribution across the six Treatment Arms, computed separately for each worker-ethnicity assignment. The K–W test fails to reject equality on the estimated mean belief from the PMF (Hispanic $p = 0.23$, Asian $p = 0.11$) and delivers the same conclusion on the self-reported mean belief ($p = 0.11$ and $p = 0.20$, respectively). Appendix Figures [1.10A](#) and [1.10B](#) visualize the same balance check: empirical CDFs of prior beliefs overlap closely across information conditions within each worker ethnicity assignment.

Next, I examine bias in prior beliefs, specifically whether employers hold systematically different expectations about worker performance before receiving any information about the group. [Figure 1.5A](#) displays employers’ prior beliefs about the average productivity of their assigned worker group, pooled across all information treatments. The results show a clear pattern: employers assigned to the Asian worker group tend to overestimate average performance, while those assigned to the Hispanic group tend to underestimate it. Both deviate from the true average productivity (indicated by the dotted horizontal line at 7.1), which is identical for both groups by design.

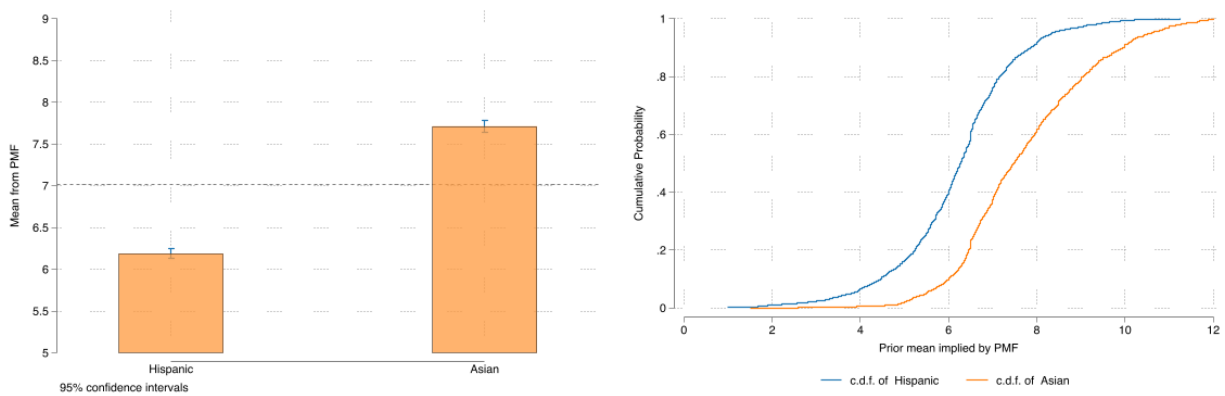
The Asian worker group main effect of 1.682 in [Table 1.1](#) is the test statistic for [Hypothesis 1](#): employers believe that Asian workers scored 1.682 points higher than Hispanic workers on average before receiving any performance information. The estimate is highly significant ($p < 0.01$), and its magnitude corresponds to roughly one standard deviation of the prior belief distribution, representing a large and economically meaningful bias against Hispanic workers in employer prior beliefs.

TABLE 1.1. Prior Beliefs by Treatment Arm

	Prior (PMF)	Prior (Self-reported)
Asian Worker	1.682*** (0.134)	1.834*** (0.153)
Voluntary	-0.164 (0.132)	-0.368** (0.151)
ExogenousMean	-0.056 (0.136)	0.048 (0.156)
Exogenous20	0.111 (0.133)	0.029 (0.152)
Commitment	-0.166 (0.131)	-0.179 (0.149)
Match	-0.050 (0.137)	-0.171 (0.156)
Asian $\times$ Voluntary	-0.097 (0.188)	0.089 (0.215)
Asian $\times$ ExogenousMean	-0.070 (0.194)	-0.320 (0.222)
Asian $\times$ Exogenous20	-0.249 (0.188)	-0.190 (0.214)
Asian $\times$ Commitment	-0.108 (0.188)	-0.230 (0.214)
Asian $\times$ Match	0.037 (0.195)	-0.054 (0.222)
Constant (Hispanic, Baseline)	6.229*** (0.092)	6.370*** (0.105)
Observations	2906	2906

Notes: This table reports OLS estimates of prior beliefs about the average score of the assigned worker group, regressed on the Asian worker group indicator, treatment-arm indicators, and their interactions, with Hispanic-Baseline as the omitted reference cell. The estimated mean belief from the PMF (Column 1) is the primary measure per the project's design; the self-reported mean belief (Column 2) is reported as a companion measure. The Voluntary, ExogenousMean, Exogenous20, Commitment, and Match indicators capture treatment-arm effects relative to Baseline. Sample: analysis sample, employer-level cross-section, $N = 2,906$. Standard errors are heteroskedasticity-robust. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

FIGURE 1.5. Prior Beliefs, Pooled Across Information Conditions



A. Mean

B. Empirical Distribution

Notes: $N = 3734$. Pooled across all information treatments within each worker group ethnicity treatment arm.

This pattern is further confirmed in [Figure 1.5B](#), which plots the empirical cumulative distribution functions (CDFs) of prior beliefs. The CDF for the Asian group first-order stochastically dominates that of the Hispanic group, indicating that employers systematically expect Asian workers to perform better. A Kolmogorov-Smirnov test rejects the null hypothesis of equal distributions ($p < 0.001$), providing strong evidence of biased prior beliefs based on worker group assignment. The next result follows:

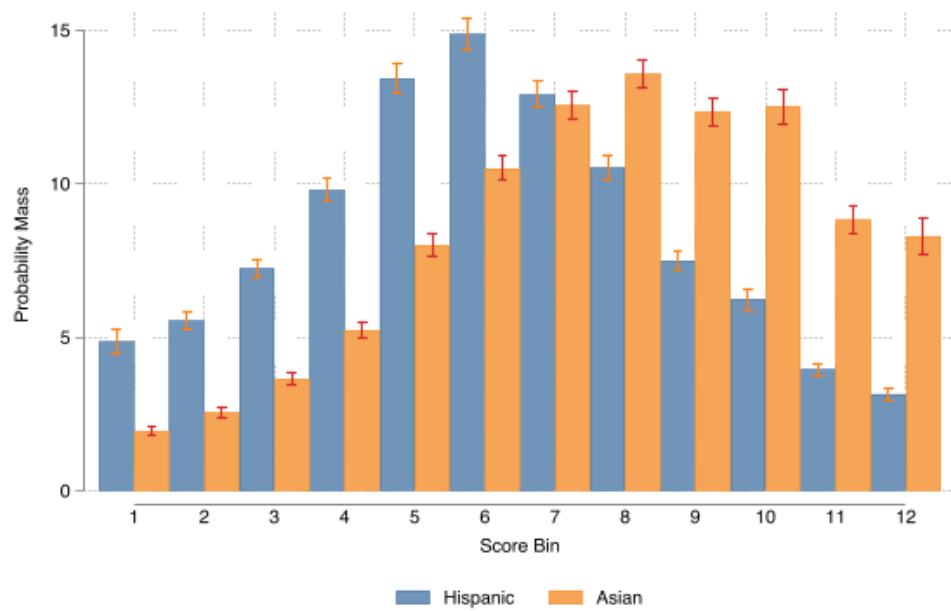
RESULT 1 (Bias in Prior Beliefs). *Before receiving performance information, employers assigned the Asian worker group hold higher prior beliefs about average worker productivity than employers assigned the Hispanic worker group.*

In addition to examining beliefs about average worker score, I also analyze the full subjective probability mass function over worker scores reported by employers. While average estimates offer a summary measure of perceived group performance, they may obscure important difference in the shape of the underlying distribution—particularly the expected proportions of high- and low-performing workers. To capture this, I elicit employers’ beliefs about the number of worker in each possible score category in their assigned group. [Figure 1.6](#) presents the results by worker group assignment. The pattern is asymmetric: the subjective belief distribution for Asian workers is strongly right-skewed, with employers expecting a high concentration of top scores, whereas the belief distribution for Hispanic workers is relatively symmetric, with employers expecting similar proportions of high- and low-performers. To test for statistical differences in these belief distributions, I conduct an Energy Distance Test (Székely & Rizzo, 2004), a nonparametric test for equality of multivariate distributions that is sensitive to both location and shape. The results indicate a statistically significant difference between the two distributions ($p < 0.001$), showing that bias in prior beliefs stems from the differential expectations about the prevalence of high-performing workers in each group.

1.5.3. Employer Posterior Belief Bias

After the information acquisition stage, employers report their updated beliefs about worker performance in their assigned group. Here I present results on the posterior belief bias in the information treatments: Voluntary, Exogenous Mean, and Exogenous 20.

FIGURE 1.6. Prior Subjective Belief Distributions



Notes: Prior subjective belief distributions of employers by worker group race, pooled across all treatments.

I begin with the Voluntary treatment in isolation. [Table 1.2](#) reports the cross-sectional regression of posterior beliefs on the Asian worker group indicator, restricted to Voluntary employers. Despite having the opportunity to sample performance information by drawing workers from the group, Voluntary employers remain significantly biased against Hispanic workers: on average, they believe Asian workers scored 0.537 points higher than Hispanic workers ($p < 0.01$), a disparity equivalent to roughly half a standard deviation of the posterior distribution. The Voluntary residual gap is approximately 32% of the prior gap of 1.682, indicating that voluntary information acquisition closes about two-thirds of the prior gap but leaves a substantial residual.

TABLE 1.2. Posterior Beliefs in Voluntary

	Posterior (PMF)	Posterior (Self-reported)
Asian Worker Group	0.537*** (0.116)	0.482*** (0.137)
Constant	5.735*** (0.946)	6.096*** (1.123)
Hispanic Mean Posterior	6.711	6.879
Asian Mean Posterior	7.245	7.352
Observations	503	503
Employer Controls	Yes	Yes

Notes: This table reports OLS estimates of posterior beliefs on the Asian worker group indicator, restricted to the Voluntary treatment. The estimated mean belief from the PMF (Column 1) is the primary measure; the self-reported mean belief (Column 2) is reported as a companion measure. The Hispanic Mean Posterior and Asian Mean Posterior rows report raw posterior means by worker group assignment. Sample: Voluntary employers in the analysis sample, $N = 503$. Employer controls include age, race/ethnicity, education level, and gender. Standard errors are heteroskedasticity-robust. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

RESULT 2 (Voluntary Residual Posterior Gap). *Within the Voluntary treatment, the Asian-Hispanic posterior gap is positive and statistically significant.*

I now turn to the cross-treatment comparison. To test how exogenously assigned information acquisition affects the posterior belief gap relative to Voluntary, I estimate the difference-in-differences specification in [Equation 1.21](#) from [Section 1.4](#), with Voluntary as the omitted reference cell. [Table 1.3](#) reports the regression. The Voluntary residual posterior gap on the Asian worker group main effect is 0.532 ($p < 0.01$), virtually identical to the Voluntary-only estimate above. The Asian $\times$ ExogenousMean interaction is -0.273 ($p < 0.10$), implying a posterior gap of 0.259 in

ExogenousMean: the gap narrows by 0.273 points relative to Voluntary, a reduction of about 51%. The Asian $\times$ Exogenous20 interaction is -0.425 ($p < 0.01$), implying a posterior gap of 0.107 in Exogenous20, an even greater reduction of about 80% relative to Voluntary. The Asian $\times$ Commitment interaction is -0.204 but not statistically significant, though it is marginally significant for the self-reported mean belief measure; I return to Commitment as a mechanism arm in Section 1.6.

TABLE 1.3. Treatment Effects on Posterior Belief Bias

	Posterior (PMF)	Posterior (Self-reported)
Asian Worker Group	0.532*** (0.096)	0.490*** (0.117)
ExogenousMean	0.195* (0.100)	0.118 (0.122)
Exogenous20	0.253*** (0.097)	0.307*** (0.119)
Commitment	0.028 (0.096)	0.083 (0.117)
Asian $\times$ ExogenousMean (ϕ_2)	-0.273* (0.140)	-0.361** (0.171)
Asian $\times$ Exogenous20 (ϕ_3)	-0.425*** (0.135)	-0.445*** (0.165)
Asian $\times$ Commitment (ϕ_5)	-0.204 (0.135)	-0.315* (0.165)
$H_0 : \phi_2 = \phi_3 = \phi_5$ (joint Wald)	0.245	0.725
$H_0 : \phi_2 = \phi_5$ (ExogMean = Commitment)	0.622	0.783
$H_0 : \phi_2 = \phi_3$ (ExogMean = Exog20)	0.276	0.624
$H_0 : \phi_3 = \phi_5$ (Exog20 = Commitment)	0.101	0.428
SD of Posterior	1.087	1.324
Observations	1967	1967
Employer Controls	Yes	Yes

Notes: This table reports OLS estimates of Equation 1.21, with Voluntary as the omitted reference cell. The estimated mean belief from the PMF (Column 1) is the primary measure; the self-reported mean belief (Column 2) is reported as a companion measure. The interaction coefficients ϕ_2 , ϕ_3 , and ϕ_5 are the difference-in-differences estimates of the Asian-Hispanic posterior gap under ExogenousMean, Exogenous20, and Commitment relative to Voluntary. The joint Wald test $H_0 : \phi_2 = \phi_3 = \phi_5$ tests strict equality of all three treatment effects on the posterior gap. Sample: employers in Voluntary, ExogenousMean, Exogenous20, and Commitment, $N = 1,967$. Employer controls include age, race/ethnicity, education level, and gender. Standard errors are heteroskedasticity-robust. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

The point estimates trace out the strict ordering predicted by Corollary 1: the residual posterior

gap is largest in Voluntary, smaller in ExogenousMean, and smallest in Exogenous20. The reduction in belief bias from Voluntary to ExogenousMean is driven by the removal of endogenous control over information search rather than simply having more information, since the average quantity of information is identical across these two arms by design. Exogenous20, which provides a larger fixed sample of 20 random workers, closes more of the gap because employers receive both more information and information whose content is independent of their priors. The joint Wald test of strict equality across the three interaction coefficients does not reject ($p = 0.245$), so the data are underpowered to distinguish among the three reductions; the support is direction-only on the strict ordering.

RESULT 3 (Ordering of Posterior Belief Gaps). *The Asian-Hispanic posterior gap is smaller in ExogenousMean than in Voluntary, and smallest in Exogenous20. The point estimates support the strict ordering $\text{Voluntary} > \text{ExogenousMean} > \text{Exogenous20}$ predicted by Corollary 1, although the joint Wald test does not reject equality of the three treatment effects.*

To corroborate this regression-based ordering with a nonparametric distributional measure, I compute energy distances between the posterior belief distributions of Asian-arm and Hispanic-arm employers, separately within each treatment arm. Table 1.4 reports the results. The energy distance is largest in Voluntary ($\mathcal{E} = 0.900$, $p < 0.01$), intermediate in Commitment ($\mathcal{E} = 0.622$, $p < 0.01$), substantially smaller in ExogenousMean ($\mathcal{E} = 0.239$, $p = 0.026$), and indistinguishable from zero in Exogenous20 ($\mathcal{E} = 0.042$, $p = 0.247$). The full distributional ordering $\text{Voluntary} > \text{Commitment} > \text{ExogenousMean} > \text{Exogenous20}$ confirms the regression-based pattern: the wedge between Asian-arm and Hispanic-arm posterior distributions monotonically shrinks as endogeneity in information acquisition is removed.

I also examine within-Voluntary belief accuracy, defined as the squared deviation of the posterior from the true average productivity. The squared-error regression of posterior accuracy on the Asian worker group indicator, restricted to Voluntary employers, yields a coefficient of 0.353 that is not statistically significant (Appendix Table 1.18). Within Voluntary, accuracy improves slightly relative to priors, but there is no race-differential improvement: posterior accuracy is similar across worker ethnicity assignments, even though the gap in posterior levels survives. This pattern is consistent

TABLE 1.4. Energy Distance Between Posterior Distributions, by Treatment Arm

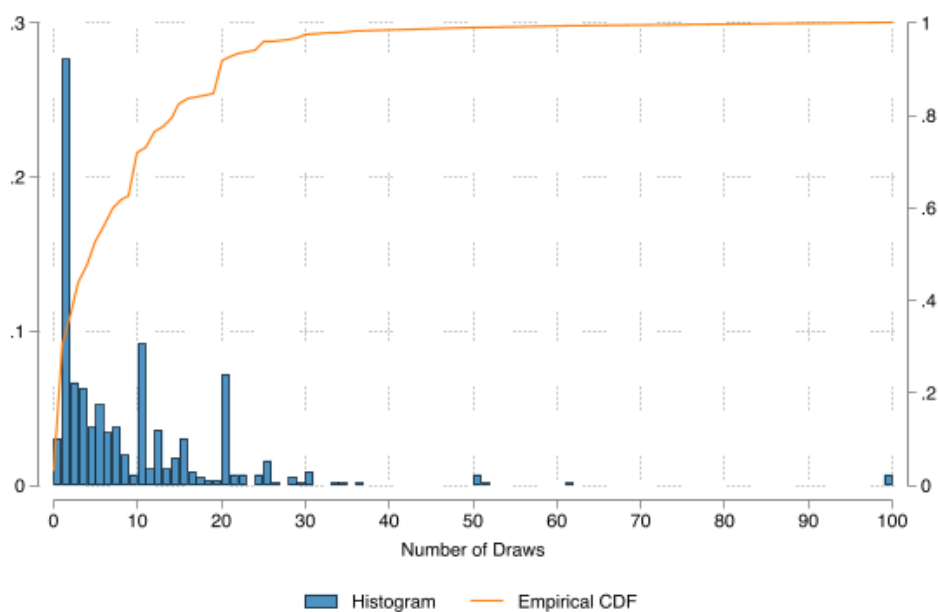
Comparison	Energy stat ($\mathcal{E}$)	Perm. p
Asian vs Hispanic, Voluntary ($n_A = 256$, $n_H = 247$)	0.9004***	0.001
Asian vs Hispanic, Commitment ($n_A = 251$, $n_H = 258$)	0.6216***	0.001
Asian vs Hispanic, ExogenousMean ($n_A = 225$, $n_H = 220$)	0.2388**	0.026
Asian vs Hispanic, Exogenous20 ($n_A = 270$, $n_H = 240$)	0.0421	0.247
Asian vs Hispanic, Match ($n_A = 224$, $n_H = 219$)	1.1889***	0.001

Notes: This table reports the energy distance test (Székely & Rizzo, 2004) between posterior belief distributions of employers assigned the Asian worker group and those assigned the Hispanic worker group, computed separately within each treatment arm. The energy statistic $\mathcal{E}$ measures the multivariate distance between two empirical distributions; permutation p -values are computed under the null of equal distributions. Sample: estimated mean posterior belief from the PMF, analysis sample. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

with realization-based stopping affecting the bias channel without inducing a separate accuracy gap.

I now examine how much information employers actually gather when they have full control over the information search process. Figure 1.7 displays both the histogram and the empirical cumulative distribution of the number of draws in the Voluntary treatment, pooled across employers assigned to Asian and Hispanic worker groups. On average, employers make 8.7 draws, a small fraction of the 100-worker pool. The median number of draws is even lower, at 5, indicating a skewed distribution with a long right tail. There is clear clustering at salient round numbers such as 10 and 20, suggesting the use of intuitive or heuristic stopping rules. Over a quarter of employers (28%) stop after a single draw, indicating that information acquisition is limited for a substantial share of the sample.

FIGURE 1.7. Number of Draws in Voluntary



When disaggregated by worker group, employers assigned to evaluate Hispanic workers make slightly more draws on average (mean = 9) than those assigned to Asian workers (mean = 8). This is shown in Figure 1.11, which plots the kernel density functions of number of draws by worker group assignment. The difference is mainly driven by the larger share of employers who make few draws for Asian workers. Employers appear more certain about their prior beliefs for Asian workers; stronger priors predict less intensive search.

1.5.4. Discriminatory Behavior

Having established how different modes of information acquisition affect belief formation, I now turn to the consequences of these beliefs for decision-making. In particular, I examine how biased beliefs translate into wage discrimination. Since worker groups are designed to be objectively identical in productivity, any systematic difference in wage offers between the Asian and Hispanic groups reflects discrimination, whether rooted in biased beliefs or preference. This allows me to assess not only the behavioral relevance of belief distortions, but also whether improved belief accuracy can meaningfully reduce discriminatory behavior.

I begin by examining wage offers in the Baseline condition, where employers make hiring decisions without receiving any additional information about the productivity distribution of their assigned worker group beyond what they were told in the instructions. In this setting, employers evaluating Hispanic workers offer significantly lower wages on average than those evaluating Asian workers, despite the fact that both groups have identical underlying performance. This pattern of behavior is consistent with belief-based discrimination: in the absence of objective information, employers rely on prior expectations that favor Asian workers over Hispanic ones. The baseline findings show that biased prior beliefs drive discriminatory outcomes in this no-information setting, providing a benchmark against which the effects of subsequent information interventions can be evaluated.

To formally assess whether employers discriminate against Hispanic workers in the Baseline condition, I estimate [Equation 1.22](#) in [Section 1.4](#) for employers in the Baseline condition, reproduced below:

$$Wage_{i,j} = \beta_0 + \beta_1 Asian_i + \delta W_{i,j} + \gamma X_i + \epsilon_i$$

This specification controls for both the observable characteristics of worker j (captured by $W_{i,j}$) and the demographic characteristics of employer i (captured by X_i). Standard errors are clustered at the employer level to account for within-employer correlation, since each employer evaluates and makes wage offers to ten workers. If the coefficient $\beta_1 > 0$, this would indicate that Asian workers are offered systematically higher wages than Hispanic workers with equivalent observable profiles, which is evidence of wage discrimination against Hispanics.

TABLE 1.5. Wage Discrimination in Baseline

	Wage Offer in Baseline
Asian Worker Group	0.604*** (0.161)
Female Worker	0.294*** (0.106)
45+ years old Worker	-0.161** (0.082)
Constant	5.913*** (0.367)
Hispanic Mean Wage	6.452
Asian Mean Wage	7.035
SD of Wage	2.242
Observations	5490
Unique Employers	549
Employer Controls	Yes
Worker Controls	Yes
Clustered SE	Yes

Notes: This table reports OLS estimates of [Equation 1.22](#), and the data consists of panel data from employers in the Baseline condition, with 10 observations per employer. Asian Worker Group is a binary indicator equal 1 if the employer's assigned worker group is Asian. Employer controls include age, race/ethnicity, education level, and gender. Worker controls include the worker's age (18-45 or 45+), gender (Male or Female), race of the name (white-sounding name or Hispanic/Asian name), and favorite color. Coefficient estimates for worker age and gender are displayed in the table, the rest are omitted. Standard errors, clustered at the employer level, are displayed in parentheses. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

RESULT 4 (Baseline Wage Discrimination). *In the Baseline condition, employers offer significantly lower wages to Hispanic workers than to Asian workers with similar observable profiles.*

The results, presented in Table 1.5, show a statistically significant wage gap of approximately 0.60 experimental tokens, or about 27% of a standard deviation (2.24), favoring Asian workers. Notably, the estimates also suggest that female workers receive slightly higher wage offers on average, pointing to a potential *wage premium* for women in this setting.

This wage gap is consistent with the distorted prior beliefs documented earlier. In the Baseline condition, employers hold systematically lower beliefs about the average productivity of Hispanic workers, despite both groups having identical performance distributions. Since employers rely on these biased priors in the absence of additional information, the observed wage discrimination likely stems from incorrect beliefs rather than taste-based preferences. This interpretation is supported by the finding that belief gaps in Baseline are large and favor Asian workers. Thus, the evidence shows that biased beliefs translate directly into wage discrimination in this no-information setting.

Next, I investigate how different modes of information acquisition affect wage discrimination by estimating the wage-level regression in Table 1.6, with Voluntary as the omitted reference cell. The Asian worker group main effect is 0.303 ($p < 0.05$): Voluntary employers offer Asian workers about 0.30 tokens more than Hispanic workers on average, a wage premium roughly half the size of the Baseline gap and consistent with information acquisition partially attenuating belief-driven discrimination. In contrast, the Asian $\times$ ExogenousMean interaction is -0.428 ($p < 0.05$) and the Asian $\times$ Exogenous20 interaction is -0.429 ($p < 0.05$), implying wage gaps in the two forced-sampling arms of approximately -0.13 in each case. Forced sampling closes the Voluntary wage premium, and the point estimates lean slightly negative, although neither falls outside conventional confidence bounds for zero.

RESULT 5 (Wage Discrimination After Learning). *The Asian-Hispanic wage gap in the Exogenous Mean and Exogenous 20 treatments is smaller than the gap in Voluntary: Voluntary employers retain a wage premium for Asian workers of about 0.30 tokens ($p < 0.05$), while the wage gap under forced sampling is statistically indistinguishable from zero.*

TABLE 1.6. Wage Offers by Treatment Arm

	Wage
Asian Worker Group	0.303** (0.124)
ExogenousMean	0.329*** (0.119)
Exogenous20	0.339*** (0.120)
Commitment	0.006 (0.123)
Asian $\times$ ExogenousMean	-0.428** (0.185)
Asian $\times$ Exogenous20	-0.429** (0.184)
Asian $\times$ Commitment	0.111 (0.187)
Observations	19670
Employer Controls	Yes
Worker Controls	Yes
Clustered SE	Yes

Notes: This table reports OLS estimates of wage offers on the Asian worker group indicator, treatment-arm indicators, and their interactions, with Voluntary as the omitted reference cell. The interaction coefficients on Asian $\times$ ExogenousMean, Asian $\times$ Exogenous20, and Asian $\times$ Commitment are the difference-in-differences estimates of the Asian-Hispanic wage gap under each information treatment relative to Voluntary. Sample: long-format wage panel for employers in Voluntary, ExogenousMean, Exogenous20, and Commitment, $N = 19,670$. Employer controls include age, race/ethnicity, education level, and gender. Worker controls include age category, gender, race of the name, and favorite color. Standard errors are clustered at the employer level. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

These wage patterns closely mirror the patterns observed in belief formation: voluntary information acquisition fails to fully eliminate biased beliefs and biased wage offers, whereas exogenous information acquisition both reduces posterior bias and removes the Asian-Hispanic wage premium. Section 1.6 formally decomposes the wage gap into belief-driven and preference-driven components and quantifies how much of the residual race effect operates through the posterior belief channel.

Taken together, the results demonstrate that endogeneity in information acquisition plays a central role in the persistence of belief biases and wage discrimination. When employers are free to choose both the quantity and content of information, as in the Voluntary treatment, their posterior beliefs remain biased and their wage offers continue to discriminate against Hispanic workers. In contrast, when the learning process is externally constrained, as in the Exogenous Mean and Exogenous 20 treatments, employers place more weight on performance data, form more accurate beliefs, and offer fairer wages. The next section investigates the mechanisms that drive the differences in belief bias and discriminatory outcomes between Voluntary and Exogenous information acquisition modes.

1.6. Mechanisms

The preceding results reveal the following pattern: employers in the Voluntary treatment exhibit the most biased posterior beliefs, despite having access to informative signals through their own sampling. Even when the average quantity of information is held equal to that of Voluntary, belief bias is substantially reduced in the Exogenous Mean condition. This raises the question: what features of endogenous information are driving the sustained residual posterior gap?

This section proceeds in four steps. First, I consider an alternative mechanism that could plausibly sustain the Voluntary residual gap, namely a within-employer correlation between prior bias and prior confidence, and use the Matched Sample treatment to rule it out (Section 1.6.1). Second, with that alternative foreclosed, I identify realization-based stopping, the ability to stop search conditional on signal realizations, as the channel through which Voluntary employers' biased priors are preserved in their posterior beliefs (Section 1.6.2). Third, I decompose wage offers into belief-driven and preference-driven components and show that wage discrimination in this setting

is almost entirely belief-driven (Section 1.6.3). Finally, I close with descriptive evidence on belief updating dynamics that complement the mechanism (Section 1.6.4).

1.6.1. Ruling Out Bias-Confidence Correlation: Evidence from the Matched Sample Treatment

Before identifying realization-based stopping as the channel sustaining the Voluntary residual gap, I consider an alternative mechanism that could plausibly produce the same pattern. Suppose employers with more biased priors, that is, priors farther from the truth, also hold those priors with higher confidence, in the sense that their belief distributions are more concentrated (with lower variance). Under voluntary information acquisition, employers with high-confidence priors stop after fewer signal draws, because their perceived posterior precision crosses the optimal-stopping threshold sooner. Such employers would update less, and their prior bias would be carried into their posterior. Under this account, the Voluntary residual Asian-Hispanic gap is a downstream consequence of within-employer correlation between prior bias magnitude and prior confidence: the bias would live in the within-employer pairing between prior characteristics and stopping behavior, not in the realized signal sequences themselves.

A first piece of descriptive evidence against this mechanism comes from the observed sampling behavior of Voluntary employers. Figure 1.8 plots the binned scatter plot of prior beliefs against Voluntary sample sizes, with separate regression lines for employers assigned the Asian and Hispanic worker groups. Both lines are essentially flat: employers with priors farther from the truth draw a similar number of signals as employers with priors near the truth, in both worker ethnicity assignments. The chain (high prior bias $\rightarrow$ high prior confidence $\rightarrow$ fewer draws) breaks somewhere, since extreme priors do not predict fewer draws within Voluntary.

The descriptive pattern provides suggestive evidence; next, I test the bias-confidence correlation channel directly with the Matched Sample treatment. In Matched Sample, employers are randomly assigned the exact sequence of signal realizations drawn by a counterpart in the Voluntary treatment with the same worker ethnicity assignment. Random assignment makes the receiving employer's own prior bias and confidence statistically independent of the length and composition of the inherited sequence: the within-employer pairing between prior characteristics and stopping behavior that

FIGURE 1.8. Prior Belief and Voluntary Sample Size



Notes: Binned scatter of employer prior beliefs (estimated mean from the PMF) against Voluntary sample size, with separate linear fits for employers assigned the Asian and Hispanic worker groups. Sample: Voluntary employers, analysis sample. The flat slopes indicate that within Voluntary, employers with priors far from the true average ($\mu = 7.02$) draw approximately the same number of signals as employers with priors near the truth, in both worker ethnicity assignments.

operates in Voluntary cannot operate in Matched Sample, because the Matched Sample employer did not choose when to stop. If the Voluntary residual gap is sustained by this within-employer pairing, random assignment destroys the covariance between prior bias and prior weight in the posterior, and the Asian-Hispanic gap should shrink toward the prior gap times the average prior weight. If, on the contrary, the bias is carried by the realized signal sequence itself, then applying the inherited sequence to any prior reproduces the bias, and the Matched Sample gap should persist at roughly the Voluntary level.

To test the bias-confidence channel, I estimate the following regression on the pooled Voluntary and Matched Sample sample:

$$\begin{aligned} Posterior_i = & \beta_0 + \beta_1 Asian_i + \rho_1 MatchedSample_i \\ & + \phi_1 Asian_i \times MatchedSample_i + \beta_2 X_i + \epsilon_i \end{aligned} \tag{1.24}$$

Here *MatchedSample_i* is an indicator for assignment to the Matched Sample treatment, *Asian_i* is the worker ethnicity assignment indicator, and *X_i* is the vector of employer demographic controls. The coefficient of interest is ϕ_1 , the difference-in-differences estimate of the Asian-Hispanic gap in Matched Sample relative to Voluntary.

Table 1.7 reports the regression. The Voluntary residual posterior gap is 0.538 ($p < 0.01$), and the Asian $\times$ Matched Sample interaction is 0.170 ($SE = 0.170$, not statistically significant), implying a Matched Sample residual gap of 0.708, slightly larger than Voluntary in point estimate. Both estimates are decisively far from the ExogenousMean residual gap of 0.259 documented in Section 1.5.3. The bias-confidence correlation account is rejected: the bias is preserved through the Voluntary $\rightarrow$ Matched Sample transfer even though random assignment destroys the within-employer pairing the alternative account relies on. The bias must therefore live in the signal sequence itself rather than in any within-employer link between prior characteristics and stopping.

The Matched Sample comparison narrows the set of viable mechanisms: the bias does not live in within-employer prior-confidence correlation, nor in rational precision-driven sampling. With both alternatives foreclosed, I now identify the channel through which the bias enters the signal sequence in Voluntary.

TABLE 1.7. Belief Bias in Matched Sample versus Voluntary

	Posterior (PMF)	Posterior (Self-reported)
Asian Worker Group	0.538*** (0.116)	0.483*** (0.138)
Match (vs. Voluntary)	-0.077 (0.121)	-0.109 (0.144)
Asian $\times$ Match	0.170 (0.170)	0.065 (0.202)
Observations	946	946
Employer Controls	Yes	Yes

Notes: This table reports OLS estimates of [Equation 1.24](#) on the pooled Voluntary and Matched Sample sample. The estimated mean belief from the PMF (Column 1) is the primary measure; the self-reported mean belief (Column 2) is reported as a companion measure. Asian Worker Group is an indicator for the Asian worker ethnicity assignment, and Match (vs. Voluntary) is an indicator for assignment to the Matched Sample treatment. The interaction Asian $\times$ Match is the difference-in-differences estimate of the Asian-Hispanic posterior gap under Matched Sample relative to Voluntary. Sample: $N = 946$, analysis sample. Employer controls include age, race/ethnicity, education level, and gender. Standard errors are heteroskedasticity-robust. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

1.6.2. Biased Information from Realization-Based Stopping

With the bias-confidence-correlation alternative ruled out, I now identify what selects the biased signal sequences. In the Voluntary treatment, employers have full endogenous control over the sampling process, which manifests in two key dimensions:

- *Realization-based stopping*: agents decide whether to stop or continue after observing signal realizations.
- *Endogenous sample size*: agents can decide how many signals to draw.

Endogenous sample size affects only the *amount* of information acquired. Realization-based stopping can alter the *content* of the information set itself, because employers may terminate search once they encounter signals consistent with their priors, or continue until they do so, and the resulting evidence can become systematically skewed toward prior beliefs. Even when the underlying signal source is unbiased, realization-based stopping can produce information sets that disproportionately reinforce prior beliefs, leading to posteriors that are biased in the same direction.

Three complementary tests of this channel are reported below: a within-Voluntary hazard model that identifies the cognitive-consistency signature in stopping behavior, a Commitment-versus-Voluntary comparison that breaks realization-based stopping while preserving sample-size endogeneity, and a four-arm synthesis of the Asian-Hispanic posterior gap predicted by Corollary 1.

Empirical Stopping Rule of Voluntary Employers

If Voluntary employers stop selectively in response to signal realizations, then conditional on the current draw, the probability of stopping should depend on the cumulative direction in which realized signals have moved the posterior away from the prior. To test this hypothesis, I investigate the empirical data of Voluntary employers' sampling behavior to see how they make decisions on when to stop sampling. Specifically, I estimate a logit hazard model on the draw-level Voluntary panel, with the indicator for stopping at draw t as the dependent variable:

$$\Pr(stop_{i,t} = 1) = \Lambda(\alpha + \beta_1|\mu_t - \mu_0|_{i,t} + \gamma\mathbf{1}[t = 10] + \gamma'\mathbf{1}[t = 20] + \delta Asian_i + \mathbf{X}'_i\boldsymbol{\theta}) \quad (1.25)$$

where $|\mu_t - \mu_0|_{i,t}$ is the absolute distance between the running posterior mean after t draws and the prior mean, and the round-number indicators $\mathbf{1}[t = 10]$ and $\mathbf{1}[t = 20]$ absorb the heuristic stopping spikes documented in Section 1.5.3. Cluster-robust standard errors at the employer level account for the multiple draws each employer makes.

Table 1.8 reports the estimates. The coefficient on $|\mu_t - \mu_0|$ in the pooled specification is 0.122 ($p < 0.01$): each one-point increase in the running posterior’s distance from the prior raises the log-odds of stopping by 0.122. Employers stop more readily once the realized signals push their beliefs in either direction away from the prior, the cognitive-consistency signature for realization-based stopping. The round-number indicators absorb large spikes at $t = 10$ and $t = 20$ (1.220 and 2.242 in log-odds, both $p < 0.01$), consistent with the heuristic stopping documented in the Voluntary draw-count distribution. The Asian $\times |\mu_t - \mu_0|$ interaction in Column 2 is 0.045 ($SE = 0.080$, not statistically significant), indicating that the cognitive-consistency stopping rule operates similarly across worker ethnicity assignments; the residual gap arises not because Asian-arm and Hispanic-arm employers stop differently in response to signals, but because they enter the sampling process with priors of different magnitudes that are then preserved by a common stopping rule.

Cross-Arm Comparison: Commitment versus Voluntary

The within-Voluntary hazard estimates identify the stopping rule but not the magnitude of the gap closure when realization-based stopping is removed. To estimate that magnitude, I introduce the Commitment treatment, which preserves endogenous sample size but eliminates the ability to condition stopping on signal realizations. In Commitment, employers choose an ex ante sample size n and then must draw exactly that many workers sequentially, regardless of what the realized signals look like. By construction, Commitment removes the endogenous stopping channel while preserving sample-size endogeneity, creating a middle ground between fully endogenous Voluntary and fully exogenous ExogenousMean.

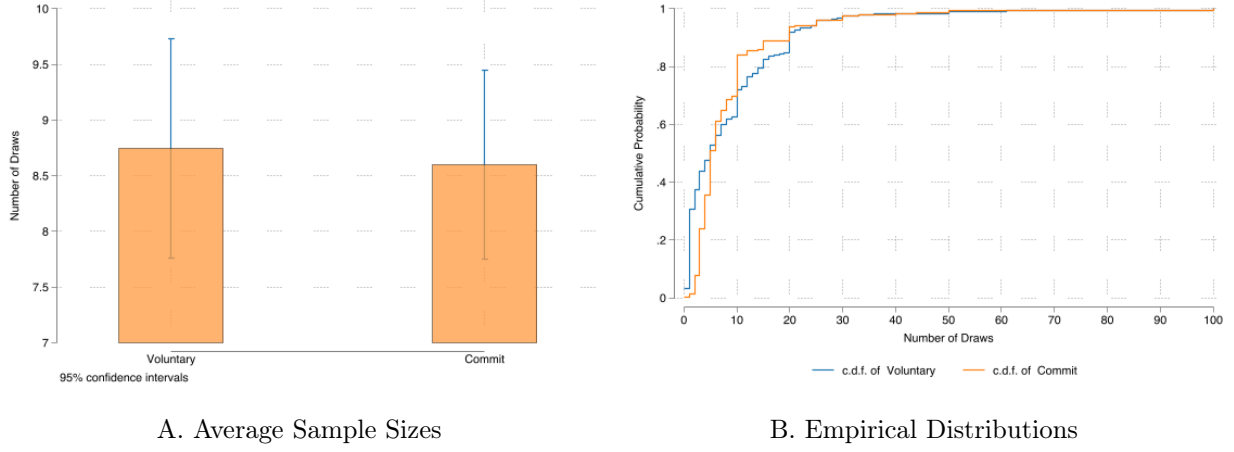
The theoretical framework in Section 1.3 predicts that posterior beliefs in Commitment exhibit less Asian-Hispanic gap than Voluntary but more than ExogenousMean, since Commitment removes the path-dependent component of the bias but retains the prior’s influence on the chosen sample size. Proposition 5 establishes that average sample sizes in Commitment are weakly smaller than in

TABLE 1.8. Hazard Model of Stopping Behavior in Voluntary

	Logit (1)	Logit (2) w/ race interaction
$ \mu_t - \mu_0 $	0.122*** (0.040)	0.097* (0.057)
Asian Worker Group	0.187 (0.133)	0.102 (0.203)
Asian $\times \mu_t - \mu_0 $		0.045 (0.080)
$t = 10$	1.220*** (0.178)	1.221*** (0.179)
$t = 20$	2.242*** (0.231)	2.240*** (0.231)
Observations (draws)	4535	4535
Clustered SE	Yes	Yes

Notes: This table reports logit estimates of Equation 1.25, with the indicator for stopping at draw t as the dependent variable. The coefficient on $|\mu_t - \mu_0|$ is the realization-based-stopping signature: a positive value indicates that employers stop more readily once their running posterior moves away from the prior. Round-number indicators $\mathbf{1}[t = 10]$ and $\mathbf{1}[t = 20]$ absorb heuristic stopping spikes. Sample: draw-level panel for Voluntary employers, analysis sample, $N = 4,535$ draws. Standard errors are clustered at the employer level. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$). A robustness check varying the prior precision parameter N_0 in the Bayesian-updated $|\mu_t - \mu_0|$ construction is reported in Appendix Table 1.27; the sign and significance of the cognitive-consistency signature are sensitive to extreme values of N_0 , which I treat as a boundary condition on the within-Voluntary identification.

FIGURE 1.9. Sample Sizes in Voluntary and Commitment Treatments



Voluntary, $\mathbb{E}[n_{commit}] \leq \mathbb{E}[n_{voluntary}]$, and Proposition 7 establishes the corresponding bias ranking.

Figure 1.9 compares the empirical sample-size distributions in the two treatments. Employers in Commitment draw slightly fewer workers on average, though the difference is not statistically significant. The distributional pattern is more revealing: Commitment reduces the proportion of both very small and very large sample sizes, consistent with the cognitive-consistency interpretation. Some Voluntary employers stop early once realized signals contradict their priors, in order to avoid revising beliefs; others continue when realizations align. Commitment imposes ex ante discipline and produces a tighter, more centered distribution.

Monotonic Belief Gap Closure and Distributional Analysis

Table 1.3 in Section 1.5.3 reports the full four-arm ordering of Asian-Hispanic posterior gaps. The Asian $\times$ Commitment interaction is -0.204 ($SE = 0.135$, not statistically significant), implying a Commitment residual gap of approximately 0.328, between the Voluntary gap of 0.532 and the ExogenousMean gap of 0.259. The point estimates trace out the strict ordering Voluntary $>$ Commitment $>$ ExogenousMean $>$ Exogenous20 predicted by Corollary 1, although the joint Wald test of strict equality across the three forced-and-committed interactions does not reject ($p = 0.245$). The data are underpowered to distinguish among Commitment, ExogenousMean, and Exogenous20; the support for the strict ordering is direction-only.

The energy-distance evidence in [Table 1.4](#) corroborates the regression-based ordering with a nonparametric distributional measure. The full four-arm energy-distance ordering is Voluntary ($\mathcal{E} = 0.900$, $p < 0.01$) > Commitment ($\mathcal{E} = 0.622$, $p < 0.01$) > ExogenousMean ($\mathcal{E} = 0.239$, $p = 0.026$) > Exogenous20 ($\mathcal{E} = 0.042$, $p = 0.247$), strictly monotone in the degree of endogeneity removed. The wedge between Asian-arm and Hispanic-arm posterior distributions shrinks as the realization-based-stopping channel is shut down, first by removing the path-dependent component (Commitment), then by fixing the sample size (ExogenousMean), and finally by enlarging it (Exogenous20).

A complementary observation comes from the within-Voluntary accuracy results documented in [Section 1.5.3](#): posterior accuracy improves under exogenous information acquisition, with the squared-error reduction reported in [Appendix Table 1.28](#) attributable to the removal of realization-based stopping rather than to any race-differential effect. The accuracy gap closes under forced sampling without inducing a separate accuracy differential across worker ethnicity assignments, consistent with the realization-based-stopping account.

Taken together, the within-Voluntary hazard signature, the Commitment-to-Voluntary gap-closure pattern, the four-arm Corollary 1 ordering, and the supporting accuracy evidence all point to realization-based stopping as the channel through which Voluntary employers' biased priors are sustained in their posterior beliefs. The bias enters the information set, not the interpretation of it.

1.6.3. Discrimination: Biased Beliefs or Preferences?

Having identified the mechanism that sustains belief bias under endogenous information acquisition, I now ask how those biased beliefs translate into wage discrimination. Two channels can in principle generate the residual Asian-Hispanic wage gap documented in [Section 1.5.4](#): biased beliefs, in which employers offer lower wages to Hispanic workers because they believe Hispanic workers are less productive; or preferences, in which employers offer lower wages to Hispanic workers conditional on belief, reflecting a taste-based component. Because the Asian and Hispanic worker groups have identical underlying productivity distributions by design, any wage gap that survives netting out beliefs is attributable to preferences rather than to accurate beliefs.

To decompose the wage gap, I estimate the following saturated specification separately for each

information regime:

$$Wage_{i,j} = \beta_0 + \beta_1 Asian_i + \beta_2 Prior_i + \beta_3 Posterior_i + \delta W_{i,j} + \gamma X_i + \epsilon_i \quad (1.26)$$

The residual coefficient β_1 on Asian Worker Group, after netting out both prior and posterior beliefs, is the preference-driven component of the wage gap; if $\beta_1 = 0$, the wage gap is fully explained by beliefs. The coefficients β_2 and β_3 describe how much weight employers place on prior versus posterior beliefs when forming wage offers in each regime.

Table 1.9 reports the four regime-specific decompositions. The main finding is that wage discrimination in this setting is almost entirely belief-driven. Across all four regimes, once both prior and posterior beliefs are netted out, the residual coefficient on Asian Worker Group is no longer the source of the observed wage gap. In Voluntary, the residual is 0.131 ($SE = 0.136$), and in Commitment it is -0.023 ($SE = 0.141$); both are statistically indistinguishable from zero. In ExogenousMean and Exogenous20 the residuals are -0.362 ($p < 0.05$) and -0.269 ($p < 0.10$), opposite in sign to the Voluntary raw wage premium and without a coherent mechanism story; I treat them as a likely fluke and report the full nested decomposition by regime in Appendix Section 1.10. The wage gap is a downstream consequence of the belief gap, not a separate taste-based discrimination channel.

TABLE 1.9. Wage Decomposition: Beliefs versus Preferences

	Voluntary	ExogMean	Exog20	Commitment
Asian Worker (residual)	0.131 (0.136)	-0.362** (0.148)	-0.269* (0.138)	-0.023 (0.141)
Prior (PMF)	0.072 (0.057)	0.068 (0.055)	0.060 (0.049)	0.222*** (0.054)
Posterior (PMF)	0.141** (0.065)	0.343*** (0.084)	0.346*** (0.074)	0.182*** (0.068)
Prior share of belief weight	0.337	0.166	0.148	0.549
Observations	5030	4450	5100	5090
Employer Controls	Yes	Yes	Yes	Yes
Worker Controls	Yes	Yes	Yes	Yes
Clustered SE	Yes	Yes	Yes	Yes

Notes: This table reports OLS estimates of [Equation 1.26](#), run separately for each information regime. Each column is one regression. The coefficient on Asian Worker (residual) is the residual Asian-Hispanic wage gap after netting out both the prior and posterior estimated mean beliefs from the PMF. The Prior share of belief weight row reports the implied weight that the prior receives in employers' wage offers, computed as $\beta_2/(\beta_2 + \beta_3)$. Sample: long-format wage panel, ten observations per employer in each regime. Employer controls include age, race/ethnicity, education level, and gender. Worker controls include age category, gender, race of the name, and favorite color. Standard errors are clustered at the employer level. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

A secondary finding emerges from the prior and posterior coefficients themselves. In Voluntary, ExogenousMean, and Exogenous20, the prior coefficient β_2 is small and not statistically distinguishable from zero (0.072, 0.068, and 0.060 respectively), while the posterior coefficient β_3 is the dominant predictor of wages (0.141, 0.343, and 0.346, all $p < 0.05$ or stronger). Employers in these regimes form wage offers primarily on the basis of their updated beliefs after observing signals. In Commitment, by contrast, the prior coefficient is 0.222 ($p < 0.01$), substantively larger than in any other regime, and the posterior coefficient is 0.182 ($p < 0.01$). The implied prior share of belief weight ($\beta_2/(\beta_2 + \beta_3)$) is 0.549 in Commitment, compared to roughly one-third or less in the three other regimes. Pre-committing to a sample size before seeing any signal realizations appears to anchor employers on the prior even after the committed sample is drawn, in a way that the optional stopping regimes do not. The Commitment prior anchoring is a finding about which belief drives wage decisions across regimes, not about whether beliefs drive discrimination, but it is consistent with the cognitive-consistency framework: Commitment removes the path-dependent stopping channel while preserving and possibly amplifying the influence of the original prior on downstream decisions.

1.6.4. Belief Updating Dynamics

The mechanism developed in Sections 1.6.1–1.6.2 attributes the Voluntary residual gap to realization-based stopping rather than to differential updating from the same information. A natural complementary question is whether employers’ updating, conditional on the signals they receive, departs from the rational benchmark in any systematic way. I examine this question using a regression of posterior beliefs on prior beliefs and the realized sample mean, separately by treatment arm and worker ethnicity assignment:

$$Posterior_i = \alpha_0 + \alpha_2 Prior_i + \alpha_4 SampleMean_i + \alpha_X X_i + \epsilon_i \quad (1.27)$$

Under the rational Bayesian benchmark with a Gaussian prior and signal, the coefficient α_2 equals the prior precision share $\sigma_d^2/(n\sigma_0^2 + \sigma_d^2)$, which is strictly positive at any finite sample size n . A non-zero α_2 on its own does not establish departures from Bayesian updating; the relevant comparison is to the rational benchmark prediction at the realized sample size.

Table 1.10 reports the estimates separately for the two forced-information regimes. In Exogenous-Mean, the prior coefficient is 0.173 ($p < 0.01$) for Hispanic-arm employers and 0.196 ($p < 0.01$) for Asian-arm employers; in Exogenous20, $\alpha_2 = 0.108$ ($p < 0.01$) for Hispanic-arm and 0.189 ($p < 0.01$) for Asian-arm. The sample-mean coefficients α_4 are large and significant in all four cells, ranging from 0.574 to 0.728. Table 1.11 reports the analogous specification on the Commitment subsample, where the prior coefficient is 0.318 ($p < 0.01$) for Hispanic-arm employers and 0.307 ($p < 0.01$) for Asian-arm employers, larger than in the forced-information regimes. The Commitment estimates are consistent with smaller endogenous sample sizes in that arm placing more weight on the prior at the updating stage, in line with Proposition 5's sample-size ranking and the wage-decomposition result that the prior carries higher weight in Commitment wage offers.

TABLE 1.10. Belief Updating in Forced-Information Regimes

	ExogMean H	ExogMean A	Exog20 H	Exog20 A	Pooled H	Pooled A
Prior (α_2)	0.173*** (0.034)	0.196*** (0.036)	0.108*** (0.036)	0.189*** (0.029)	0.176*** (0.034)	0.188*** (0.034)
Sample mean (α_4)	0.673*** (0.058)	0.574*** (0.053)	0.654*** (0.084)	0.728*** (0.070)	0.674*** (0.057)	0.573*** (0.050)
Prior $\times$ Exog20 (α_3)					-0.063 (0.048)	0.007 (0.045)
Sample mean $\times$ Exog20 (α_5)					-0.014 (0.100)	0.159* (0.088)
Exog20					0.475 (0.765)	-1.320* (0.719)
Observations	220	225	240	270	460	495
Employer Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports OLS estimates of Equation 1.27, with the estimated mean posterior belief from the PMF as the dependent variable. Columns 1–2 use the ExogenousMean Hispanic-arm and Asian-arm subsamples; Columns 3–4 use the Exogenous20 subsamples; Columns 5–6 use the pooled forced-information sample with Exog20 indicator interactions. α_2 is the weight on the prior; α_4 is the weight on the realized sample mean. A non-zero α_2 does not establish departures from Bayesian updating, since the rational Bayesian also places non-zero weight on the prior at any finite sample size. Sample: forced-information regimes, analysis sample. Employer controls include age, race/ethnicity, education level, and gender. Standard errors are heteroskedasticity-robust. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

A final descriptive observation: the prior coefficient is slightly larger for Asian-arm employers than for Hispanic-arm employers in the forced regimes, a difference that is consistent across cells but small in magnitude and only partially statistically significant. This asymmetry is best read as a minor descriptive observation about updating heterogeneity across worker ethnicity assignments

TABLE 1.11. Belief Updating in the Commitment Regime

	PMF, H	PMF, A	Self-reported, H	Self-reported, A
Prior (PMF) [α_2]	0.318*** (0.040)	0.307*** (0.035)		
Sample mean [α_4]	0.518*** (0.047)	0.447*** (0.042)	0.754*** (0.055)	0.806*** (0.049)
Prior (Self-reported) [α_2]			0.183*** (0.041)	0.163*** (0.037)
Observations	258	250	258	250
Employer Controls	Yes	Yes	Yes	Yes

Notes: This table reports OLS estimates of [Equation 1.27](#) on the Commitment subsample, separately by worker ethnicity assignment. Columns 1–2 use the estimated mean belief from the PMF as the prior and posterior measures; Columns 3–4 use the self-reported mean belief. Sample: Commitment employers, analysis sample. Employer controls include age, race/ethnicity, education level, and gender. Standard errors are heteroskedasticity-robust. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

rather than as a load-bearing claim about non-Bayesian updating per se. It would take a sharper test, comparing the prior coefficient to the rational benchmark $\sigma_d^2/(n\sigma_0^2 + \sigma_d^2)$ at each cell’s realized sample size, to support a substantive non-Bayesian-updating claim. I leave that comparison for future work.

1.7. Robustness

This section reports two robustness checks on the empirical findings developed in Sections 1.5 and 1.6. The first uses a nonparametric distributional measure to confirm that the residual posterior closure pattern documented in Section 1.5.3 is not an artifact of differential prior distributions across Treatment Arms (Section 1.7.1). The second formalizes the agreement between the estimated mean belief from the PMF and the self-reported mean belief, the two belief measures elicited side-by-side from each employer (Section 1.7.2). The substantive findings of Sections 1.5 and 1.6 carry through under both checks.

1.7.1. Distributional Balance of Priors across Treatment Arms

The randomization balance check reported in Section 1.5.2 compares prior means across Treatment Arms using the regression-based balance test in Table 1.1: priors do not differ systematically across the six Treatment Arms, conditional on worker ethnicity assignment, and the Asian–Hispanic prior gap appears in every cell. A regression-based check on means is sensitive to differences in the first moment of the prior distribution but does not constrain higher-order distributional features. If, for example, the Asian-arm prior distribution were more dispersed in Voluntary than in Exogenous20, the residual closure wedge documented in Section 1.5.3 could in principle be partly attributable to a pre-existing distributional asymmetry rather than to the realization-based-stopping channel.

I complement the K–W balance check (Section 1.5.2) with a multi-sample energy-distance test on the elicited prior PMFs within each worker-ethnicity assignment. Unlike the K–W test, which has power primarily against location alternatives, the energy-distance test has power against alternatives in any moment of the multivariate distribution. Table 1.12 reports the omnibus statistics: for Hispanic-arm employers, the six-arm energy distance is -0.86 with permutation $p = 0.54$, and for Asian-arm employers, -1.84 with $p = 0.62$. Both omnibus tests fail to reject distributional

equality across arms, indicating that prior belief distributions are equal not only in mean but also in their full 12-bin shape within each worker-ethnicity cell. The residual closure wedge documented in Section 1.5.3 therefore cannot be attributed to a pre-existing distributional asymmetry.

TABLE 1.12. Within-Group Distributional Balance of Priors across Treatment Arms

Comparison	Energy stat ($\mathcal{E}$)	Perm. p
Hispanic workers: 6-arm omnibus ($N = 1443$)	-0.8583	0.539
Asian workers: 6-arm omnibus ($N = 1463$)	-1.8416	0.621

Notes: Multi-sample energy-distance test (Székely–Rizzo 2013) on the elicited prior PMFs, computed separately for each worker-ethnicity assignment. The null hypothesis is that prior distributions are identical across the six paper Treatment Arms (Baseline, Voluntary, ExogenousMean, Exogenous20, Commitment, Match). U-statistic estimator; 1,000 permutations; seed = 20260507.

Beyond this within-group balance check, I test the parallel between-group question by computing the energy distance between the Asian-arm and Hispanic-arm prior PMF distributions within each Treatment Arm. Under the null of distributional equality, the permutation p -value equals the share of random reassignments of the worker ethnicity label that produce a larger energy distance than the observed value. Table 1.13 reports the results.

TABLE 1.13. Distributional Balance of Priors across Treatment Arms

Treatment arm	Energy stat ($\mathcal{E}$)	Perm. p
Baseline	3.5728***	0.001
Voluntary	3.9979***	0.003
ExogenousMean	3.7651***	0.003
Exogenous20	2.7963***	0.008
Commitment	3.2104***	0.003
Match	3.8885***	0.001

Notes: This table reports the energy distance between the Asian-arm and Hispanic-arm prior PMF distributions within each treatment arm, computed on the 12-bin prior PMF support. The permutation p -value is the share of 999 random reassignments of the worker ethnicity label (within each arm) that produce a larger energy distance than the observed value. Sample: analysis sample, prior PMFs only. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

Turning to the between-group comparison, the Asian–Hispanic prior energy distance is large and significant in every Treatment Arm: $\mathcal{E} = 3.998$ in Voluntary, $\mathcal{E} = 3.211$ in Commitment, $\mathcal{E} = 3.765$ in ExogenousMean, $\mathcal{E} = 2.796$ in Exogenous20, $\mathcal{E} = 3.573$ in Baseline, and $\mathcal{E} = 3.889$ in Match (all

with permutation $p \leq 0.008$). The pattern reproduces the central prior-bias finding of Section 1.5.2 at the distributional level: priors differ between Asian-arm and Hispanic-arm employers within every Treatment Arm, in a way that holds across the full 12-bin PMF support and not just the first moment. The magnitudes are roughly comparable across treatment arms, with no systematic relationship between an arm’s Asian-Hispanic prior energy distance and that arm’s residual posterior gap. Voluntary, where the residual posterior gap is the largest at 0.532, exhibits a prior energy distance of 3.998; Exogenous20, where the residual posterior gap closes to 0.107, exhibits a prior energy distance of 2.796. The differential closure pattern in Section 1.5.3 is not an artifact of differential prior distributions across Treatment Arms. The realization-based-stopping channel identified in Section 1.6.2 operates on prior distributions that are similar in distributional shape across the four sampling-based regimes.

1.7.2. Self-Reported Mean Belief: Additional Evidence from Alternative Belief Measure

The empirical results in Sections 1.5 and 1.6 use the estimated mean belief from the PMF as the primary belief measure, computed by summing each PMF cell’s midpoint weighted by its reported probability mass. Each employer also reports a separate self-reported mean belief, elicited as a text box response on the same scale, alongside the PMF elicitation. The two measures are companion elicitations of the same latent belief: under exact internal consistency, the estimated mean belief from the PMF and the self-reported mean belief should coincide. In practice they differ, both because employers may anchor the self-reported text box to round numbers and because the PMF and text box draw on slightly different cognitive representations of uncertainty. The estimated mean belief from the PMF is the primary measure throughout this paper; the self-reported mean belief is reported as a companion. This subsection shows that the two belief measures yield qualitatively similar results: the main-text coefficients reported in Sections 1.5 and 1.6 reproduce when the regressions are re-estimated with the self-reported mean belief as the dependent variable.

The main-text regression tables in Sections 1.5 and 1.6 (Tables 1.1, 1.2, 1.3, 1.7, and 1.11) report the self-reported mean belief in a companion column alongside the estimated mean belief from the PMF. The pattern across the two measures is consistent: the Asian-Hispanic posterior

gap is statistically significant under both measures at the same significance level, the rank ordering of the residual gaps across the four sampling-based regimes is preserved at the level of the global pattern, and the Match-vs-Voluntary interaction is statistically indistinguishable from zero under both measures. Two specific observations further strengthen the robustness of the main results.

First, the self-reported mean belief produces more strongly significant coefficients on Voluntary versus Commitment in the four-arm Corollary 1 regression (Table 1.3, Column 2): the Asian $\times$ Commitment interaction is -0.315 ($p < 0.10$) under the self-reported measure, compared to -0.204 (not statistically significant) under the estimated mean belief from the PMF. The forced-regime interactions ϕ_2 (Asian $\times$ ExogenousMean) and ϕ_3 (Asian $\times$ Exogenous20) become more strongly significant under the self-reported measure as well: $\phi_2 = -0.361$ ($p < 0.05$) and $\phi_3 = -0.445$ ($p < 0.01$), compared to -0.273 ($p < 0.10$) and -0.425 ($p < 0.01$) under the estimated mean belief from the PMF. The point estimates are similar in magnitude across the two measures; the self-reported measure produces tighter standard errors on the cross-arm comparisons, consistent with the text box eliciting a belief representation more directly tied to employers' downstream decisions.

Second, the strict rank ordering predicted by Corollary 1, Voluntary $>$ Commitment $>$ ExogenousMean $>$ Exogenous20, is preserved under the self-reported measure. The Voluntary residual gap is the omitted-baseline coefficient at 0.490; the Commitment residual gap is $0.490 - 0.315 = 0.175$; the ExogenousMean residual gap is $0.490 - 0.361 = 0.129$; and the Exogenous20 residual gap is $0.490 - 0.445 = 0.045$. Under the estimated mean belief from the PMF, the corresponding values are 0.532, 0.328, 0.259, and 0.107, and the rank ordering Voluntary $>$ Commitment $>$ ExogenousMean $>$ Exogenous20 holds identically. Point estimates differ slightly across measures, but the four-arm ordering is the same. The pattern is preserved under the self-reported measure: sign, extremes, and rank ordering hold without exception.

Table 1.14 reports the self-reported-measure version of the belief updating regression in Table 1.10. The prior coefficient α_2 is in the range 0.108–0.196 under the estimated mean belief from the PMF, and 0.136–0.166 under the self-reported mean belief; the sample-mean coefficient α_4 is in the range 0.574–0.728 under the primary measure and 0.696–0.814 under the self-reported measure. The substantive interpretation is unchanged: posterior beliefs in the forced-information regimes load on

both prior and sample mean, with the sample mean carrying the larger weight, and the relative magnitudes of α_2 and α_4 are preserved.

TABLE 1.14. Belief Updating in Forced-Information Regimes (Self-Reported Mean Belief)

	ExogMean H	ExogMean A	Exog20 H	Exog20 A	Pooled H	Pooled A
Prior (α_2)	0.137*** (0.037)	0.139*** (0.036)	0.166*** (0.039)	0.136*** (0.033)	0.149*** (0.038)	0.137*** (0.038)
Sample mean (α_4)	0.716*** (0.080)	0.766*** (0.054)	0.814*** (0.126)	0.696*** (0.087)	0.699*** (0.083)	0.765*** (0.057)
Prior $\times$ Exog20 (α_3)					0.010 (0.053)	-0.002 (0.049)
Sample mean $\times$ Exog20 (α_5)					0.145 (0.145)	-0.073 (0.100)
Exog20					-0.953 (1.067)	0.590 (0.807)
Observations	220	225	240	270	460	495
Employer Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports OLS estimates of Equation 1.27 with the self-reported mean belief as the dependent variable, replacing the estimated mean belief from the PMF used in Table 1.10. Columns are arranged identically to Table 1.10: Columns 1–2 use the ExogenousMean subsamples; Columns 3–4 the Exogenous20 subsamples; Columns 5–6 the pooled forced-information sample with Exogenous20 interactions. Sample: forced-information regimes, analysis sample. Employer controls include age, race/ethnicity, education level, and gender. Standard errors are heteroskedasticity-robust. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

Taken together, the two belief measures further support the substantive findings in Sections 1.5 and 1.6: the results do not depend on the choice of belief measure. The Asian-Hispanic prior and posterior gaps appear under both elicitations; the rank ordering of residual posterior gaps across the four sampling-based regimes is preserved; the Match-vs-Voluntary interaction is statistically indistinguishable from zero under both measures; and the belief updating coefficients in the forced-information regimes carry the same qualitative pattern. Within-regime ordering differences in the interior of the rank ordering, when they appear, are consistent with text box-anchoring measurement error in the self-reported elicitation rather than with substantive heterogeneity in beliefs across the two measures.

1.8. Conclusion

“We can be blind to the obvious, and we are also blind to our blindness.”

This paper shows that the mode through which individuals acquire information shapes belief formation and subsequent behavior. Using a theory-guided experiment in the context of hiring, I show that even when individuals receive similar amounts of information, the endogeneity of the learning process shapes how beliefs evolve. Employers begin with biased prior beliefs that favor Asian over Hispanic workers by about 1.682 points, and in the Voluntary treatment, where employers engage in sequential information search, they tend to stop once the running posterior has stayed close to the prior. Realization-based stopping leaves a residual Asian-Hispanic posterior gap of 0.537, which then produces a wage premium for Asian workers. In contrast, when information is acquired through exogenously assigned samples or pre-committed sample sizes, the residual posterior gap shrinks substantially and the wage premium is no longer statistically distinguishable from zero. These differences arise despite holding the amount of information constant, since Voluntary and ExogenousMean deliver the same average sample size by design.

These findings clarify why biased beliefs survive despite informative signals: not because decision makers ignore evidence, but because they control when and how to seek it. By identifying realization-based stopping as a key mechanism, the paper complements explanations centered on non-Bayesian updating, motivated reasoning, and cognitive imprecision in the encoding of signals.

These findings speak directly to policymakers and practitioners seeking to combat bias. Many existing efforts to reduce discrimination focus on exposure to information, but my results show that exposure alone is insufficient if people retain discretion over stopping; the remedy is to structure the learning process. In hiring, for instance, organizations can set fixed, pre-committed review budgets: evaluate a predetermined number of resumes and complete a standardized set of interview questions before any decisions. Moving from rolling, real-time decisions to batched evaluation windows with a common decision date further limits opportunistic stopping. Similar logic applies to promotion and performance review, where fixed-horizon evidence packets, pre-specified time windows, and standardized requirements can replace ad hoc “stop-when-satisfied” sampling. And in evidence-generating contexts such as program pilots and clinical research, pre-specified sample sizes or stopping boundaries prevent realization-based stopping from biasing conclusions. These

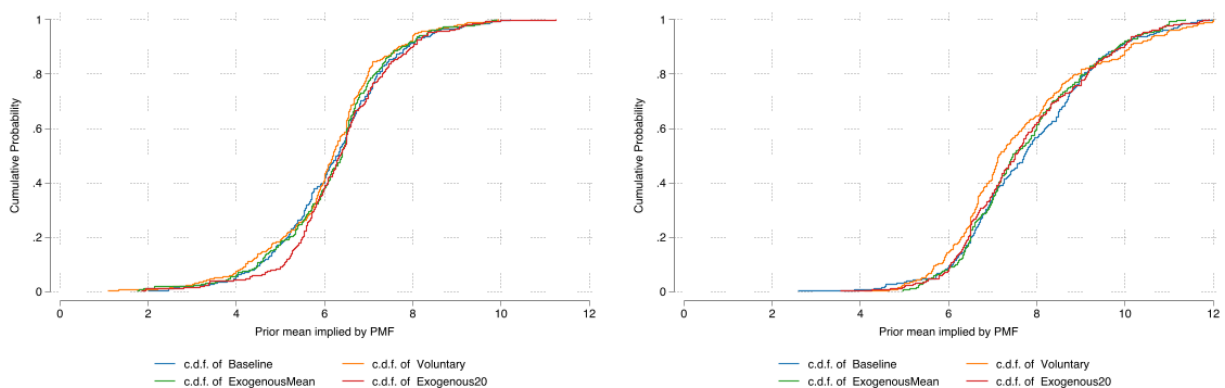
designs improve belief accuracy and decision quality without changing information content, and are inexpensive to implement. The wage decomposition adds one qualification: in the Commitment regime, the prior receives substantively higher weight on wage offers than in the optional-stopping regimes. This is consistent with pre-commitment to a sample size anchoring the wage decision on the original prior even after the committed sample is drawn. Pre-committed sampling rules close the realization-based-stopping channel and reduce the posterior gap, but they do not by themselves correct the prior bias that survives in the wage decision.

This paper also lays the foundation for future work exploring belief formation in value-neutral settings. In an upcoming project, I study how endogenous and exogenous modes of information acquisition affect beliefs in value-neutral domains. By removing potential motivated reasoning tied to group identity, this follow-up study will help isolate the pure cognitive channels through which information structure shapes belief accuracy. Together, these lines of research clarify how people learn about the world, and how the design of learning environments can support better judgments and more equitable outcomes.

1.9. Sample Composition and Balance

This appendix documents the demographic composition of the analysis sample referenced in Section 1.5.1.

FIGURE 1.10. Empirical Distributions of Prior Belief by Information Condition



A. Employers assigned Hispanic workers

B. Employers assigned Asian workers

Notes: Empirical cumulative distribution functions of prior beliefs about the average worker score, plotted separately for each information condition within each worker ethnicity treatment arm. $N = 1,056$ employers assigned the Hispanic worker group (left panel) and $N = 1,089$ assigned the Asian worker group (right panel).

TABLE 1.15. Sample Composition

	N	%
<i>Panel A: Employer race</i>		
White	2,207	75.9
Black	380	13.1
Native American	48	1.7
Asian	264	9.1
Pacific Islander	7	0.2
<i>Panel B: Employer Hispanic indicator</i>		
Not Hispanic	2,659	91.5
Hispanic	247	8.5
<i>Panel C: Employer gender</i>		
Male	1,099	37.8
Female	1,728	59.5
Non-binary	74	2.5
Other	5	0.2
<i>Panel D: Employer age group</i>		
18-24 years old	323	11.1
25-34 years old	903	31.1
35-44 years old	721	24.8
45-54 years old	524	18.0
55-64 years old	299	10.3
65+ years old	136	4.7
<i>Panel E: Treatment cell N (employer count)</i>		
Baseline	496	17.1
Voluntary	503	17.3
ExogenousMean	445	15.3
Exogenous20	510	17.5
Commit	509	17.5
RandomSample	443	15.2
Total	2,906	100.0

Notes: This table reports the demographic composition of the analysis sample. Panel A reports employer race; Panel B reports the employer Hispanic indicator; Panel C reports employer gender; Panel D reports employer age group; Panel E reports the employer count in each Treatment Arm. The Match arm is labeled RandomSample in the underlying analysis-pipeline output. Sample: analysis sample, $N = 2,906$ employers.

TABLE 1.16. Employer Demographics by Information Condition, Hispanic Workers

	Baseline	Voluntary	ExogenousMean	Exogenous20
Age	40.350 (13.270)	40.619 (13.577)	36.277 (12.156)	39.593 (13.333)
Male	0.357 (0.480)	0.424 (0.495)	0.311 (0.464)	0.424 (0.495)
Female	0.618 (0.487)	0.554 (0.498)	0.651 (0.478)	0.558 (0.498)
Non-Binary	0.021 (0.145)	0.022 (0.148)	0.034 (0.181)	0.019 (0.135)
White	0.743 (0.438)	0.773 (0.420)	0.710 (0.455)	0.755 (0.431)
Black	0.132 (0.339)	0.119 (0.324)	0.143 (0.351)	0.141 (0.349)
Asian	0.096 (0.296)	0.082 (0.275)	0.130 (0.337)	0.089 (0.286)
Hispanic	0.093 (0.291)	0.086 (0.280)	0.101 (0.302)	0.108 (0.311)
Education	4.532 (1.244)	4.535 (1.250)	4.466 (1.262)	4.487 (1.357)
Current Student	0.190 (0.394)	0.135 (0.342)	0.217 (0.414)	0.173 (0.379)
Employed Full-Time	0.336 (0.473)	0.353 (0.479)	0.294 (0.457)	0.320 (0.467)
Employed Part-Time	0.100 (0.301)	0.145 (0.353)	0.097 (0.296)	0.093 (0.291)
Unemployed / Not in paid work	0.089 (0.286)	0.119 (0.324)	0.118 (0.323)	0.126 (0.333)
Born in US	0.939 (0.239)	0.918 (0.274)	0.937 (0.244)	0.968 (0.176)
Subjects	280	269	238	269

Notes: This table reports the mean of each demographic variable across information treatments (Baseline, Voluntary, Exogenous Mean, Exogenous 20) for employers randomly assigned the Hispanic worker group, with the corresponding standard deviation in parentheses. Age is a numeric variable with integer values; the remaining demographic variables are binary indicators. Age, student status, employment status, and country of birth were provided by Prolific. Gender and ethnicity were recoded from categorical variables. Education level is a categorical variable with values 0 (Prefer not to say) through 6 (graduate or professional degree).

TABLE 1.17. Employer Demographics by Information Condition, Asian Workers

	Baseline	Voluntary	ExogenousMean	Exogenous20
Age	40.282 (13.106)	40.267 (13.701)	38.806 (12.980)	39.189 (13.408)
Male	0.387 (0.488)	0.408 (0.492)	0.393 (0.489)	0.377 (0.486)
Female	0.580 (0.494)	0.563 (0.497)	0.579 (0.495)	0.609 (0.489)
Non-Binary	0.030 (0.170)	0.028 (0.166)	0.028 (0.166)	0.014 (0.117)
White	0.755 (0.431)	0.757 (0.430)	0.794 (0.406)	0.723 (0.448)
Black	0.160 (0.367)	0.137 (0.345)	0.105 (0.308)	0.163 (0.370)
Asian	0.067 (0.250)	0.092 (0.289)	0.085 (0.279)	0.104 (0.306)
Hispanic	0.063 (0.244)	0.063 (0.244)	0.077 (0.267)	0.093 (0.292)
Education	4.361 (1.390)	4.396 (1.312)	4.433 (1.320)	4.581 (1.182)
Current Student	0.195 (0.397)	0.188 (0.392)	0.213 (0.411)	0.214 (0.411)
Employed Full-Time	0.342 (0.475)	0.313 (0.465)	0.344 (0.476)	0.388 (0.488)
Employed Part-Time	0.112 (0.315)	0.106 (0.308)	0.093 (0.291)	0.097 (0.296)
Unemployed / Not in paid work	0.112 (0.315)	0.155 (0.362)	0.097 (0.297)	0.107 (0.310)
Born in US	0.964 (0.187)	0.926 (0.262)	0.964 (0.188)	0.953 (0.213)
Subjects	269	284	247	289

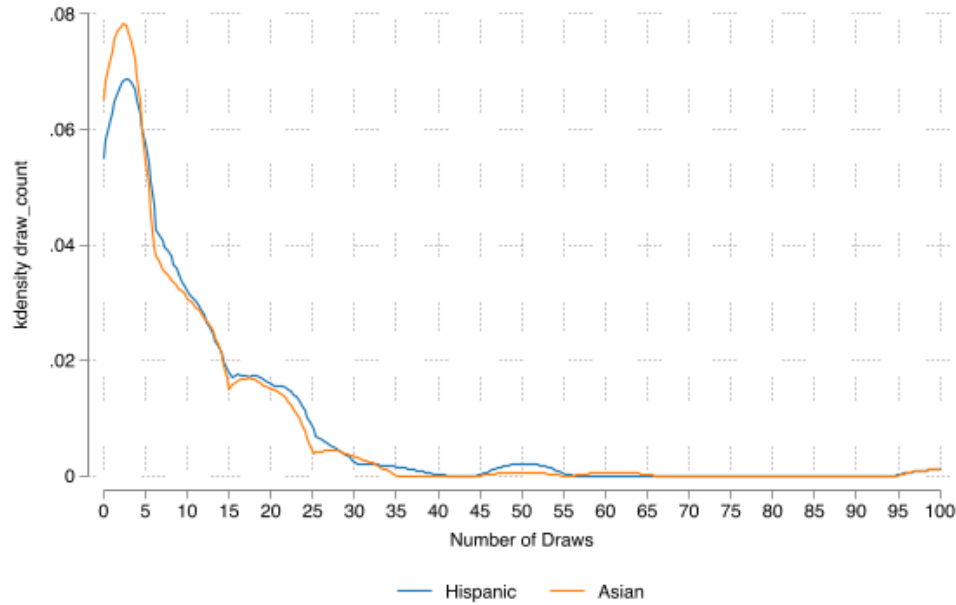
Notes: This table reports the mean of each demographic variable across information treatments (Baseline, Voluntary, Exogenous Mean, Exogenous 20) for employers randomly assigned the Asian worker group, with the corresponding standard deviation in parentheses. Age is a numeric variable with integer values; the remaining demographic variables are binary indicators. Age, student status, employment status, and country of birth were provided by Prolific. Gender and ethnicity were recoded from categorical variables. Education level is a categorical variable with values 0 (Prefer not to say) through 6 (graduate or professional degree).

TABLE 1.18. Within-Voluntary Posterior Accuracy

	Sq Err Post (PMF)	Sq Err Post (Self-reported)
Asian Worker Group	0.353 (0.346)	0.666 (0.419)
Constant	1.547*** (0.247)	2.113*** (0.299)
Hispanic Mean Sq Err	1.547	2.113
Asian Mean Sq Err	1.900	2.778
Observations	503	503

Notes: This table reports OLS estimates of the squared deviation of posterior beliefs from the true average productivity, regressed on the Asian worker group indicator, restricted to Voluntary employers. The estimated mean belief from the PMF (Column 1) is the primary measure; the self-reported mean belief (Column 2) is reported as a companion measure. Sample: Voluntary employers in the analysis sample, $N = 503$. Standard errors are heteroskedasticity-robust. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

FIGURE 1.11. Number of Draws in Voluntary, by Worker Group



Notes: Kernel density of the number of draws in the Voluntary treatment, plotted separately for employers assigned the Asian and Hispanic worker groups.

1.10. Wage Decomposition Panels

This appendix reports the regime-by-regime wage decompositions that underlie the saturated specification in Section 1.6.3. Each regime appears in two panels: Panel A nets the wage gap against the posterior estimated from the prior PMF; Panel B nets the wage gap against the self-reported posterior. Within each panel, the four columns correspond to a nested controls schedule, beginning with no controls and adding the posterior, then worker characteristics, and finally employer characteristics.

The pattern across regimes is consistent with the main-text finding that wage discrimination is belief-driven. In Voluntary (Tables 1.19 and 1.20) the raw Asian Worker Group coefficient is 0.283 ($p < 0.05$) and falls to between 0.187 and 0.224 once the posterior is netted, statistically indistinguishable from zero in the intermediate-controls specifications and only marginally significant in the saturated specification of either measure. In Commitment (Tables 1.25 and 1.26) the raw coefficient of 0.401 ($p < 0.01$) declines to roughly 0.29–0.35 across specifications and remains statistically significant ($p < 0.05$); when the prior is also added in the saturated specification reported in the main text, the residual coefficient becomes statistically indistinguishable from zero.

In ExogenousMean (Tables 1.21 and 1.22) and Exogenous20 (Tables 1.23 and 1.24), the residual Asian Worker Group coefficients are negative and reach -0.362 and -0.269 in the saturated main-text specification. Without a coherent mechanism story for why forced sampling would produce a negative Asian residual, I report these panels but do not flag them as a substantive finding. The pattern is more likely a fluke than a separate channel of discrimination operating in the forced-information regimes.

TABLE 1.19. Wage Decomposition: Voluntary, Panel A (Posterior from PMF)

	(1)	(2)	(3)	(4)
Asian Worker Group	0.283** (0.125)	0.187 (0.125)	0.192 (0.125)	0.224* (0.126)
Posterior (PMF)		0.179*** (0.052)	0.179*** (0.052)	0.183*** (0.052)
Controls	None	+ posterior	+ worker	+ employer
Observations	5030	5030	5030	5030
Clustered SE	Yes	Yes	Yes	Yes

Notes: OLS estimates with the wage offer as the dependent variable, restricted to the Voluntary regime. Column 1 reports the raw Asian Worker Group coefficient; Columns 2–4 add the posterior estimated mean belief from the PMF, then worker controls, then employer controls. Worker controls include age category, gender, race of the name, and favorite color. Employer controls include age, race/ethnicity, education level, and gender. Sample: long-format wage panel for Voluntary employers, analysis sample. Standard errors are clustered at the employer level. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

TABLE 1.20. Wage Decomposition: Voluntary, Panel B (Self-Reported Posterior)

	(1)	(2)	(3)	(4)
Asian Worker Group	0.283** (0.125)	0.204* (0.123)	0.208* (0.124)	0.231* (0.124)
Posterior (Self-reported)		0.168*** (0.046)	0.168*** (0.046)	0.189*** (0.046)
Controls	None	+ posterior	+ worker	+ employer
Observations	5030	5030	5030	5030
Clustered SE	Yes	Yes	Yes	Yes

Notes: Same specifications as Table 1.19, with the self-reported mean posterior in place of the PMF-derived measure. Sample: long-format wage panel for Voluntary employers, analysis sample. Standard errors are clustered at the employer level. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

TABLE 1.21. Wage Decomposition: ExogenousMean, Panel A (Posterior from PMF)

	(1)	(2)	(3)	(4)
Asian Worker Group	-0.144 (0.137)	-0.229* (0.134)	-0.223* (0.133)	-0.261* (0.133)
Posterior (PMF)		0.352*** (0.082)	0.351*** (0.081)	0.373*** (0.082)
Controls	None	+ posterior	+ worker	+ employer
Observations	4450	4450	4450	4450
Clustered SE	Yes	Yes	Yes	Yes

Notes: OLS estimates with the wage offer as the dependent variable, restricted to the ExogenousMean regime. Column 1 reports the raw Asian Worker Group coefficient; Columns 2–4 add the posterior estimated mean belief from the PMF, then worker controls, then employer controls. Worker controls include age category, gender, race of the name, and favorite color. Employer controls include age, race/ethnicity, education level, and gender. Sample: long-format wage panel for ExogenousMean employers, analysis sample. Standard errors are clustered at the employer level. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

TABLE 1.22. Wage Decomposition: ExogenousMean, Panel B (Self-Reported Posterior)

	(1)	(2)	(3)	(4)
Asian Worker Group	-0.144 (0.137)	-0.164 (0.135)	-0.159 (0.135)	-0.190 (0.136)
Posterior (Self-reported)		0.192*** (0.064)	0.192*** (0.064)	0.220*** (0.063)
Controls	None	+ posterior	+ worker	+ employer
Observations	4450	4450	4450	4450
Clustered SE	Yes	Yes	Yes	Yes

Notes: Same specifications as Table 1.21, with the self-reported mean posterior in place of the PMF-derived measure. Sample: long-format wage panel for ExogenousMean employers, analysis sample. Standard errors are clustered at the employer level. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

TABLE 1.23. Wage Decomposition: Exogenous20, Panel A (Posterior from PMF)

	(1)	(2)	(3)	(4)
Asian Worker Group	-0.113 (0.134)	-0.151 (0.132)	-0.158 (0.131)	-0.187 (0.133)
Posterior (PMF)		0.344*** (0.072)	0.345*** (0.072)	0.374*** (0.070)
Controls	None	+ posterior	+ worker	+ employer
Observations	5100	5100	5100	5100
Clustered SE	Yes	Yes	Yes	Yes

Notes: OLS estimates with the wage offer as the dependent variable, restricted to the Exogenous20 regime. Column 1 reports the raw Asian Worker Group coefficient; Columns 2–4 add the posterior estimated mean belief from the PMF, then worker controls, then employer controls. Worker controls include age category, gender, race of the name, and favorite color. Employer controls include age, race/ethnicity, education level, and gender. Sample: long-format wage panel for Exogenous20 employers, analysis sample. Standard errors are clustered at the employer level. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

TABLE 1.24. Wage Decomposition: Exogenous20, Panel B (Self-Reported Posterior)

	(1)	(2)	(3)	(4)
Asian Worker Group	-0.113 (0.134)	-0.126 (0.132)	-0.132 (0.131)	-0.152 (0.134)
Posterior (Self-reported)		0.261*** (0.079)	0.261*** (0.080)	0.262*** (0.081)
Controls	None	+ posterior	+ worker	+ employer
Observations	5100	5100	5100	5100
Clustered SE	Yes	Yes	Yes	Yes

Notes: Same specifications as Table 1.23, with the self-reported mean posterior in place of the PMF-derived measure. Sample: long-format wage panel for Exogenous20 employers, analysis sample. Standard errors are clustered at the employer level. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

TABLE 1.25. Wage Decomposition: Commitment, Panel A (Posterior from PMF)

	(1)	(2)	(3)	(4)
Asian Worker Group	0.401*** (0.143)	0.297** (0.138)	0.298** (0.139)	0.293** (0.138)
Posterior (PMF)		0.310*** (0.063)	0.309*** (0.063)	0.303*** (0.064)
Controls	None	+ posterior	+ worker	+ employer
Observations	5090	5090	5090	5090
Clustered SE	Yes	Yes	Yes	Yes

Notes: OLS estimates with the wage offer as the dependent variable, restricted to the Commitment regime. Column 1 reports the raw Asian Worker Group coefficient; Columns 2–4 add the posterior estimated mean belief from the PMF, then worker controls, then employer controls. Worker controls include age category, gender, race of the name, and favorite color. Employer controls include age, race/ethnicity, education level, and gender. Sample: long-format wage panel for Commitment employers, analysis sample. Standard errors are clustered at the employer level. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

TABLE 1.26. Wage Decomposition: Commitment, Panel B (Self-Reported Posterior)

	(1)	(2)	(3)	(4)
Asian Worker Group	0.401*** (0.143)	0.348** (0.139)	0.352** (0.140)	0.349** (0.137)
Posterior (Self-reported)		0.256*** (0.052)	0.256*** (0.052)	0.264*** (0.052)
Controls	None	+ posterior	+ worker	+ employer
Observations	5090	5090	5090	5090
Clustered SE	Yes	Yes	Yes	Yes

Notes: Same specifications as Table 1.25, with the self-reported mean posterior in place of the PMF-derived measure. Sample: long-format wage panel for Commitment employers, analysis sample. Standard errors are clustered at the employer level. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

1.11. Robustness Appendix Tables

This appendix collects three supplementary tables referenced in Section 1.6.2. Table 1.27 reports the Bayesian-updated hazard track. Table 1.28 reports the forced-regime posterior accuracy regression that supports the realization-based-stopping interpretation. A third table, Table 1.29, reports the standalone Commitment-versus-Voluntary residual gap regression, subsumed by the four-arm Corollary 1 synthesis but reported separately.

Bayesian-Updated Hazard Track

The within-Voluntary hazard estimates in Section 1.6.2 construct the running posterior μ_t as the running mean of realized signals. An alternative construction updates the posterior in Bayesian fashion, weighting the prior mean and the realized sample mean by their respective precisions. Implementing the Bayesian update requires a value for the prior precision parameter N_0 , which scales the weight that the prior receives relative to the signals. Because N_0 is not separately identified from the data, I report the hazard regression for a range of values: $N_0 \in \{0, 1, 5, 10\}$, where $N_0 = 0$ recovers the running-mean specification as a limit and larger values place increasing weight on the prior in the running posterior.

Table 1.27 reports the four estimates. The coefficient on $|\mu_t - \mu_0|$ is positive and significant at $N_0 = 0$ ($\hat{\beta} = 0.121$, $p < 0.01$), declines to a small and statistically insignificant value at $N_0 = 1$ ($\hat{\beta} = 0.016$), and turns negative and significant at $N_0 \in \{5, 10\}$. The round-number coefficients on $\mathbf{1}[t = 10]$ and $\mathbf{1}[t = 20]$ are stable across N_0 values, as is the Asian Worker Group coefficient. The sign-flip on $|\mu_t - \mu_0|$ across N_0 is consistent with a boundary condition on the within-Voluntary identification: at low N_0 the running posterior responds primarily to the realized signals and the cognitive-consistency signature is recovered; at high N_0 the running posterior is anchored on the prior, $|\mu_t - \mu_0|$ varies less across draws, and the signature collapses. The running-mean specification reported in the main text avoids this dependence on the prior-precision parameter and produces a positive, significant cognitive-consistency signature.

TABLE 1.27. Hazard Model with Bayesian-Updated Running Posterior, Sensitivity to Prior Precision

	$N_0 = 0$	$N_0 = 1$	$N_0 = 5$	$N_0 = 10$
$ \mu_t - \mu_0 $	0.121*** (0.042)	0.016 (0.055)	-0.184** (0.076)	-0.304*** (0.095)
Asian Worker Group	0.184 (0.133)	0.178 (0.137)	0.157 (0.137)	0.147 (0.136)
$t = 10$	1.217*** (0.178)	1.188*** (0.176)	1.215*** (0.176)	1.226*** (0.176)
$t = 20$	2.236*** (0.230)	2.176*** (0.229)	2.204*** (0.234)	2.240*** (0.237)
Observations (draws)	4535	4535	4535	4535
Clustered SE	Yes	Yes	Yes	Yes

Notes: Logit estimates of Equation 1.25 with the running posterior μ_t constructed as a Bayesian update of the prior mean and the realized sample mean, with the prior receiving weight proportional to N_0 and the signals receiving weight proportional to the number of draws so far. Each column corresponds to a different value of the prior precision parameter N_0 . Sample: draw-level panel for Voluntary employers, analysis sample, $N = 4,535$ draws. Standard errors are clustered at the employer level. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

Forced-Regime Posterior Accuracy

Table 1.28 reports an OLS regression of squared posterior error on the Asian Worker Group indicator and treatment-arm indicators, restricted to the three sampling regimes (Voluntary, ExogenousMean, Exogenous20). Squared posterior error is constructed as $(\mu_{post} - \theta)^2$, with the true average productivity $\theta = 7.02$ and the posterior measured from the PMF in Column 1 and from the self-reported text box in Column 2. Voluntary is the omitted reference cell. The ExogenousMean and Exogenous20 main effects are negative and significant in both columns (-0.687 and -0.924 in Column 1, both $p < 0.01$), indicating that forced sampling reduces squared posterior error relative to Voluntary. The Asian $\times$ ExogenousMean and Asian $\times$ Exogenous20 interactions are not statistically distinguishable from zero in Column 1, indicating that the accuracy improvement is not differential across worker ethnicity assignments in the PMF-derived measure. The pattern supports the realization-based-stopping interpretation: forced sampling closes the accuracy gap by removing endogenous stopping rather than by altering how Asian-arm and Hispanic-arm employers process information.

Commitment versus Voluntary

Table 1.29 reports a standalone two-arm regression of posterior beliefs on the Asian Worker Group indicator, the Commitment indicator, and their interaction, restricted to the Voluntary and Commitment regimes. The Voluntary residual gap is 0.526 ($p < 0.01$) in Column 1; the Commitment-versus-Voluntary main effect is statistically indistinguishable from zero; and the Asian $\times$ Commitment interaction is -0.195 ($SE = 0.155$, not statistically significant), implying a Commitment residual gap of approximately 0.331 . The pattern is fully consistent with the four-arm Corollary 1 synthesis reported in Section 1.6.2, which presents the strict ordering Voluntary $>$ Commitment $>$ ExogenousMean $>$ Exogenous20 in point estimates with direction-only Wald support.

1.11.1. Survey Interface Screenshots

TABLE 1.28. Forced-Regime Posterior Accuracy

	Sq Err Post (PMF)	Sq Err Post (Self-reported)
Asian Worker Group	0.343 (0.222)	0.640** (0.324)
ExogenousMean	-0.687*** (0.233)	-0.725** (0.339)
Exogenous20	-0.924*** (0.226)	-0.724** (0.329)
Asian $\times$ ExogenousMean	-0.183 (0.325)	-0.719 (0.473)
Asian $\times$ Exogenous20	-0.190 (0.314)	-1.005** (0.457)
Observations	1458	1458
Employer Controls	Yes	Yes

Notes: OLS estimates with the squared deviation of posterior beliefs from the true average productivity ($\theta = 7.02$) as the dependent variable, restricted to the three sampling regimes (Voluntary, ExogenousMean, Exogenous20). Column 1 uses the estimated mean belief from the PMF; Column 2 uses the self-reported mean belief. Voluntary is the omitted reference cell. Sample: analysis sample, $N = 1,458$ employers in the three sampling regimes. Employer controls include age, race/ethnicity, education level, and gender. Standard errors are heteroskedasticity-robust. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

TABLE 1.29. Posterior Belief Gap: Voluntary versus Commitment

	Posterior (PMF)	Posterior (Self-reported)
Asian Worker Group	0.526*** (0.109)	0.483*** (0.133)
Commitment (vs. Voluntary)	0.023 (0.109)	0.083 (0.133)
Asian $\times$ Commitment	-0.195 (0.155)	-0.327* (0.188)
Observations	1012	1012
Employer Controls	Yes	Yes

Notes: OLS estimates with the posterior mean belief as the dependent variable, restricted to the Voluntary and Commitment regimes. Column 1 uses the estimated mean belief from the PMF; Column 2 uses the self-reported mean belief. Voluntary is the omitted reference cell. The Asian $\times$ Commitment interaction captures the change in the Asian-Hispanic posterior gap moving from Voluntary to Commitment. Sample: Voluntary and Commitment regimes, analysis sample, $N = 1,012$ employers. Employer controls include age, race/ethnicity, education level, and gender. Standard errors are heteroskedasticity-robust. * ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

FIGURE 1.12. Elicitation for Belief Distribution

Out of the 100 Hispanic workers, how many do you think have a score of... (the numbers should add up to 100)

1?	0
2?	0
3?	0
4?	0
5?	0
6?	0
7?	0
8?	0
9?	0
10?	0
11?	0
12?	0
Total	0

FIGURE 1.13. Elicitation for Belief on Average Worker Score

What do you think is the **average score** of these workers? (out of 12)

FIGURE 1.14. Sampling Interface

Press the button below to draw workers.

Draw a Worker

Worker 1 score: between 7 to 9

Worker 2 score: between 7 to 9

Worker 3 score: between 10 to 12

FIGURE 1.15. Example Worker Profiles

A. Example Hispanic Profile

	
Nickname:	Marcos
Gender:	Male
Country:	United States
Age:	18-45
Favorite color:	purple

B. Example Asian Profile

	
Nickname:	Tian
Gender:	Male
Country:	United States
Age:	18-45
Favorite color:	purple

FIGURE 1.16. Wage Offer Interface

What is your wage offer to this worker? (any integer from 0 to 12)

1.12. Theoretical Framework Proofs

1.12.1. Proof of Proposition 1 (Commitment Optimum)

Step 1: Differentiate σ_n^2 in n .

From (1.4),

$$\sigma_n^2 = \frac{\sigma_0^2 \sigma_d^2}{n \sigma_0^2 + \sigma_d^2}.$$

Treating n as continuous,

$$\frac{d\sigma_n^2}{dn} = -\frac{\sigma_0^2 \sigma_d^2 \cdot \sigma_0^2}{(n \sigma_0^2 + \sigma_d^2)^2} = -\frac{\sigma_0^4 \sigma_d^2}{(n \sigma_0^2 + \sigma_d^2)^2} < 0.$$

Step 2: First-order condition.

The Commitment objective is $V^{\text{stop}}(n) = -\sigma_n^2 - cn$, with first-order condition

$$-\frac{d\sigma_n^2}{dn} = c \iff \frac{\sigma_0^4 \sigma_d^2}{(n \sigma_0^2 + \sigma_d^2)^2} = c. \quad (1.28)$$

Step 3: Solve for n^* .

Taking square roots (the square is positive on both sides):

$$\frac{\sigma_0^2 \sigma_d}{n \sigma_0^2 + \sigma_d^2} = \sqrt{c} \implies n \sigma_0^2 + \sigma_d^2 = \frac{\sigma_0^2 \sigma_d}{\sqrt{c}}.$$

Solving for n :

$$n^* = \frac{1}{\sigma_0^2} \left[\frac{\sigma_0^2 \sigma_d}{\sqrt{c}} - \sigma_d^2 \right] = \frac{\sigma_d}{\sqrt{c}} - \frac{\sigma_d^2}{\sigma_0^2} = \frac{\sigma_d}{\sigma_0^2} \left[\frac{\sigma_0^2}{\sqrt{c}} - \sigma_d \right],$$

which gives (1.8). The solution is interior whenever $\sigma_d/\sqrt{c} > \sigma_d^2/\sigma_0^2$, i.e., $\sigma_0^2 > \sigma_d \sqrt{c}$, equivalently $\sigma_0^2/\sqrt{c} > \sigma_d$.

Step 4: Second-order condition (interior maximum).

From Step 1, $\frac{d\sigma_n^2}{dn} = -\sigma_0^4 \sigma_d^2 \cdot (n\sigma_0^2 + \sigma_d^2)^{-2}$. Differentiating once more in n via the chain rule (note that $\frac{d}{dn}(n\sigma_0^2 + \sigma_d^2)^{-2} = -2(n\sigma_0^2 + \sigma_d^2)^{-3} \cdot \sigma_0^2$):

$$\frac{d^2\sigma_n^2}{dn^2} = -\sigma_0^4 \sigma_d^2 \cdot \left[-2\sigma_0^2 (n\sigma_0^2 + \sigma_d^2)^{-3} \right] = +\frac{2\sigma_0^6 \sigma_d^2}{(n\sigma_0^2 + \sigma_d^2)^3} > 0.$$

So σ_n^2 is convex in n , hence $V^{\text{stop}}(n) = -\sigma_n^2 - cn$ is concave in n (as the negative of a convex function plus a linear function). The critical point n^* is therefore a global maximum.

Step 5: Value at the optimum.

Substituting $n = n^*$ into $V^{\text{stop}}(n) = -\sigma_n^2 - cn$. At the optimum, $n^* \sigma_0^2 + \sigma_d^2 = \sigma_0^2 \sigma_d / \sqrt{c}$, so

$$\sigma_{n^*}^2 = \frac{\sigma_0^2 \sigma_d^2}{\sigma_0^2 \sigma_d / \sqrt{c}} = \sigma_d \sqrt{c}.$$

Therefore

$$V^{\text{stop}}(n^*) = -\sigma_d \sqrt{c} - cn^* = -\sigma_d \sqrt{c} - c \left(\frac{\sigma_d}{\sqrt{c}} - \frac{\sigma_d^2}{\sigma_0^2} \right) = -2\sigma_d \sqrt{c} + \frac{c\sigma_d^2}{\sigma_0^2},$$

which is (1.9).

Step 6: Comparative statics.

From (1.8), $n^* = \sigma_d / \sqrt{c} - \sigma_d^2 / \sigma_0^2$.

- $\partial n^* / \partial \sigma_0^2 = +\sigma_d^2 (\sigma_0^2)^{-2} > 0$.
- $\partial n^* / \partial c = -\frac{\sigma_d}{2c^{3/2}} < 0$.
- $\partial n^* / \partial \sigma_d^2$: rewriting $n^* = \sigma_d / \sqrt{c} - \sigma_d^2 / \sigma_0^2$ as a function of σ_d , $\partial n^* / \partial \sigma_d = 1/\sqrt{c} - 2\sigma_d / \sigma_0^2$. So $\partial n^* / \partial \sigma_d > 0 \iff \sigma_d < \sigma_0^2 / (2\sqrt{c})$. Since by Step 3 $n^* > 0$ requires $\sigma_d < \sigma_0^2 / \sqrt{c}$, the relevant range can include both signs but typically (with calibrated parameters) σ_d is small enough that the derivative is positive.

■

1.12.2. Proof of Proposition 2 (Voluntary $\equiv$ Commitment)

We show by backward induction on the Bellman (1.10) that the value function $V_n(\mu)$ is independent of μ at every n , and that the optimal stopping rule under Voluntary search is the deterministic threshold n^* (the same as Commitment).

Step 1: Base case.

At the terminal date T , the employer must stop: $V_T(\mu) = -\sigma_T^2$. This is constant in μ (does not depend on the realized posterior mean).

Step 2: Induction step.

Suppose $V_{n+1}(\mu') = \tilde{V}_{n+1}$ is constant in μ' for some date $n+1 \leq T$. Then the continuation value at n is

$$C_n = \mathbb{E}[V_{n+1}(\mu_{n+1}) \mid \mu_n] - c = \tilde{V}_{n+1} - c,$$

which is also constant in μ_n . The stopping value is $-\sigma_n^2$, also constant in μ_n . So

$$V_n(\mu_n) = \max\{-\sigma_n^2, \tilde{V}_{n+1} - c\}$$

is constant in μ_n . This completes the induction.

Step 3: Optimal stopping rule is n -only.

Since V_n depends on n but not on μ_n , the decision “stop or continue” is a function of n alone. The optimal *Voluntary* stopping rule therefore takes the form $\tau^{\text{vol}} = \min\{n : -\sigma_n^2 \geq \tilde{V}_{n+1} - c\}$, i.e., a deterministic threshold on n .

Step 4: This deterministic threshold equals n^* .

The Commitment problem is to choose n to maximize $V^{\text{stop}}(n) = -\sigma_n^2 - cn$, with optimum n^* from Proposition 1. The Voluntary problem under the deterministic-threshold policy maximizes the same objective, since the path of cost cn accumulated from period 1 to τ equals $c\tau$. Hence $\tau_{\text{rat}}^{\text{vol}} = n^*$.

Step 5: Posterior distribution equivalence.

Conditional on stopping at any sample size n , the posterior $\mu_n \mid \theta$ is normal with mean $\omega_n \theta + (1 - \omega_n) \mu_0$ and variance $\omega_n^2 \sigma_d^2 / n$, since $\bar{s}_n \mid \theta \sim \mathcal{N}(\theta, \sigma_d^2 / n)$ and μ_n is linear in $\bar{s}_n$ with slope ω_n , so the conditional variance scales as ω_n^2 . Under the deterministic $n^* = n$ in both regimes, $\mu_{n^*} \mid \theta$ has the same conditional distribution under Voluntary and Commitment, and integrating over the prior $\theta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ gives the same unconditional distribution under both regimes. ■

1.12.3. Auxiliary: Posterior-mean evolution as a martingale

LEMMA 1 (Expected posterior equals current posterior). *The posterior-mean process $\{\mu_n\}_{n=0}^\infty$ is a martingale with respect to the information filtration: $\mathbb{E}[\mu_{n+1} \mid \mu_n] = \mu_n$, with conditional variance*

$$\text{Var}(\mu_{n+1} \mid \mu_n) = \omega_{n+1}^2 (\sigma_{n+1}^2 + \sigma_d^2). \quad (1.29)$$

PROOF. The posterior mean $\mu_n = \mathbb{E}[\theta \mid D_n]$ is a sufficient statistic for θ under conjugate normal-normal updating, so conditioning on μ_n is equivalent to conditioning on the full signal history D_n . By the tower property of conditional expectation,

$$\mathbb{E}[\mu_{n+1} \mid \mu_n] = \mathbb{E}[\mathbb{E}[\theta \mid D_{n+1}] \mid D_n] = \mathbb{E}[\theta \mid D_n] = \mu_n,$$

which is the martingale property. The recursive update $\mu_{n+1} = \mu_n + \omega_{n+1}(s_{n+1} - \mu_n)$ (which follows from substituting $(n+1)\bar{s}_{n+1} = n\bar{s}_n + s_{n+1}$ into the conjugate-update formulas) gives $\text{Var}(\mu_{n+1} \mid \mu_n) = \omega_{n+1}^2 \text{Var}(s_{n+1} - \mu_n \mid \mu_n)$, and the predictive variance of $s_{n+1} - \mu_n$ is $\sigma_{n+1}^2 + \sigma_d^2$ (posterior variance of θ at $n+1$ plus signal noise), giving (1.29). ■ □

1.12.4. Auxiliary: Unconditional variance identity $\text{Var}_0(\mu_t) = \sigma_0^2 - \sigma_t^2$

This identity is used in §1.3.4 to simplify the ex-ante Commitment objective under cognitive consistency. We work unconditionally (under the prior at $t = 0$).

Step 1: Decomposition of θ .

Under the prior, $\theta \sim \mathcal{N}(\mu_0, \sigma_0^2)$. The posterior mean at t is $\mu_t = \omega_t \bar{s}_t + (1 - \omega_t)\mu_0$.

Step 2: Apply law of total variance.

By the law of total variance (conditioning on θ),

$$\text{Var}_0(\mu_t) = \mathbb{E}_0[\text{Var}(\mu_t \mid \theta)] + \text{Var}_0(\mathbb{E}[\mu_t \mid \theta]).$$

Conditional on θ , $\bar{s}_t$ is normal with mean θ and variance σ_d^2/t . So $\text{Var}(\mu_t \mid \theta) = \omega_t^2 \sigma_d^2/t$, and $\mathbb{E}[\mu_t \mid \theta] = \omega_t \theta + (1 - \omega_t)\mu_0$.

Step 3: Substitute.

$$\begin{aligned} \text{Var}_0(\mu_t) &= \omega_t^2 \cdot \frac{\sigma_d^2}{t} + \omega_t^2 \text{Var}(\theta) \\ &= \omega_t^2 \left(\frac{\sigma_d^2}{t} + \sigma_0^2 \right). \end{aligned}$$

Using $\omega_t = t\sigma_0^2/(t\sigma_0^2 + \sigma_d^2)$:

$$\begin{aligned} \omega_t^2 \left(\frac{\sigma_d^2}{t} + \sigma_0^2 \right) &= \frac{t^2 \sigma_0^4}{(t\sigma_0^2 + \sigma_d^2)^2} \cdot \frac{t\sigma_0^2 + \sigma_d^2}{t} \\ &= \frac{t\sigma_0^4}{t\sigma_0^2 + \sigma_d^2}. \end{aligned}$$

Step 4: Simplify to $\sigma_0^2 - \sigma_t^2$.

We have $\sigma_t^2 = \sigma_0^2 \sigma_d^2 / (t\sigma_0^2 + \sigma_d^2)$, hence

$$\sigma_0^2 - \sigma_t^2 = \sigma_0^2 \left(1 - \frac{\sigma_d^2}{t\sigma_0^2 + \sigma_d^2} \right) = \frac{t\sigma_0^4}{t\sigma_0^2 + \sigma_d^2},$$

matching Step 3. Therefore

$$\text{Var}_0(\mu_t) = \sigma_0^2 - \sigma_t^2, \tag{1.30}$$

which is (1.12). ■

1.12.5. Proof of Proposition 3 (Commitment optimum under cognitive consistency)

We expand the Commitment objective and derive the marginal-value expression (1.14).

Step 1: Substitute $\text{Var}_0(\mu_t)$.

The Commitment objective (1.13) is

$$\tilde{V}^{\text{commit}}(\tau) = -\sigma_\tau^2 - \gamma \sum_{t=1}^{\tau} (\sigma_0^2 - \sigma_t^2) - c\tau,$$

using (1.30) for the second term.

Step 2: Marginal value of an additional signal.

Compute $\tilde{V}^{\text{commit}}(\tau + 1) - \tilde{V}^{\text{commit}}(\tau)$:

$$\begin{aligned} & [-\sigma_{\tau+1}^2 - \gamma \sum_{t=1}^{\tau+1} (\sigma_0^2 - \sigma_t^2) - c(\tau + 1)] - [-\sigma_\tau^2 - \gamma \sum_{t=1}^{\tau} (\sigma_0^2 - \sigma_t^2) - c\tau] \\ &= (\sigma_\tau^2 - \sigma_{\tau+1}^2) - \gamma(\sigma_0^2 - \sigma_{\tau+1}^2) - c, \end{aligned}$$

which is (1.14).

Step 3: $\tilde{n}_\gamma^* \leq n^*$.

The rational marginal value at τ is $(\sigma_\tau^2 - \sigma_{\tau+1}^2) - c$ (Proposition 1 and the proof of Proposition 2); this equals zero at $\tau = n^*$. The cognitive-consistency marginal value is the rational marginal value minus the non-negative term $\gamma(\sigma_0^2 - \sigma_{\tau+1}^2)$, so the cognitive-consistency marginal value is weakly less. The cognitive-consistency optimum $\tilde{n}_\gamma^*$ is therefore reached at a weakly smaller τ : $\tilde{n}_\gamma^* \leq n^*$.

Step 4: Boundary case ($\gamma = 0$).

At $\gamma = 0$, the cumulative belief-cost term is zero and (1.14) reduces to the rational marginal value $(\sigma_\tau^2 - \sigma_{\tau+1}^2) - c$. The optimum $\tilde{n}_0^* = n^*$. ■

1.12.6. Proof of Proposition 4 (Voluntary $\neq$ Commitment under cognitive consistency)

The argument proceeds in three steps. We show that the cognitive-consistency continuation value $\tilde{C}_t(\mu)$ is a non-trivial function of μ for any $\gamma > 0$ and any $t < T$, that this μ -dependence propagates to $\tilde{V}_t(\mu)$, and that the optimal Voluntary stopping rule $\tilde{\tau}_\gamma^{\text{vol}}$ is therefore not a deterministic function of t alone.

Step 1: Continuation value carries μ_t -dependence.

By the law of total variance applied to μ_{t+1} conditional on μ_t ,

$$\mathbb{E}\left[(\mu_{t+1} - \mu_0)^2 \mid \mu_t\right] = \text{Var}(\mu_{t+1} \mid \mu_t) + (\mathbb{E}[\mu_{t+1} \mid \mu_t] - \mu_0)^2 = \text{Var}(\mu_{t+1} \mid \mu_t) + (\mu_t - \mu_0)^2,$$

where the second equality uses the martingale property of Lemma 1: $\mathbb{E}[\mu_{t+1} \mid \mu_t] = \mu_t$. We record this as

$$\mathbb{E}\left[(\mu_{t+1} - \mu_0)^2 \mid \mu_t\right] = \text{Var}(\mu_{t+1} \mid \mu_t) + (\mu_t - \mu_0)^2. \quad (1.31)$$

The first term $\text{Var}(\mu_{t+1} \mid \mu_t) = \omega_{t+1}^2(\sigma_{t+1}^2 + \sigma_d^2)$ is a function of t alone. The second term $(\mu_t - \mu_0)^2$ is non-trivially a function of the realized μ_t . Substituting (1.31) into the cognitive-consistency continuation value:

$$\tilde{C}_t(\mu_t) = \mathbb{E}[\tilde{V}_{t+1}(\mu_{t+1}) \mid \mu_t] - \gamma \text{Var}(\mu_{t+1} \mid \mu_t) - \gamma (\mu_t - \mu_0)^2 - c.$$

The third term in this expression depends on μ_t through the location $(\mu_t - \mu_0)^2$, with coefficient $-\gamma < 0$ for $\gamma > 0$.

Step 2: $\tilde{V}_t(\mu)$ inherits μ -dependence by backward induction.

We argue by backward induction, propagating μ -dependence from the latest non-terminal date $t = T - 1$ back to $t < T - 1$. The terminal value $\tilde{V}_T(\mu) = -\sigma_T^2$ is constant in μ and serves only as the boundary condition for the recursion; the claim concerns $\tilde{V}_t$ for $t < T$, where the location-term coefficient $-\gamma$ in the continuation value of Step 1 is active.

Base case ($t = T - 1$). Substituting $\mu_t = \mu$ and the constant terminal value into the continuation-

value expression of Step 1 gives

$$\tilde{C}_{T-1}(\mu) = \mathbb{E}[\tilde{V}_T(\mu_T) \mid \mu_{T-1} = \mu] - \gamma \text{Var}(\mu_T \mid \mu_{T-1}) - \gamma(\mu - \mu_0)^2 - c.$$

Since $\tilde{V}_T$ is constant, the first term contributes a μ -independent constant; the second term is a function of $T - 1$ alone; the third term, $-\gamma(\mu - \mu_0)^2$, is a non-trivial function of μ . Hence $\tilde{C}_{T-1}$ depends on μ , and $\tilde{V}_{T-1}(\mu) = \max\{-\sigma_{T-1}^2, \tilde{C}_{T-1}(\mu)\}$ also depends on μ (the max of a constant and a μ -dependent function is μ -dependent on the set where the latter strictly dominates). The base case is established.

Inductive step (backward, from $t + 1$ to t). Fix $t < T - 1$ and assume $\tilde{V}_{t+1}(\mu)$ is a non-trivial function of μ (the inductive hypothesis at $t + 1$). Then

$$\tilde{C}_t(\mu) = \mathbb{E}[\tilde{V}_{t+1}(\mu_{t+1}) \mid \mu_t = \mu] - \gamma \text{Var}(\mu_{t+1} \mid \mu_t) - \gamma(\mu - \mu_0)^2 - c.$$

Even if the conditional expectation in the first term were identically constant in μ (which it generically is not under the inductive hypothesis), the location term $-\gamma(\mu - \mu_0)^2$ contributes μ -dependence directly to $\tilde{C}_t(\mu)$. Hence $\tilde{V}_t(\mu) = \max\{-\sigma_t^2, \tilde{C}_t(\mu)\}$ depends on μ . Stepping backward through $t = T - 2, T - 3, \dots, 1$ propagates the μ -dependence of $\tilde{V}_t$ to every $t < T$.

Step 3: Stopping rule is realization-conditional.

The optimal Voluntary stopping rule is

$$\tilde{\tau}_\gamma^{\text{vol}} = \min\{t : -\sigma_t^2 \geq \tilde{C}_t(\mu_t)\}.$$

Because $\tilde{C}_t(\mu_t)$ is a non-trivial function of μ_t (Step 1), the inequality $-\sigma_t^2 \geq \tilde{C}_t(\mu_t)$ holds at some realizations and fails at others. Concretely, fix any $\gamma > 0$ and any $t < T$ with $\tilde{C}_t$ continuous in μ at μ_0 . Choose two realized histories D_t, D'_t with $\mu_t(D_t)$ near μ_0 and $\mu_t(D'_t)$ far from μ_0 . Then the term $-\gamma(\mu_t - \mu_0)^2$ takes values ≈ 0 and $\ll 0$ respectively, so $\tilde{C}_t(D_t) > \tilde{C}_t(D'_t)$. There exist parameters such that $-\sigma_t^2$ falls between these two values: at D_t the employer continues ($\tilde{C}_t(D_t) > -\sigma_t^2$) and at D'_t she stops ($\tilde{C}_t(D'_t) < -\sigma_t^2$). The set of such realizations has positive probability under the prior

$\theta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ with $\sigma_0^2 > 0$, which makes the distribution of μ_t non-degenerate at $t > 0$. Hence $\tilde{\tau}_\gamma^{\text{vol}}$ is not a deterministic function of t , and Proposition 2 fails. $\blacksquare$

Discussion.

The contrast with the rational case (Proposition 2) is direct. In the rational benchmark, $\mathbb{E}[(\mu_{t+1} - \mu_0)^2 \mid \mu_t]$ is never weighted into the payoff, so the location term in (1.31) is invisible to the employer's optimization, and the deterministic-in- t structure survives. Cognitive consistency activates the location term by attaching a non-zero coefficient $-\gamma$ to it.

1.12.7. Proof of Proposition 5 (Commitment vs. Sequential: expected sample size)

The Commitment employer chooses $\tilde{n}_\gamma^*$ ex ante to maximize (1.13). The Voluntary employer, conditional on μ_t , faces the per-step continuation value

$$\tilde{C}_t(\mu_t) = \mathbb{E}[\tilde{V}_{t+1}(\mu_{t+1}) \mid \mu_t] - \gamma \text{Var}(\mu_{t+1} \mid \mu_t) - \gamma (\mu_t - \mu_0)^2 - c.$$

The Voluntary employer is willing to continue as long as $\tilde{C}_t(\mu_t) > -\sigma_t^2$, i.e., as long as the realized continuation value (with realized μ_t entering the location term) exceeds the stopping value. The Commitment employer must internalize the ex-ante expected location term $-\gamma \mathbb{E}_0[(\mu_t - \mu_0)^2] = -\gamma(\sigma_0^2 - \sigma_t^2)$, which is non-zero at every $t > 0$. The Voluntary employer, by contrast, can stop early on histories where the realized μ_t is close to μ_0 (so the location term is small), avoiding the cumulative ex-ante cost. Therefore the Voluntary employer's expected stopping time is weakly larger than the Commitment employer's deterministic sample size.

Formally: at any candidate stopping period $t < \tilde{n}_\gamma^*$, the rational ex-ante marginal value (from Proposition 3(2)) is strictly positive. The Voluntary employer at $t < \tilde{n}_\gamma^*$ faces the same instrumental gain $\sigma_t^2 - \sigma_{t+1}^2 - c > 0$ but the realized location-term cost $\gamma(\mu_t - \mu_0)^2$ varies. On histories where $(\mu_t - \mu_0)^2$ is small, the Voluntary employer's per-step marginal value exceeds the Commitment employer's ex-ante marginal value, so she continues. On histories where $(\mu_t - \mu_0)^2$ is large, she stops. Averaging over histories, the expected Voluntary stopping time is at least $\tilde{n}_\gamma^*$. The strict-inequality refinement follows from a standard option-value argument: $\max\{a, x\}$ is a convex function of x , so

Jensen's inequality gives $\mathbb{E}_0[\max\{a, x\}] \geq \max\{a, \mathbb{E}_0[x]\}$, with strict inequality when x has nonzero variance and the constraint set straddles a . The realization-dependent stopping rule (LHS) therefore strictly exceeds the pre-committed deterministic rule (RHS): the value of being able to condition the stop/continue decision on the realized $x = \gamma(\mu_t - \mu_0)^2$ strictly exceeds the value of pre-committing at $\mathbb{E}_0[x]$, so the expected per-step continuation value under the realization-dependent stopping rule strictly exceeds the ex-ante expected continuation value evaluated at the mean, and the Voluntary employer continues strictly more often than the Commitment employer's deterministic threshold prescribes. This holds at any $t > 0$ under the model's prior $\theta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ with $\sigma_0^2 > 0$, which makes the distribution of μ_t non-degenerate. ■

1.12.8. Proof of Proposition 6 (Gap Closure)

Under equal prior variance $\sigma_0^H = \sigma_0^L \equiv \sigma_0^2$, $\omega_n^H = \omega_n^L \equiv \omega_n$, so (1.18) gives $\Delta_n = (1 - \omega_n)\Delta_0$.

The shrinkage factor is

$$\frac{\Delta_n}{\Delta_0} = 1 - \omega_n = 1 - \frac{n\sigma_0^2}{n\sigma_0^2 + \sigma_d^2} = \frac{\sigma_d^2}{n\sigma_0^2 + \sigma_d^2},$$

which is (1.18). The comparative statics:

$$\begin{aligned} \frac{\partial}{\partial n} \left(\frac{\sigma_d^2}{n\sigma_0^2 + \sigma_d^2} \right) &= -\frac{\sigma_d^2 \cdot \sigma_0^2}{(n\sigma_0^2 + \sigma_d^2)^2} < 0, \\ \frac{\partial}{\partial \sigma_0^2} \left(\frac{\sigma_d^2}{n\sigma_0^2 + \sigma_d^2} \right) &= -\frac{\sigma_d^2 \cdot n}{(n\sigma_0^2 + \sigma_d^2)^2} < 0, \\ \frac{\partial}{\partial \sigma_d^2} \left(\frac{\sigma_d^2}{n\sigma_0^2 + \sigma_d^2} \right) &= \frac{(n\sigma_0^2 + \sigma_d^2) \cdot 1 - \sigma_d^2 \cdot 1}{(n\sigma_0^2 + \sigma_d^2)^2} = \frac{n\sigma_0^2}{(n\sigma_0^2 + \sigma_d^2)^2} > 0. \end{aligned}$$
■

1.12.9. Proof of Proposition 7 (Voluntary bias toward prior: selection channel)

We show the chain $A \leq B = C$, where $A = \mathbb{E}_0[(\mu_{\tau_Y^{\text{cons}}}^{\text{cons}} - \mu_0)^2]$ is the cognitive-consistency Voluntary employer's expected squared deviation from prior, $B = \mathbb{E}_0[(\mu_{\tau_Y^{\text{rat}}}^{\text{rat}} - \mu_0)^2]$ is the unselected rational Bayesian's expected squared deviation evaluated at the same stopping distribution, and $C =$

$\mathbb{E}_0[\sigma_0^2 - \sigma_{\tilde{\tau}_Y^{\text{vol}}}^2]$ is the algebraic simplification of B via the unconditional variance identity. Step 1 establishes the substantive inequality $A \leq B$ via the selection channel; Step 2 establishes the algebraic identity $B = C$ via Lemma 1's unconditional variance formula.

Step 1: Selection inequality $A \leq B$.

By Proposition 4, the Voluntary employer's stopping rule is realization-conditional: $\tilde{\tau}_Y^{\text{vol}} = \min\{t : -\sigma_t^2 \geq \tilde{C}_t(\mu_t)\}$, where $\tilde{C}_t(\mu_t)$ is a decreasing function of $(\mu_t - \mu_0)^2$. The location term $-\gamma(\mu_t - \mu_0)^2$ in (1.17) pushes the continuation value down as μ_t moves away from μ_0 , so the employer is more likely to stop when μ_t is far from μ_0 and more likely to continue when μ_t is close to μ_0 . Economically, the employer with strong prior bias finds it costly to keep updating away from her prior, so she stops earlier when realized signals push μ_t far from μ_0 , while she continues when μ_t stays close to her starting point.

Conditioning on the realized stopping time $\tilde{\tau}_Y^{\text{vol}} = n$, the distribution of μ_n^{cons} is the distribution of μ_n under the prior, restricted to the event $\{\tilde{\tau}_Y^{\text{vol}} = n\}$, the event that the stopping inequality $-\sigma_n^2 \geq \tilde{C}_n(\mu_n)$ first holds at n . This event selects against μ_n values that are far from μ_0 , so the conditional distribution is truncated toward μ_0 relative to the unconditional Bayesian distribution μ_n^{rat} at sample size n . Therefore, for every n in the support of $\tilde{\tau}_Y^{\text{vol}}$,

$$\mathbb{E}\left[(\mu_n^{\text{cons}} - \mu_0)^2 \mid \tilde{\tau}_Y^{\text{vol}} = n\right] \leq \mathbb{E}\left[(\mu_n^{\text{rat}} - \mu_0)^2\right].$$

Taking expectation over $\tilde{\tau}_Y^{\text{vol}}$ on both sides preserves the ordering by the law of iterated expectations:

$$A = \mathbb{E}_0\left[(\mu_{\tilde{\tau}_Y^{\text{vol}}}^{\text{cons}} - \mu_0)^2\right] \leq \mathbb{E}_0\left[(\mu_{\tilde{\tau}_Y^{\text{vol}}}^{\text{rat}} - \mu_0)^2\right] = B.$$

Step 2: Variance identity $B = C$.

By the unconditional variance identity (1.30), at every fixed n , $\mathbb{E}_0[(\mu_n^{\text{rat}} - \mu_0)^2] = \text{Var}_0(\mu_n^{\text{rat}}) = \sigma_0^2 - \sigma_n^2$, since $\mathbb{E}_0[\mu_n^{\text{rat}}] = \mu_0$ by the martingale property of Lemma 1. Averaging over the Voluntary

stopping distribution,

$$B = \mathbb{E}_0[(\mu_{\tilde{\tau}_\gamma^{\text{vol}}}^{\text{rat}} - \mu_0)^2] = \mathbb{E}_0[\sigma_0^2 - \sigma_{\tilde{\tau}_\gamma^{\text{vol}}}^2] = C.$$

Combining Steps 1 and 2 yields $A \leq B = C$, the chain stated in Proposition 7.

Equality $A = B$ holds at $\gamma = 0$ because the location-term cost in $\tilde{C}_t$ vanishes, so the stopping rule no longer selects on $(\mu_t - \mu_0)^2$ and the Step 1 inequality becomes an equality pointwise. ■

1.12.10. Proof of Corollary 1 (Gap-closure ranking under cognitive consistency)

We work conditional on Propositions 4, 3, and 7–5. The corollary states the ranking of expected residual gaps: $\mathbb{E}_0[\Delta_n^{\text{vol}}/\Delta_0] \geq \mathbb{E}_0[\Delta_n^{\text{commit}}/\Delta_0] \geq \mathbb{E}_0[\Delta_n^{\text{ExogMean}}/\Delta_0]$. Equivalently the expected closure rates satisfy the reverse ordering (Voluntary closes least; ExogenousMean closes most).

Step 1: Voluntary > Commitment in residual gap.

By Proposition 7 (selection channel), the Voluntary employer’s terminal posterior is biased toward her prior relative to the unselected Bayesian benchmark at the same stopping time: $\mathbb{E}_0[(\mu_{\tilde{\tau}_\gamma^{\text{vol}}}^{\text{cons}} - \mu_0)^2] \leq \mathbb{E}_0[\sigma_0^2 - \sigma_{\tilde{\tau}_\gamma^{\text{vol}}}^2]$. The Commitment employer’s posterior at $\tilde{n}_\gamma^*$ is unselected on realizations, so $\mathbb{E}_0[(\mu_{\tilde{n}_\gamma^*}^{\text{commit}} - \mu_0)^2] = \sigma_0^2 - \sigma_{\tilde{n}_\gamma^*}^2$. Translating into the cross-group residual gap, the per-group deviation $\delta_n^r := \mu_n^r - \mu_0^r$ inherits the bias toward μ_0^r in the expected squared deviation $\mathbb{E}_0[(\delta_n^r)^2] = \mathbb{E}_0[(\mu_n^r - \mu_0^r)^2]$ for each group r under the model’s stated prior $\theta_r \sim \mathcal{N}(\mu_0^r, \sigma_0^2)$ and equal prior variance across groups; this is the model’s idealized regime in which the variance identity $\text{Var}_0(\mu_n) = \sigma_0^2 - \sigma_n^2$ applies.

Step 1a: Per-group deviation in ω_n . Under the prior assumption $\theta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ from equation (1.1) and the variance identity $\sigma_0^2 - \sigma_n^2 = \omega_n \sigma_0^2$, the unselected Bayesian benchmark satisfies $\mathbb{E}_0[(\delta_n^r)^2] = \omega_n \sigma_0^2$. Proposition 7 bounds the Voluntary-under-cognitive-consistency per-group deviation strictly below this benchmark when $\gamma > 0$, equivalently selecting against histories with large effective ω_n .

Step 1b: Cross-group residual gap in ω_n . Under equal prior variance, $\omega_n^H = \omega_n^L \equiv \omega_n$, and

equation (1.18) gives $\mathbb{E}_0[\Delta_n/\Delta_0] = 1 - \omega_n$ at any realization-independent n . Selection against large ω_n in the per-group deviation maps directly to selection toward larger $1 - \omega_n$ in the cross-group residual gap. The Commitment posterior at $\tilde{n}_\gamma^*$ is unselected and tracks the benchmark $1 - \omega_{\tilde{n}_\gamma^*}$ exactly; Voluntary's selection under cognitive consistency shifts the expected residual gap strictly upward.

The cross-group residual gap therefore satisfies

$$\mathbb{E}_0 \left[\frac{\Delta_n^{\text{vol}}}{\Delta_0} \right] \geq \mathbb{E}_0 \left[\frac{\Delta_n^{\text{commit}}}{\Delta_0} \right],$$

with strict inequality at $\gamma > 0$.

Step 2: Commitment > ExogenousMean in residual gap.

By Proposition 3(2), $\tilde{n}_\gamma^* \leq n^*$. By Proposition 5, $\mathbb{E}_0[\tilde{\tau}_\gamma^{\text{vol}}] \geq \tilde{n}_\gamma^*$. By the constraint $n^{\text{ExogMean}} = \mathbb{E}_0[\tilde{\tau}_\gamma^{\text{vol}}]$ defining the ExogenousMean regime, $n^{\text{ExogMean}} \geq \tilde{n}_\gamma^*$.

The gap-closure rate is $g(n) = \sigma_d^2/(n\sigma_0^2 + \sigma_d^2)$, which is strictly decreasing in n . Therefore $g(n^{\text{ExogMean}}) \leq g(\tilde{n}_\gamma^*) = g(n_\gamma^{\text{commit}})$, i.e.,

$$\mathbb{E}_0 \left[\frac{\Delta_n^{\text{commit}}}{\Delta_0} \right] \geq \mathbb{E}_0 \left[\frac{\Delta_n^{\text{ExogMean}}}{\Delta_0} \right],$$

with strict inequality at $\gamma > 0$ when $n^{\text{ExogMean}} > \tilde{n}_\gamma^*$.

Step 3: Convergence as $\gamma \rightarrow 0$.

As $\gamma \rightarrow 0$, three things happen: (i) the location-term cost in $\tilde{C}_t$ vanishes, so the realization-dependence in the Voluntary stopping rule disappears (the wedge from Proposition 4 closes); (ii) $\tilde{n}_\gamma^* \rightarrow n^*$ (the Commitment optimum collapses to the rational benchmark by Proposition 3(1)); (iii) $\mathbb{E}_0[\tilde{\tau}_\gamma^{\text{vol}}] \rightarrow n^*$ as well, since the rational Voluntary equilibrium has $\tilde{\tau}_\gamma^{\text{vol}} = n^*$ deterministically. Hence all three regimes' closure rates converge to $g(n^*) = 1 - \omega_{n^*}$, and the inequalities collapse to equality. ■

Immigrant Peers and the Long-Run Outcomes of US-Born Students

2.1. Introduction

¹Immigration to the U.S. has increased significantly over the past several decades, causing dramatic economic (Peri, 2012), social (Cascio & Lewis, 2012) and political (Mayda et al., 2022) change. Specifically, the number of immigrants in the U.S. has increased from nearly 4.7% of the overall population in 1970 to 14% today, with large concentrations in urban areas and southern border states. As a result, public schools across the country have become increasingly diverse, with immigrant students comprising a growing share of enrollment.² Yet we know remarkably little about whether this exposure has lasting effects on native-born students' educational attainment and labor market outcomes. Most of the literature examining immigrant exposure in schools has focused on contemporaneous effects, with rare exceptions (Gould et al., 2009) that follow students only through high school. This paper provides the first U.S. evidence on this question. Using a novel longitudinal dataset linking Texas K–12 records to higher education and earnings, we exploit within-family variation in immigrant peer exposure across siblings to address endogenous sorting. Our central finding is that immigrant peer exposure has no meaningful long-run impact: effects on college completion, employment, and earnings are economically negligible, and a small negative effect on college enrollment disappears under an instrumental variables strategy.

¹This chapter is co-authored with Briana Ballis (University of California, Merced; bballis@ucmerced.edu) and Derek Rury (School of Public Policy, Oregon State University; ruryd@oregonstate.edu). This work has been supported by grant #G-2210-39878 from the Russell Sage Foundation. Any opinions expressed are those of the author alone and should not be construed as representing the opinions of the funder. The conclusions of this research do not necessarily reflect the opinions or official position of the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas.

²As of 2021, roughly 3 percent of children younger than 17 in the U.S. were foreign-born, and almost one in four come from an immigrant household. In Texas, the setting of our study, roughly 5% of youth are foreign-born and 34.4% grew up in immigrant families. For additional details see https://tea.texas.gov/sites/default/files/enroll_2019-20.pdf and <https://www.migrationpolicy.org/programs/data-hub/charts/children-immigrant-families>.

Understanding these long-run effects is critical for policy and welfare analysis for at least two reasons. First, if short-run test score impacts do not persist into adulthood, or if long-run impacts differ in sign and magnitude, then responses based solely on short-run evidence may be misguided, with important implications for economic mobility and inequality. Second, the long-run effects themselves are theoretically ambiguous, as immigrant peers could enhance labor market outcomes through expanded social networks and cultural exposure, or diminish them through increased competition for jobs and resources.

Moreover, even the short-run evidence is limited. While many studies find small negative or null effects, most are limited in their ability to account for endogenous sorting across schools with different immigrant compositions. Native families often respond to rising immigrant enrollment by switching schools or neighborhoods (a phenomenon referred to as “native flight”), meaning students with greater immigrant exposure may differ systematically in unobserved characteristics. Traditional peer effects models address static selection into schools through school fixed effects (e.g., Hoxby, 2000), but cannot address dynamic sorting in response to changing school composition. In recent influential work Figlio et al., 2021 uses family fixed effects to find modest positive spillover effects from immigrant peers in Florida. That said, Florida’s immigrant population differs substantially from Texas’s in national origin, income, and education levels (see Section 2.2.1), raising natural questions about the external validity of existing estimates. Texas provides a complementary setting that substantially broadens the evidence base.³

Identifying long-run impacts of immigrant peers poses even greater challenges. First, most administrative datasets that track childhood immigrant peer exposure do not follow students into adulthood.⁴ Second, long-run identification must confront the same sorting concerns as short-run estimates, but this challenge is amplified. Families do not respond to immigrant inflow just once, but repeatedly over a child’s entire schooling. As a result, long-run outcomes reflect the *accumulated effects* of years of dynamic family-driven sorting. Standard peer effects strategies cannot resolve this

³This distinction matters because migration motivations differ across these groups: Cuban (and more highly educated) immigrants have predominantly sought political asylum (Pedraza, 2007), while Mexican immigrants primarily migrate for economic opportunities (Massey et al., 2002) and Central Americans increasingly flee violence (Clemens, 2021). Cortes, 2004 shows that refugees integrate into the U.S. labor market faster than economic migrants, suggesting spillover effects may differ between these immigrant populations. See Appendix Table 2.10 for details.

⁴Few U.S. states link K-12 records through college to earnings. Texas is unique in providing this comprehensive linkage through unemployment insurance records.

selection, and even datasets with rich long-run outcomes are insufficient without a way to net out these family-level responses. Yet few U.S. states maintain administrative school records that allow families to be linked.

We address these challenges by constructing a novel dataset that links administrative education and labor market records for the universe of Texas births between 1976 and 1997. Critically, the birth records allow us to link siblings together, allowing us to follow the strategy of Figlio et al., 2021 and compare how students within the same family respond to different levels of immigrant peer exposure. This approach controls for families' decisions about school choice and neighborhood selection in response to immigrant inflows. Remaining variation in immigrant peer exposure across siblings reflects idiosyncratic cohort level differences within schools.⁵ This strategy provides a powerful and credible approach to isolating causal effects of immigrant peer exposure from confounding family preferences.

When not considering the role of families, we find significant sorting along two dimensions. First, native students with higher immigrant peer exposure are systematically more disadvantaged and more likely to be minorities than those with lower exposure. A one standard deviation increase in immigrant exposure is associated with natives being more likely to be free and reduced price lunch (FRL) eligible (26 percent of the mean), Hispanic (34 percent of the mean), and Black (25 percent of the mean). Second, we find significant "native flight" in response to immigrant exposure, confirming results from prior work (Betts & Fairlie, 2003; Cascio & Lewis, 2012), with higher exposure predicting increased district switching and reduced continued enrollment within Texas. Without family fixed effects we find negative and economically meaningful effects on natives' short and long-run outcomes. However, these estimates conflate true peer effects with the fact that students with higher exposure would likely have performed worse regardless.

Accounting for sorting with the inclusion of family fixed effects substantially attenuates the estimated impacts of immigrant exposure on all outcomes. Relative to the estimates without family fixed effects, the coefficients fall by 84-98%, underscoring the magnitude of the bias when family level

⁵This assumes parents choose schools similarly across siblings. While we believe this assumption is plausible as most children during this period attended their residentially assigned schools, we formally test for child-specific sorting in Section 2.5.1 and find it is unlikely to bias our results.

selection is not addressed. With family fixed effects, the estimates for college graduation, employment and earnings seven years after graduating high school are no longer statistically significant. We continue to find statistically significant negative effects on academic outcomes and college enrollment, but these effects are economically very small. For example, a one standard deviation increase in immigrant exposure reduces math test scores by less than one percentage point of a standard deviation. This represents approximately 1.6% of the impact of a highly effective teacher for one year Rockoff, 2004. The impacts on college enrollment are similarly small. A one standard deviation increase in immigrant exposure decreases college enrollment by 0.4 percentage points (0.67%). This effect is substantially smaller than those found in other educational intervention studies.⁶

We find modest heterogeneity across native and immigrant characteristics. The negative impacts are driven by White and non-FRL natives, while Black and Hispanic natives tend to benefit from immigrant exposure, though effects are modest and only statistically significant for high school graduation among Black students and college enrollment among Hispanic students, at the 5 percent and 1 percent levels, respectively. Effects vary with the characteristics of immigrant peers: exposure to White, Asian, and non-FRL immigrant peers is associated with small positive effects, while exposure to Hispanic and FRL immigrant peers produces small negative effects, though all magnitudes remain small. Finally, to the extent any effects appear by school locale, they are primarily concentrated in city schools, consistent with Getik et al. (2024).

We further validate these findings using an instrumental variables strategy that predicts immigrant exposure based on typical school transition patterns rather than parental choices, following Figlio et al., 2021. The IV estimates confirm the family fixed effects results: across all outcomes, the IV point estimates are close to zero and statistically insignificant, including for college enrollment, the one outcome where family fixed effects yield a significant negative estimate.

This paper makes several contributions. First, it provides the first evidence from the U.S. on the long-run effects of childhood exposure to immigrant peers, moving beyond the contemporaneous labor market impacts that dominate existing immigration research (Boje-Kovacs et al., 2024; Dustmann et al., 2013; Ottaviano & Peri, 2012). Prior work has largely focused on short-run consequences for

⁶Other studies have shown impacts on college enrollment ranging from 6 to 9 percentage points for interventions such as Head Start participation, and same-race teacher assignment Deming, 2009; Gershenson et al., 2022.

native workers following inflows of migrant labor, whereas far less is known about how immigrant presence in schools shapes native children’s own educational and economic trajectories. By tracing effects into college enrollment and labor market outcomes, we provide evidence directly relevant to debates over how immigrant inflows may alter the long-run economic opportunities of U.S.-born children.

Second, we advance research on how immigration affects public school students. Assessing the impacts of immigrant peer exposure is notoriously challenging due to well-documented native flight. Only a handful of studies overcome this challenge by linking siblings across administrative data, with Figlio et al. (2021) providing the seminal example in Florida. Similar approaches in Europe using rich population registries (Getik et al., 2024) replicate the modest positive gains Figlio et al. (2021) documents. Researchers have estimated causal effects of refugee students on native academic outcomes (Figlio & Ozek, 2019; Morales, 2022; Van der Werf, 2021). However, research shows that refugees (who migrate as a result of a sudden shock) differ in important ways from immigrants (who consciously decide to migrate).⁷ We provide the first long-run evidence by tracing immigrant peer exposure from kindergarten through 12th grade to adult college and labor market outcomes using linked administrative data from Texas. This matters because the existing short-run evidence cannot assess long-run welfare effects. Moreover, our focus on Texas with predominantly low-income Hispanic communities, which are at the center of U.S. immigration debates, makes the findings directly policy relevant.

Third, our findings contribute more broadly to the literature on how diversity affects public school students. A specific literature studies the effect of desegregation among Black and White students across the 20th century using changes in busing policies (Angrist & Lang, 2004; Guryan, 2004; Johnson, 2011; Tuttle, 2019). Results from this literature find little impact of desegregation programs on the long-run outcomes of White students. We contribute to this literature by studying immigrants as another dimension of diversity in American schools, extending the evidence on the long-run effects of diversity beyond the Black-White context.

Finally, we contribute to the broader peer effects literature in education, which has focused

⁷In addition, because of their hardships, refugees benefit from several government programs which help economically and socially, where immigrants are not able to access these types of programs.

primarily on short-run academic outcomes (Cools & Patacchini, 2021; Sacerdote, 2011). The few existing studies examining long-run outcomes focus on peer exposure either in early grades (e.g., Carrell et al., 2018; Denning et al., 2023) or in later grades (e.g., Brenøe & Zölitz, 2020), but not throughout, and do not address selective residential mobility.⁸ We push this literature forward by examining cumulative immigrant peer exposure from kindergarten through high school while being the first to leverage family linkages to address selective mobility. Our results improve our understanding of the long-run impacts of peer composition which is essential for making policy recommendations, since credible welfare analysis requires measures of actual productivity such as earnings or employment.

2.2. Data and Setting

2.2.1. Immigration in Texas

The U.S. has experienced a substantial rise in immigration over the past few decades. Immigration patterns vary substantially across states (Appendix Table 2.10). Texas stands out as having the immigrant population with the lowest average income and education levels among major receiving states: in 2010, approximately 50% of immigrants lived below 200% of the federal poverty line (vs. 37% nationally), only 56% had a high school degree (vs. 67% nationally), and 36% reported difficulty speaking English (vs. 20% nationally). Immigration from Mexico dominates flows to Texas (54% in 2010). These patterns position Texas as particularly policy-relevant, as its immigrant population exemplifies the profile most central to current debates (Center, 2019; Hanson, 2010).

2.2.2. Data Sources and Sample Construction

The core data for this project come from a restricted access dataset that links administrative school records for the universe of Texas public school students from the Texas Education Agency (TEA) to key adult outcomes, including college records from the Texas Higher Education Coordinating Board (THECB) and labor market outcomes from the Texas Workforce Commission (TWC). These data span 1993 to the present and provide a comprehensive view of students' educational trajectories and

⁸We acknowledge that selective mobility concerns may vary by context and peer characteristic. Our family linkage approach addresses well-documented sorting in response to immigrant peer composition, though it is plausible that families may not respond as strongly to other peer characteristics, such as disruptive peers.

adult outcomes in the second largest state in the U.S., which serves over 5 million students annually.

The data are rich in both short- and long-run outcomes. Short-run outcomes include performance on English and Math exams, as well as disciplinary actions such as suspensions, attendance, and high school graduation. Long-run outcomes include enrollment in all public and private universities, college major choice, college graduation and degree attainment, employment, and quarterly earnings. The schooling records also contain detailed demographic information on students (such as FRL participation, race, and gender), allowing us to explore heterogeneous impacts of immigrant peer exposure.

One limitation of these data is that we are not able to follow those who leave Texas. Fortunately, out-migration from Texas is low. Most people born in Texas remain in the state Aisch et al., 2014 and only 1.7 percent of Texas residents leave the state each year S. White et al., 2016. In addition, we are able to link a subset of our sample to the National Student Clearinghouse (NSC) data which track college enrollment outside of the state of Texas to determine how often students attend college out of state.

Constructing Family Links

While comprehensive, these data lack family identifiers, which are a key component of our empirical design. We overcome this limitation by merging records from the Texas Birth Index (TBI) from 1976 to 1997.⁹ The TBI records the universe of births in Texas along with parent information (their full names, including middle names and maiden names). We match siblings based on them having parents with the same name. To account for the possibility of misspelled names, we developed a matching technique that incorporates adjustments for nicknames and employs phonetic matching algorithms.¹⁰

Using both parents' full names, we are very unlikely to incorrectly classify two children as siblings when they are not. To further ensure the accuracy of our sibling matches, we leverage the TEA data

⁹The TBI goes back to the 1940s; however, the TEA data begin in 1993, and the 1976 cohort is the first birth cohort observable in the TEA data.

¹⁰We provide further details on the sibling identification procedure in Appendix 2.8, where we also verify the stability of our findings under alternative matching specifications, such as requiring exact name matches or applying nickname and phonetic corrections separately.

and retain children as siblings only if they attended schools in the same or contiguous counties for the majority of the time they spent in school together. However, our algorithm may classify some true siblings as only children if the name matching algorithm cannot correct for misspellings or if siblings were born outside the window for which we observe TBI data.

Our sibling matching approach leads to the exclusion of two groups. We exclude children from single-parent households, since our match requires both parents' names to be present. We also do not include siblings from blended families whose parents may have divorced and remarried. For these reasons, our matched sibling sample will be more positively selected than the overall Texas population. Importantly, this does not threaten internal validity but may affect external validity. As discussed in more detail in Section 2.2.4, this selection is modest, and we discuss the implications for generalizability in Appendix 2.8.

Definition of Immigrants

Importantly, the schooling data includes an immigrant flag making it possible to measure immigrant peer exposure. Each year, the TEA flags students who were not born in the U.S. and have been attending school in the U.S. for less than three full academic years. We define immigrants as students who were ever flagged as an immigrant during their tenure in Texas public schools. This approach leverages the panel nature of our data to identify all immigrant students, not just recent arrivals in a given year. Because Texas public schools require birth certificates or other official proof of identity for enrollment, we are confident in the accuracy of this immigrant status flag (Texas Education Agency, 2017).

In robustness checks, we examine alternative definitions of immigrant students. First, we use a narrow definition focusing on recent immigrants flagged in the data, who may face the greatest integration challenges. Second, we use a broader definition that includes all students not born in Texas (as determined by their absence from the Texas Birth Index) whose home language is not English. While this broader measure may overestimate the immigrant student population, it captures spillovers from both immigrant students and children growing up in immigrant households. Our results are similar across both definitions.

We measure immigrant peer exposure using *cumulative* peer exposure which accounts for both contemporaneous exposure to immigrant peers and all previous exposure. We first define *contemporaneous* immigrant peer exposure in each year at the campus grade level as follows:

$$ShareImmigrant_{igst} = \frac{\sum_{k \neq i} Immigrant_{kgst}}{n_{gst} - 1} \quad (2.1)$$

where the numerator counts the number of immigrants in a campus (s), grade (g), and year (t) using one of the three immigrant definitions discussed above, as a proportion of all other students enrolled in a campus, grade, and year. We then define *cumulative* peer exposure from school entry up to a specific grade as follows:

$$ShareImmigrant_{igst} = \frac{1}{g} \sum_{g' \leq g} ShareImmigrant_{ig'st} \quad (2.2)$$

where $ShareImmigrant_{ig'st}$ is contemporaneous immigrant peer exposure for student i as defined in Equation 2.1 during grade g' , where $g' \leq g$. We define this cumulative measure from the start of school entry up until the grade the outcome is being measured in. This cumulative measure varies at the individual level and captures differences in immigrant peer exposure as students progress through school and potentially move across campuses.

2.2.3. Sample Construction

We focus our analysis on 1st grade cohorts between 1994 and 2000. The 1994 cohorts are the first we can fully observe in primary school, as the schooling data begin in 1994.¹¹ We restrict the final cohort to 2000 to ensure we can observe long-run outcomes at least seven years after expected high school graduation. Students entering 1st grade in 2000 are expected to graduate high school in 2012, which allows us to measure their outcomes through 2019, just before the COVID-19 pandemic began. In total, we track seven cohorts of Texas students, encompassing approximately 400,000 students per cohort.¹²

¹¹While the data technically start in 1993, some key demographics are missing in those years.

¹²Although short-run outcomes could be measured for more recent cohorts, we limit our analysis to these 1st grade cohorts so that the same set of students contributes to both short- and long-run outcome analyses.

We define US-born students as those born in Texas who were never classified as English Language Learners (ELL). We impose the Texas born restriction because we only observe siblings from the Texas Birth Index (which includes the universe of births *only* in Texas) and controlling for sibling fixed effects is a key part of our empirical approach. The ELL restriction ensures that we are omitting students who may be growing up in immigrant households, and allows us to isolate a sample of students without direct immigrant family ties, ensuring we measure spillovers rather than direct effects. These two sample restrictions follow Figlio et al., 2021, who examines immigrant peer spillovers in Florida.¹³

2.2.4. Summary Statistics

Table 2.1 presents summary statistics for outcomes and covariates for the full sample of Texas public school students, those matched to the TBI, those with at least one sibling, our final analysis sample (students with siblings who were born in Texas and never classified as an ELL), and students who were ever flagged as an immigrant. The overall Texas population shown in Column 1 is 40% Hispanic, 42% White, 14% Black and 3% Asian. Students born in Texas (Column 2) look very similar in terms of demographics and outcomes, though they are less likely to be classified as an ELL. This is an expected pattern since the full sample includes immigrant students who arrived after birth.

The subset of Texas-born children with siblings (Column 3) is slightly more positively selected, with lower FRL enrollment and a higher share of White students. This positive selection stems from our sibling identification strategy, which requires both parents listed on the birth certificate. During our analysis period, 16% of birth certificates have a missing father, mechanically excluding single-parent families. This selection is also reflected in higher test scores, college enrollment, and earnings. Our final analysis sample of non-ELL speakers (Column 4) exhibits even stronger positive selection, as we exclude students ever classified as an ELL. This restriction further increases average achievement and long-run outcomes. These patterns highlight that our estimates apply to a somewhat advantaged subset of Texas students.¹⁴

Finally, Column 5 reports summary statistics for students ever classified as immigrants. Relative

¹³In Section 2.5.3, we show that our results are similar if we do not restrict the sample to English speakers.

¹⁴Appendix Table 2.16 shows that our results are robust to including ELL students, with effects if anything slightly more negative in the full sample.

TABLE 2.1. Summary Statistics - 1st Grade Cohorts Between 1994-2000

Variable	All (1)	Texas Born (2)	Texas Born with Siblings (3)	Final Family Sample (4)	Immigrant (5)
Female	.490	.471	.474	.468	.490
Special Education	.121	.131	.135	.126	.053
Asian	.027	.017	.014	.010	.090
Black	.142	.150	.172	.097	.027
Hispanic	.400	.391	.403	.281	.834
White	.425	.438	.407	.607	.046
ELL	.116	.084	.088	0	.573
Free Lunch	.381	.373	.418	.278	.642
Math Test Scores	.025	-.000	-.022	.190	-.063
Reading Test Scores	.021	-.001	-.037	.172	-.321
Discipline Flag	.162	.166	.172	.140	.173
Any College Enrollment	.509	.544	.529	.592	.178
Any College Degree	.166	.169	.163	.209	.068
Employed 7 Years	.644	.707	.707	.711	.217
Income Rank 7 Years	54.747	57.800	57.612	59.010	33.479
Cumulative Immigrant Exposure	.045	.040	.041	.026	.143
N	25,122,192	17,729,988	10,749,480	3,691,473	1,363,494
Distinct Students	2,933,968	1,744,597	1,046,855	448,695	229,231

Notes: This table presents summary statistics for students in first grade cohorts from 1994 to 2000. Column 1 shows statistics for all Texas public school students. Columns 2-4 show progressively restricted samples used to construct our analysis sample: Column 2 includes students born in Texas (matched to the Texas Birth Index), Column 3 includes those with siblings, and Column 4 presents our final family sample that additionally excludes students who were ever English Language Learners. Column 5 shows immigrant students for comparison. Test scores are standardized. Cumulative Immigrant Exposure measures the proportion of immigrant students in a student's school-grade-year and is calculated using all years the student was observed. Short-run outcomes (test scores and discipline) are also calculated using all years the student was observed. Income Rank 7 Years represents an individual's percentile rank in the income distribution measured seven years after their expected high school graduation year, where individuals with zero income are included in the distribution. The Final Family Sample includes students with linked family identifiers used in family fixed effects specifications.

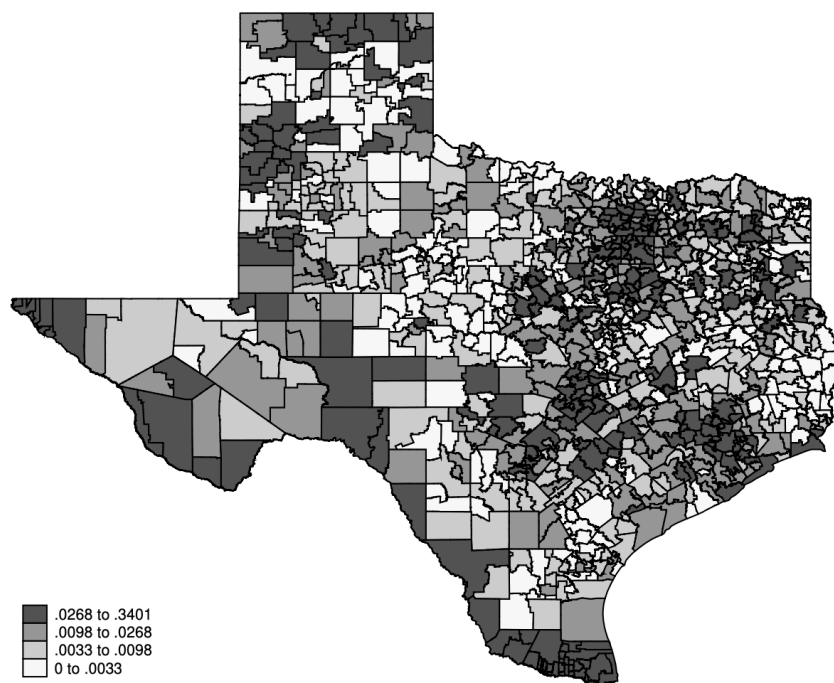
to Texas-born students, immigrants are disproportionately Hispanic (83 vs. 40 percent), more likely to be ELL (57 vs. 12 percent) and FRL (64 vs. 38 percent), relative to natives. They also have lower academic achievement, educational attainment, and adult earnings.

Immigrant Exposure Patterns and Variation

We next characterize immigrant exposure in our data. Figure 2.1 shows that immigrants concentrate geographically in large urban areas (Houston, Dallas-Fort Worth) and border counties, generating substantial cross-district variation. However, Figure 2.2 reveals that cumulative immigrant exposure (as of 9th grade) remains relatively stable over time, hovering around 7 percent across cohorts, with

Hispanic immigrants comprising 5-6 percent.¹⁵ This stability suggests our strategy is unlikely to reflect secular immigration trends or policy changes, and instead captures idiosyncratic year-to-year variation within schools.

FIGURE 2.1. Geographic Distribution of Immigrant Enrollment in Texas Public K–12 Schools



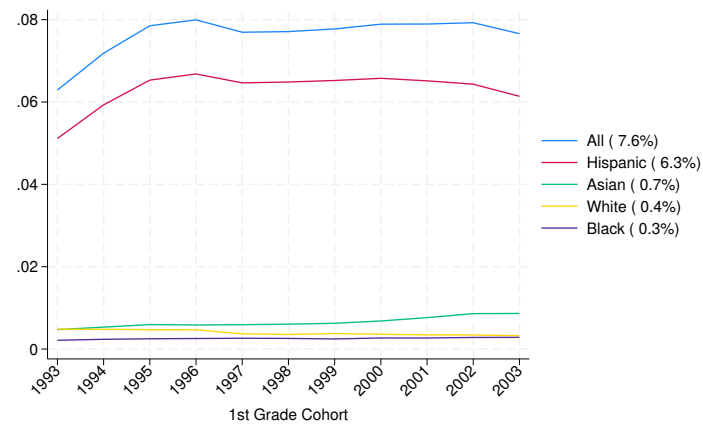
Note: This figure displays immigrant enrollment intensity across Texas public school districts between 1994 and 2004. Darker shading indicates higher shares of immigrant students.

Despite these stable aggregate trends, individual cohorts within the same school experience varying levels of immigrant exposure due to the random timing of immigrant arrivals to specific neighborhoods and schools, grade-level enrollment fluctuations, and the fact that immigrant students enter U.S. schools at varying ages. Appendix Figure 2.9 illustrates this last point, showing that immigrant enrollment increases at higher grades.

Appendix Figure 2.10 quantifies the magnitude of within-school variation for the 1,000 largest districts, plotting mean immigrant exposure against the interquartile range across cohorts within each district. This illustrates that within districts, individual cohorts face meaningfully different exposure levels. This generates differential exposure for siblings. For instance, two siblings attending the same

¹⁵Consistent with this, Appendix Table 2.11 shows that the immigrant population is overwhelmingly Spanish-speaking, with 76.4 percent of immigrant students reporting Spanish as their home language, reflecting the predominantly Latin American composition of the immigrant population in Texas during this period.

FIGURE 2.2. Immigrant Share of Total Enrollment in Texas Public Schools, by Race



Note: This figure plots cumulative immigrant exposure through 9th grade for native students who entered 1st grade between 1993 and 2003, separately by race. Immigrant exposure is measured as the average cumulative share of peers classified as immigrants through 9th grade across all students in a given 1st grade cohort, as defined in Equation 2.2.

school and living in the same neighborhood may face substantially different immigrant shares. An older sibling entering 9th grade in one year may encounter 5 percent immigrant classmates, while their younger sibling entering three years later faces 8 percent. Our family fixed effects design exploits precisely this variation, comparing outcomes across siblings while holding constant all time-invariant family characteristics, school quality, and neighborhood attributes.

2.3. Empirical Strategy

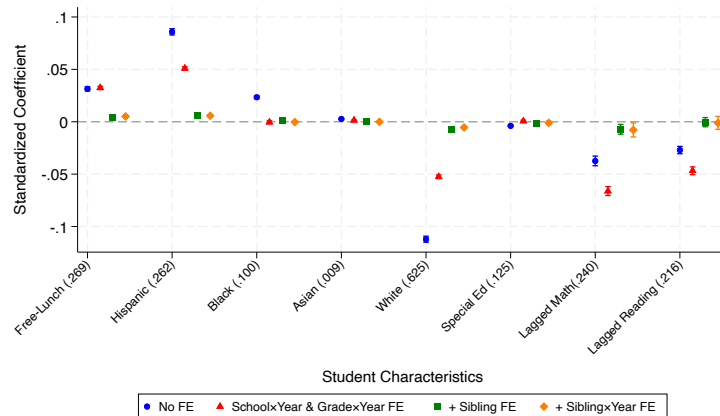
Two key challenges are present in any peer effects study: reflection and selection (Manski, 1993). The reflection problem arises when it is difficult to distinguish whether peers affect a student's outcomes or whether the student influences their peers' outcomes. In our setting, this concern is minimal because immigration status is determined by place of birth. A native-born student cannot cause their classmates to become immigrants, making reflection less relevant in our context than in studies examining behavioral peer effects.

Selection, however, remains a central challenge. Selection occurs when parents choose schools based on characteristics such as peer composition or neighborhood quality. If families sort into schools with specific immigrant concentrations, and this sorting is correlated with unobserved determinants of student outcomes, our estimates of immigrant peer effects could be biased. Most peer effect studies address this by exploiting quasi-random variation in exposure across students in the same school based on idiosyncratic differences across grade-cohorts (e.g. Hoxby, 2000). This approach accounts for selection into schools by comparing students within the same school who differ in immigrant exposure based on entry cohort.

However, standard peer effects models may still be biased if more advantaged parents move their children across schools in response to an influx of immigrant peers. Indeed, there is a large literature documenting that an influx of immigrant peers leads affluent families to move from their schools or districts (Betts & Fairlie, 2003; Cascio & Lewis, 2012). We find direct evidence of this in our setting. Without family fixed effects, native students with higher immigrant peer exposure are systematically more disadvantaged and more likely to be minorities than those with lower exposure, suggesting estimates of immigrant peer effects using standard peer effects models will be biased downward

(Figure 2.3).¹⁶ Higher immigrant exposure also predicts increased district switching and reduced continued enrollment within Texas (Figure 2.4).¹⁷ It is only after including family fixed effects that all evidence of sorting disappears across both dimensions, motivating our family fixed effects design.

FIGURE 2.3. Impact of Immigrant Exposure on Baseline Characteristics



Note: These figures plot regression coefficients and 95% confidence intervals from Equation 2.3 where the dependent variables are the baseline demographic or lagged test score characteristics. The mean of these outcomes are shown in parentheses on the x-axis labels. Each dot represents the coefficient estimate from a separate regression. Blue dots show models with no fixed effects, red triangles add school-by-year and grade-by-year fixed effects, green squares additionally include family fixed effects, and orange diamonds additionally include family-by-year fixed effects. Standard errors are clustered at the school level.

Following Figlio et al., 2021, we compare outcomes of US-born *siblings* who experienced different levels of immigrant peer exposure during their schooling, with identification coming from cohort-level variation in immigrant peer composition within families. Specifically, we estimate:

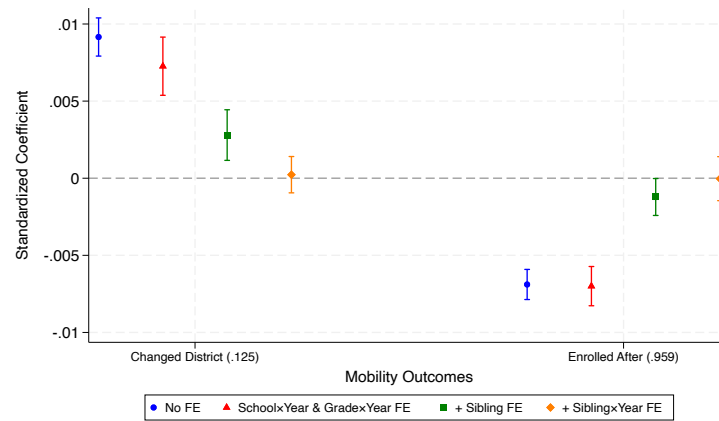
$$y_{igst} = \beta_0 + \beta_1 \text{ShareImmigrant}_{igst} + \beta_2 X_i + \lambda_{st} + \lambda_{gt} + \lambda_f + u_{igst} \quad (2.3)$$

where y_{igst} is a short-run outcome such as test-scores for individual i in grade g in school s in year t . The term $\text{ShareImmigrant}_{igst}$ is cumulative immigrant peer exposure as defined above in Equation 2.2, defined as cumulative peer exposure from the time of school entry up to the specific grade that the

¹⁶The magnitude of these effects are meaningful. Without family fixed effects, a one standard deviation increase in immigrant peer exposure, equivalent to moving from the mean of 3 percent to roughly 7 percent immigrant peers, is associated with natives being more likely to be free lunch eligible (26 percent of the mean), Hispanic (34 percent of the mean), and Black (25 percent of the mean).

¹⁷For mobility, a one standard deviation increase in immigrant peer exposure predicts a 1 percentage point increase in the probability of switching districts, representing roughly 8 percent of the mean switching rate, and a 0.7 percentage point reduction in continued enrollment, representing roughly 1 percent of the mean.

FIGURE 2.4. Impact of Immigrant Exposure on Mobility



Note: These figures plot regression coefficients and 95% confidence intervals from Equation 2.3 where the dependent variables are whether the student changes schools or left the dataset (representing transferring out of the public school system or leaving the state). Each dot represents the coefficient estimate from a separate regression. Blue dots show models with no fixed effects, red triangles add school-by-year and grade-by-year fixed effects, green squares additionally include family fixed effects, and orange diamonds additionally include family-by-year fixed effects. Standard errors are clustered at the school level.

outcome is being measured. λ_{st} are school-by-year fixed effects to control for unobservable differences across schools over time (that may relate to immigrant peer exposure).¹⁸ λ_{gt} are grade-by-year to account for trends in long-run outcomes for all Texas students. X_i include individual characteristics that include an indicator for gender, race, FRL status, whether students participate in special education, and age as well as birth-order, which research shows to predict student academic outcomes (Black et al., 2005; Breining et al., 2020). Finally, λ_f are family fixed effects and their inclusion in our model allows us to control for unobservable characteristics of family preferences. We also estimate models replacing family fixed effects with family-by-year fixed effects to account for time-varying changes in family preferences and circumstances.¹⁹ The main coefficient of interest, β_1 , represents the average impact of immigrant peer exposure on natives' short run outcomes.²⁰

For long-run outcomes such as high school completion, college enrollment, and labor market outcomes, we use the same specification but make an adjustment given these outcomes are observed only once per student. Because students can still contribute multiple observations through the weighting scheme, we apply inverse-frequency weights. As a robustness check, we show results are similar when each student enters the regression only once, regressing long-run outcomes on cumulative immigrant peer exposure as of 12th grade. In this specification, we replace grade-by-year fixed effects with high school cohort fixed effects and include 9th grade school fixed effects.

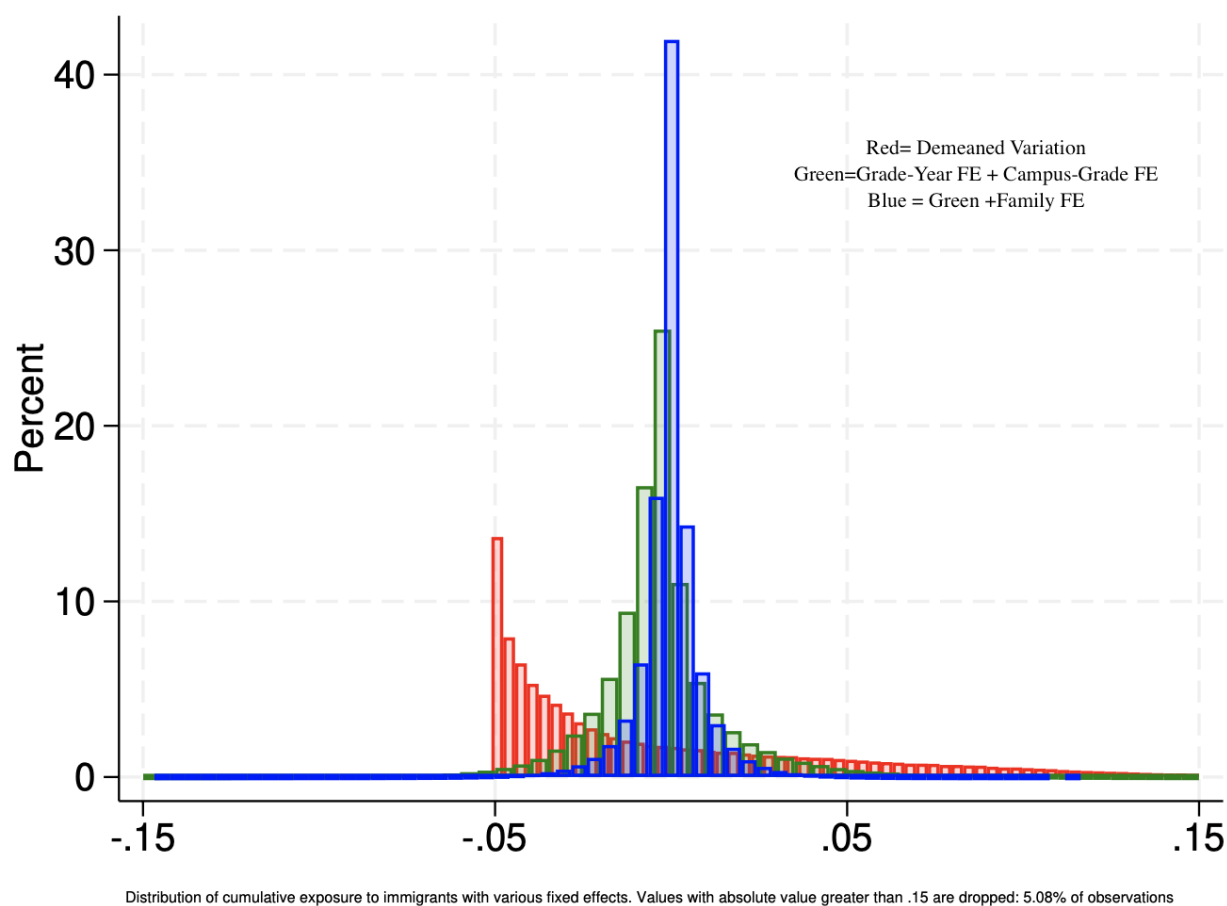
The key assumption of our identification strategy is that families select schools similarly across their children. This assumption could be violated if parents have preferences for immigrant peer exposure that differ by child. For instance, if parents prefer lower immigrant shares for children they perceive as lower-achieving, this would lead to a downward bias in our estimates. Conversely, if parents prefer that their higher-achieving children be exposed to more immigrants, this would lead to an upward bias. Given the prevalence of school choice in Texas during this period, this is a

¹⁸School-by-year fixed effects allows us to control for compositional changes in a school overtime, perhaps driven by an influx of immigrant peers. This fixed effect approach also handles the reality that immigrants tend to migrate to specific places school-year fixed effects control for choices in locations by immigrant families over time.

¹⁹Figure 2.5 plots the distribution of residuals for cumulative immigrant exposure across our different versions of fixed effects, illustrating that while family fixed effects absorb a meaningful share of the variation, substantial variation remains across all models. This indicates that meaningful variation in immigrant exposure remains after conditioning on our fixed effects.

²⁰Given that individual students will contribute to the estimation multiple times, we weight all regressions by the inverse of the number of times a student is observed in the sample.

FIGURE 2.5. Residual Variation in Immigrant Peer Exposure after accounting for fixed effects



Note: This figure shows the residuals of immigrant exposure after including a series of control variables. The red bars display immigrant exposure after variation in the exposure is demeaned. The green bars show residuals in immigrant exposure after including campus-by-year and grade-by-year fixed effects. The blue bar shows residuals in immigrant exposure after then adding family fixed effects.

legitimate concern.²¹ Reassuringly, we address this potential bias using an instrumental variables approach that uses predicted immigrant peer exposure as an instrument for actual immigrant peer exposure. As we discuss in more detail in Section 2.5.1, this IV strategy yields very similar results, suggesting that differential school selection across siblings is not driving our findings.

2.4. Results

We first study the effect of cumulative exposure to immigrant peers on the short-run outcomes of natives, including reading and math test scores from 3rd through 8th grade. Table 2.2 presents results for reading (Panel A) and math (Panel B). The treatment variable is measured on a 0 to 100 scale, meaning the raw coefficient represents the effect of moving from 0 to 100 percent immigrant peers, a shift far outside the range of variation observed in the data where the mean is just 3 percent. We therefore present standardized coefficients in brackets throughout, capturing the effect of a one standard deviation increase in immigrant peer exposure, equivalent to moving from the mean of 3 percent to roughly 7 percent immigrant peers, on outcomes. Column 1 excludes race and free-lunch status, as well as family fixed effects; column two adds race and free-lunch controls; column 3 further adds family fixed effects; and column 4 replaces family fixed effects with family-by-year fixed effects. Each column reports the outcome mean in standard deviation units, standardized using the entire population of public school students, including immigrants. The mean is positive rather than zero, as native students with siblings who were never identified as English language learners are positively selected relative to the broader student population.

Focusing on Panel A, we see negative impacts on natives' reading scores as immigrant exposure increases across all specifications. The most sparse specification in Column 1 yields a meaningful negative coefficient, which attenuates by more than half once race and free-lunch controls are added in Column 2, reflecting the importance of accounting for the socioeconomic composition of exposed students. Coefficients decline substantially further once family fixed effects are included in Column 3, and diminish further still and lose statistical significance entirely in Column 4 with family-by-year fixed effects. This pattern of attenuation across specifications is consistent with our balance checks,

²¹A history of legislation on charter schools in Texas can be found at <https://tea.texas.gov/texas-schools/texas-schools-charter-schools/charter-schools-history-of-charter-schools>.

confirming that failure to account for family-level sorting leads to an overstatement of negative effects. In our most sparse specification in Column 1, a one standard deviation increase in immigrant peer exposure is associated with a 0.042 standard deviation reduction in reading scores, roughly 40 percent of the effect of a one standard deviation improvement in teacher quality for one year (Chetty et al., 2014). Relative to well-identified educational interventions, this magnitude lies at the lower end of meaningful effects and falls between the effects of peer disruption and teacher quality, at roughly one-quarter to one-third the size of a class size reduction.²² Once family-level sorting is accounted for, however, this shrinks to a 0.005 standard deviation decrease in our fully specified model, rendering the impact of immigrant peers on native reading scores economically negligible.

Panel B reveals a similar pattern for math scores. Column 1 again yields statistically significant negative standardized coefficients, which falls by nearly a half once race and free-lunch fixed effects are included, and shrinks further with the addition of family and family-by-year fixed effects. Examining the magnitude of the standardized coefficient in column 4, we again see that, despite being statistically significant, it is economically extremely small.

Tables 2.3 and 2.4 extend our analysis to long-run outcomes. Even in the absence of test score effects, immigrant peer exposure could affect long-run outcomes through non-cognitive channels such as changes in aspirations or social networks. We therefore examine whether immigrant exposure has lasting consequences for educational attainment and labor market success. Across all panels of Table 2.3, without family fixed effects, increased immigrant exposure significantly reduces high school graduation, college enrollment, and receipt of any college degree and four-year degree completion.²³ Once family or family-by-year fixed effects are included, these effects collapse to essentially zero, with only college enrollment remaining marginally significant.

Before the inclusion of family fixed effects, a one standard deviation increase in immigrant peers lowers high school completion by 2.2 percentage points (or 2.7%) for high school completion, college enrollment by 2.8 percentage points (or 5%), receipt of any degree by 1.8 percentage points (or

²²For comparison, a one standard deviation improvement in teacher quality raises test scores by 0.10–0.15 standard deviations (Chetty et al., 2014; Rockoff, 2004); class size reductions raise scores by 0.15–0.25 standard deviations (Krueger, 1999); and exposure to a disruptive peer reduces scores by 0.02–0.05 standard deviations (Carrell et al., 2018). All estimates are per-year.

²³College enrollment is measured within 4 years of expected high school completion. The college completion outcomes are measured within 6 years of expected high school graduation.

TABLE 2.2. Effects of Immigrant Peers on Academic Achievement

	(1)	(2)	(3)	(4)
Panel A: Reading Test Scores				
Cumulative Exposure	-1.067*** (0.075) [-0.042]	-0.436*** (0.062) [-0.017]	-0.161* (0.067) [-0.006]	-0.125 (0.096) [-0.005]
Mean (Y)	0.205	0.205	0.205	0.205
Panel B: Math Test Scores				
Cumulative Exposure	-1.441*** (0.088) [-0.056]	-0.734*** (0.069) [-0.029]	-0.234** (0.074) [-0.009]	-0.237* (0.109) [-0.009]
Mean (Y)	0.232	0.232	0.232	0.232
Individual Controls	Yes	Yes	Yes	Yes
Race FE		Yes		
Free Lunch Status		Yes		
Campus-Year FE	Yes	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes	Yes
Family FE			Yes	
Family-Year FE				Yes
R Squared	0.209	0.244	0.590	0.712
Mean (X)	0.028	0.028	0.028	0.028
N	1,635,630	1,635,630	1,635,630	1,635,630
Distinct Students	341,649	341,649	341,649	341,649

Notes: This table reports estimates of the impact of immigrant exposure on math and reading test scores. Each column within a panel presents estimates of β_1 from a separate regression of Equation 2.3. The sample includes first grade cohorts from 1994 to 2000 who were never classified as an English Language Learner. Cumulative immigrant exposure is calculated as the average immigrant share across all grades up to the current grade, with students reweighted by their frequency of appearance in the regression. Individual controls include gender, age as of September 1, special education status, and birth order. Standardized coefficients are shown in brackets. Coefficients are standardized by multiplying the estimated coefficient by the standard deviation of the treatment variable. Standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

9%), and receipt of a four-year degree by 1.6 percentage points (or 9.6%). These magnitudes are similar to the effect of exposure to disruptive peers (Carrell et al., 2018), comparable to moving a child out of high-poverty public housing (Chetty et al., 2016), and roughly four times larger than a one standard deviation improvement in teacher quality for one year (Chetty et al., 2014).²⁴ Once family or family-by-year fixed effects are included, however, the point estimates collapse to essentially zero, shrinking by roughly 84-98% across the different outcomes. This pattern is consistent with the desegregation literature, which finds little impact of desegregation programs on the long-run outcomes of White students (Angrist & Lang, 2004; Guryan, 2004).

Table 2.4 turns to labor market outcomes measured seven years after expected high school graduation, including an employment indicator (Panel A), income rank (Panel B) and income rank conditional on being employed (Panel C). The pattern is consistent with what we have found throughout. For all three outcomes, we observe negative effects of increased immigrant exposure absent family fixed effects. For being employed, a one standard deviation increase in immigrant exposure reduces employment by 1-1.2 percentage points and income rank by approximately 0.8 percentile points. To put these in context, a one standard deviation improvement in teacher quality raises annual earnings by approximately 1 percent (Chetty et al., 2014), and exposure to a single disruptive peer throughout elementary school reduces earnings at age 24–28 by 3 percent (Carrell et al., 2018). Once family or family-by-year fixed effects are included, however, these relationships disappear entirely.

2.4.1. Heterogeneous Impacts

In this section, we study how immigrant exposure impacts natives' outcomes based on natives' characteristics as well as *immigrant* characteristics. These include students racial and ethnic background (White, Black, Hispanic and Asian) as well as whether they qualify for FRL or not. Because FRL status changes from year to year, for this analysis, we characterize a student as being

²⁴Specifically, Carrell et al., 2018 find that a 4 percentage point increase in disruptive peers leads to declines in college enrollment of approximately 2 percent and completion of 8–9 percent; Chetty et al., 2016 find that moving to a lower-poverty neighborhood increases college attendance by 2.5 percentage points; and Chetty et al., 2014 find that a one standard deviation improvement in teacher value-added for one year raises college attendance by just 0.8 percentage points.

TABLE 2.3. Effects of Immigrant Peers on College Enrollment and Degree Attainment

	(1)	(2)	(3)	(4)
Panel A: High School Completion				
Exposure	-0.557*** (0.028) [-0.022]	-0.437*** (0.027) [-0.017]	0.015 (0.022) [0.001]	0.040 (0.038) [0.002]
Mean (Y)	0.809	0.809	0.809	0.809
Panel B: College Enrollment				
Exposure	-0.716*** (0.034) [-0.028]	-0.496*** (0.031) [-0.019]	-0.102*** (0.027) [-0.004]	-0.099* (0.046) [-0.004]
Mean (Y)	0.595	0.595	0.595	0.595
Panel C: Any Degree				
Exposure	-0.456*** (0.028) [-0.018]	-0.287*** (0.024) [-0.011]	-0.005 (0.021) [-0.000]	0.015 (0.037) [0.001]
Mean (Y)	0.211	0.211	0.211	0.211
Panel D: 4-Year Degree				
Exposure	-0.413*** (0.026) [-0.016]	-0.259*** (0.023) [-0.010]	-0.010 (0.019) [-0.000]	0.004 (0.033) [0.000]
Mean (Y)	0.167	0.167	0.167	0.167
Individual Controls	Yes	Yes	Yes	Yes
Race FE		Yes		
Free Lunch Status		Yes		
Campus-Year FE	Yes	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes	Yes
Family FE			Yes	
Family-Year FE				Yes
R Squared	0.147	0.174	0.692	0.710
Mean (X)	0.025	0.025	0.025	0.025
N	3,545,720	3,545,720	3,545,720	3,545,720
Distinct Students	431,522	431,522	431,522	431,522

Notes: This table reports estimates of the impact of immigrant exposure on college outcomes. Each column within a panel presents estimates of β_1 from a separate regression of Equation 2.3. The sample includes first grade cohorts from 1994 to 2000 who were never classified as an English Language Learner. Cumulative immigrant exposure is calculated as the average immigrant share across all grades up to the current grade, with students reweighted by their frequency of appearance in the regression. College enrollment is measured within 4 years of expected high school graduation, and college completion is measured within 6 years of expected high school graduation. Individual controls include gender, age as of September 1, special education status, and birth order. Standardized coefficients are shown in brackets. Coefficients are standardized by multiplying the estimated coefficient by the standard deviation of the treatment variable. Standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

TABLE 2.4. Effects of Immigrant Peers on Labor Force Outcomes (Seven Years Post HS Graduation)

	(1)	(2)	(3)	(4)
Panel A: Employed 7 Years				
Exposure	-0.253*** (0.025) [-0.010]	-0.304*** (0.025) [-0.012]	-0.049 (0.029) [-0.002]	-0.038 (0.050) [-0.001]
Mean (Y)	0.713	0.713	0.713	0.713
Panel B: Income Rank 7 Years				
Exposure	-20.496*** (1.476) [-0.798]	-19.814*** (1.459) [-0.771]	-0.346 (1.613) [-0.013]	0.897 (2.801) [0.035]
Mean (Y)	59.069	59.069	59.069	59.069
Panel C: Income Rank 7 Years Employed Only				
Exposure	-25.544*** (3.541) [-0.937]	-14.117*** (3.443) [-0.518]	-0.102 (4.676) [-0.004]	-2.445 (8.698) [-0.090]
Mean (Y)	56.355	56.355	56.355	56.355
Individual Controls	Yes	Yes	Yes	Yes
Race FE		Yes		
Free Lunch Status		Yes		
Campus-Year FE	Yes	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes	Yes
Year Post Graduation FE	Yes	Yes	Yes	Yes
Family FE			Yes	
Family-Year FE				Yes
R Squared	0.045	0.048	0.565	0.583
Mean (X)	0.025	0.025	0.025	0.025
N	3,545,720	3,545,720	3,545,720	3,545,720
Distinct Students	431,522	431,522	431,522	431,522

Notes: This table reports estimates of the impact of immigrant exposure on labor market outcomes. Each column within a panel presents estimates of β_1 from a separate regression of Equation 2.3. The sample includes first grade cohorts from 1994 to 2000 who were never classified as an English Language Learner. Cumulative immigrant exposure is calculated as the average immigrant share across all grades up to the current grade, with students reweighted by their frequency of appearance in the regression. Employment outcomes are measured 7 years post expected high school graduation. Income Rank 7 Years represents an individual's percentile rank in the income distribution measured seven years after their expected high school graduation year, where Panel B includes individuals with zero income and Panel C focuses on only those who are employed. Individual controls include gender, age as of September 1, special education status, and birth order. Standardized coefficients are shown in brackets. Coefficients are standardized by multiplying the estimated coefficient by the standard deviation of the treatment variable. Standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

FRL if they ever were eligible for FRL.²⁵ To do this, we estimate the same equation as in equation 2.3 but restricting the sample to students of different characteristics.

Tables 2.5 and 2.6 present heterogeneity analysis by native student characteristics, with all models including family-by-year fixed effects. Table 2.5 focuses on short-run test score outcomes. The modest overall declines in reading scores are driven primarily by White and non-FRL students. For White students, a one standard deviation increase in immigrant peer exposure reduces reading scores by 0.007 standard deviations and math scores by 0.008 standard deviations, both significant at the 5 percent level. Non-FRL students show a similar pattern, with reductions of 0.004 and 0.006 standard deviations for reading and math respectively, although neither coefficient is statistically significant. All other subgroups show positive coefficients, with Hispanic students displaying a statistically significant positive effect on reading scores of 0.011 standard deviations. Asian students are the only subgroup with an economically meaningful magnitude, of approximately 0.034 standard deviation in both math and reading, though the coefficient does not reach statistical significance at conventional levels given the considerably smaller sample size. Math scores follow a very similar pattern, except the significant effect for Hispanic students disappears and the sign for FRL students flips, though the effect size remains essentially zero.

Table 2.6 turns to long-run outcomes, and the patterns broadly mirror those for test scores, though with some notable exceptions. White and non-FRL students experience the most consistently negative impacts on educational attainment, including a 0.5 percentage point (1%) decline in high school graduation and a 0.3 percentage point decline in receipt of any postsecondary degree. Black students, by contrast, show a significant positive effect on high school graduation (0.6 percentage points or 1%), suggesting immigrant peer exposure may benefit this group on at least one key attainment margin. Hispanic students show a significant positive effect on college enrollment of 0.5 percentage points (or 3.6%), suggesting that immigrant peers may open pathways to postsecondary education for this group. Turning to labor market outcomes, employment seven years after expected high school graduation is positive across all subgroups but largely statistically insignificant. The one exception is White students, for whom the effect is statistically significant, amounting to a

²⁵We also observe students' races changing across the study period. In this analysis, we define someone's race as their race when they first enter Texas Public Schools.

TABLE 2.5. Impact of Immigrant Peer on Test Scores of Natives by Native Characteristics

	Race/Ethnicity				Income	
	White	Black	Hispanic	Asian	FRL	Non-FRL
<i>Panel A: Reading Test Scores</i>						
Exposure	−0.228*	0.099	0.230*	0.836	0.016	−0.126
	(0.092)	(0.199)	(0.115)	(0.599)	(0.110)	(0.081)
	[−0.007]	[0.004]	[0.011]	[0.034]	[0.001]	[−0.004]
Mean (Y)	0.389	−0.248	−0.097	0.779	−0.183	0.407
<i>Panel B: Math Test Scores</i>						
Exposure	−0.267*	0.066	0.180	0.841	−0.060	−0.185
	(0.111)	(0.231)	(0.131)	(0.772)	(0.123)	(0.097)
	[−0.008]	[0.003]	[0.009]	[0.035]	[−0.003]	[−0.006]
Mean (Y)	0.438	−0.276	−0.086	0.913	−0.171	0.448
Mean (X)	0.020	0.038	0.043	0.039	0.036	0.024
Distinct Students (Panel A)	210,685	30,723	87,465	2,448	148,793	254,028
Observations (Panel A)	1,019,341	142,467	421,803	10,053	548,948	1,092,108
Distinct Students (Panel B)	207,969	30,156	86,161	2,442	145,370	250,414
Observations (Panel B)	992,495	138,193	408,653	9,977	528,528	1,067,769

Columns show separate regressions for subgroups of native-born students defined by race/ethnicity or free/reduced-price lunch (FRL) status, both measured in first grade. Each column within a panel presents estimates of β_1 from a separate regression of Equation 2.3. The sample includes first grade cohorts from 1994 to 2000 who were never classified as an English Language Learner. Immigrant exposure is calculated as the average immigrant share across all grades through the current grade, reweighting students by appearance frequency. The model includes family-by-year fixed effects. See Table 2.2 for the full set of controls. Standard errors in parentheses are clustered at the campus level. Standardized coefficients in brackets are computed by multiplying the estimated coefficient by the standard deviation of the treatment variable. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

0.2 percentage point increase in employment likelihood. This may reflect the modest reduction in educational attainment for this group, as students who do not continue to graduate school are more likely to be in the labor force. Income rank follows a similar pattern, with negative coefficients across most subgroups that are not statistically significant, and a positive but noisy estimate for Asian students. Taken together, the heterogeneity results suggest that the modest negative average effects documented in the main analysis are concentrated among White and non-FRL students, while minority and lower-income students are largely unaffected or modestly benefited.

To better understand how each subgroup is affected by the inclusion of family fixed effects, Figures 2.6 and 2.7 plot these results graphically with fixed effects included iteratively. Effects attenuate across all subgroups as more demanding fixed effects are added, with somewhat larger attenuation for White and non-FRL students once family fixed effects are included. Black and FRL students show comparatively less attenuation, consistent with family-level sorting playing a smaller role in driving their estimates.

If peer effects operate through specific characteristics of immigrant students, impacts should vary depending on the types of immigrants to whom natives are exposed. Tables 2.7 and 2.8 examine this by replacing overall immigrant exposure with exposure to immigrants with specific characteristics (e.g., Hispanic, Black, or FRL). Table 2.7 reports results for short-run outcomes. Across different definitions of immigrant peers, reading and math scores follow similar patterns. Exposure to White and Asian immigrant peers is associated with small positive effects (roughly 2 to 4 percent of the outcome mean), although the coefficient for White immigrants on math scores is only marginally significant. In contrast, exposure to Black and Hispanic immigrant peers is associated with small negative effects, with estimates for Hispanic immigrants more than twice as large in absolute value as those for Black immigrants. By economic disadvantage, exposure to FRL immigrants is associated with small negative effects, while exposure to non-FRL immigrants yields modest positive effects.

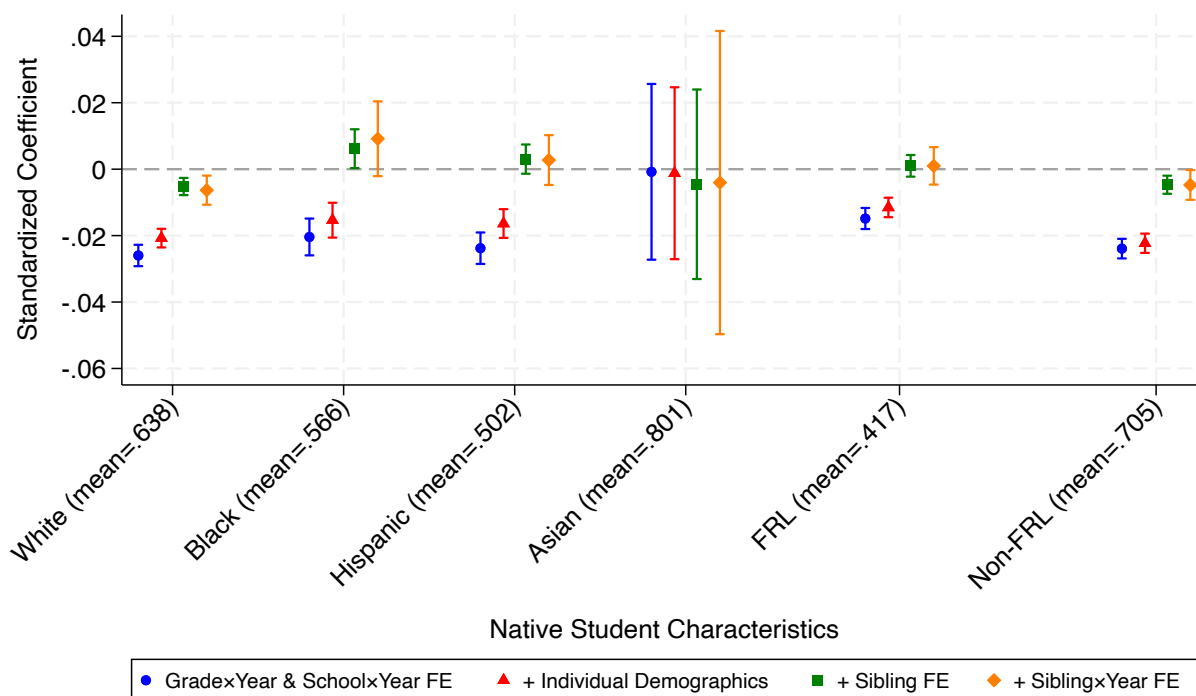
Table 2.8 presents results for long-run outcomes. Exposure to White, Asian, and non-FRL immigrant peers is associated with small increases in high school graduation, while exposure to Hispanic and FRL immigrant peers is linked to modest declines in college enrollment. White immigrant peers also generate a positive effect on college graduation. All of these attainment effects

TABLE 2.6. Effects of Immigrant Peers on Long-Run Outcomes by Native Student Characteristics

	Race/Ethnicity				Income	
	White	Black	Hispanic	Asian	FRL	Non-FRL
<i>Panel A: High School Graduation</i>						
Exposure	-0.175*** (0.044) [-0.005]	0.141* (0.068) [0.006]	0.061 (0.046) [0.003]	-0.108 (0.346) [-0.005]	0.022 (0.037) [0.001]	-0.139*** (0.041) [-0.005]
Mean (Y)	0.639	0.566	0.503	0.801	0.418	0.706
<i>Panel B: College Enrollment 4 Years Post-Graduation</i>						
Exposure	-0.001 (0.038) [-0.000]	-0.005 (0.049) [-0.000]	0.107*** (0.032) [0.005]	-0.236 (0.377) [-0.010]	0.027 (0.024) [0.001]	0.006 (0.038) [0.000]
Mean (Y)	0.257	0.136	0.129	0.479	0.084	0.290
<i>Panel C: Any Postsecondary Degree 6 Years Post-Graduation</i>						
Exposure	-0.097* (0.047) [-0.003]	-0.097 (0.072) [-0.004]	-0.011 (0.051) [-0.001]	0.098 (0.466) [0.004]	0.017 (0.044) [0.001]	-0.096* (0.044) [-0.003]
Mean (Y)	0.701	0.727	0.744	0.597	0.712	0.715
<i>Panel D: Employed 7 Years Post-Graduation</i>						
Exposure	0.078* (0.033) [0.002]	0.074 (0.060) [0.003]	0.075 (0.040) [0.004]	0.379 (0.256) [0.016]	0.040 (0.035) [0.002]	0.038 (0.030) [0.001]
Mean (Y)	0.837	0.778	0.756	0.928	0.700	0.878
<i>Panel E: Income Rank 7 Years Post-Graduation</i>						
Exposure	-0.556 (2.648) [-0.017]	-6.828 (3.833) [-0.298]	-0.106 (2.852) [-0.005]	26.864 (25.978) [1.130]	0.899 (2.304) [0.040]	-0.719 (2.569) [-0.024]
Mean (Y)	59.832	56.223	58.789	54.143	56.205	60.884
Mean (X)	0.018	0.035	0.039	0.037	0.031	0.021
N	263,340	39,891	106,001	3,173	216,038	331,955
Distinct Students	2,146,170	328,958	869,067	21,167	1,271,244	2,173,343

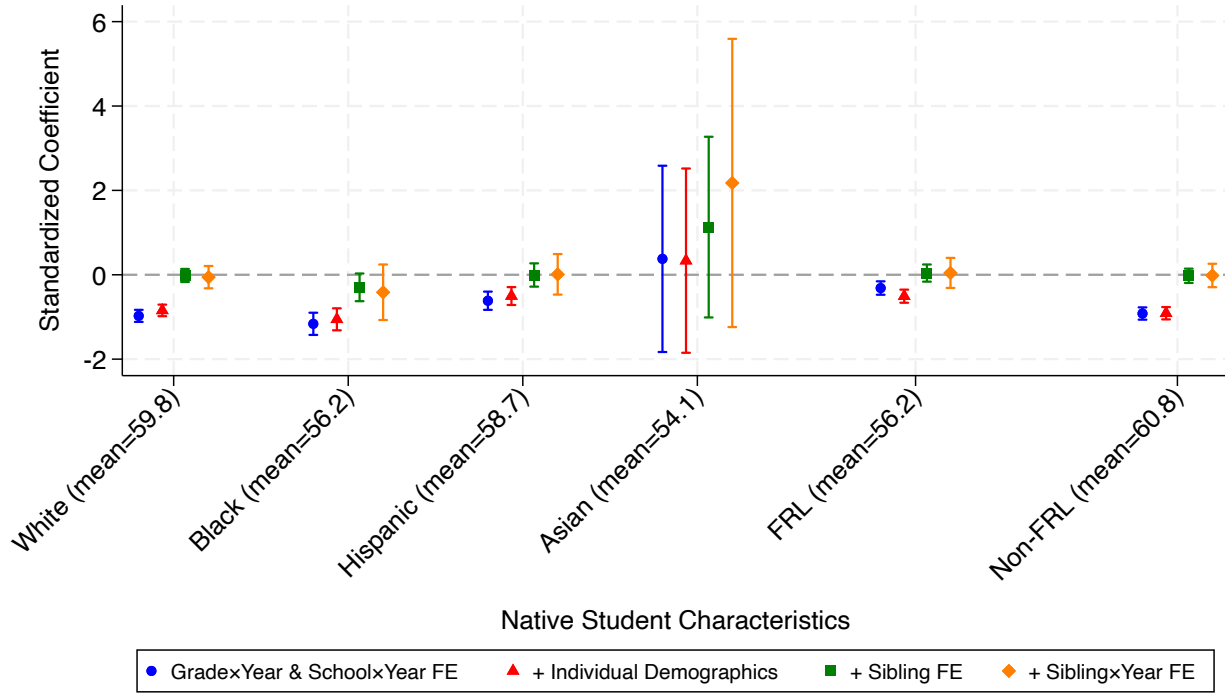
Notes: This table presents heterogeneous treatment effects by native student characteristics. Columns show separate regressions for subgroups defined by race/ethnicity or free/reduced-price lunch (FRL) status, both measured in first grade. Each column within a panel presents estimates of β_1 from a separate regression of Equation 2.3. The sample includes first grade cohorts from 1994 to 2000 who were never classified as an English Language Learner. The model includes family-by-year fixed effects. See Tables 2.3 and 2.4 for more details on the full list of controls and how the long-run outcomes are defined. Standardized coefficients are shown in brackets and are computed by multiplying the estimated coefficient by the standard deviation of the treatment variable. Standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

FIGURE 2.6. Impact of Immigrant Exposure on College Enrollment by Native Student Characteristics



Note: This figure plots regression coefficients and 95% confidence intervals from Equation 2.3 where the dependent variable is college attendance within four years of expected high school graduation. Each dot represents the coefficient estimate from a separate regression estimated on a specific subgroup of native-born students defined by either race/ethnicity or free or reduced-price lunch (FRL) eligibility. Blue circles show models with grade-by-year and school-by-year fixed effects, red triangles additionally include individual demographics, green squares additionally include family fixed effects, and orange diamonds additionally include family-by-year fixed effects. The mean outcome for each subgroup is shown in parentheses below the characteristic label. Standard errors are clustered at the school level. The sample includes first grade cohorts from 1994 to 2000 who were never classified as an English Language Learner.

FIGURE 2.7. Impact of Immigrant Exposure on Earnings Percentile by Native Student Characteristics



Note: This figure plots regression coefficients and 95% confidence intervals from Equation 2.3, where the dependent variable is an individual's percentile rank in the income distribution, measured seven years after their expected high school graduation, with zero earners included. Each dot represents the coefficient estimate from a separate regression estimated on a specific subgroup of native-born students defined by either race/ethnicity or free or reduced-price lunch (FRL) eligibility. Blue circles show models with grade-by-year and school-by-year fixed effects, red triangles additionally include individual demographics, green squares additionally include family fixed effects, and orange diamonds additionally include family-by-year fixed effects. The mean outcome for each subgroup is shown in parentheses below the characteristic label. Standard errors are clustered at the school level. The sample includes first grade cohorts from 1994 to 2000 who were never classified as an English Language Learner.

TABLE 2.7. Impact of Immigrant Peer on Test Scores of Natives by Immigrant Characteristics

	Race/Ethnicity				Income	
	White	Black	Hispanic	Asian	FRL	Non-FRL
<i>Panel A: Reading Test Scores</i>						
Exposure	0.575** (0.194) [0.004]	-1.311* (0.543) [-0.004]	-0.367*** (0.074) [-0.013]	0.870*** (0.210) [0.007]	-0.378*** (0.073) [-0.013]	0.563*** (0.136) [0.007]
Mean (Y)	0.200	0.200	0.200	0.200	0.200	0.200
Mean (X)	0.002	0.001	0.022	0.003	0.022	0.007
N	1,696,169	1,696,169	1,696,169	1,696,169	1,696,169	1,696,169
Students	348,134	348,134	348,134	348,134	348,134	348,134
<i>Panel B: Math Test Scores</i>						
Exposure	0.584* (0.267) [0.004]	-1.819** (0.650) [-0.005]	-0.434*** (0.084) [-0.015]	0.727** (0.257) [0.006]	-0.433*** (0.084) [-0.015]	0.405* (0.173) [0.005]
Mean (Y)	0.232	0.232	0.232	0.232	0.232	0.232
Mean (X)	0.002	0.001	0.022	0.003	0.022	0.007
N	1,649,024	1,649,024	1,649,024	1,649,024	1,649,024	1,649,024
Students	343,360	343,360	343,360	343,360	343,360	343,360

Notes: Columns show separate regressions where immigrant exposure is decomposed by race/ethnicity or free/reduced-price lunch (FRL) status of immigrant peers. Each column within a panel presents estimates of β_1 from a separate regression of Equation 2.3. The sample includes first grade cohorts from 1994 to 2000 who were never classified as an English Language Learner. Immigrant exposure is calculated as the average immigrant share (by subgroup) across all grades through the current grade, reweighting students by appearance frequency. The model includes family-by-year fixed effects. See Table 2.2 for the full set of controls. Standardized coefficients in brackets are computed by multiplying the estimated coefficient by the standard deviation of the treatment variable. Standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

correspond to roughly 1 percent of the mean. Labor market outcomes show little systematic variation across immigrant characteristics. Employment effects are generally small and imprecise, though exposure to Asian and non-FRL immigrant peers is associated with statistically significant reductions in employment. Again, this may reflect increases in educational attainment, as more individuals may be enrolled in graduate school and therefore not participating in the labor force. For earnings, we find statistically significant positive effects from exposure to Black immigrant peers and statistically significant negative effects from exposure to non-FRL immigrant peers. But across the board, these effects are imprecise. Taken together, these results suggest that while the characteristics of immigrant peers appear to matter for the direction of estimated effects, the magnitudes remain modest overall and do not point to large systematic impacts on native students.

Finally, Appendix Table 2.12 examines heterogeneity by school locale; effects, to the extent they appear, are primarily concentrated in city schools, consistent with Getik et al., 2024.

2.5. Robustness

This section assesses the robustness of our main results along three dimensions. First, we examine sensitivity to within-family sorting, addressing the possibility that parents make schooling decisions based on each child’s characteristics. Second, we investigate whether our results reflect immigrant status itself or the broader peer characteristics with which immigrant exposure is correlated. Finally, we assess sensitivity to sample construction and our definition of immigrant peers. Across all checks, our main findings remain stable.

2.5.1. Assessing Sensitivity to Within Family Sorting

Our main identification assumption requires that families select schools similarly across their children. However, if parents make schooling decisions based on each child’s characteristics, siblings may experience different levels of immigrant exposure in ways that correlate with those characteristics. For example, if parents believe immigrants negatively affect academic performance, they might send the higher achieving child to a school with lower immigrant peer exposure to protect that child’s academic advantage. This would generate downward bias in our estimates, since the child with less immigrant exposure would have performed better anyway. Conversely, if parents compensate for

TABLE 2.8. Impact of Immigrant Peer on Long-Run Outcomes of Natives by Immigrant Characteristics

	Race/Ethnicity				Income	
	White	Black	Hispanic	Asian	FRL	Non-FRL
<i>Panel A: High School Graduation</i>						
Exposure	0.212** (0.073) [0.002]	0.110 (0.190) [0.000]	-0.030 (0.026) [-0.001]	0.260** (0.088) [0.002]	-0.015 (0.026) [-0.001]	0.156** (0.052) [0.002]
Mean (Y)	0.809	0.809	0.809	0.809	0.809	0.809
<i>Panel B: College Enrollment 4 Years Post-Graduation</i>						
Exposure	-0.101 (0.100) [-0.001]	-0.014 (0.261) [-0.000]	-0.126*** (0.030) [-0.004]	0.018 (0.105) [0.000]	-0.140*** (0.030) [-0.005]	0.016 (0.075) [0.000]
Mean (Y)	0.596	0.596	0.596	0.596	0.596	0.596
<i>Panel C: Any Postsecondary Degree 6 Years Post-Graduation</i>						
Exposure	0.262*** (0.078) [0.002]	0.055 (0.212) [0.000]	-0.027 (0.024) [-0.001]	-0.134 (0.094) [-0.001]	-0.034 (0.024) [-0.001]	0.111 (0.058) [0.001]
Mean (Y)	0.211	0.211	0.211	0.211	0.211	0.211
<i>Panel D: Employed 7 Years Post-Graduation</i>						
Exposure	-0.083 (0.109) [-0.001]	0.224 (0.263) [0.001]	-0.040 (0.033) [-0.001]	-0.228* (0.108) [-0.002]	-0.018 (0.034) [-0.001]	-0.211** (0.072) [-0.003]
Mean (Y)	0.713	0.713	0.713	0.713	0.713	0.713
<i>Panel E: Income Rank 7 Years Post-Graduation</i>						
Exposure	1.408 (6.090) [0.011]	29.652* (14.403) [0.080]	-0.531 (1.835) [-0.018]	-7.014 (6.648) [-0.059]	1.506 (1.859) [0.050]	-8.546* (4.065) [-0.106]
Mean (Y)	59.069	59.069	59.069	59.069	59.069	59.069
Mean (X)	0.002	0.001	0.019	0.003	0.019	0.006
N	3,545,720	3,545,720	3,545,720	3,545,720	3,545,720	3,545,720
Students	431,522	431,522	431,522	431,522	431,522	431,522

Notes: Columns show separate regressions where immigrant exposure is decomposed by race/ethnicity or free/reduced-price lunch (FRL) status of immigrant peers. Each column within a panel presents estimates of β_1 from a separate regression of Equation 2.3. The sample includes first grade cohorts from 1994 to 2000 who were never classified as an English Language Learner, with outcomes observed through 2019. Immigrant exposure is calculated as the average immigrant share (by subgroup) across all grades through the current grade, reweighting students by appearance frequency. See Tables 2.3 and 2.4 for more details on the full set of controls and how the long-run outcomes are defined. Standardized coefficients in brackets are computed by multiplying the estimated coefficient by the standard deviation of the treatment variable. Standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

ability differences by sending their lower-achieving child to a school with lower immigrant share, this would lead to an upward bias, since the child with less exposure would have performed worse anyway. While most families during this period likely enrolled all their children in the same residentially assigned school, the expansion of school choice options (open enrollment and charter schools) made it increasingly feasible for parents to select different schools for different siblings. Texas was an early adopter of school-choice policies, so this is a legitimate concern.²⁶

To address whether within-family selection is a concern in our setting, we employ an instrumental variable strategy following (Figlio et al., 2021) who also face a similar issue identifying the effect of immigrant exposure on native academic outcomes. Within-family sorting can occur in two ways. First, parents may enroll siblings in the same initial school but choose different schooling paths thereafter (e.g., transferring one child but not another). Second, parents may send siblings to different initial schools. We employ two separate IV strategies to address these concerns, both of which exploit population-level patterns of school transitions across grades to predict immigrant exposure under different assumptions about parental sorting. Specifically, we construct a transition matrix that captures population-level school-to-school transition probabilities across grades using data from 1994–2000. By combining these transition probabilities with the observed immigrant composition of each school-grade-year cell, we predict each student’s immigrant peer exposure under the assumption that they followed the *modal pathways* of their cohort rather than making endogenous school choices. Full details are in Appendix 2.9.

Addressing different schooling paths after same initial school

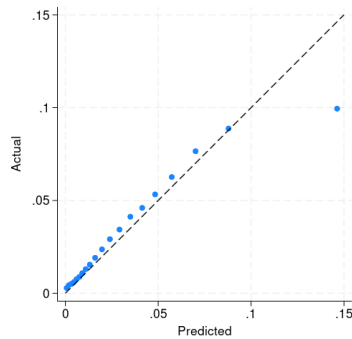
First, we examine the importance of different schooling paths after starting at the same school. For this analysis, we restrict the sample to siblings who attended the same initial school. Then we use predicted immigrant exposure (based on their 1st grade cohort and initial school) as an instrument for actual immigrant exposure. Within families, the only difference in predicted immigrant peer exposure comes from cohort variation. This approach identifies the local average treatment effect (LATE) of immigrant peer exposure for families who did not move in response to immigrant peers, precisely the group for whom differences in immigrant shares is most exogenous.

²⁶Texas was an early adopter of school-choice policies, becoming the fourth state in the nation to authorize charter schools in 1995.

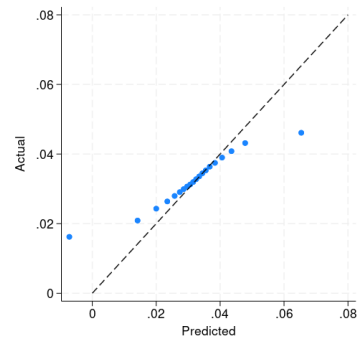
This IV approach relies on two key identifying assumptions. First, predicted exposure must be correlated with actual immigrant exposure (the relevance or first-stage assumption). As demonstrated by the first-stage F-statistics reported in Table 2.9 (ranging from 46 to 59 across samples), predicted immigrant peer exposure strongly predicts actual exposure. Figure 2.8 illustrates this relationship visually. Panel A plots binned means of actual against predicted cumulative immigrant exposure. The strong positive relationship confirms instrument relevance, while the fact that points fall systematically below the 45-degree line at higher exposure levels is consistent with families sorting away from high-immigrant schools. Panel B repeats this exercise after absorbing family fixed effects. The positive relationship persists, confirming that the instrument has power even within families, and the continued attenuation below the 45-degree line is consistent with child-specific sorting that the IV is designed to address.

FIGURE 2.8. First Stage: Actual vs. Predicted Immigrant Peer Exposure

Panel A: Without Fixed Effects



Panel B: Absorbing Family Fixed Effects



Note: This figure plots binned means of actual cumulative immigrant peer exposure against predicted cumulative exposure from the transition matrix, for siblings who attended the same initial school. The dashed line represents the 45-degree line where actual equals predicted exposure. Panel A shows the unconditional relationship. Panel B shows the relationship after absorbing family fixed effects, isolating within-family variation. The sample includes first grade cohorts from 1994 to 2000 who were never classified as an English Language Learner.

Second, the exclusion restriction must hold. That is, predicted exposure should affect student outcomes only through actual exposure. This assumption is plausible because siblings' entry cohorts are determined by birth timing, which predates and is independent of subsequent immigrant inflows to their school. Conditional on school fixed effects, cohort variation in predicted exposure has no direct pathway to affect outcomes other than through actual peer composition.

Table 2.9 presents the IV estimates. Column 1 reports IV estimates that instrument actual cumulative exposure with predicted exposure from the transition matrix. Column 2 presents the

reduced-form effect of predicted exposure on student outcomes. For comparison, Column 3 reports OLS estimates of the effect of actual immigrant exposure on outcomes on the same sample. If parents differentially select schools for their children based on individual characteristics, the OLS estimates will be biased. The IV estimates in Column 1 address this concern. Across most outcomes, the IV point estimates are statistically insignificant and the standardized coefficients are close to zero, suggesting economically negligible effects. Notably, the negative OLS estimate for college enrollment (-0.165 , $p < 0.001$) shrinks to near zero in the IV specification (-0.027), suggesting that the OLS estimate reflects residual sorting on child-specific characteristics rather than a causal effect of immigrant exposure. The one exception is reading test scores, where the IV estimate is negative and statistically significant (-0.633 , $p < 0.05$), though the standardized coefficient remains small (-0.034). Taken together, these results suggest that within-family sorting based on child-specific characteristics is unlikely to substantially bias our main family fixed effects results.

Addressing differences in initial school

Finally, we relax the restriction that siblings must start at the same school. We expand the sample to include all siblings regardless of initial school and use predicted immigrant exposure based on each child's cohort and the initial school of the eldest sibling in the family. Again, within families, the only source of variation in predicted immigrant exposure comes from cohort differences. The same exclusion restriction must hold, and for the reasons discussed above, it is likely satisfied in this setting as well.

Appendix Table 2.13 presents these findings. Consistent with the estimates in Table 2.9, the IV estimates are close to zero and smaller than the OLS estimates, suggesting that neither initial school sorting nor subsequent school changes substantially bias our main results.

Taken together, these IV results provide strong evidence that within-family sorting does not drive our main findings. If anything, the IV estimates suggest that the true causal effect of immigrant peer exposure is even smaller than our family fixed effects estimates imply.

TABLE 2.9. Transition Matrix IV: Same Initial School Sample

	(1) IV	(2) Reduced Form	(3) OLS
<i>Panel A: Reading Test Scores</i>			
Cumulative Exposure	−0.633* (0.268) [−0.034]		−0.129 (0.075) [−0.007]
Predicted Exposure		−0.200* (0.088) [−0.009]	
Mean (Y)	0.163	0.163	0.163
SD (Y)	0.886	0.886	0.886
<i>Panel B: Math Test Scores</i>			
Cumulative Exposure	0.655 (0.398) [0.036]		−0.018 (0.088) [−0.001]
Predicted Exposure		0.207 (0.118) [0.009]	
Mean (Y)	0.228	0.228	0.228
SD (Y)	0.921	0.921	0.921
<i>Panel C: Any College Enrollment</i>			
Cumulative Exposure	−0.027 (0.066) [−0.001]		−0.165*** (0.025) [−0.009]
Predicted Exposure		−0.009 (0.022) [−0.000]	
Mean (Y)	0.635	0.635	0.635
SD (Y)	0.481	0.481	0.481
<i>Panel D: Income Rank 7 Years Post-Graduation</i>			
Cumulative Exposure	6.622 (4.043) [0.367]		−1.102 (1.545) [−0.061]
Predicted Exposure		2.192 (1.358) [0.094]	
Mean (Y)	59.775	59.775	59.775
SD (Y)	27.061	27.061	27.061
Individual Controls	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes
Family × Initial School FE	Yes	Yes	Yes
N (Panels A–B)	1,251,532	1,251,532	1,251,532
Distinct Students (Panels A–B)	168,639	168,639	168,639
N (Panels C–D)	1,768,127	1,768,127	1,768,127
Distinct Students (Panels C–D)	185,021	185,021	185,021
Mean (X)	0.045	0.039	0.045
SD (X)	0.054	0.043	0.054
R-Squared (Panel A)	—	0.616	0.616
First Stage F (Panels A–B)	46.1		
First Stage F (Panels C–D)	59.2		

Notes: This table reports two-stage least squares (Column 1), reduced form (Column 2), and OLS (Column 3) estimates of the effect of cumulative immigrant peer exposure on student outcomes. Predicted immigrant peer exposure is constructed using a transition matrix that captures population-level school-to-school transition probabilities across grades. For each student, predicted exposure is computed by combining their starting school and cohort with the transition matrix and the observed immigrant composition of each destination school-grade-year cell. The sample is restricted to siblings who attended the same initial (first grade) school. Predicted exposure is based on each student's own cohort and their shared initial school, so within-family variation in predicted exposure comes solely from cohort differences. All specifications include family-by-initial-school and grade-by-year fixed effects, following Figlio et al., 2021. Panels A–B use the short-run sample (students with non-missing test scores); Panels C–D use the long-run sample (students with non-missing college and labor market outcomes). The sample includes first grade cohorts from 1994 to 2000 who were never classified as an English Language Learner. Individual controls include gender, age as of September 1, special education status, and birth order. Standardized coefficients are shown in brackets and are computed by multiplying the estimated coefficient by the standard deviation of the treatment variable. Standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

2.5.2. Immigrant Status or Other Characteristics?

Next, we examine whether our results reflect immigrant status itself or the broader peer characteristics with which immigrant exposure is correlated. As a first step, Appendix Table 2.14 documents these correlations descriptively. In models with family or family-by-year fixed effects, immigrant exposure is strongly positively correlated with exposure to Hispanic and ELL peers, though notably uncorrelated with FRL exposure. We also find a positive correlation with Asian peers. These patterns reflect the underlying composition of immigrant students in Texas and confirm that immigrant exposure is not orthogonal to other dimensions of peer composition, motivating the augmented specifications we turn to next.

To assess whether our main results are driven by these correlated characteristics, we augment our baseline specification by sequentially adding cumulative exposure to race, FRL, and ELL peers alongside immigrant exposure. Appendix Table 2.15 presents the results. For test scores, including these additional controls eliminates and reverses the estimated effect of immigrant exposure, which shifts from negative and significant (-0.226) to positive and insignificant (0.108). This reversal is driven by correlated peer characteristics — particularly ELL status. Notably, immigrant exposure is most strongly correlated with ELL peer shares within families (coefficient 0.139 , Appendix Table 2.14), while the correlation with FRL peer shares is negligible. This suggests that the language acquisition needs of immigrant peers, rather than their socioeconomic status per se, explain the composition effects we document for test scores. Consistent with this, adding ELL peer exposure controls is what primarily drives the reversal of the test score coefficient.

For longer-run outcomes, coefficients are largely stable. College enrollment and high school graduation retain significance even in the fully saturated specification, but these effects attenuate to zero under the IV strategy (Table 2.9), suggesting they reflect residual within-family sorting rather than a true causal effect. For degree attainment, employment, and earnings, effects remain insignificant throughout, consistent with Figlio et al., 2021.

2.5.3. Alternative Sample and Immigrant Definitions

As a final robustness check, we examine the sensitivity to sample construction and immigrant definitions. Our baseline specification restricts the native sample to students never classified as ELL, isolating a population unlikely to include first or second generation immigrants. Appendix Table 2.16 shows in Column (2) that results are similar when we relax this restriction. As this broader sample likely includes second generation immigrants, the more negative estimates may reflect resource competition between native ELL students and their immigrant peers for limited ELL services.

We next explore sensitivity to alternative immigrant definitions, finding that conclusions are similar across all specifications, though exact magnitudes and significance vary slightly. Column 3 restricts to recent immigrants only, defined as those who arrived within the last three years, and yields slightly more negative estimates, consistent with these children likely requiring more school resources upon arrival. Column 4 considers a broader definition that classifies as immigrant any student not born in Texas who was classified as ELL at school entry, to capture cases where our baseline immigrant measure may miss some children, and finds similar results. Finally, a small number of students in our data are flagged as immigrants despite being born in Texas, which may reflect matching errors, imprecision in the immigrant flag, or cases where individuals were born in Texas, left the country, and later returned to Texas public schools. Our main estimates drop these ambiguous cases, but reassuringly, Column 5 shows that our conclusions are unaffected if we instead include them.

Across all specifications, the central conclusion remains unchanged: immigrant peer exposure has negligible effects on native students' long-run outcomes.

2.6. Conclusion

Immigration in the U.S. has increased significantly over the past several decades. As a result, public schools enroll an increasing number of immigrant students. This paper examined the effects of exposure to immigrant peers on the educational and labor market trajectories of U.S.-born students. Leveraging rich administrative data from Texas that link K–12 records to higher education, employment, and earnings outcomes, we provided one of the first comprehensive analyses of immigrant

peer spillovers extending well into adulthood. By exploiting sibling fixed effects within families, we overcome the long-standing challenge of endogenous sorting—where native families self-select into or out of schools in response to changing immigrant concentrations. Our findings help clarify how demographic change in schools translates into longer-term human capital and labor market outcomes for native students.

Our results highlight that the initial negative correlations between immigrant exposure and native achievement are largely explained by selection. Without controlling for family, greater exposure to immigrant peers is associated with lower test scores and slightly reduced college enrollment. Yet once we account for family fixed effects these negative effects diminish substantially. The attenuation of these effects suggests that the negative correlations between immigrant exposure and native outcomes reflect selection rather than a causal effect of immigrant peers. These conclusions are reinforced by our instrumental variables analysis, which accounts for within-family sorting by predicting exposure based on typical schooling pathways rather than parental choices. The IV estimates are consistently close to zero across all outcomes, including college enrollment, confirming that residual within-family sorting does not drive our findings.

Heterogeneity analyses reveal that the effects of immigrant peer exposure vary by both native and immigrant characteristics. Negative impacts are concentrated among White and non-FRL natives, while Black and Hispanic natives tend to benefit, with small but statistically significant positive effects on high school graduation and college enrollment, respectively. Exposure to White, Asian, and non-FRL immigrant peers is associated with small positive spillovers, while exposure to Hispanic and FRL immigrant peers is linked to modest negative effects. This heterogeneity underscores that “immigrant exposure” is not a uniform experience, but all magnitudes remain small and the overall picture is one of limited long-run benefit or harm to native-born students.

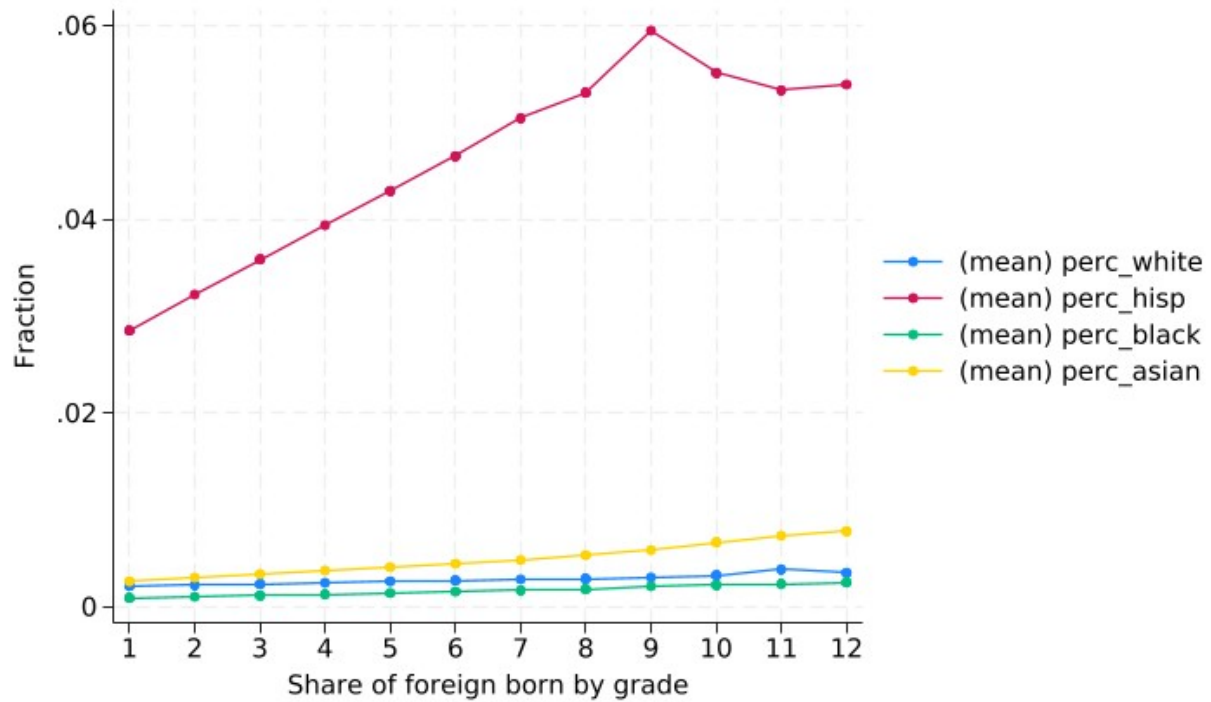
From a policy perspective, these results provide reassurance that increasing immigrant enrollment in public schools does not undermine the long-run prospects of native students. The absence of meaningful negative effects, even in a setting with large and rapidly growing immigrant populations, suggests that concerns about native harm from immigrant peer exposure may be overstated.

Finally, this work contributes to broader discussions about diversity, social integration, and

mobility in the United States. Just as earlier periods of racial desegregation revealed both challenges and long-term gains, the demographic transformation brought about by immigration can yield benefits that extend beyond test scores or wages—shaping tolerance, cooperation, and civic cohesion in future generations. Understanding these broader social spillovers remains an important avenue for future research.

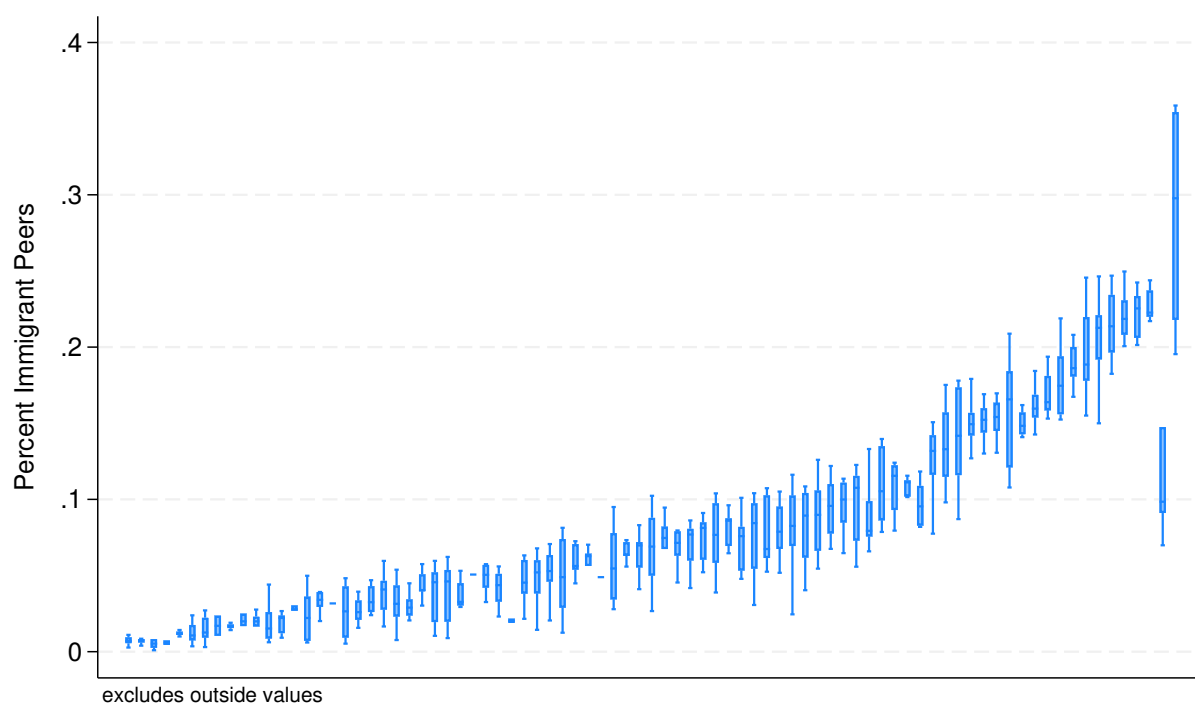
2.7. Appendix Tables and Figures

FIGURE 2.9. Immigrant Share of Total Enrollment in Texas Public Schools, by Race and Grade



Note: This figure plots cumulative immigrant exposure by grade and race, pooling all 1st grade cohorts. Immigrant exposure is measured as the average cumulative share of peers classified as immigrants through each grade, as defined in Equation 2.2.

FIGURE 2.10. Within-School Variation in Immigrant Exposure Across Years



Note: This figure displays the distribution of cumulative immigrant exposure through 9th grade across Texas public school districts with enrollment above 1,000 students. Each box represents the interquartile range of cumulative immigrant exposure within a district across 1st grade cohorts, with whiskers extending to the 5th and 90th percentiles. Immigrant exposure is measured as defined in Equation 2.2.

TABLE 2.10. Immigrant Population Composition and Socioeconomic Characteristics by State, 1990–2010

State	Year	Origin Region (%)					Family Characteristics (%)		
		Mexico	Central Am.	Caribbean	Asia	Other	Below 200% FPL	≤HS Edu.	LEP
Texas	1990	48.6	4.5	1.5	14.3	31.1	55.9	47.0	28.2
	2000	56.4	5.6	1.4	14.9	21.7	50.1	49.2	33.5
	2010	54.3	7.0	1.3	16.4	21.0	49.8	56.0	35.7
Florida	1990	2.5	6.0	31.1	5.9	54.6	39.3	63.9	19.1
	2000	5.5	7.1	31.5	7.3	48.5	37.8	69.1	21.7
	2010	6.4	7.7	31.1	8.1	46.7	39.2	74.4	23.7
California	1990	33.5	7.4	1.1	28.0	30.0	43.6	54.8	28.6
	2000	40.4	7.5	0.8	30.2	21.2	42.8	58.6	30.5
	2010	40.0	7.6	0.7	32.5	19.2	39.5	64.7	32.8
New York	1990	1.3	3.7	19.1	16.2	59.8	38.7	60.6	20.0
	2000	3.7	4.6	22.5	20.1	49.2	40.1	66.4	22.8
	2010	5.0	5.2	21.7	23.0	45.1	38.1	71.3	24.9
New Jersey	1990	1.1	2.9	12.1	18.1	65.8	27.7	62.3	19.3
	2000	4.0	5.2	13.7	23.7	53.3	28.7	68.8	22.2
	2010	6.4	6.1	13.4	27.1	47.1	28.9	74.4	23.2
Illinois	1990	21.7	2.0	1.9	17.6	56.8	32.9	60.2	19.7
	2000	33.0	2.4	1.7	19.4	43.6	31.6	62.7	25.0
	2010	33.7	2.4	1.2	21.8	40.9	34.4	69.8	25.7
US	1990	11.2	3.0	5.0	13.7	67.1	36.8	58.1	13.6
	2000	17.8	4.0	5.7	16.3	56.2	36.0	61.8	17.9
	2010	19.2	4.8	5.9	18.0	52.2	36.9	66.7	19.8

Notes: This table reports statistics on immigration to the United States and the six largest immigrant-receiving states. Estimates for 1990 and 2000 are based on the 5% Decennial Census samples; estimates for 2010 come from the 5-year American Community Survey. All figures are weighted to produce nationally representative estimates. We categorize someone as an immigrant if they were born outside of the U.S. To create the LEP or limited English proficiency variable, we use self-reported English-speaking ability from the Census/ACS.

TABLE 2.11. Language Spoken at Home Among Immigrant Students

Language	Share	Language	Share	Language	Share
1. Spanish	0.764	6. Mandarin	0.004	11. Farsi	0.001
2. English	0.130	7. Arabic	0.003	12. German	0.001
3. Vietnamese	0.019	8. French	0.002	13. Gujarati	0.001
4. Korean	0.008	9. Filipino	0.002	14. Other	0.047
5. Urdu	0.006	10. Cantonese	0.002		
Observations	1,363,503	Distinct Students	229,228		

Notes: This table reports the share of immigrant students by language spoken at home, ranked by prevalence. The sample includes immigrant students in first grade cohorts from 1994 to 2003. Languages with a share of less than 0.001 are excluded. Shares sum to one.

TABLE 2.12. Effects of Immigrant Peer Exposure by Locale

	City	Suburb	Town
<i>Panel A: Reading Test Scores</i>			
Exposure	−0.378*** (0.089) [−0.020]	0.094 (0.361) [0.006]	0.309 (0.238) [0.010]
Mean (Y)	0.114	0.011	0.105
Distinct Students	341,583	36,998	144,090
Observations	1,512,054	126,276	544,946
<i>Panel B: Math Test Scores</i>			
Exposure	−0.280** (0.092) [−0.015]	0.149 (0.341) [0.009]	0.275 (0.266) [0.009]
Mean (Y)	0.152	0.056	0.123
Distinct Students	349,314	37,942	147,947
Observations	1,567,591	130,730	566,489
<i>Panel C: College Attendance</i>			
Exposure	−0.177*** (0.036) [−0.010]	0.195 (0.127) [0.012]	−0.189* (0.086) [−0.006]
Mean (Y)	0.579	0.557	0.560
Distinct Students	431,257	52,350	194,587
Observations	2,985,791	252,174	1,056,766
<i>Panel D: Income Rank 7 Years</i>			
Exposure	−2.572 (2.183) [−0.143]	8.526 (7.418) [0.510]	6.401 (5.124) [0.218]
Mean (Y)	58.413	58.039	60.279
Distinct Students	431,284	52,321	194,619
Observations	2,985,860	252,192	1,056,900
Mean (X)	0.046	0.049	0.016

Notes: This table presents heterogeneous treatment effects by locale. Columns show separate regressions for subgroups defined by locale. Each column within a panel presents estimates of β_1 from a separate regression of Equation 2.3. The sample includes first grade cohorts from 1994 to 2000. The model includes family-by-year fixed effects. See Tables 2.3 and 2.4 for more details on the full list of controls and how the long-run outcomes are defined. Standardized coefficients are shown in brackets and are computed by multiplying the estimated coefficient by the standard deviation of the treatment variable. Standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

TABLE 2.13. Transition Matrix IV: Expanded Sample

	(1) IV	(2) Reduced Form	(3) OLS
<i>Panel A: Reading Test Scores</i>			
Cumulative Exposure	−0.409 (0.272) [−0.022]		−0.030 (0.070) [−0.002]
Predicted Exposure		−0.121 (0.083) [−0.005]	
Mean (Y)	0.139	0.139	0.139
SD (Y)	0.896	0.896	0.896
<i>Panel B: Math Test Scores</i>			
Cumulative Exposure	0.682 (0.403) [0.037]		−0.004 (0.083) [−0.000]
Predicted Exposure		0.202 (0.111) [0.009]	
Mean (Y)	0.198	0.198	0.198
SD (Y)	0.928	0.928	0.928
<i>Panel C: Any College Enrollment</i>			
Cumulative Exposure	0.054 (0.068) [0.003]		−0.139*** (0.020) [−0.008]
Predicted Exposure		0.017 (0.021) [0.001]	
Mean (Y)	0.621	0.621	0.621
SD (Y)	0.485	0.485	0.485
<i>Panel D: Income Rank 7 Years Post-Graduation</i>			
Cumulative Exposure	5.641 (4.205) [0.313]		−0.575 (1.249) [−0.032]
Predicted Exposure		1.732 (1.308) [0.073]	
Mean (Y)	59.578	59.578	59.578
SD (Y)	26.924	26.924	26.924
Individual Controls	Yes	Yes	Yes
Grade-Year FE	Yes	Yes	Yes
Family × Initial School FE	Yes	Yes	Yes
N (Panels A–B)	1,907,194	1,907,194	1,907,194
Distinct Students (Panels A–B)	258,439	258,439	258,439
N (Panels C–D)	2,704,659	2,704,659	2,704,659
Distinct Students (Panels C–D)	285,296	285,296	285,296
Mean (X)	0.046	0.040	0.046
SD (X)	0.054	0.043	0.054
R-Squared (Panel A)	—	0.641	0.641
First Stage F (Panels A–B)	47.9		
First Stage F (Panels C–D)	65.8		

Notes: Two-stage least squares (Column 1), reduced form (Column 2), and OLS (Column 3) estimates of cumulative immigrant peer exposure on student outcomes. Predicted exposure is constructed by combining each student's starting school and cohort with a population-level school-to-school transition matrix and the observed immigrant composition of destination school-grade-year cells. All siblings are included regardless of initial school; predicted exposure uses the oldest sibling's initial school, so within-family variation comes solely from cohort differences—relaxing the same-school restriction of Table 2.9. All specifications include family-by-initial-school and grade-by-year fixed effects, following Figlio et al., 2021. Panels A–B use the short-run sample; Panels C–D the long-run sample; first-grade cohorts 1994–2000, never classified as an English Language Learner. Individual controls: gender, age as of September 1, special-education status, and birth order. Standardized coefficients are in brackets (coefficient × SD of treatment); standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

TABLE 2.14. Correlation Between Immigrant Peer Exposure and Other Peer Characteristics

Peer Characteristic	(1)	(2)	(3)	(4)
Hispanic Share	−0.013*** (0.002) [−0.004]	0.005** (0.002) [0.002]	0.026*** (0.002) [0.009]	0.030*** (0.003) [0.010]
Black Share	0.031*** (0.002) [0.005]	0.003* (0.001) [0.001]	0.004* (0.002) [0.001]	0.002 (0.002) [0.000]
Asian Share	0.132*** (0.012) [0.005]	0.067*** (0.006) [0.003]	0.077*** (0.005) [0.003]	0.073*** (0.007) [0.003]
Other Share	−0.645 (2.886) [−0.000]	1.212 (5.158) [0.000]	−1.970 (1.723) [−0.000]	6.045 (6.847) [0.000]
FRL Share	−0.011*** (0.002) [−0.003]	−0.003* (0.001) [−0.001]	−0.003 (0.002) [−0.001]	−0.003 (0.003) [−0.001]
LEP Share	0.217*** (0.004) [0.048]	0.194*** (0.003) [0.043]	0.158*** (0.003) [0.035]	0.139*** (0.004) [0.031]
Campus-Year FE	No	Yes	Yes	Yes
Grade-Year FE	No	Yes	Yes	Yes
Family FE			Yes	
Family-Year FE				Yes
Mean (Y)	0.039	0.039	0.039	0.039
R Squared	0.636	0.888	0.945	0.959
Distinct Students	566,158	566,158	566,158	566,158
Observations	4,664,796	4,664,796	4,664,796	4,664,796

Notes: This table examines the correlation between immigrant peer exposure and other peer characteristics. The outcome variable is immigrant peer share, and the regressors include peer shares for free/reduced-price lunch (FRL) status, race/ethnicity, and English language learner (LEP) status. Each column presents coefficient estimates from a single regression, with each row reporting the estimate for the corresponding peer characteristic. Column (1) includes no fixed effects. Column (2) adds campus-by-year and grade-by-year fixed effects. Column (3) further adds family fixed effects. Column (4) further adds family-by-year fixed effects. The sample includes first grade cohorts from 1994 to 2000. See Table 2.3 for more details on the full set of controls. Standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

TABLE 2.15. Robustness: Saturated Model with Alternative Fixed Effects Specifications

	Additional Peer Exposure Variables Included				
	None	FRL	Race	ELL	All
<i>Panel A: Average Test Scores</i>					
Exposure	−0.226** (0.073)	−0.026 (0.073)	−0.018 (0.077)	0.097 (0.080)	0.108 (0.080)
Mean (Y)	0.097	0.097	0.097	0.097	0.097
<i>Panel B: High School Graduation</i>					
Exposure	−0.078** (0.027)	−0.065* (0.027)	−0.065* (0.028)	−0.063* (0.029)	−0.064* (0.029)
Mean (Y)	0.793	0.793	0.793	0.793	0.793
<i>Panel C: College Enrollment</i>					
Exposure	−0.159*** (0.031)	−0.131*** (0.031)	−0.135*** (0.032)	−0.139*** (0.033)	−0.139*** (0.034)
Mean (Y)	0.573	0.573	0.573	0.573	0.573
<i>Panel D: Any Postsecondary Degree</i>					
Exposure	−0.007 (0.024)	0.015 (0.024)	0.017 (0.025)	0.008 (0.026)	0.010 (0.026)
Mean (Y)	0.195	0.195	0.195	0.195	0.195
<i>Panel E: Employed 7 Years Post-Graduation</i>					
Exposure	−0.049 (0.034)	−0.050 (0.034)	−0.044 (0.036)	−0.052 (0.037)	−0.051 (0.037)
Mean (Y)	0.712	0.712	0.712	0.712	0.712
<i>Panel F: Income Rank 7 Years Post-Graduation</i>					
Exposure	−0.124 (1.858)	0.337 (1.882)	0.701 (1.958)	−0.303 (2.027)	−0.118 (2.033)
Mean (Y)	58.847	58.847	58.847	58.847	58.847
Mean (X)	0.039	0.039	0.039	0.039	0.039
Distinct Students (Panel A)	490,179	490,179	490,179	490,179	490,179
Observations (Panel A)	2,550,365	2,550,365	2,550,365	2,550,365	2,550,365
Distinct Students (Panels B-F)	566,123	566,123	566,123	566,123	566,123
Observations (Panels B-F)	4,664,813	4,664,813	4,664,813	4,664,813	4,664,813

Notes: This table examines the robustness of main results to the inclusion of other peer characteristics. Column (1) includes no peer exposure measures beyond immigrant peer exposure. Columns (2)–(5) progressively add peer shares for free/reduced-price lunch (FRL) status, race/ethnicity, English language learner (ELL) status, and all simultaneously. Immigrant exposure is calculated as the average immigrant share (by subgroup) across all grades through the current grade, reweighting students by appearance frequency. The sample includes first grade cohorts from 1994 to 2000. The model includes family-by-year fixed effects. See Tables 2.3 and 2.4 for more details on the full set of controls and how the long-run outcomes are defined. Standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

TABLE 2.16. Sensitivity of Main Results to Alternative Sample and Immigrant Definitions

	Baseline	Sensitivity to US Born Definition	Sensitivity to Immigrant Definition		
	(1)	(2)	(3)	(4)	(5)
		No ELL Restriction	Recent Immigrant	ESL Not TX-Born	Immigrant With TX-Born
<i>Panel A: Average Test Scores</i>					
Exposure	-0.093 (0.095)	-0.287*** (0.080)	-0.323* (0.154)	-0.145* (0.069)	-0.095 (0.054)
Mean (Y)	0.219	0.106	0.219	0.219	0.219
<i>Panel B: College Enrollment</i>					
Exposure	-0.100* (0.046)	-0.159*** (0.031)	-0.094 (0.066)	-0.062* (0.031)	-0.058 (0.030)
Mean (Y)	0.596	0.573	0.596	0.596	0.596
<i>Panel C: Any Degree</i>					
Exposure	0.013 (0.037)	-0.008 (0.024)	0.010 (0.052)	-0.044 (0.024)	0.018 (0.021)
Mean (Y)	0.211	0.195	0.211	0.211	0.211
<i>Panel D: Employed 7 Years</i>					
Exposure	-0.040 (0.050)	-0.064 (0.034)	-0.056 (0.073)	-0.034 (0.035)	-0.024 (0.027)
Mean (Y)	0.711	0.712	0.711	0.711	0.711
<i>Panel E: Income Rank 7 Years</i>					
Exposure	1.132 (2.800)	-0.711 (1.861)	-1.893 (4.063)	-1.503 (1.951)	0.450 (1.519)
Mean (Y)	59.059	58.853	59.059	59.059	59.059
Mean (X)	0.025	0.039	0.012	0.056	0.030
Students	365,343	474,317	365,343	365,343	365,343
Observations	3,545,593	4,664,353	3,545,593	3,545,593	3,545,593

Notes: This table examines the sensitivity of main results to alternative definitions of the US-born sample and immigrant exposure. Column (1) reports baseline estimates using the main specification, in which US-born students are defined as Texas-born students who were never classified as English Language Learners (ELL), and immigrant peers are defined as immigrants not born in Texas. Column (2) relaxes the US-born sample restriction to include all Texas-born students regardless of ELL status, keeping the baseline immigrant definition. Columns (3)–(5) maintain the baseline US-born sample definition but vary the definition of immigrant peers: Column (3) restricts to recent immigrants; Column (4) defines immigrants as ESL students not born in Texas; Column (5) includes immigrants even if we classify them as Texas born. The sample includes first grade cohorts from 1994 to 2000. Immigrant exposure is calculated as the average immigrant share (by subgroup) across all grades through the current grade, reweighting students by appearance frequency. See Tables 2.3 and 2.4 for more details on the full set of controls and how the long-run outcomes are defined. Standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

2.8. Sibling Linkage: Details and External Validity

2.8.1. Algorithmic Matching Procedure

Appendix Table 2.17 documents our iterative matching strategy. We start with exact matches across all parental name fields, then progressively introduce flexibility through spelling standardization, nickname corrections, and finally phonetic algorithms (Soundex). We also vary middle name requirements, testing specifications that require full middle names, middle initials only, or no middle name at all.

This approach minimizes false positives because the name combinations of parents are very unique. Analyzing 1966–2019 Texas marriage records reveals that 98.9% of couples possess unique first-last name pairs statewide. This suggests that even common individual names become highly distinctive when combined across both parents, making it unlikely that our procedure links unrelated children. As noted in Section 2.2.2 we also exclude pairs whose schools fall in non-adjacent counties more than 25% of the time as an additional precaution.

To examine the robustness of our results to alternative definitions of sibling groups, we re-estimate the baseline model using a sequence of increasingly restrictive specifications, summarized in Appendix Table 2.19. The estimates are broadly consistent across columns, though they tend to be somewhat more negative under less restrictive matching criteria. This is consistent with the broader pattern in the paper, whereby failing to account for family sorting leads to more negative estimates driven by selection.

2.8.2. Incomplete Linkage and Sample Composition

While we are unlikely to link unrelated siblings, two factors prevent us from capturing all siblings. First, paternal information is absent from roughly 18% of birth certificates during 1976–1997, mechanically excluding these children. Second, name spelling variations of parents across siblings’ birth records may not always be fully reconciled by our algorithm, even after applying phonetic corrections.

These linkage constraints produce observable differences from population benchmarks. Appendix

Table 2.18 compares family size distributions in our data versus American Community Survey estimates for Texas-born children. Our sample exhibits notably elevated rates of single-child families, consistent with our inability to connect all true sibling groups.

2.8.3. External Validity Considerations

Our identification strategy restricts the analysis to children from relatively stable family structures in which both biological parents are observed, raising questions about external validity beyond this selected population.

Despite this mechanical selection, the sibling sample closely resembles the full Texas public school population on observable characteristics. As shown in Table 2.1 and discussed in Section 2.2.4, distributions of demographics, academic achievement, and long-run outcomes are similar across samples, suggesting that the analysis sample is not substantially unrepresentative along measured dimensions.

Extrapolation to different family contexts (such as single-parent or blended households) should be interpreted with caution, as resource allocation and household dynamics may differ in ways our design cannot capture.

TABLE 2.17. Distribution of Sibling Linkage Versions Using Name Matching Algorithms

Sibling Linkage Version (1-12)	Freq.	Percent	Cum.
Perfect name with full middle name	5,239,330	91.40	91.40
Cleaned name with full middle name	1,197	0.02	91.42
Nickname-corrected name with full middle name	43,759	0.76	92.18
Soundex name with full middle name	103,652	1.81	93.99
Perfect name with middle initial	114,760	2.00	95.99
Cleaned name with middle initial	45	0.00	95.99
Nickname-corrected name with middle initial	3,626	0.06	96.06
Soundex name with middle initial	13,240	0.23	96.29
Perfect name without middle name	163,967	2.86	99.15
Cleaned name without middle name	26	0.00	99.15
Nickname-corrected name without middle name	5,868	0.10	99.25
Soundex name without middle name	42,862	0.75	100.00
Total	5,732,332	100.00	

Notes: This table summarizes the number and share of sibling linkages made across 12 iterative versions of name-matching algorithms. "Perfect" refers to exact string matches of full names. "Cleaned" versions apply case-insensitive and character-stripped normalization (e.g., removing punctuation or accents). "Nickname-corrected" uses mappings from ethnically appropriate nickname dictionaries that considers the sex of the individual (e.g., "Sam" → "Samuel" or "Sam" → "Samantha"), as in Abramitzky et al. (2020). "Soundex" applies a phonetic encoding algorithm designed to capture similarly pronounced names despite spelling differences. Each name version is attempted using full middle names, middle initials, and no middle names in descending order of matching precision.

TABLE 2.18. Family Size comparing the ACS to TBI

Siblings	ACS Percentage	TBI Percentage
1	35%	62%
2	36%	27%
3	19%	8%
4	7%	2%
5+	2%	1%

Notes: This table compares the distribution of sibling group sizes from two sources: the American Community Survey (ACS) and the Texas Birth Index (TBI). Percentages represent the share of children living in families with the indicated number of siblings. The ACS percentages are calculated from children born in Texas between 1976 and 1997, while the TBI figures reflect counts of children born to the same mother and father in the birth records that matched to the TEA data between 1976 and 1997.

TABLE 2.19. Effects of Immigrant Peer Exposure: Robustness to Family Linking Algorithm

	Baseline	All Siblings	Contiguous Counties	Strict Name Match
<i>Panel A: Reading Test Scores</i>				
Exposure	-0.467 (0.437) [-0.018]	-0.025 (0.112) [-0.001]	-0.261 (0.268) [-0.011]	-0.208 (0.467) [-0.008]
Mean (Y)	0.205	-0.192	0.065	0.233
Distinct Students	34,179	91,587	46,342	29,261
Observations	152,572	490,151	206,697	129,942
<i>Panel B: Math Test Scores</i>				
Exposure	-0.457 (0.574) [-0.017]	-0.087 (0.101) [-0.004]	-0.464 (0.310) [-0.019]	-0.643 (0.534) [-0.024]
Mean (Y)	0.227	-0.241	0.061	0.260
Distinct Students	34,391	92,075	46,685	29,396
Observations	153,723	492,312	208,478	130,980
<i>Panel C: College Attendance</i>				
Exposure	-0.013 (0.171) [-0.001]	-0.306*** (0.039) [-0.014]	-0.235* (0.097) [-0.010]	0.027 (0.194) [0.001]
Mean (Y)	0.575	0.434	0.513	0.586
Distinct Students	43,157	111,611	60,011	36,945
Observations	302,859	942,983	432,072	257,767
<i>Panel D: Income Rank 7 Years</i>				
Exposure	-9.243 (11.615) [-0.364]	-5.834** (2.153) [-0.274]	-13.690* (6.233) [-0.584]	-7.693 (13.389) [-0.299]
Mean (Y)	57.965	54.634	56.327	58.209
Distinct Students	43,157	111,611	60,011	36,945
Observations	302,859	942,983	432,072	257,767
Mean (X)	0.026	0.035	0.029	0.025

Notes: This table presents robustness checks on the family linking algorithm used to construct sibling groups. The baseline column uses the main specification from Table 2.3. Column (2) includes all identified siblings without geographic restrictions. Column (3) further restricts to siblings linked across contiguous counties only. Column (4) applies the strictest matching criteria, requiring exact matches on first, last, and middle names of both parents. Each column within a panel presents estimates of β_1 from a separate regression of Equation 2.3. The sample includes first grade cohorts from 1994 to 2000 who were never classified as an English Language Learner. Immigrant exposure is calculated as the average immigrant share across all grades through the current grade, reweighting students by appearance frequency. The model includes family-by-year fixed effects. See Table 2.3 for more details on the full list of controls and how the long-run outcomes are defined. Standardized coefficients are shown in brackets and are computed by multiplying the estimated coefficient by the standard deviation of the treatment variable. Standard errors in parentheses are clustered at the campus level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$.

2.9. Mathematics of Transition Matrix IV

To create our two instrumental variables used in Section 2.5.1, we first calculate transitions of students to subsequent schools in our data. Specifically, for each grade g , we construct a square matrix $M(g+1|g)$, where rows index the set of schools attended by students in grade g , and columns index the set of schools attended in grade $g+1$:

$$M(g+1|g) = \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} & \dots & \pi_{1N} \\ \pi_{21} & \pi_{22} & \pi_{23} & \dots & \pi_{2N} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \pi_{N1} & \pi_{N2} & \pi_{N3} & \dots & \pi_{NN} \end{bmatrix} \quad (2.4)$$

In this transition matrix, each element, $\pi_{k,j} = Pr(\text{attend school } j \text{ in } g+1, t+1 | \text{attend school } k \text{ in } g, t)$, represents the population share of students who transition from school i to school j as they advance to the next grade, i.e. the Markov probability that a student in school k at grade g going to school j in grade $g+1$. The dimension of the matrix, N , is the total number of schools in the data. Therefore, each row of the matrix represents a probability distribution across all possible destination schools.

We then generate the predicted exposure by multiplying the transition matrix with the observed immigrant shares in each potential school. Let $w_j(g, t)$ denote the observed immigrant share of school j in grade g and year t , then a student currently enrolled in school k in grade g, t would, under the transition process, face an expected exposure in the next grade equal to:

$$z_k(g+1, t+1) = \sum_j \pi_{kj} \cdot w_j(g+1, t+1). \quad (2.5)$$

Iterating this process across grades for each school in our data, we can compute the predicted

exposure vector for all schools at time $(\tilde{g}, \tilde{t})$ based on the Markov chain for any given (g_0, t_0) :

$$\mathbf{z}(\tilde{g}, \tilde{t}) = \prod_{g=g_0}^{\tilde{g}-1} M(g+1|g)\mathbf{w}(\tilde{g}, \tilde{t}) \quad (2.6)$$

where $\mathbf{z}(\tilde{g}, \tilde{t}) = (z_1(\tilde{g}, \tilde{t}), z_2(\tilde{g}, \tilde{t}), \dots, z_N(\tilde{g}, \tilde{t}))$ and $\mathbf{w}(\tilde{g}, \tilde{t}) = (w_1(\tilde{g}, \tilde{t}), w_2(\tilde{g}, \tilde{t}), \dots, w_N(\tilde{g}, \tilde{t}))$.

LGBTQ+ Students' High School Experiences and Academic Plans

3.1. Introduction

¹ A critical choice students make is whether to pursue a postsecondary education. This choice is shaped by students' experiences in high school (Kurlaender et al., 2019; Reed et al., 2023; Uwah et al., 2008), which can influence their educational aspirations and expectations (Bergerson, 2009)—important predictors of later educational attainment (Jacob & Wilder, 2010; Sewell et al., 1969). Moreover, research has long documented differences in aspirations and expectations across student groups (Bergerson, 2009), including race/ethnicity, gender, and first-generation status (Kao & Tienda, 1998; Mau & Bikos, 2000; McCarron & Inkelas, 2006; Solorzano, 1992; Trusty, 2002), with recent studies drawing on more expansive definitions of gender identity and sexual orientation (Aragon et al., 2014; Feldman et al., 2022).

Despite this, gaps are still present in research examining the educational aspirations of different gender and sexual identities, including intended field of study in college. Given that lesbian, gay, bisexual, transgender, and queer (and/or questioning) (LGBTQ+) students are distinct groups (e.g., Beemyn, 2022; Watson et al., 2020), each of which likely experience school and the college-going process differently, it is important to interrogate the extent to which college aspirations and expectations differ by gender identity and sexual orientation. Moreover, LGBTQ+ students frequently report experiencing harassment and discrimination in high school, often related to their

¹This chapter is coauthored with Alexandria Hurtt (University of California, Davis; amhurtt@ucdavis.edu) and Michal Kurlaender (University of California, Davis; mkurlaender@ucdavis.edu). The data used in this research was collected through a partnership between the California Education Lab and the California Student Aid Commission. The authors are grateful for ongoing grant support from the College Futures Foundation and the Stuart Foundation. The findings and conclusions here are those of the authors alone and do not necessarily reflect the positions or policies of the funders or of the California Student Aid Commission. The authors thank Kairo Weber, Baiyu Zhou, Elizabeth Friedmann, Sherrie Reed, Paco Martorell, Scott Carrell, Maria Charles, and Laurel Westbrook for their expertise and support in the construction of the survey, analysis, and/or review of findings. We also thank Kerith Jane Conron for helpful feedback on an earlier draft of the paper.

sexual orientation, gender identity, and/or how they express their gender (Kosciw et al., 2022). This may impact their specific academic aspirations and broader considerations about college, as school victimization can lead to lower levels of school belonging and self-esteem, high rates of absenteeism, and diminished academic outcomes, including college enrollment (Aragon et al., 2014; Heck et al., 2014; Sansone, 2019). How differences in LGBTQ+ students' secondary experiences may affect their educational trajectories is also important to examine, particularly as federal protections for these students have been rolled back amidst the current federal landscape (Schultz, 2025). This paper thus aims to examine associations between gender and sexual identities, academic aspirations, and college concerns among a recent cohort of college-bound 12th graders in California. In addition, we evaluate whether differences in students' aspirations and expectations may be attributable, at least in part, to differences in high school experiences, including whether a student faced harassment or bullying.

This study explores the following research questions: (1) Do students' college aspirations and expectations differ by gender identity and sexual orientation? and (2) If so, to what extent are these observed differences attributable to students' high school experiences? Drawing on the results of a large-scale survey conducted in partnership with the state agency responsible for administering financial aid in California, we explore patterns in high school experiences and college expectations to investigate what LGBTQ+ students consider ahead of the college transition. In this case, we specifically highlight subgroups of the LGBTQ+ community in effort to analyze potential inequalities students may experience and how this might influence the college transition process. Additionally, this study focuses on California students given the state's broad diversity across demographics, geography, and postsecondary opportunities. Moreover, as the state has historically adopted more protective policies for LGBTQ+ students compared to other states (e.g., the FAIR Education Act, comprehensive anti-discrimination protections), California is a particularly important context for this analysis.

3.1.1. Defining Gender Identity and Sexual Orientation

Students use a variety of labels to describe both their gender identity and sexual orientation (Callis, 2014; Galupo et al., 2015), reflecting the fact that traditional binary labels (e.g., woman, man, heterosexual, homosexual) are often constraining descriptors for the complex ways individuals

experience gender and sexuality (Hammack et al., 2022; Morgan, 2013). Research suggests students are constructing their own identities through different labels (Watson et al., 2020; A. White et al., 2018); although individuals continue to identify with traditional binary labels, many students now use a more expansive set of labels for both gender (e.g., nonbinary, genderfluid) and sexual orientation (e.g., asexual, queer) (Beemyn, 2022; Hammack et al., 2022). Additionally, not every individual may feel comfortable selecting a label, or may be unsure about which labels to select, with *questioning* and *unsure* commonly reported terms for both gender identity and sexual orientation (Watson et al., 2020; A. White et al., 2018). In this study, we utilize a novel dataset that acknowledges the spectrum of labels students may use to define their identities outside a binary construct and spotlight LGBTQ+ students in our analysis to enrich understanding of their unique experiences navigating the high school to college transition.

This study is organized as follows: We first review prior literature and the theoretical frameworks guiding our analysis. Next, we describe our data sources, key measures, and analytic approach, followed by our findings. We then discuss limitations before considering implications for policy and practice.

3.2. Theoretical Background and Prior Literature

The transition from high school to college is shaped by a confluence of individual and institutional factors that can affect students' educational expectations and later attainment (Kurlaender et al., 2019). Individual differences, relationships, the school one attends, and one's broader context all play a role during this critical period (Luh et al., 2023; Pike & Robbins, 2016). At the student level, attributes such as motivation, self-efficacy, and academic performance moderate students' educational trajectory (Bandura, 1981; Casillas et al., 2012; Lazowski & Hulleman, 2016). Moreover, high school academic performance is associated with one's aspirations to both two- and four-year colleges (DesJardins et al., 2019), suggesting that students' perceptions of their chances of admission and success also guide their postsecondary plans.

Additionally, various forms of social and cultural capital play a role in shaping students' educational pathways (Bourdieu, 1986; Dumais & Ward, 2010; Merolla & Jackson, 2014; Plagens,

2011; Wells, 2008). For example, social capital (the resources available through one's social networks) shapes college-going outcomes; research indicates that relationships with peers, family members, and mentors can support students' acquisition of valuable information about the college application and enrollment process and help develop stronger college-going attitudes and behaviors (Alvarado & An, 2015; Hill, 2008; Morton et al., 2018; Perna, 2000; Skobba et al., 2018). Moreover, students that can navigate the socially constructed norms surrounding the college process, and whose networks provide access to academic and social resources for this transition, are more likely to enroll in college (Choi et al., 2008; Fletcher, 2015). As LGBTQ+ students often experience hostile social environments in schools and at home, caring adults and social support more broadly may be particularly important to the college-going process. Beyond social capital, research suggests LGBTQ+ students living in cis-normative environments exhibit particular strengths, possessing several types of queer cultural capital (Pennell, 2016). This includes navigational capital, or the ability to navigate schools and other institutional systems that are not always designed with LGBTQ+ students in mind. In this case, students must learn to navigate cis-normative school systems and spaces (e.g., state policies about gendered restrooms, sports participation, or social activities), potentially in creative ways, in order to access the psychological, social, cultural, and physical resources they need for their well-being (Pennell, 2016).

At the institutional level, schools shape students' expectations of college and future educational attainment through the environments they cultivate (Eccles & Roeser, 2011; Uwah et al., 2008). A school culture characterized by high values and expectations for learning is associated with increased academic outcomes (Hattie, 2009), and school safety and climate are also critical to engagement, belongingness, and achievement (Osterman, 2000; Walton & Cohen, 2011). As schools may be unsafe learning environments for LGBTQ+ students, who often experience higher rates of bias-based bullying and victimization (Day et al., 2018), LGBTQ+ students' perception of school climates may be inherently different from those of their cisgender heterosexual peers. Moreover, students' schooling experiences are mediated by stressors that underlie societal conditions, including external (e.g., discrimination and violence) and internal stressors (e.g., expectations of prejudice, concealment, internalized stigmas). These stressors, heightened for LGBTQ+ students, can affect psycho-social outcomes (Hunter et al., 2021), ultimately influencing educational expectations and the transition

from high school to college more broadly. For example, prior research indicates that transgender youth often experience gender-related victimization, particularly in secondary schools, which may lead to higher levels of emotional distress and lower levels of school belonging (McBride, 2021). Moreover, victimization, reduced feelings of school belonging, and emotional distress can have a negative effect on school attendance, academic performance, and expectations to enroll in a four-year college for all LGBTQ+ youth (Aragon et al., 2014). Altogether, research suggests students' high school experiences influence their future educational behavior (Gong & Toutkoushian, 2024).

Our work is thus motivated by several frameworks to understand how LGBTQ+ students navigate the transition from high school to college. First, we draw on a newer model of college choice (Klasik et al., 2025) that builds upon prior models (e.g., Hossler & Gallagher, 1987; Perna, 2006) to include expectations as a critical component in students' predisposition for college. We examine the ways in which high school experiences influence students' college expectations (including academic plans, concerns, and excitements) in the spring of their senior year and focus on an understudied and marginalized group: LGBTQ+ students.

As college expectations are context-dependent, this work is also framed by the social ecological theory of resilience (Ungar, 2011) to understand the potential variation in the college expectations of LGBTQ+ students, who must often navigate hostile environments and cis-normative policies and norms at school and within society more generally (Kosciw et al., 2022). LGBTQ+ youth that faced harassment or social exclusion in high school may look towards college as a new opportunity to fit in, while those who did not struggle as much may have other concerns. Similarly, LGBTQ+ students that plan to live away from home while in college may feel increased trepidation (i.e., leaving familial support) or, in contrast, relief (i.e., escaping familial discomfort or hostility) (O'Neill & Conron et al., 2022). As college transitions vary across demographic differences, including gender (Means et al., 2016; Toutkoushian et al., 2018), students' identities concurrently influence their lived experiences and expectations for their pathway to college (Feldman et al., 2022; Fotopoulou, 2012). Therefore, it is important to consider different elements of students' identities in addition to how those identities are shaped by students' networks and environment. As such, the differences we observe in LGBTQ+ students' expectations for college will reflect students' experiences and the ways in which they navigate their own situated context.

3.3. Data and Methods

3.3.1. Survey Instrument and Sample Characteristics

In May 2023, in partnership with the California Student Aid Commission (CSAC)—the state agency responsible for administering financial aid programs for students attending public and private universities, colleges, and vocational postsecondary institutions in California—we administered a web-based survey to document the experiences of high school seniors planning to enroll in college. The survey was distributed to all seniors statewide that submitted either the Free Application for Federal Student Aid (FAFSA) or for undocumented students, the California Dream Act Application (CADAA) for the 2023-24 academic year.² A total of 323,555 students were sent an invitation to participate in the survey through CSAC; 10,221 completed surveys, including 9,230 respondents who indicated they planned to go to college, 170 who did not plan to attend, 294 that were unsure, and 527 who did not respond about their enrollment plans. Our response rate was 3.2%, which is low but not unexpected with census level surveys of young adults. Students reported on varying dimensions of their high school experiences; the support they received in preparation for college; their academic interests and plans for postsecondary enrollment; and their concerns about college. Detailed information on gender identity and sexual orientation was also collected. After excluding respondents who were missing information about gender identity ($n = 1,708$), sexual orientation ($n = 1,749$), race/ethnicity ($n = 1,617$), or parental education ($n = 1,647$), our final analytic sample included 7,522 participants (Table 3.1).

Compared to California’s 12th grade enrollment in 2022-23 (see Appendix Table 3.11), our sample is broadly, though imperfectly, representative of the state’s college-bound senior population. In terms of race and ethnicity, our sample mirrors statewide enrollment, though Hispanic/Latinx students are modestly underrepresented (49.8% vs. 55.7% statewide), while Black/African American and Asian students are slightly overrepresented (7.0% vs. 4.9% and 15.1% vs. 9.5%, respectively). A direct comparison of gender is difficult to draw given that the state reports enrollment using only male, female, and nonbinary categories. With that caveat in mind, our sample skews more female

²This research was approved by the Institutional Review Board (IRB) at the University of California, Davis under protocol number 2041454-1 and considered minimal risk. Informed consent was obtained from all participants prior to survey administration.

than the statewide enrollment figures would suggest (57.0% cisgender women vs. 48.2% female statewide).

To collect data on gender identity, we adopted the two-step approach recommended by the GenIUSS group given its high sensitivity and specificity in accurately discriminating between transgender and cisgender respondents (Herman, 2014; Saperstein & Westbrook, 2021). This approach includes two questions: (1) What is your sex assigned at birth, on your original birth certificate?, with *male* and *female* as response options; and (2) What is your current gender identity?, with *woman*, *man*, *nonbinary*, *prefer not to say*, and *other* as response options. Respondents who chose *other* also had the option to write in a response. Using these survey items, we classify respondents into one of the following umbrella categories: *cisgender*, *transgender* and *gender expansive* (Robinson & Cooper, 2024), and *prefer not to say*. Respondents who reported that their sex assigned at birth aligned with their current gender identity (e.g., male and man; female and woman) were categorized as *cisgender*, while respondents who reported their sex assigned at birth differed from their current gender identity were categorized as *trans(gender)* and *gender expansive*. The Human Rights Campaign uses the term *trans(gender)* and *gender expansive* to describe “the collective community of individuals whose gender identities, gender presentations and expressions, and/or lived experiences differ from those typically associated with the sex they were assigned at birth” (Robinson & Cooper, 2024). Here, we adopt this collective umbrella term to refer to all respondents who identified differently than their sex assigned at birth while recognizing that “trans” and “gender expansive” are not necessarily distinct or mutually exclusive identities, and that some may not use these labels to describe their personal identity.

We further classify trans and gender expansive students into the following sub-categories, which we utilize in our analysis: *transgender men*, *transgender women*, *nonbinary*, and *gender diverse/questioning*. For example, students who reported their sex assigned at birth as female or male and who identified as a man or a woman were categorized as *binary-identifying transgender men* and *women*, respectively. Students who selected *nonbinary* or wrote in responses that included the term “nonbinary” were categorized as *nonbinary* regardless of sex assigned at birth.³ Students who wrote in other diverse gender identities (e.g., genderfluid) are captured by the term *gender*

³Of the 266 students who identified as nonbinary, 87% (*n* = 231) were assigned female at birth.

TABLE 3.1. Demographic Characteristics of 12th Grade Participants

	n	Percent
<u>Panel A. Gender identity</u>		
Cisgender man	2,675	35.6
Cisgender woman	4,289	57.0
Transgender man	59	0.8
Transgender woman	20	0.3
Nonbinary	267	3.6
Gender diverse/questioning	71	0.9
Prefer not to say	141	1.8
Total reported gender identity	7,522	100.0
<u>Panel B. Sexual orientation</u>		
Straight/heterosexual	5,190	69.4
Gay or lesbian	330	4.4
Bisexual/pansexual/omnisexual	1,139	15.2
Asexual/aromantic/demisexual	182	2.4
Other/queer/questioning	131	1.8
Prefer not to say	509	6.8
Total reported sexual orientation	7,481	100.0
<u>Panel C. Race/ethnicity</u>		
American Indian/Alaskan Native	163	2.1
Black/African American	532	7.0
Hispanic/Latinx	3,793	49.8
Pacific Islander	64	0.8
Filipino	235	3.1
Asian	1,146	15.1
Other	182	2.4
White/non-Hispanic	1,498	19.7
Total reported race/ethnicity	7,613	100.0
<u>Panel D. Highest level of education among parents</u>		
Did not complete high school	1,373	18.1
High school diploma	1,578	20.8
Some college, no college degree	975	12.9
Associate degree	452	6.0
Bachelor's degree	1,363	17.9
Graduate/professional degree	1,455	19.2
Don't know	387	5.1
Total reported parental education	7,583	100.0

diverse/questioning. Our analysis revealed some significant differences between the experiences of nonbinary and gender diverse/questioning students, which provides additional support for these distinctions. Finally, students who chose *prefer not to say* or wrote in malicious responses (e.g., racial and gender slurs) were classified as *prefer not to say*. Among the trans and gender expansive students in our sample ($n = 417$), about 19% identified as transgender, 64% as nonbinary, and 17% as gender diverse/questioning. Beyond predefined options, students also noted 26 unique terms to describe their gender identity, including *genderfluid*, *agender*, *genderqueer*, and *two-spirit*, among many others.⁴ We also examined associations between gender identity and sexual orientation and found that trans and gender expansive students were more likely to select a non-heterosexual sexual orientation than cisgender students (see Table 3.10 in Appendix 3.10).

To measure sexual orientation, we adapted the survey instrument recommended by the Sexual Minority Assessment Research Team (SMART) to capture the breadth of students' identities (Badgett et al., 2009). Students were asked to self-identify among the following options: *lesbian or gay*; *straight (not gay or lesbian)*; *bisexual*; *asexual*; *prefer not to say*; and *other*, which also included the opportunity to write in a response. The option *asexual* was added to proffer more expansive options for students. Within our sample, nearly one-quarter ($n = 1,782$) of students identified as LGB+, with most students identifying as bisexual/pansexual/omnisexual (15.2%), while about 4% of students identified as gay or lesbian, 2.4% as asexual/aromantic/demisexual, and 1.8% as other/queer/questioning. Similar to gender identity, students used a variety of terms to describe their sexual orientation beyond those provided,⁵ such as *pansexual* and *queer*, while others used multiple terms or did not want to label their sexual orientation.⁶

Survey items related to bullying and harassment were also adapted from the California Healthy Kids Survey (CHKS) (Austin et al., 2023) to examine students' experiences with identity-based victimization in high school. Specifically, we asked students to report the frequency with which they were bullied or harassed at school in the past 12 months, with response options ranging from zero

⁴In the survey sample, of all students who responded to the gender identity questions ($N = 7,551$), 1.1% ($n = 86$) selected "other" and wrote in a response.

⁵A full descriptive examination of the diverse labels students used to describe their gender identities and sexual orientation in the open response option, beyond those provided, is included in a companion paper currently under review (Hurt et al., 2026, forthcoming).

⁶In the survey sample, of all students who responded to the sexual orientation questions ($N = 7,507$), 3.1% ($n = 231$) selected "other" and wrote in a response.

times to four or more times. Students who reported being bullied at least once were then asked to identify the reasons for their victimization, including race or ethnicity, religion, gender identity, and sexual orientation.

In addition, we include responses from LGBTQ+ students for two open-ended survey questions: (1) What do you think is the biggest challenge you'll face in college?; and (2) What are you most excited about for college? Data was imported into MaxQDA 2022 and questions were thematically coded to identify common challenges and excitements about college. Here, we highlight two key themes related to the present analysis.⁷

3.4. Data Analysis

In the present analysis, we examine three main areas: high school experiences, including bullying and harassment; concerns about college; and postsecondary educational plans (i.e., enrollment intentions, intended field of study, and degree aspirations). To measure high school experience, we construct an additive index (equally weighted) from six survey items given their internal consistency (Cronbach's $\alpha = 0.74$): (1) students' ratings of their high school academic performance, (2) social experience, (3) sense of belonging in their school community, (4) whether teachers and staff cared about their future, (5) the quality of college advising, and (6) feelings of preparedness for college.⁸ All items range from -2 to 2, with -2 indicating a very negative experience, 2 indicating a very positive experience, and 0 indicating a neutral experience. The index for high school experience ranges from -12 to 12, with higher values indicating more positive high school experiences. To explore differences, we first regress the high school experience index on both gender identity and sexual orientation in an unconditional model and then add controls for race/ethnicity and parental education.⁹ Additionally, we descriptively examine bullying and harassment in high school disaggregated by gender identity and sexual orientation. For students who reported being bullied at least once, we further explore whether students attributed the bullying to their gender identity or sexual orientation.

⁷Detailed information about the qualitative methods and additional results is included in a companion paper currently under review (Hurt et al., 2026, forthcoming). Summary statistics of the qualitative sample can be found in Appendix 3.14.

⁸See Appendix 3.11 for high school experience survey items and summary statistics.

⁹For sexual orientation, we also estimate an additional specification that controls for assigned sex at birth. See Appendix 3.15 for results.

In terms of college concerns, we analyze a battery of 12 survey items in which students were asked to rate how worried they were about their upcoming college experience on a four-point Likert scale (0 = *not at all worried*, 1 = *slightly worried*, 2 = *somewhat worried*, and 3 = *very worried*). Eight items captured students' general concerns: affording tuition, living expenses, performing well academically, balancing work obligations, balancing family obligations, finding community, living away from home, and having adequate support for emotional and mental health needs. Four items captured concerns about discrimination: discrimination based on race, gender identity, sexual orientation, and religion (see Appendices 3.12 and 3.13 for disaggregation of each item by gender identity and sexual orientation, respectively). We assessed the internal consistency of these items using Cronbach's alpha (0.80) and performed principal component analysis (PCA)¹⁰ to identify underlying dimensions of worries about college. We retained components with eigenvalues greater than 1.0, which identified three principal components. We then constructed a composite worry index for each retained component for every student as a weighted sum of the student's responses to the 12 worry items, with weights given by the eigenvector of the corresponding component. The first component, which we label as *general worries*, loads broadly across most items and captures overall anxiety about the college transition. The second component, *worries about discrimination*, loads most heavily on the four discrimination items. The third component, *worries about financial burdens*, loads primarily on the tuition and living expense items. Each index is a continuous measure, with higher values indicating greater concern along that dimension.

We then use regression analysis to explore potential differences in the level of each type of worry by gender identity and sexual orientation. We estimate three different specifications in our analysis: (1) an unconditional model with gender identity/sexual orientation alone; (2) a model controlling for student demographics (i.e., the full set of race/ethnicity and highest parental education indicators); and (3) a model that controls for student demographics and the high school experience index. As coefficient estimates from models 2 and 3 are very similar, we report results for models 1 and 3 alone. For robustness, we additionally estimate a model that controls for demographics, the high school experience index, and sex assigned at birth (for differences by sexual orientation). Results for this

¹⁰PCA is a statistical technique that reduces the dimensionality of complex data by identifying underlying patterns and grouping correlated variables into distinct components/constructs. This approach is particularly useful when analyzing multiple survey items that may measure overlapping concepts as it allows us to identify the core dimensions of students' concerns while reducing dimensionality in the data.

additional model are reported in Appendix 3.16.

We examine postsecondary educational plans along three dimensions for both gender identity and sexual orientation, including type of college, intended field of study, and degree aspirations. For college type, we estimate logit models for intent to attend a four-year (versus two-year) college, while we estimate ordered logit models for students' degree aspirations, where the outcome is an ordinal measure with five categories: vocational/technical certificate, associate degree, bachelor's degree, master's degree, and doctoral degree. In both cases, we estimate an unconditional model and a model that controls for race/ethnicity and parental education, with cisgender men and straight/heterosexual students serving as the reference categories. For sexual orientation, we estimate an additional specification that controls for sex assigned at birth. We report odds ratios with 95% confidence intervals and present predicted probabilities from the ordered logit models as a stacked bar chart evaluated at sample means.

3.5. Findings

High School Experiences

We first compare average high school experiences by gender identity (Table 3.2). Cisgender men report the most positive high school experiences (mean = 4.3), followed by cisgender women (mean = 3.8). Comparatively, trans and gender expansive students report substantially less positive experiences, with gender diverse/questioning students reporting the lowest scores (mean = 1.1). We plot these differences as regression coefficients in Figure 3.1 and find them statistically significant and robust to the inclusion of demographic controls.

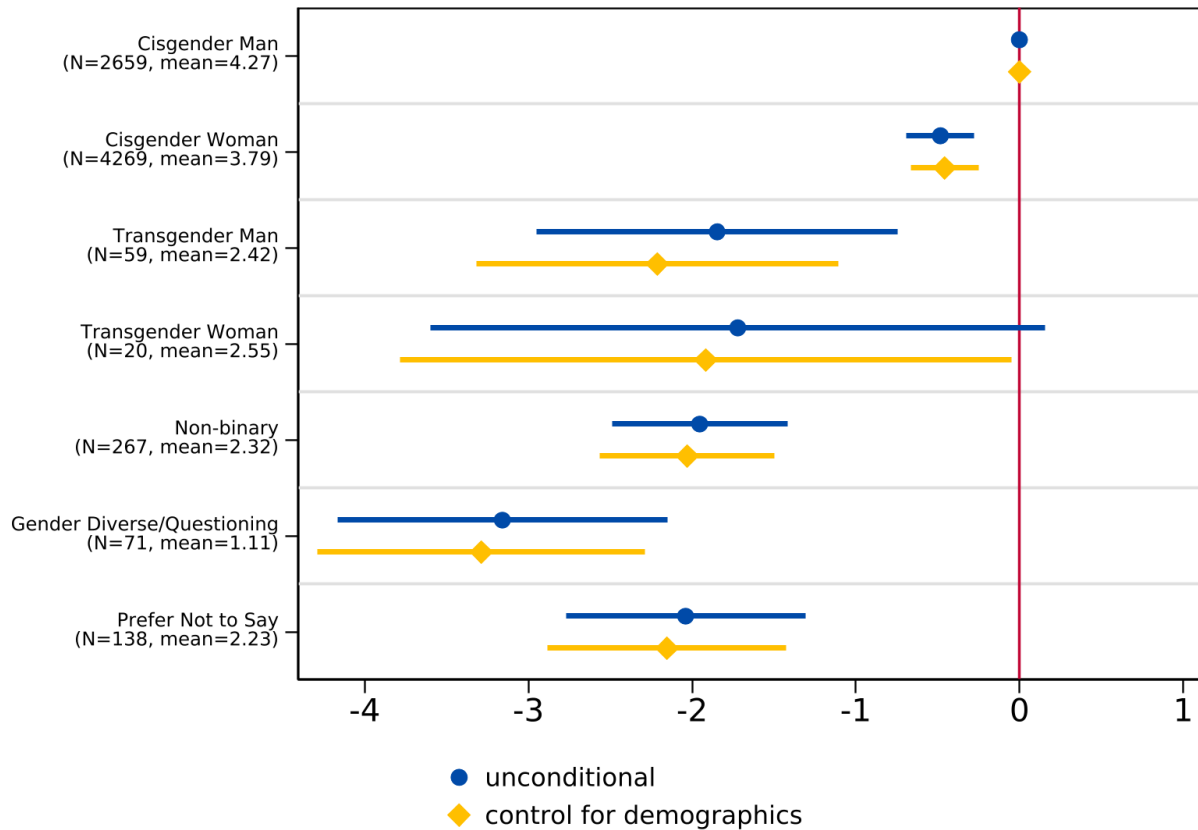
We observe a similar pattern by sexual orientation (Table 3.3). Relative to straight/heterosexual students (mean = 4.1), all other sexual orientation groups report significantly lower high school experience scores. Students who prefer not to disclose their sexual orientation (mean = 2.9) and bisexual/pansexual/omnisexual students (mean = 3.1) report the largest gaps, while asexual/aromantic/demisexual students (mean = 3.7) and gay or lesbian students (mean = 3.6) report somewhat smaller but still significant differences. These differences remain statistically significant after controlling for race/ethnicity, parental education, as well as assigned sex at birth (Figure 3.2).

TABLE 3.2. Summary Statistics of High School Experience Index by Gender Identity

	count	mean	sd
Cisgender Man	2659	4.27	4.25
Cisgender Woman	4269	3.79	4.27
Transgender Man	59	2.42	3.85
Transgender Woman	20	2.55	4.82
Non-binary	267	2.32	4.45
Gender Diverse/Questioning	71	1.11	4.32
Prefer Not to Say	138	2.23	4.41
Total	7483	3.84	4.30

Notes: The general high school experience index only applies to students who reported non-missing values for all six high school experience items and ranges from -12 to 12 (negative values indicate a negative experience, positive values indicate a positive experience, and zero indicates a neutral experience). High school experience items are available in Appendices B and C.

FIGURE 3.1. Regression of the High School Experience Index on Gender Identity



Notes: For the unconditional model, $n = 7,483$. For the model controlling for demographics (race/ethnicity and parental level of education), $n = 7,464$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each gender identity are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (cisgender man). See the Data Analysis section for details on the construction of the high school experience index.

While not causal, these results suggest an association between both gender identity/sexual orientation and students’ experiences in high school, which we consider in later analyses.

TABLE 3.3. General High School Experience by Sexual Orientation

	count	mean	sd
Straight/Heterosexual	5159	4.13	4.27
Gay or Lesbian	329	3.55	4.26
Bisexual/Pansexual/Omnisexual	1136	3.12	4.35
Asexual/Aromantic/Demisexual	182	3.68	3.95
Other/Queer/Questioning	131	3.17	4.03
Prefer Not to Say	505	2.91	4.42
Total	7442	3.84	4.30

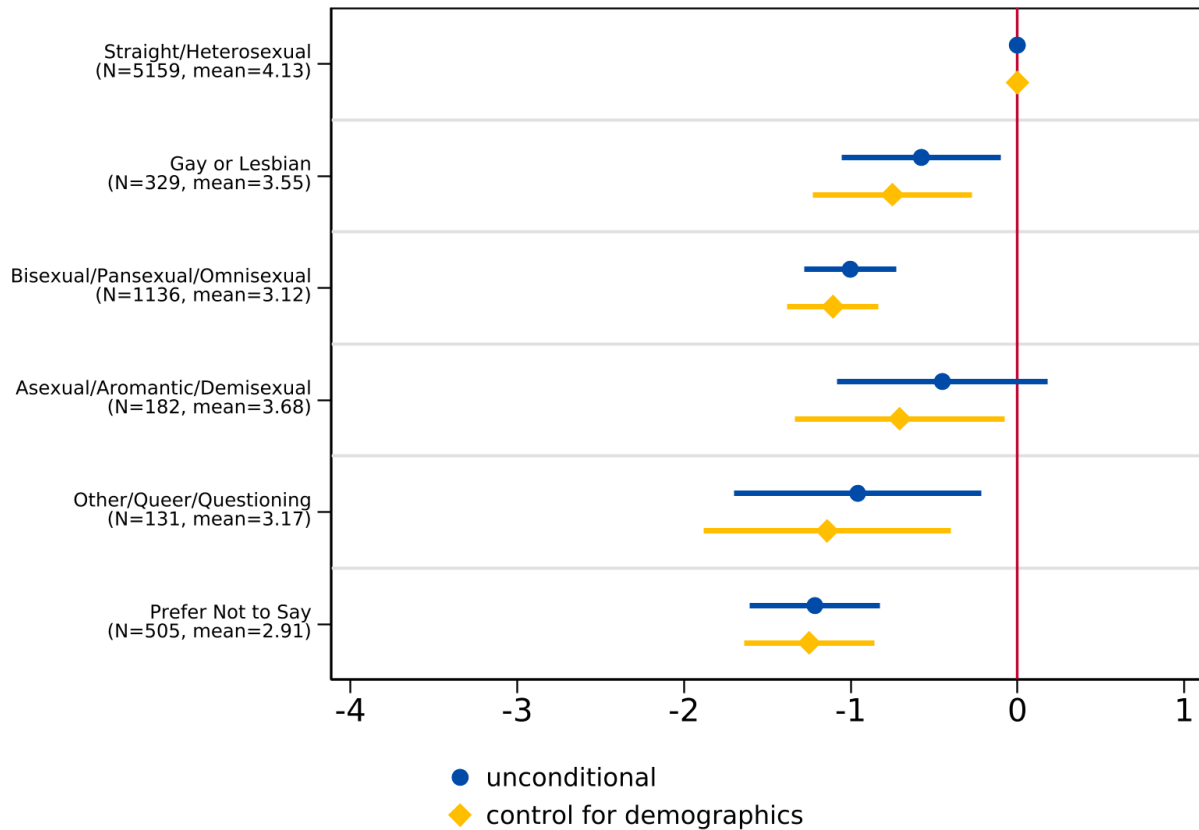
Notes: The general high school experience index only applies to students who reported non-missing values for all six high school experience items and ranges from -12 to 12 (negative values indicate a negative experience, positive values indicate a positive experience, and zero indicates a neutral experience). High school experience items are available in Appendices B and C.

3.5.1. Bullying in High School

Figure 3.3 presents the frequency of bullying students reported experiencing in high school by gender identity and sexual orientation. Across all respondents, approximately 77% of students reported not being bullied or harassed on school property in the past year, while roughly 10% reported being bullied once, 9% reported being bullied two to three times, and about 4% were bullied four or more times. However, these rates varied considerably by gender identity. Among cisgender men, nearly 84% reported no bullying compared to only about 35% of transgender students and 54% of nonbinary and gender diverse/questioning students. Moreover, transgender students reported the highest rates of repeated victimization, as approximately 27% were bullied two to three times and nearly 16% experienced bullying four or more times — rates far exceeding those of any other gender identity. Nonbinary and gender diverse/questioning students also reported elevated rates of bullying relative to their cisgender peers, with roughly 23% reporting two to three incidents and about 10% reporting four or more.

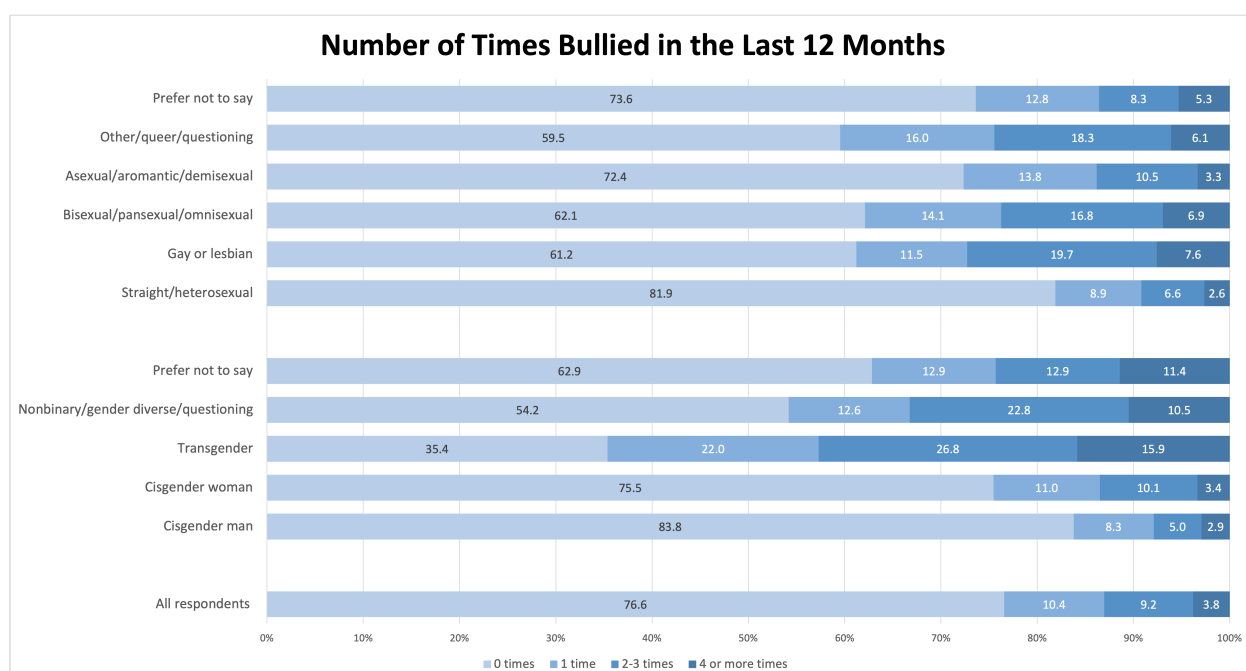
Patterns by sexual orientation revealed similar disparities. While approximately 82% of straight/heterosexual students reported no bullying in the past year, this figure dropped to about 61% for gay or lesbian students and 62% for bisexual, pansexual, and omnisexual students. Students

FIGURE 3.2. Regression of the High School Experience Index on Sexual Orientation



Notes: For the unconditional model, $n = 7,483$. For the model controlling for demographics (race/ethnicity and parental level of education), $n = 7,464$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each sexual orientation are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (straight/heterosexual). See the Data Analysis section for details on the construction of the high school experience index.

FIGURE 3.3. Frequency of Bullying/Harassment in the Past Year



Notes: Rates depicted in the figure represent the sample of students that responded to questions about bullying and harassment ($n = 7,512$ for disaggregation by gender, and $n = 7,471$ for disaggregation by sexual orientation).

identifying as other, queer, or questioning also reported lower rates of no bullying (about 60%) and higher rates of repeated victimization compared to their heterosexual peers. Asexual, aromantic, and demisexual students reported rates closer to the overall average, though still below those of heterosexual students.

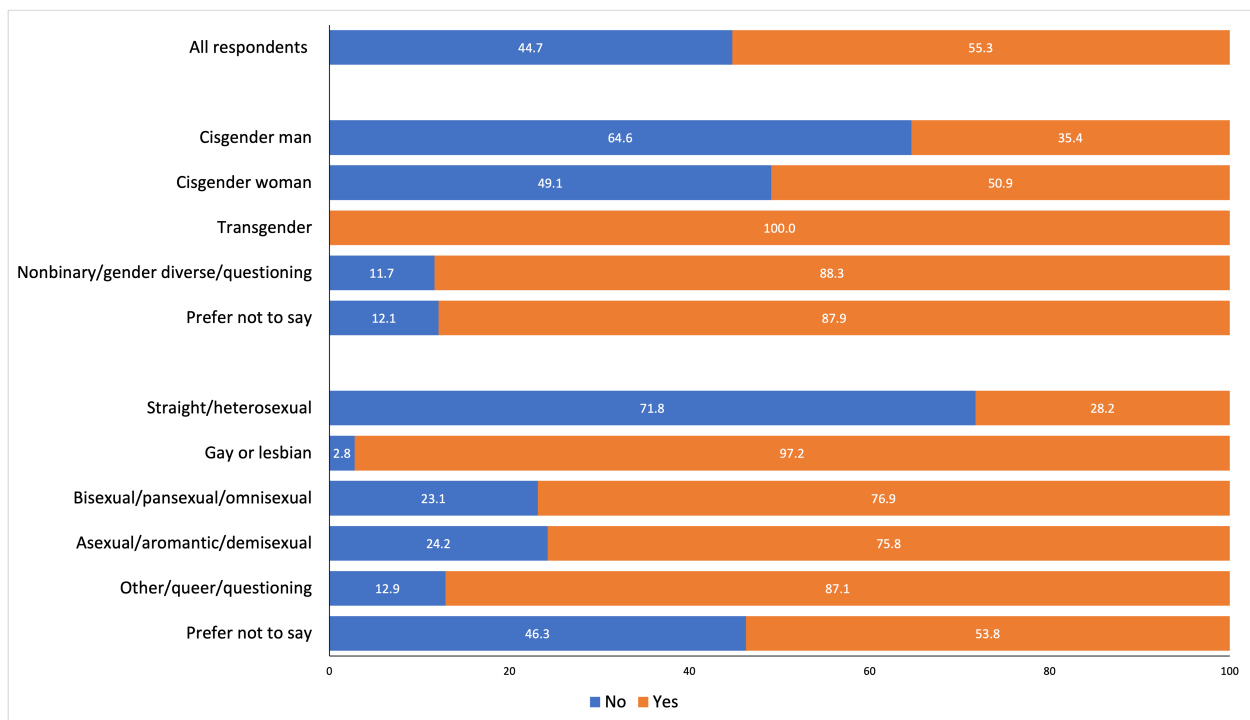
Among respondents who reported being bullied one or more times, we further examined whether students attributed the bullying to their gender identity or sexual orientation (Figure 3.4). Overall, about 55% of students that were bullied indicated they had been harassed because of their gender or sexual orientation. This figure, however, masks stark differences across identity groups. Among transgender students who reported being bullied ($n = 42$), all attributed the bullying to their gender identity or sexual orientation, and nearly 88% of nonbinary and gender diverse/questioning students who were bullied reported the same. In contrast, only about 35% of cisgender men and 51% of cisgender women who were bullied attributed it to gender or sexual orientation. Disparities were just as pronounced across sexual orientation: approximately 97% of gay or lesbian students who experienced bullying reported it was related to their gender or sexual orientation, compared to 28% of straight/heterosexual students. These patterns underscore that LGBTQ+ students not only experience higher rates of bullying overall but are also disproportionately targeted because of their identity.

3.5.2. Bullying as a Potential Mediator of High School Experiences

To investigate whether identity-based victimization helps to explain the differences we observe in high school experience by gender identity and sexual orientation, we estimated additional models that control for the frequency of bullying or harassment in the past year (full results presented in Appendix 3.18). Adding bullying as a covariate attenuates the coefficients for gender identity; for example, the gap between gender diverse/questioning students and cisgender men narrows from -3.29 to -2.46 index points, and the gap for nonbinary students narrows from -2.03 to -1.44, though both remain significant. Additionally, across all gender identity categories, the R-squared increases from 0.034 (demographics only) to 0.072 (demographics plus bullying), indicating that bullying frequency explains a meaningful share, but not all, of the variation in students' high school experiences.

A similar pattern emerges by sexual orientation (Appendix Figure 3.21). After controlling for

FIGURE 3.4. Student Responses to “During the Past 12 Months, Were You Harassed or Bullied for Your Gender Identity and Sexual Orientation?”



Notes: Rates depicted in the figure represent the sample of students that reported being bullied or harassed more than once ($n = 960$ for disaggregation by gender, and $n = 958$ for disaggregation by sexual orientation).

bullying, the coefficient for gay or lesbian students is no longer statistically significant, suggesting that the lower high school experience reported by this group is largely attributable to experiences of bullying and harassment. The gap for bisexual/pansexual/omnisexual students narrows from -1.11 to -0.72 but remains significant, indicating that bullying also explains some of the differences for these students. Further controlling for assigned sex at birth produces only modest additional attenuation. These results suggest bullying is a meaningful mediator of the relationship between gender identity, sexual orientation, and overall high school experience, though it does not fully account for the observed disparities.

3.5.3. Concerns About College

Students were also asked about their concerns ahead of the transition to college. We constructed three indices of worry using PCA eigenvectors (summary statistics of the PCA constructs are reported in Table 3.4), including: (1) general worries; (2) worries about discrimination (including gender, sexuality, race, and religion); and (3) worries about financial burdens (including tuition and living costs). For summary statistics of the survey items used to construct these dimensions by gender identity and sexual orientation, see Appendix 3.13. Together, these components explained 56.0% of the variance in overall worries about the transition to college.

In terms of broad concerns, we find trans and gender expansive students experience a higher level of general worries compared to cisgender students. When we correlated these worries to students' high school experiences, there appeared to be a negative, fairly weak correlation between high school experience and general worries about college; that is, students who reported more negative experiences in high school, on average, reported more worries about college. However, there is almost no correlation between high school experience and worries about discrimination or financial burdens, respectively.

Regression estimates for each component of worry by gender identity, with and without controls, are presented in Figures 3.5 through 3.7. In all regressions, *cisgender men* was set as the baseline comparison category. In terms of general college worries, among trans and gender expansive students, we find no statistical differences between different gender identities (Figure 3.5). That is, the general worries of transgender men, transgender women, and nonbinary or gender questioning students were

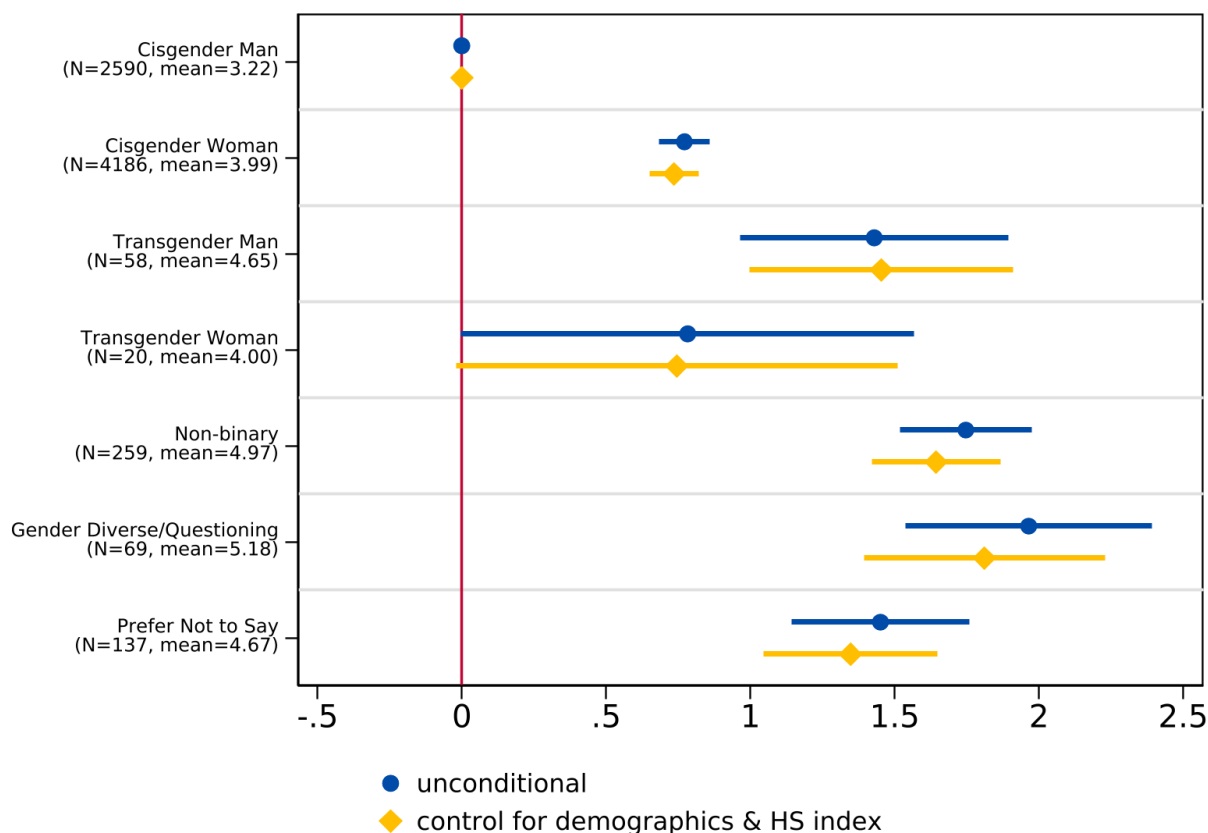
TABLE 3.4. College Concern Items and Results from Principal Component Analysis (PCA)

	count	mean
Items:		
How worried are you about...		
affording tuition fees?	7957	1.83
living expenses?	7956	1.68
performing well academically?	7958	1.50
balancing work obligations?	7943	1.56
balancing family obligations?	7941	1.06
finding community?	7934	1.32
living away from home?	7948	0.91
having adequate support for emotional and mental health needs?	7944	1.13
discrimination based on gender?	7886	0.44
discrimination based on sexual orientation?	7876	0.33
discrimination based on race?	7895	0.51
discrimination based on religion?	7891	0.23
PCA Constructs:		
General worries	7706	3.78
Worries about discrimination	7706	-1.22
Worries about financial burdens	7706	0.63

Notes: The first twelve rows are survey items related to college worries. Each item ranges from 0 (*not worried at all*) to 3 (*very worried*). The last three rows indicate the three constructs: (1) *general worries* ranges from 0 to 10.23; (2) *worries about discrimination* ranges from -3.83 to 5.12; and (3) *worries about financial burdens* ranges from -3.64 to 3.65. Constructs were only created for high school students that reported non-missing values for all twelve items. The PCA proportion explained is reported in parentheses after each construct. See the Data Analysis section for details on the construction of the worry indices.

significantly higher than cisgender men, but not significantly different from each other.

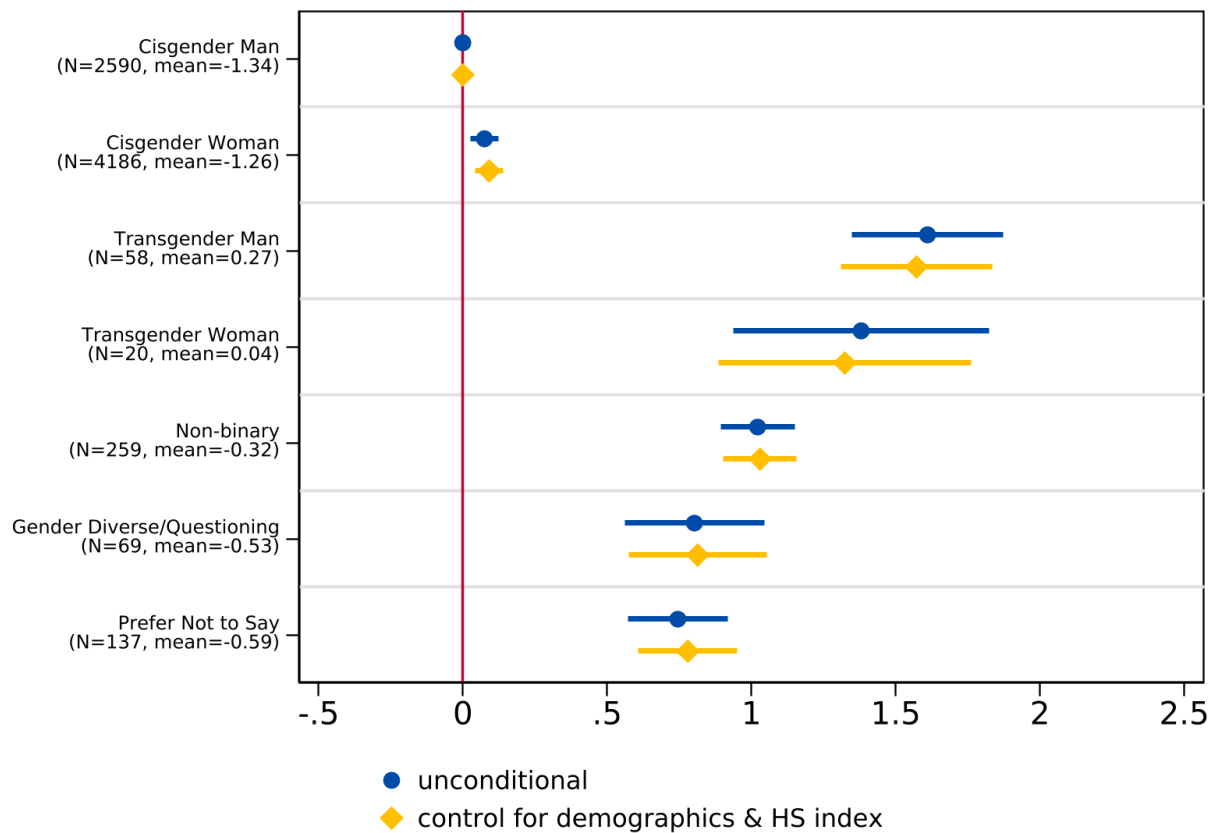
FIGURE 3.5. Regression of General Worries about College on Gender Identity



Notes: We estimate regression models with gender indicators for an unconditional model ($n = 7,319$) and include race/ethnicity indicators, parental education indicators, and a constructed high school experience index as covariates for the specification with control variables ($n = 7,276$). The lines indicate 95% confidence intervals while the sample frequency and mean for each gender identity are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (cisgender man). See the Data Analysis section for details on the construction of the worry indices.

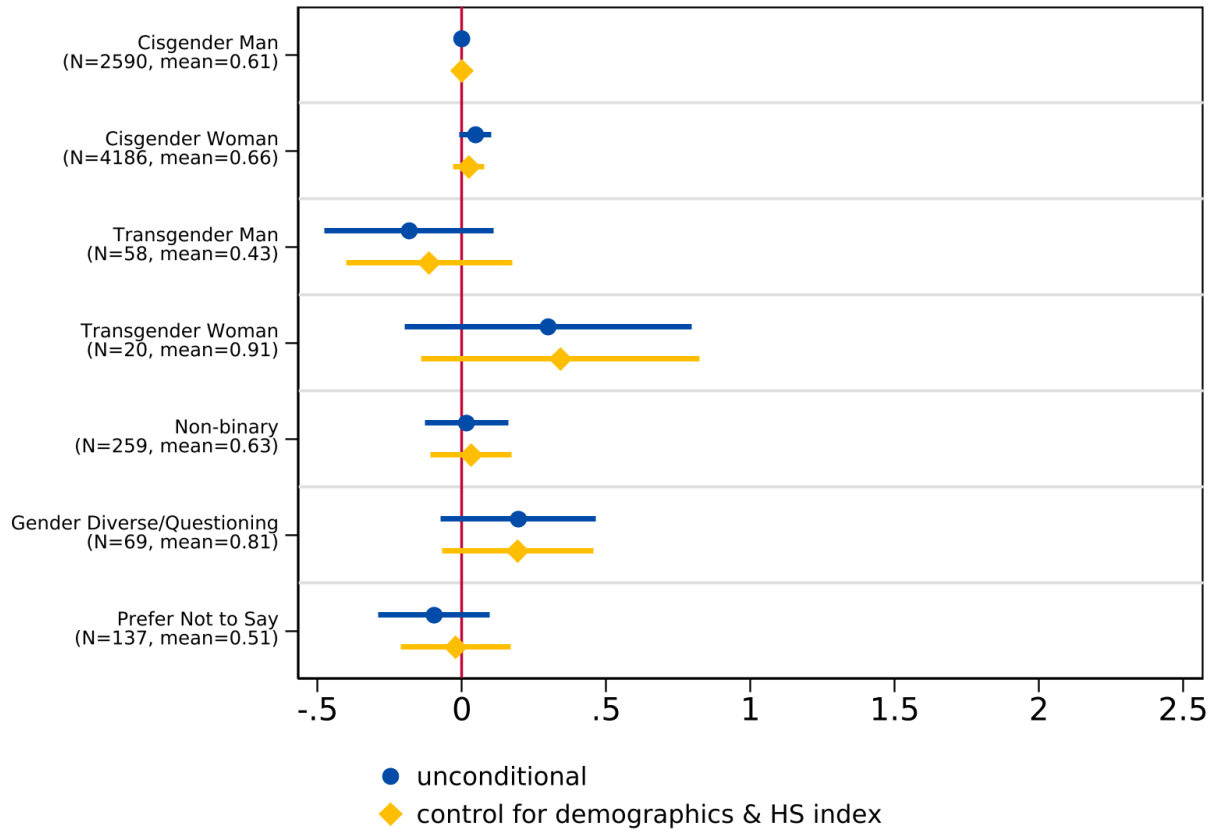
Additionally, results illustrate that trans and gender expansive students maintain higher levels of worry about discrimination compared to cisgender students (Figure 3.6), as the magnitude of the difference between transgender men and cisgender men exceeds the mean level of worry about discrimination for cisgender men. Additionally, among trans and gender expansive students, transgender students report higher levels of worry about discrimination compared to nonbinary and gender diverse/questioning students, although this estimate is somewhat noisy due to sample size. Finally, Figure 3.7 suggests that worries about financial burdens do not differ by gender identity.

FIGURE 3.6. Regression of Worries About Discrimination in College on Gender Identity



Notes: We estimate regression models with gender indicators for an unconditional model ($n = 7,319$) and include race/ethnicity indicators, parental education indicators, and a constructed high school experience index as covariates for the specification with control variables ($n = 7,276$). The lines indicate 95% confidence intervals while the sample frequency and mean for each gender identity are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (cisgender man). See the Data Analysis section for details on the construction of the worry indices.

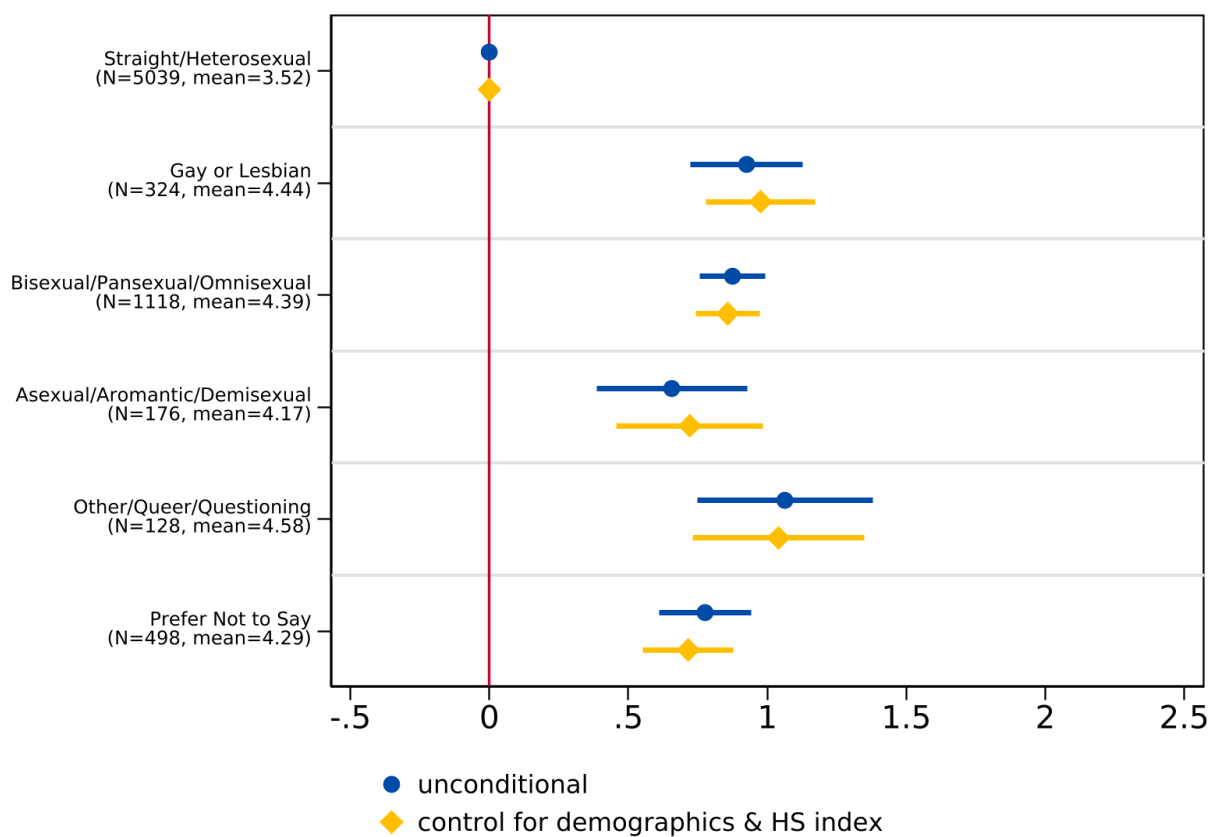
FIGURE 3.7. Regression of Worries About Financial Burdens in College on Gender Identity



Notes: We estimate regression models with gender indicators for an unconditional model ($n = 7,319$) and include race/ethnicity indicators, parental education indicators, and a constructed high school experience index as covariates for the specification with control variables ($n = 7,276$). The lines indicate 95% confidence intervals while the sample frequency and mean for each gender identity are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (cisgender man). See the Data Analysis section for details on the construction of the worry indices.

We find a similar pattern when examining college concerns by sexual orientation (Figure 3.8), as non-heterosexual students report significantly higher levels of general worries compared to straight/heterosexual students. The magnitude of these differences is fairly consistent across sexual orientation categories, with point estimates ranging from roughly 0.5 to 1.0 units above straight/heterosexual students. Other/queer/questioning students report the highest level of general worries (mean = 4.58), though the confidence interval is wider due to a smaller sample size. These differences remain statistically significant after controlling for demographics and high school experience.

FIGURE 3.8. Regression of General Worries About College on Sexual Orientation

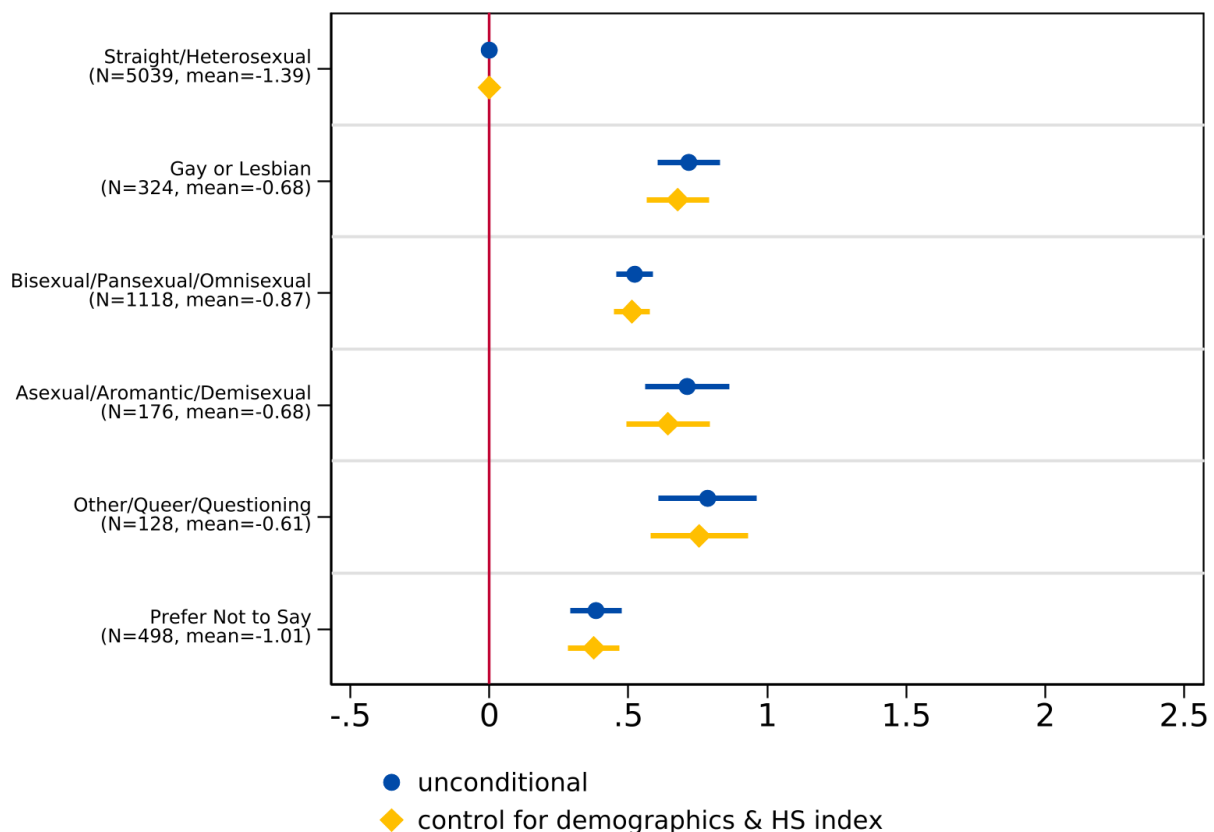


Notes: For the unconditional model, $n = 7,283$. For the model controlling for demographics (race/ethnicity and parental level of education), $n = 7,263$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each sexual orientation are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (straight/heterosexual). See the Data Analysis section for details on the construction of the worry indices.

Worries about discrimination also differ significantly by sexual orientation (Figure 3.9), as

non-heterosexual students also report higher levels of worry about discrimination relative to straight/heterosexual students (mean = -1.39). Among these students, asexual/aromantic/demisexual and other/queer/questioning students report the largest differences, while bisexual/pansexual/omnisexual students report somewhat smaller but still significant differences. These estimates are robust to the inclusion of demographic controls and the high school experience index.

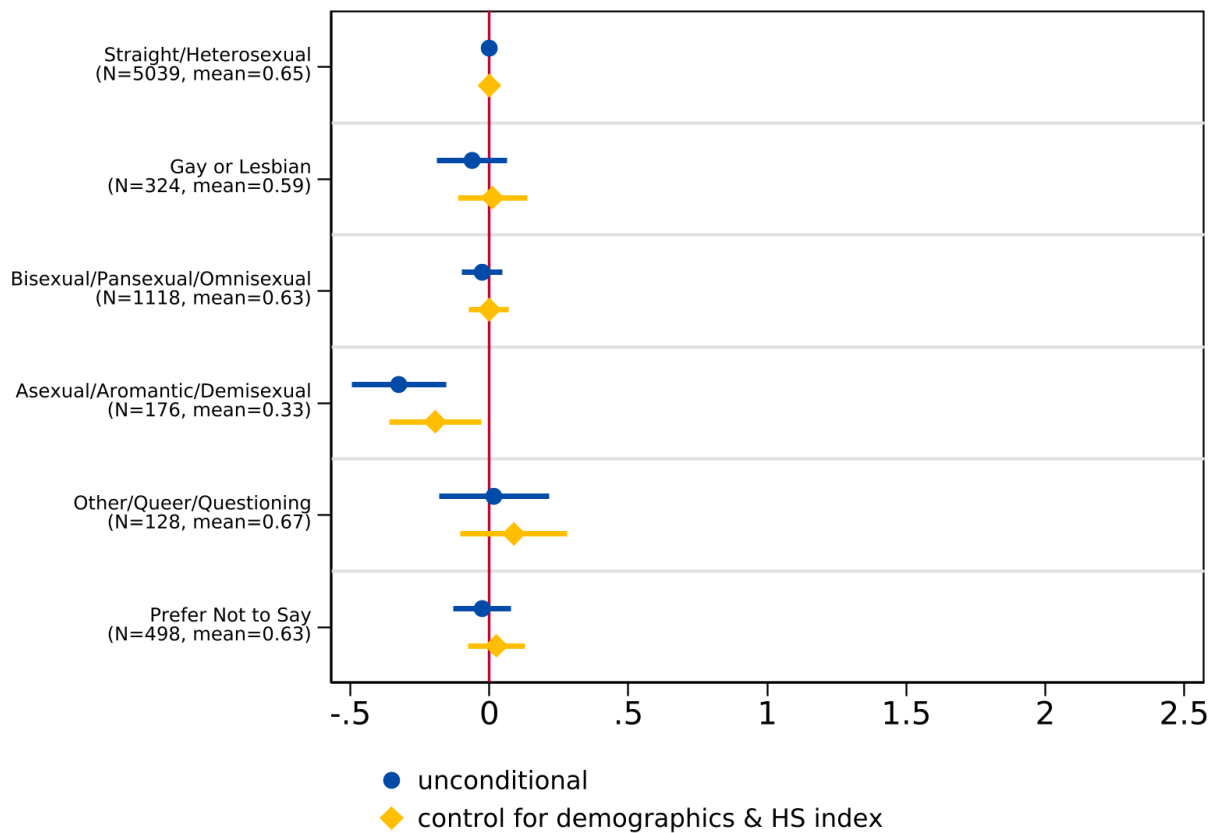
FIGURE 3.9. Regression of Worries About Discrimination in College on Sexual Orientation



Notes: For the unconditional model, $n = 7,283$. For the model controlling for demographics (race/ethnicity and parental level of education), $n = 7,263$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each sexual orientation are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (straight/heterosexual). See the Data Analysis section for details on the construction of the worry indices.

As with gender identity, worries about financial burdens also do not differ meaningfully by sexual orientation (Figure 3.10). Point estimates are close to zero for most groups and confidence intervals generally overlap with the reference line, suggesting that financial concerns about college are not strongly associated with sexual orientation.

FIGURE 3.10. Regression of Worries About Financial Burdens in College on Sexual Orientation



Notes: For the unconditional model, $n = 7,283$. For the model controlling for demographics (race/ethnicity and parental level of education), $n = 7,263$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each sexual orientation are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (straight/heterosexual). See the Data Analysis section for details on the construction of the worry indices.

3.6. Bullying as a Potential Mediator of College Worries

To investigate whether bullying experiences help to explain the observed differences in college worries, we estimated additional models that control for the frequency of bullying or harassment students experienced in the past year alongside demographic controls and the high school experience index (results are presented in Appendix 3.19).

For general worries by gender identity, controlling for bullying does not substantively change the pattern of results (Appendix Figure 3.22). For example, while the coefficient for nonbinary and gender diverse/questioning students decreases (from 1.64 to 1.49, and from 1.81 to 1.62, respectively), coefficients for trans and gender expansive students remain significantly higher than cisgender men. The bullying coefficients themselves are positive and significant, suggesting that students who experienced more frequent bullying report higher general worries about college, although this association does not fully account for the gender identity differences. We find a similar pattern for sexual orientation (Appendix Figure 3.23): in this case, all non-heterosexual student groups continue to report significantly higher general worries after controlling for bullying, with only modest attenuation of the coefficients.

For worries about discrimination, results are even more striking. After controlling for bullying, the coefficients for gender identity remain large and highly significant (Appendix Figure 3.24). Transgender men maintain the highest level of worry about discrimination relative to cisgender men ($b = 1.48$), followed by transgender women ($b = 1.22$), nonbinary students ($b = 0.97$), and gender diverse/questioning students ($b = 0.73$). The persistence of these large coefficients after controlling for bullying suggests that worry about discrimination is not simply a reflection of past bullying experiences but may reflect broader anticipation of discrimination in the college environment. By sexual orientation (Appendix Figure 3.25), all non-heterosexual students likewise maintain significantly elevated worries about discrimination after controlling for bullying, with other/queer/questioning students ($b = 0.71$) and asexual/aromantic/demisexual students ($b = 0.62$) reporting the largest differences.

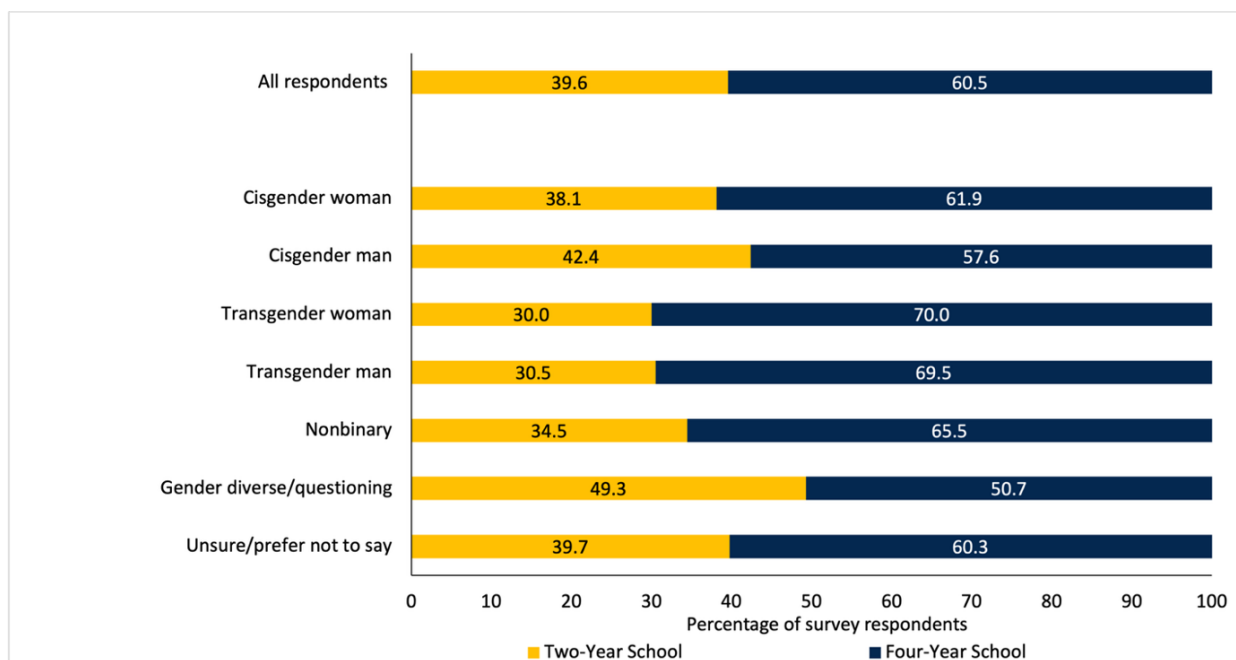
Consistent with our earlier findings, worries about financial burdens remain unrelated to either

gender identity or sexual orientation after controlling for bullying (Appendix Figures 3.26 and 3.27). The bullying coefficients themselves are small and largely non-significant in the financial burden models, confirming that financial concerns about college operate independently of both identity-based differences and bullying experiences.

3.7. Postsecondary Educational Plans

We next explore patterns in academic plans across gender identity, including college program type, intended field of study, and the highest degree students planned to pursue. Findings indicate that most students planned to enroll in a four-year college compared to a two-year college (Figure 3.11); however, more transgender and nonbinary students planned to enroll in four-year colleges compared to their cisgender peers. Conversely, students that identified as gender diverse or questioning intended to enroll in a two- or four-year college at nearly similar rates (49.3% and 50.7%, respectively).

FIGURE 3.11. Plans for College Enrollment by Gender Identity

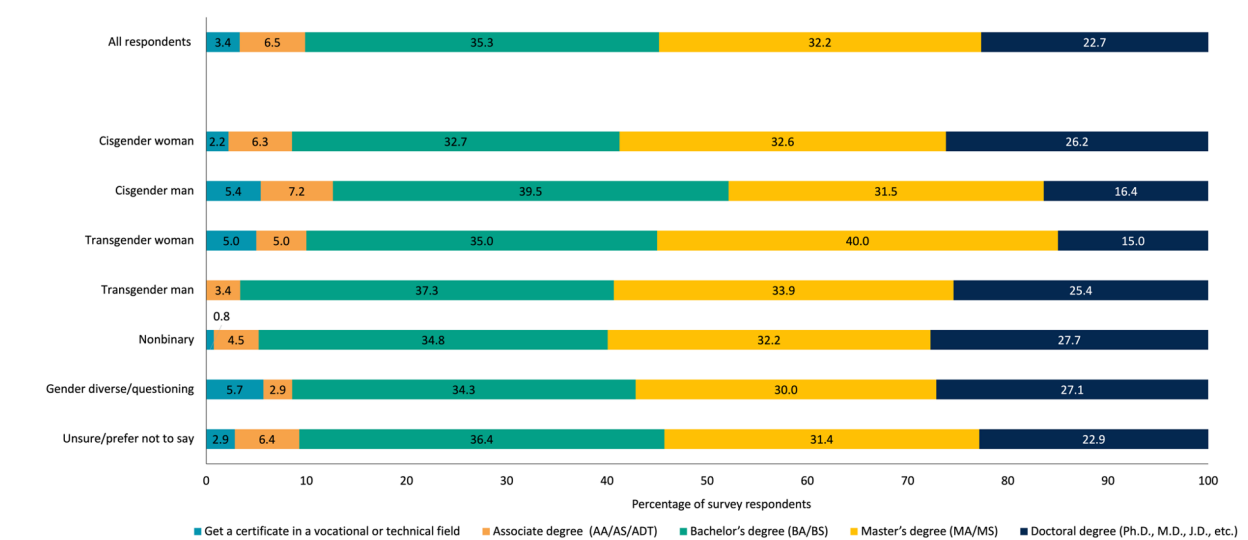


Differences also emerged in students' intended field of study (Table 3.5). Overall, most students planned to study natural sciences (16.3%), engineering (16.2%), and health sciences (14.8%), including trans and gender expansive students, who often selected these majors at higher rates than their peers. For example, transgender women planned to study engineering at nearly 9 percentage points above

the rate overall. Trans and gender expansive students were also more likely to pursue humanities and the arts, as gender diverse students (31.0%), transgender men (28.8%), nonbinary students (25.5%), and transgender women (25.0%) all indicated these fields at much higher rates than the sample overall (11.4%).

Students also indicated the highest degree they planned to complete (Figure 3.12). Across all respondents, over half (54.9%) planned to earn an advanced degree. Additionally, we find that more trans and gender expansive students indicated plans to earn a master’s or doctorate at higher rates compared to cisgender students. Among transgender women, of the 55.0% that planned to complete an advanced degree, 72.7% intended to earn a master’s degree—higher than any other gender identity we surveyed. Comparatively, transgender women, gender diverse/questioning students, and cisgender men were more likely to pursue an associate or vocational degree compared to other students, although transgender women and cisgender men appear to be outliers overall.

FIGURE 3.12. Plans for Degree Completion by Gender Identity



Notes: Bars represent percentages ($n = 7,497$).

To examine whether descriptive differences in postsecondary plans persist after accounting for student demographics, we estimated logit models for intent to attend a four-year college and ordered logit models for highest degree aspirations by both gender identity and sexual orientation. Here, we report odds ratios, where values greater than one indicate higher odds relative to the reference category. We also report predicted probabilities from the ordered logit models as stacked bar charts

evaluated at sample means in Appendix [3.17](#).

For four-year enrollment by gender identity (Table [3.6](#)), nonbinary students were significantly more likely than cisgender men to plan to attend a four-year institution, and this difference was robust to the inclusion of demographic controls. Cisgender women were also significantly more likely to plan to attend a four-year college in both specifications. Additionally, for gender diverse/questioning students, the odds ratio was less than one in both specifications, consistent with the descriptive finding that this group was roughly equally split between two- and four-year enrollment plans. For both transgender men and transgender women, the odds ratios were greater than one; however, estimates may be imprecise due to smaller sample size. In this case, we refrain from drawing conclusions about four-year enrollment for these categories based on the regression estimates alone and instead direct readers to the descriptive results reported in Figure [3.11](#).

TABLE 3.5. Intended Field of Study by Gender Identity

	Business	Engineering	Natural sciences	Social sciences	Humanities and arts	Health sciences	Education	Applied sciences	Public service	Undecided	N
All respondents	10.0	16.2	16.3	12.3	11.4	14.8	3.5	3.4	4.8	7.5	7,499
Cisgender man	12.1	30.0	12.2	7.9	8.9	8.0	1.4	7.0	4.3	8.4	2,662
Cisgender woman	9.6	7.8	18.4	14.7	11.2	19.9	4.9	1.2	5.3	7.1	4,280
Transgender man	1.7	13.6	22.0	15.3	28.8	6.8	5.1	1.7	1.7	3.4	59
Transgender woman	10.0	25.0	10.0	10.0	25.0	0.0	0.0	15.0	0.0	5.0	20
Nonbinary	3.0	12.0	19.9	15.4	25.5	7.9	3.0	3.4	3.4	6.7	267
Gender diverse/ questioning	1.4	9.9	19.7	16.9	31.0	8.5	1.4	1.4	5.6	4.2	71
Prefer not to say	5.0	19.3	18.6	12.9	17.9	9.3	2.9	2.1	3.6	8.6	140
N	749	1,211	1,220	920	852	1,106	262	255	360	564	7,499

Notes: Each cell represents the row percentage.

Results by sexual orientation were broadly consistent. For four-year enrollment (Table 3.7), gay or lesbian students and bisexual/pansexual/omnisexual students were significantly more likely to plan to attend a four-year college, and these differences persisted after controlling for demographics. The odds ratios for asexual/aromantic/demisexual and other/queer/questioning students were greater than one and significant in the unconditional models, but no longer significant with the inclusion of demographic controls. This suggests that some of the descriptive differences for these groups may reflect compositional variation in race/ethnicity or parental education.

TABLE 3.6. Logit Models: Intent to Attend 4-Year College by Gender Identity (Odds Ratios)

	(1)	(2)
Cisgender Woman	1.196*** [1.084,1.319]	1.282*** [1.154,1.424]
Transgender Man	1.676* [0.958,2.933]	1.481 [0.816,2.686]
Transgender Woman	1.717 [0.658,4.482]	1.516 [0.549,4.190]
Non-binary	1.400** [1.075,1.822]	1.341** [1.015,1.771]
Gender Diverse/Questioning	0.757 [0.472,1.213]	0.681 [0.412,1.125]
Prefer Not to Say	1.117 [0.790,1.579]	1.100 [0.754,1.603]
Observations	7522	7498
Demographic Controls	No	Yes

Notes: Exponentiated coefficients; 95% confidence intervals included in brackets.

* ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

For degree aspirations (Tables 3.8 and 3.9), ordered logit results reported as odds ratios indicate that nonbinary students had significantly higher degree aspirations than cisgender men (OR = 1.80, 95% CI [1.43, 2.26]), with the largest and most precisely estimated odds ratio among the trans and gender expansive categories, a difference that remained robust with demographic controls (OR = 1.83, 95% CI [1.45, 2.30]). Transgender men also reported significantly higher degree aspirations in both the unconditional (OR = 1.74, 95% CI [1.10, 2.75]) and controlled (OR = 1.66, 95% CI [1.03, 2.67]) models. Gender diverse/questioning students had 1.58 times the odds of aspiring to a higher

degree category in the unconditional model (95% CI [1.02, 2.45]), though this attenuated to marginal significance with controls (OR = 1.54, 95% CI [0.99, 2.41]). For transgender women, the small sample size ($n = 20$) produced estimates that were too imprecise to support meaningful inference (OR = 1.19, 95% CI [0.55, 2.61]); while the descriptive patterns suggest transgender students as a whole report lower rates of certificate and associate degree aspirations and comparable or higher rates of master's degree aspirations relative to cisgender men, the regression models lack the statistical power to confirm these patterns. We therefore direct readers to the descriptive results presented in Figure 3.12 rather than interpret these coefficients substantively.

TABLE 3.7. Logit Models: Intent to Attend 4-Year College by Sexual Orientation (Odds Ratios)

	(1)	(2)	(3)
Gay or Lesbian	1.692*** [1.327,2.156]	1.498*** [1.159,1.935]	1.496*** [1.157,1.935]
Bisexual/Pansexual/Omnisexual	1.410*** [1.233,1.614]	1.347*** [1.169,1.553]	1.264*** [1.093,1.462]
Asexual/Aromantic/Demisexual	1.608*** [1.168,2.213]	1.190 [0.845,1.675]	1.116 [0.790,1.575]
Other/Queer/Questioning	1.413* [0.979,2.039]	1.186 [0.803,1.750]	1.121 [0.758,1.657]
Prefer Not to Say	1.001 [0.832,1.204]	0.899 [0.737,1.096]	0.863 [0.707,1.054]
Observations	7481	7456	7446
Demographic Controls	No	Yes	Yes
Assigned Sex at Birth	No	No	Yes

Notes: Exponentiated coefficients; 95% confidence intervals included in brackets.

* ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

In terms of sexual orientation, gay or lesbian students (OR = 1.53, 95% CI [1.25, 1.87]) and bisexual/pansexual/omnisexual students (OR = 1.64, 95% CI [1.45, 1.84]) consistently reported higher degree aspirations across both specifications. Additionally, the odds ratios for asexual/aromantic/demisexual students (OR = 1.52, 95% CI [1.17, 1.98]) and other/queer/questioning students (OR = 1.44, 95% CI [1.06, 1.96]) were found to only be significant in our unconditional models.

TABLE 3.8. Ordered Logit Models: Highest Degree Plans by Gender Identity (Odds Ratios)

	(1)	(2)
Cisgender Woman	1.627*** [1.489,1.777]	1.734*** [1.585,1.896]
Transgender Man	1.738** [1.098,2.750]	1.661** [1.032,2.674]
Transgender Woman	1.194 [0.548,2.605]	1.050 [0.477,2.311]
Non-binary	1.798*** [1.432,2.257]	1.829*** [1.451,2.304]
Gender Diverse/Questioning	1.584** [1.023,2.453]	1.541* [0.985,2.411]
Prefer Not to Say	1.368** [1.005,1.863]	1.296 [0.945,1.778]
Observations	7497	7475
Demographic Controls	No	Yes

Notes: Exponentiated coefficients; 95% confidence intervals included in brackets.

* ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

TABLE 3.9. Ordered Logit Models: Highest Degree Plans by Sexual Orientation (Odds Ratios)

	(1)	(2)	(3)
Gay or Lesbian	1.527*** [1.250,1.867]	1.388*** [1.130,1.704]	1.353*** [1.101,1.663]
Bisexual/Pansexual/Omnisexual	1.636*** [1.454,1.840]	1.587*** [1.408,1.788]	1.386*** [1.226,1.566]
Asexual/Aromantic/Demisexual	1.522*** [1.168,1.983]	1.251 [0.955,1.638]	1.079 [0.822,1.416]
Other/Queer/Questioning	1.440** [1.056,1.963]	1.293 [0.941,1.777]	1.117 [0.811,1.537]
Prefer Not to Say	1.162* [0.986,1.370]	1.049 [0.887,1.240]	0.957 [0.808,1.132]
Observations	7456	7433	7423
Demographic Controls	No	Yes	Yes
Assigned Sex at Birth	No	No	Yes

Notes: Exponentiated coefficients; 95% confidence intervals included in brackets.

* ($p < 0.10$) ** ($p < 0.05$) *** ($p < 0.01$).

3.7.1. Qualitative Findings

The survey also asked respondents about challenges they anticipated as well as what they were the most excited about ahead of the transition to college. Although a full qualitative analysis of these data is beyond the scope of this paper, we highlight the social concerns and hopes raised by transgender, nonbinary, and gender diverse/questioning students in their responses.

Social Challenges and Discrimination

In open-ended responses, nearly one-quarter of trans and gender expansive students identified social challenges as a concern, as many worried about making new friends or whether they would make friends at all. A nonbinary student described, “Making friends has been a challenge for me in the past when I had to go somewhere new and didn’t know anyone, so it might take a little while to get close with people.” Transgender students also expressed concerns over finding community once they entered college, including worries they may not fit in, or may feel alone or overwhelmed in a new social environment. These worries also included fears of discrimination; multiple students noted concerns about transphobia, coming out to others, and seeing kids from their high school. Moreover, the most intense social concern came from one nonbinary student who shared: “I could die from hate crimes.”

Finding Friends and Community

In contrast, over 36% of trans and gender expansive students were excited about the social aspects of college, as students anticipated meeting new people and building new friendships despite some apprehension. One nonbinary student expressed, “I will be making friends on my own, which is scary, but something to look forward to.” Several students looked forward to finding community, with one gender diverse student writing: “[I am excited] to be able to build a community for myself where I truly feel that I belong.” Additionally, students anticipated meeting other students like themselves. One nonbinary student explained, “I’m most excited about the friendships I will form with people who share the same mindset as me.” Students also wanted to meet people with similar academic interests. As one gender diverse/questioning student noted: “I’m really excited to be part of a supportive and engaging community... [and] be around people with a similar drive for learning

and an appreciation of education.”

Several students were also excited to attend a college where they could access gender-affirming healthcare and housing. One gender diverse/questioning student mentioned, “I’m also going to be downtown for college and I know that the college I’m going to is very LGBTQ+ friendly.” Another nonbinary student explained, “I’m also looking forward to receiving certain medical treatment at my college’s health center that I currently don’t have access to.”

Additionally, some trans and gender expansive students looked forward to the opportunity to leave their home environment. Students described a desire to “[get] away from home and a community that I was almost constantly harassed in,” “leave my abusive household,” and “be away from home. . . [and] finally hav[e] my own autonomy.” One trans woman added that leaving home and entering college explicitly offered “the ability to transition my gender identity, as my parents are opposed to it.” Together, these responses illustrate a pattern consistent with our quantitative findings: even as trans and gender expansive students report elevated concerns about discrimination and less positive high school experiences, they simultaneously express hope that college will offer safer, more affirming environments where they can build community and access the support they need.

3.8. Limitations

An important limitation to consider in this work is that our sample is self-selected and restricted to students that indicated plans to enroll in college in the fall after their senior year. In this case, our results speak to a group of students that made it to the 12th grade, filed a FAFSA or CADAA, and volunteered to participate in our survey. As such, results do not include students who dropped out prior to 12th grade, had other plans, or for whom college seemed out of reach—a population future research should examine in more detail, particularly for LGBTQ+ students. It is also important to consider that students’ plans can change—some students, despite their survey responses, may not enroll in college after all, underscoring that expectations to enroll do not always lead to enrollment (Klasik et al., 2025). Thus, these results reflect a unique sample of intended college-goers captured at one point in time. Future research on high school completion as well as college enrollment and retention by gender and sexual identities is needed.

3.9. Discussion and Conclusion

As high school experiences can influence students' educational aspirations and later attainment, it is important to examine the extent to which these may differ by gender identity and sexual orientation. Our findings indicate that identity-based bullying explains a meaningful share of the observed disparities in the high school experiences of LGBTQ+ students. In fact, while controlling for bullying attenuates gaps in gender identity and sexual orientation, significant differences persist, suggesting that the challenges LGBTQ+ students face extend beyond victimization alone, consistent with prior research on gender-related victimization in high school (Hatchel et al., 2018; McBride, 2021).

Given these experiences, we examined students' concerns ahead of college enrollment by gender identity and sexual orientation. Overall, trans and gender expansive students worried more about discrimination compared to cisgender students. Notably, these concerns persisted after controlling for bullying in high school, indicating that worries about discrimination in college may represent anticipatory concerns about future discrimination. Research indicates that the energy LGBTQ+ students expend navigating school and trying to avoid adverse experiences often detracts from the quality of their educational experience and ability to create community (Goldberg & Kuvalanka, 2018; Kersey & Voigt, 2021). Our findings suggest that this burden may extend into the college transition more broadly, particularly for trans and gender expansive students. In contrast, general worries about college (e.g., academic performance, finding community) and worries about financial burdens (e.g., tuition, living expenses) showed little association with bullying in high school, suggesting these concerns are driven by different factors.

However, despite any concerns, LGBTQ+ students still planned to pursue postsecondary schooling at significantly higher rates than their cisgender heterosexual peers. While these results differ from prior research (Feldman et al., 2022; Sansone, 2019; Wilkinson et al., 2021), the schooling experiences of LGBTQ+ students in our sample may include supports (whether at school or at home) potentially absent to students documented in other studies. For example, students may have had access to resources that can promote the well-being of trans and gender expansive students, such as social support from peers, family, and the community; school-inclusive policies; and extracurricular

activities, among others (Fernandes et al., 2023). Moreover, barriers students might encounter (e.g., financial constraints, a lack of information) can drive them toward various degree pathways. For example, a student may perceive a mismatch between their ability level and certain colleges (Dillon & Smith, 2013), or have concerns about college (e.g., campus climate) (Dugan et al., 2012), driving them to select or avoid selecting a particular institution or pursuing a more advanced degree.

Results also indicated differences in students' planned fields of study. While most students, regardless of gender identity, planned to study various STEM fields (e.g., engineering, natural sciences), trans and gender expansive students often selected these fields at higher rates. Moreover, trans and gender expansive students were more likely to pursue humanities and the arts, aligning with prior research that suggests LGBTQ+ students often select these fields as they believe they are more affirming (Linley et al., 2018) and inclusive than STEM or business (Forbes, 2022; Kersey & Voigt, 2021). These findings offer important insight into the majors and careers trans and gender expansive students plan to pursue and may be of interest to deans, department chairs, and other personnel involved in outreach and career opportunities.

Moreover, our findings carry several implications for education policy in California. First, the stark disparities in bullying rates and identity-based victimization documented in this study suggest that generic anti-bullying programs may be insufficient to address the experiences of LGBTQ+ students. For example, that 100% of transgender students that experienced bullying attributed their victimization to their gender identity or sexual orientation points to the need for targeted interventions that specifically address identity-based harassment, including expanded support for gender-sexuality alliances (GSAs) and professional development for school personnel on the needs of LGBTQ+ students. While California's Seth's Law (AB 9, 2012) already requires schools to adopt anti-bullying policies that enumerate sexual orientation and gender identity as protected categories, these findings indicate that substantial identity-based victimization persists, suggesting school districts may benefit from additional resources and accountability mechanisms to ensure policies are actively implemented.

Second, the patterns documented here, in which trans and gender expansive students report less positive high school experiences and elevated rates of identity-based bullying yet maintain

elevated educational aspirations, are consistent with the social ecological theory of resilience framing this study (Ungar, 2011). In this case, the persistence of elevated educational aspirations in the face of documented adversity reflects the resilience and navigational capital LGBTQ+ students develop through experience in cis-normative environments (Pennell, 2016). This is underscored by our qualitative findings; even as students described fears of discrimination and hate crimes, they simultaneously expressed excitement about finding community and accessing gender-affirming resources in college. Policymakers and institutions should therefore ensure that students are supported with resources that address these concerns, which may include investments in bridge and/or mentoring programs at the high school and college level, as well as campus resources that are responsive to the needs of incoming LGBTQ+ students.

Third, this study highlights the critical need for the state to collect and disaggregate detailed data by gender identity and sexual orientation. The richness of the present data allowed us to identify disparities that would otherwise be invisible in datasets that treat LGBTQ+ students as a singular group or omit gender identity measures entirely. Presently, the California Healthy Kids Survey (CHKS) collects some information on gender and sexual identities; an anonymous survey of school climate, safety, and student wellness, the CHKS is administered biennially in grades 5, 7, 9, and 11 to students in schools that opt in with modules that collect gender identity (male, female, transgender) and sexual orientation data (gay/lesbian/bisexual, straight, not sure).¹¹ However, our findings reveal distinct differences by gender identity and sexual orientation beyond these measures. This heterogeneity reinforces the importance of detailed data collection across gender and sexual identity subgroups, underscoring that LGBTQ+ students should not be treated as a monolith in terms of their college aspirations and broader academic plans (Feldman et al., 2022). As the state considers how to monitor school climate and student well-being going forward, broadening the identities captured in the CHKS as well as mandating its administration statewide would offer opportunities to further understand and address the needs of this population. However, we also recognize that in the current political climate, students may have legitimate concerns about disclosing their identities in institutional surveys. In this case, efforts to build this kind of evidence base will need to balance the analytic power of detailed identity data with robust protections for student

¹¹For more information about the California Healthy Kids Survey, see <<https://www.cde.ca.gov/ls/he/at/chks.asp>> and <<https://calschls.org/about/the-surveys>>.

privacy and confidentiality.

The present findings take on added significance in the current policy landscape. Since survey administration in 2023, federal policy for LGBTQ+ students has shifted considerably, including executive orders withdrawing the federal recognition of transgender identities; restrictions on Title IX protections; and efforts to curtail access to gender-affirming care for minors (Exec. Order No. 14168, 2025; U.S. Department of Education, 2025). At the same time, several states have enacted laws that limit the discussion of gender identity and sexual orientation in schools (Movement Advancement Project, 2025). In contrast, California has continued to strengthen state-level protections, including AB 1955 (the SAFETY Act), which prohibits school districts from requiring staff to disclose a student's gender identity or sexual orientation to parents without the student's consent (California Legislative Information, 2023; Equality California, 2025). This divergence underscores the importance of state-level data and state-level action. Our results suggest that, even within California's comparatively protective policy environment, LGBTQ+ students experienced significant disparities in high school climate and identity-based victimization at the time of survey administration. As federal protections for LGBTQ+ students are rolled back, the role of state and local policy in addressing these disparities becomes even more critical. More broadly, our findings underscore the importance of California's continued commitment to fostering safe, affirming educational environments and ensuring that all students, regardless of gender identity or sexual orientation, are supported in realizing their educational aspirations. As California approaches a new chapter of its education policy, we hope these findings will offer a foundation for evidence-based action on behalf of some of the state's most vulnerable students.

3.10. Additional Demographics

TABLE 3.10. Gender Identity and Sexual Orientation Crosstab

	Straight/ Heterosexual	Gay or Lesbian	Bisexual/Pansexual/ Omnisexual	Asexual/Aromantic/ Demisexual	Other/Queer/ Questioning	Prefer Not to Say	Total
Cisgender Man	85.0	4.0	5.1	0.8	0.6	4.5	100.0
Cisgender Woman	68.1	3.0	17.9	2.3	1.5	7.3	100.0
Transgender Man	6.9	32.8	32.8	6.9	12.1	8.6	100.0
Transgender Woman	20.0	40.0	25.0	10.0	0.0	5.0	100.0
Non-binary	1.9	18.8	50.2	13.4	9.6	6.1	100.0
Gender Diverse/Questioning	2.9	13.0	50.7	14.5	18.8	0.0	100.0
Prefer Not to Say	7.9	6.5	33.1	10.1	4.3	38.1	100.0
Total	69.4	4.4	15.2	2.4	1.7	6.8	100.0

Notes: Each cell represents row percentages—that is, the percentage of students within each gender category of the corresponding sexual orientation.

TABLE 3.11. Demographic Characteristics of California Students Enrolled in 12th Grade (2022-2023 Academic Year)

	n	Percent
<u>Panel A. Race/ethnicity</u>		
American Indian or Alaska Native	2,349	0.5
Black/African American	23,852	4.9
Hispanic or Latino	272,323	55.7
Pacific Islander	2,205	0.5
Filipino	13,231	2.7
Asian	46,278	9.5
Two or more races	17,435	3.6
White/non-Hispanic	103,787	21.2
Not reported	7,476	1.5
<u>Panel B. Gender</u>		
Male	252,390	51.6
Female	235,802	48.2
Nonbinary	744	0.2
Total	488,936	

Notes: Data accessed through DataQuest (<https://dq.cde.ca.gov/dataquest>), which is managed by the California Department of Education.

3.11. High School Experience Items and Construct

TABLE 3.12. High School Experience Items and Construct

	N	Mean
<i>Items:</i>		
How would you rate your academic performance in high school?	7,847	0.7
How would you rate your social experience in high school?	7,825	0.1
I feel that I belong in my high school community.	7,711	0.6
Teachers and staff at my high school care about my future.	7,711	1.1
I received good advising from my high school about my college plans.	7,713	0.7
I feel prepared for college.	7,709	0.7
<i>Construct:</i>		
General high school experience	7,687	3.8

Notes: The first six rows include the Likert-scale survey items that relate to a student's high school experience. All items range from -2 to 2 , with -2 indicating a very negative experience, 2 indicating a very positive experience, and 0 indicating a neutral experience. The final row summarizes the general high school experience construct, which includes all six aforementioned high school experience items. The construct is only created for students who reported non-missing values for all six high school experience items.

3.12. Summary Statistics of Individual High School Experience Items

TABLE 3.13. High School Experience Items by Gender Identity

	(1) Cisgender Man mean/count	(2) Cisgender Woman mean/count	(3) Transgender Man mean/count	(4) Transgender Woman mean/count	(5) Non-binary mean/count	(6) Gender Diverse/ Questioning mean/count	(7) Prefer Not to Say mean/count
how do you rate HS academic experience	0.59 (2675)	0.69 (4289)	0.86 (59)	0.35 (20)	0.67 (267)	0.38 (71)	0.54 (141)
how do you rate your HS social experience	0.26 (2673)	0.08 (4286)	-0.36 (59)	-0.50 (20)	-0.30 (267)	-0.41 (71)	-0.32 (140)
I belong in my high school community	0.70 (2667)	0.56 (4280)	0.02 (59)	0.00 (20)	0.11 (267)	-0.06 (71)	0.16 (139)
teachers and staff at my HS care about my future	1.15 (2667)	1.09 (4281)	1.03 (59)	1.10 (20)	1.01 (267)	0.86 (71)	0.94 (139)
I received good advising from HS about college	0.81 (2667)	0.72 (4282)	0.29 (59)	0.75 (20)	0.40 (267)	0.39 (71)	0.44 (140)
I feel prepared for college	0.75 (2667)	0.64 (4281)	0.58 (59)	0.85 (20)	0.43 (267)	-0.06 (71)	0.38 (139)

Notes: The first six rows include the Likert-scale survey items that relate to a student's high school experience. All items range from -2 to 2, with -2 indicating a very negative experience, 2 indicating a very positive experience, and 0 indicating a neutral experience. The final row summarizes the general high school experience construct, which includes all six aforementioned high school experience items. The construct is only created for students who reported non-missing values for all six high school experience items.

TABLE 3.14. High School Experience Items by Sexual Orientation

	(1) Straight/ Heterosexual	(2) Gay or Lesbian	(3) Bisexual/ Pansexual/ Omnisexual	(4) Asexual/ Aromantic/ Demisexual	(5) Other/Queer/ Questioning	(6) Prefer Not to Say
	mean/count	mean/count	mean/count	mean/count	mean/count	mean/count
how do you rate HS academic experience	0.64 (5190)	0.78 (330)	0.61 (1139)	0.98 (182)	0.69 (131)	0.61 (509)
how do you rate your HS social experience	0.20 (5186)	0.01 (330)	-0.07 (1138)	-0.22 (182)	-0.08 (131)	-0.14 (508)
I belong in my high school community	0.67 (5174)	0.38 (330)	0.36 (1138)	0.40 (182)	0.34 (131)	0.35 (507)
teachers and staff at my HS care about my future	1.11 (5176)	1.16 (329)	1.06 (1138)	1.19 (182)	1.15 (131)	1.01 (507)
I received good advising from HS about college	0.79 (5179)	0.62 (329)	0.58 (1138)	0.65 (182)	0.55 (130)	0.58 (507)
I feel prepared for college	0.70 (5177)	0.62 (330)	0.58 (1137)	0.68 (182)	0.53 (131)	0.48 (506)

3.13. Summary Statistics of Individual College Concern Items

TABLE 3.15. College Concern Items by Gender Identity

	(1) Cisgender Man mean/count	(2) Cisgender Woman mean/count	(3) Transgender Man mean/count	(4) Transgender Woman mean/count	(5) Non-binary mean/count	(6) Gender Diverse/ Questioning mean/count	(7) Prefer Not to Say mean/count
how worried are you about affording tuition fees	1.65 (2664)	1.93 (4278)	1.71 (59)	1.55 (20)	1.98 (266)	2.24 (70)	1.88 (139)
how worried are you about living expenses	1.52 (2662)	1.76 (4279)	1.54 (59)	1.85 (20)	1.90 (267)	1.97 (71)	1.77 (139)
how worried are you about performing well academically	1.38 (2662)	1.57 (4279)	1.49 (59)	1.05 (20)	1.69 (267)	1.80 (71)	1.68 (139)
how worried are you about balancing work obligations	1.41 (2659)	1.62 (4272)	1.48 (58)	1.70 (20)	1.81 (265)	2.06 (71)	1.81 (139)
how worried are you about balancing family obligations	0.96 (2656)	1.09 (4273)	1.15 (59)	0.75 (20)	1.23 (266)	1.34 (70)	1.25 (138)
how worried are you about finding community	1.14 (2655)	1.39 (4274)	1.58 (59)	0.95 (20)	1.62 (263)	1.70 (71)	1.59 (139)
how worried are you about living away from home	0.80 (2660)	0.98 (4273)	0.81 (59)	0.45 (20)	0.91 (266)	0.82 (71)	0.88 (139)
how worried about having adequate support for emotional and mental health needs	0.85 (2660)	1.23 (4273)	1.63 (59)	1.60 (20)	1.72 (267)	1.62 (71)	1.60 (139)
how worried about discrimination based on gender	0.12 (2665)	0.50 (4260)	2.19 (59)	1.80 (20)	1.57 (267)	1.35 (71)	1.05 (138)
how worried about discrimination based on sexual orientation	0.18 (2656)	0.30 (4258)	1.46 (59)	1.35 (20)	1.42 (267)	1.32 (71)	1.14 (138)
how worried about discrimination based on race	0.38 (2660)	0.58 (4274)	0.59 (59)	0.40 (20)	0.72 (267)	0.73 (71)	0.73 (139)
how worried about discrimination based on religion	0.19 (2658)	0.25 (4273)	0.24 (59)	0.00 (20)	0.23 (267)	0.44 (71)	0.31 (138)

TABLE 3.16. College Concern Items by Sexual Orientation

	(1) Straight/ Heterosexual	(2) Gay or Lesbian	(3) Bisexual/ Pansexual/ Omnisexual	(4) Asexual/ Aromantic/ Demisexual	(5) Other/Queer/ Questioning	(6) Prefer Not to Say
	mean/count	mean/count	mean/count	mean/count	mean/count	mean/count
how worried are you about affording tuition fees	1.81 (5173)	1.82 (330)	1.93 (1136)	1.66 (181)	1.98 (130)	1.92 (507)
how worried are you about living expenses	1.64 (5173)	1.82 (330)	1.81 (1137)	1.50 (182)	1.82 (130)	1.80 (506)
how worried are you about performing well academically	1.46 (5171)	1.64 (330)	1.61 (1137)	1.55 (182)	1.55 (130)	1.64 (507)
how worried are you about balancing work obligations	1.49 (5167)	1.68 (329)	1.72 (1133)	1.69 (182)	1.70 (130)	1.70 (505)
how worried are you about balancing family obligations	1.02 (5166)	1.12 (328)	1.11 (1134)	1.06 (181)	1.21 (129)	1.16 (505)
how worried are you about finding community	1.24 (5166)	1.49 (330)	1.50 (1133)	1.47 (180)	1.53 (129)	1.52 (505)
how worried are you about living away from home	0.90 (5169)	0.87 (329)	0.94 (1134)	0.92 (182)	0.95 (130)	0.93 (505)
how worried about having adequate support for emotional and mental health needs	0.99 (5167)	1.47 (330)	1.48 (1137)	1.50 (181)	1.49 (129)	1.32 (506)
how worried about discrimination based on gender	0.26 (5157)	0.82 (330)	0.83 (1136)	1.01 (181)	1.14 (131)	0.71 (506)
how worried about discrimination based on sexual orientation	0.10 (5147)	1.33 (328)	0.83 (1138)	0.80 (182)	1.08 (131)	0.58 (504)
how worried about discrimination based on race	0.45 (5164)	0.55 (330)	0.67 (1139)	0.62 (182)	0.65 (130)	0.72 (506)
how worried about discrimination based on religion	0.22 (5158)	0.16 (330)	0.24 (1139)	0.19 (182)	0.33 (131)	0.30 (506)

Notes: The rows include survey items related to college worries. Each item ranges from 0 (*not worried at all*) to 3 (*very worried*).

3.14. Additional Demographics

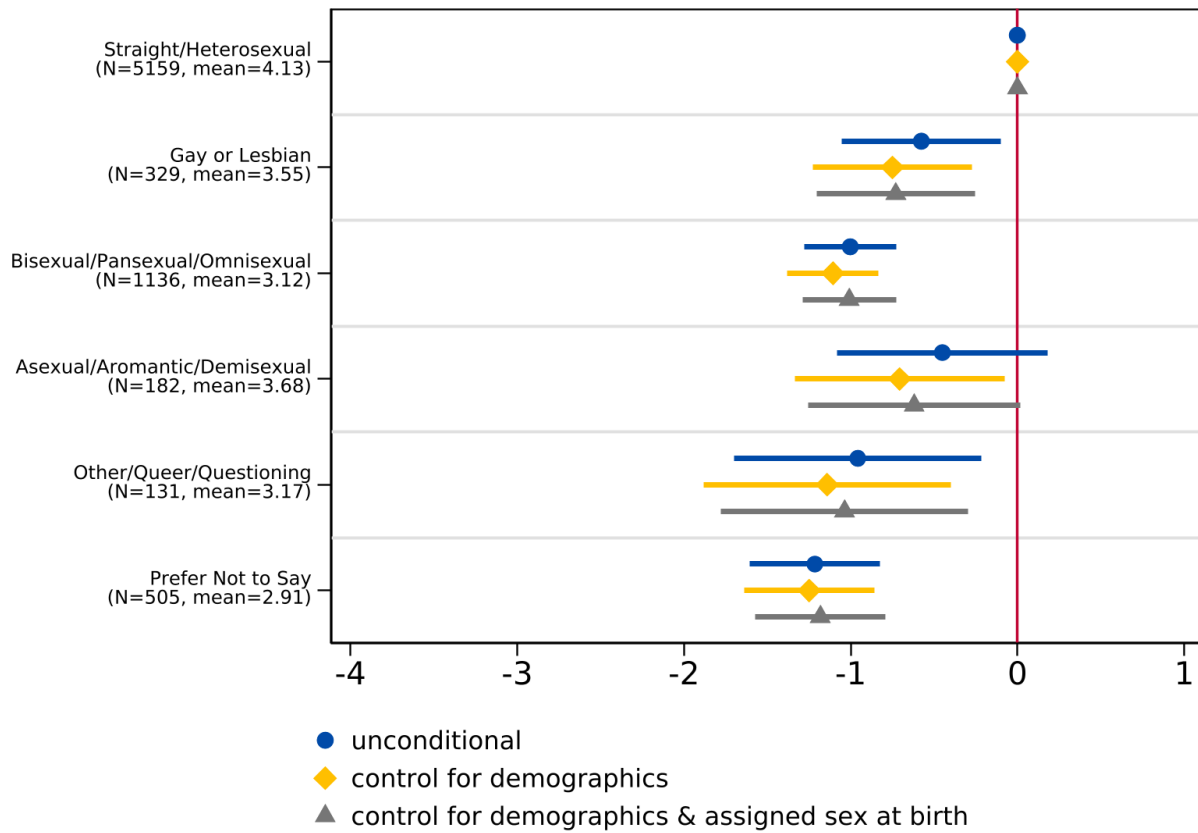
TABLE 3.17. Demographic Characteristics of Qualitative Response Participants

	n	Percent
<u>Panel A. Gender identity</u>		
Transgender man	50	14.5
Transgender woman	17	4.9
Nonbinary	215	62.1
Gender diverse/questioning	64	18.5
Total reported gender identity	346	100.0
<u>Panel B. Sexual orientation</u>		
Straight/heterosexual	12	3.5
Gay or lesbian	69	19.9
Bisexual/pansexual/omnisexual	166	47.9
Asexual/aromantic/demisexual	41	11.9
Other/queer/questioning	38	11.0
Prefer not to say	20	5.8
Total reported sexual orientation	346	100.0
<u>Panel C. Race/ethnicity</u>		
American Indian/Alaskan Native	13	3.8
Black/African American	27	7.8
Hispanic/Latinx	143	41.3
Pacific Islander	5	1.5
Filipino	18	5.2
Asian	43	12.4
Other	3	0.9
White/non-Hispanic	94	27.1
Total reported race/ethnicity	346	100.0
<u>Panel D. Highest level of education among parents</u>		
Did not complete high school	42	12.1
High school diploma	63	18.2
Some college, no college degree	59	17.1
Associate degree	25	7.2
Bachelor's degree	57	16.5
Graduate/Professional degree	86	24.9
Don't know	14	4.0
Total reported parental education	346	100.0

Notes: The rows include survey items related to college worries. Each item ranges from 0 (*not worried at all*) to 3 (*very worried*).

3.15. High School Experience by Sexual Orientation Robustness

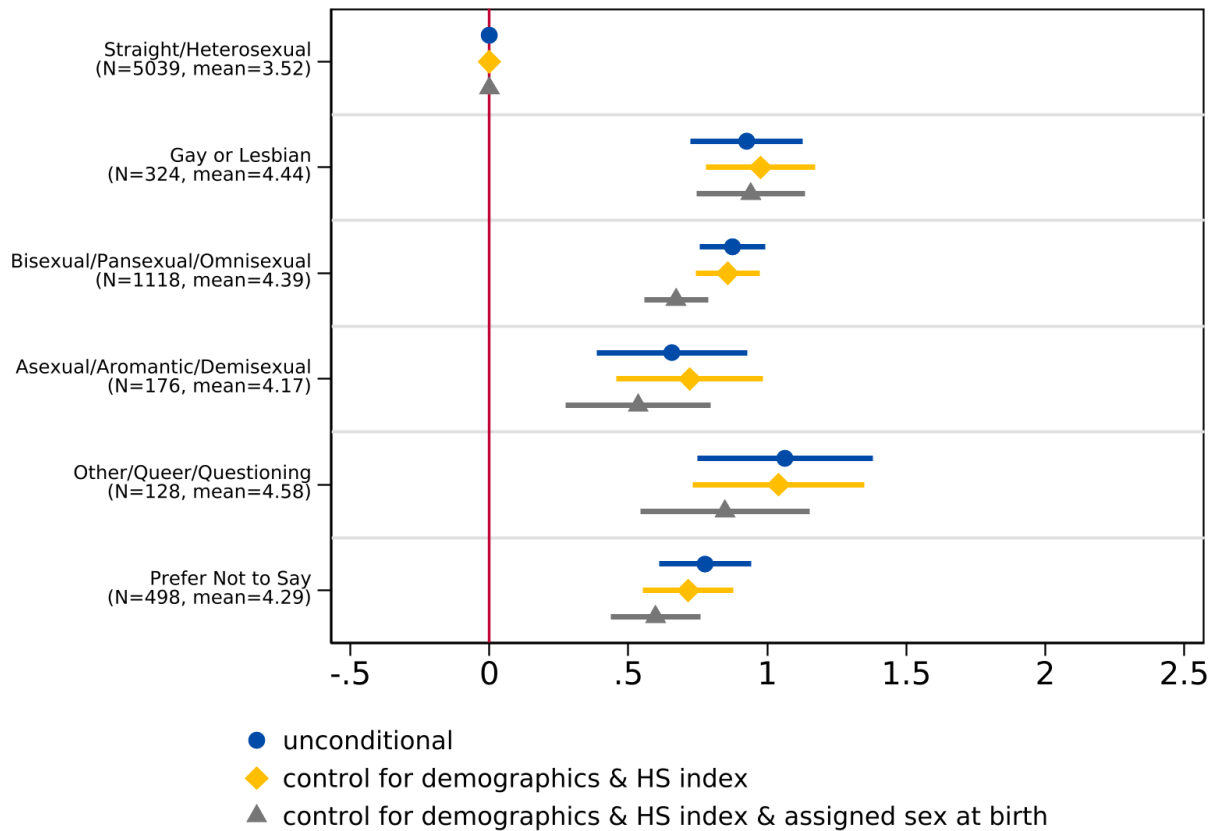
FIGURE 3.13. Regression of HS Experience Index on Sexual Orientation (with Assigned Sex at Birth Control)



Notes: For the unconditional model, $n = 7,483$. For the model controlling for demographics (race/ethnicity and parental level of education), $n = 7,464$. For the model controlling for demographics (race/ethnicity and parental level of education) and sex assigned at birth, $n = 7,412$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each sexual orientation are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (straight/heterosexual).

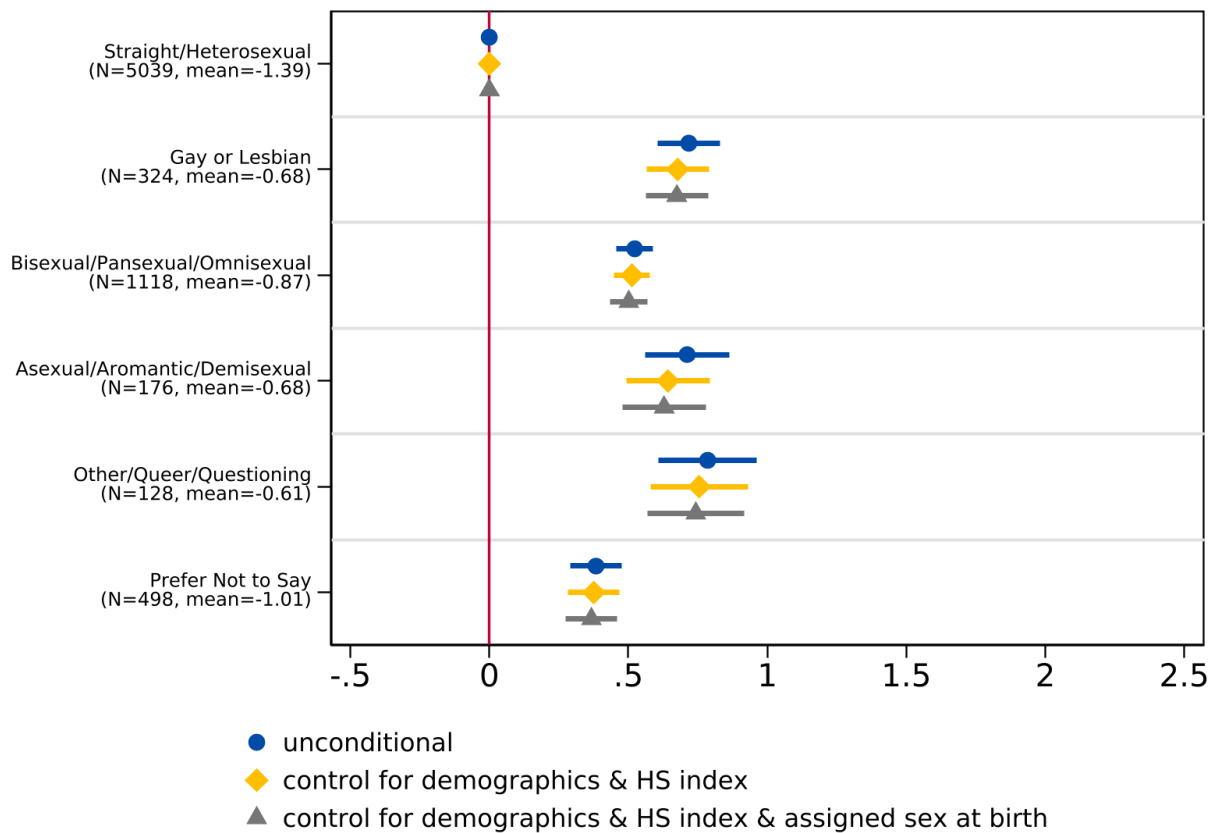
3.16. College Worries by Sexual Orientation Robustness

FIGURE 3.14. Regression of General College Worries on Sexual Orientation (with Assigned Sex at Birth Control)



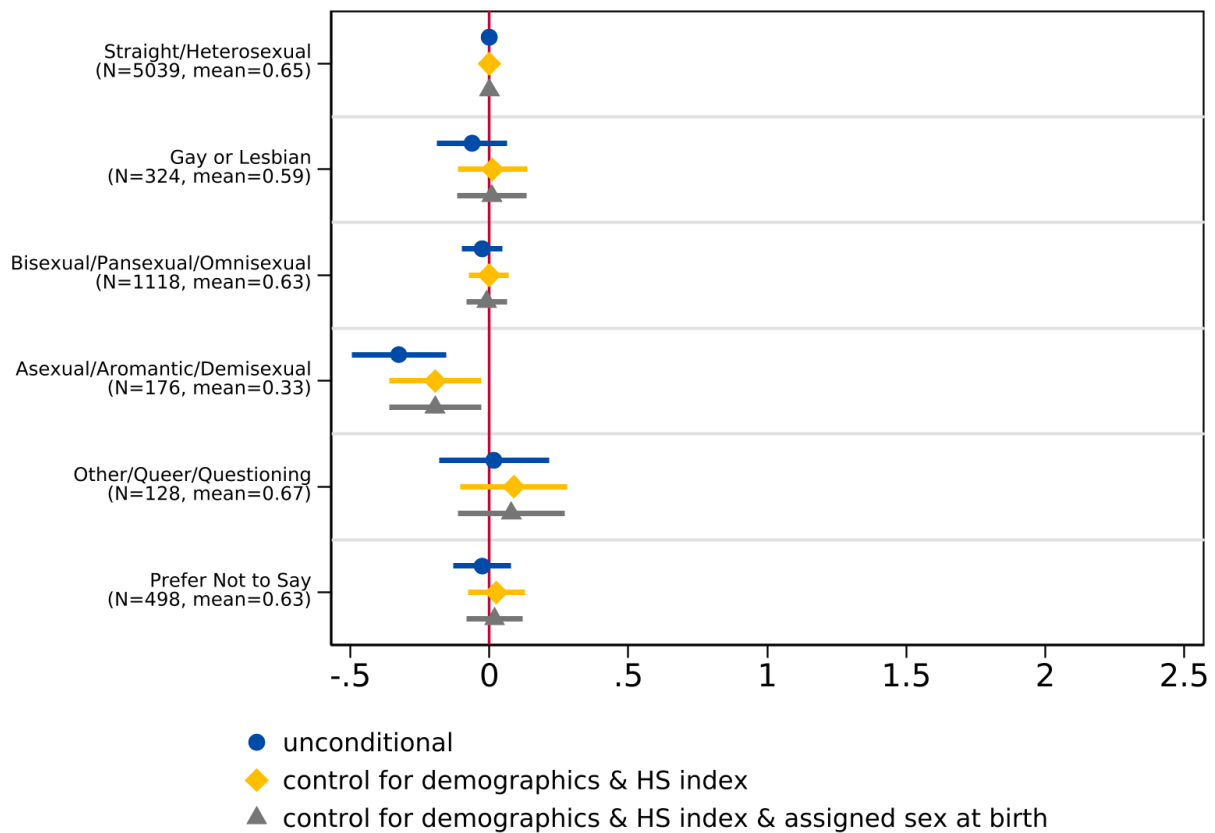
Notes: For the unconditional model, $n = 7,283$. For the model controlling for demographics (race/ethnicity and parental level of education), $n = 7,263$. For the model controlling for demographics (race/ethnicity and parental level of education) and sex assigned at birth, $n = 7,254$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each sexual orientation are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (straight/heterosexual).

FIGURE 3.15. Regression of College Worries About Discrimination on Sexual Orientation (with Assigned Sex at Birth Control)



Notes: For the unconditional model, $n = 7,283$. For the model controlling for demographics (race/ethnicity and parental level of education), $n = 7,263$. For the model controlling for demographics (race/ethnicity and parental level of education) and sex assigned at birth, $n = 7,254$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each sexual orientation are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (straight/heterosexual).

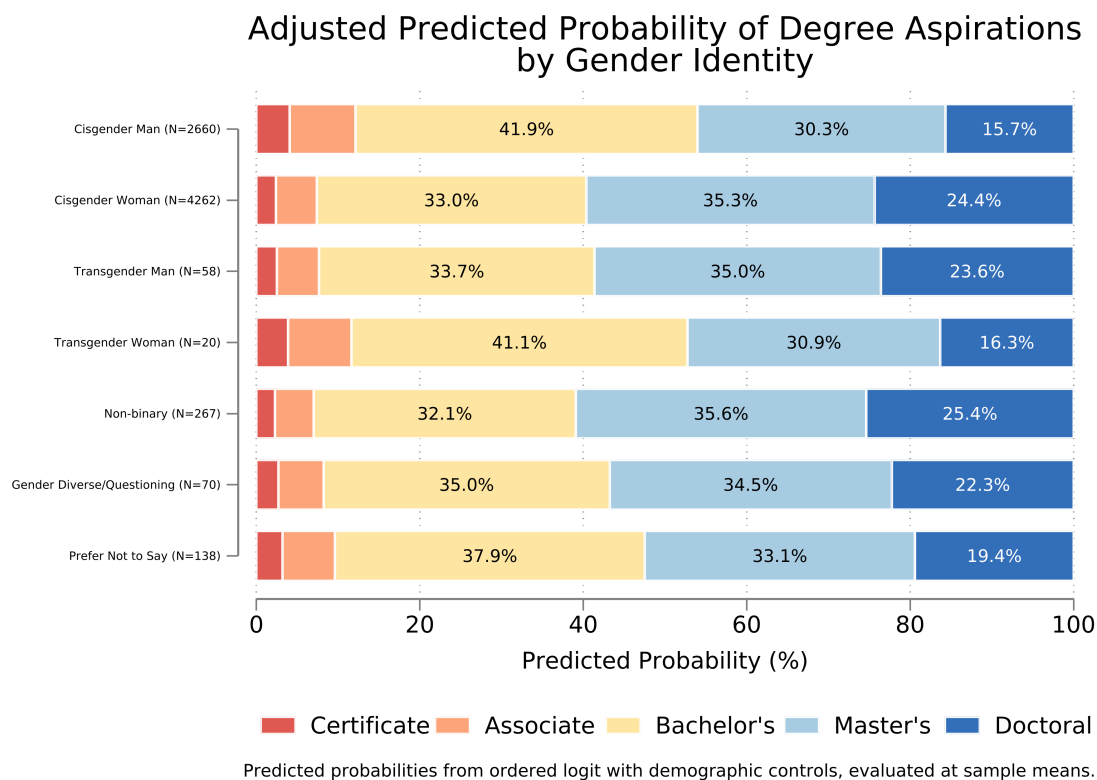
FIGURE 3.16. Regression of College Worries About Financial Burdens on Sexual Orientation (with Assigned Sex at Birth Control)



Notes: For the unconditional model, $n = 7,283$. For the model controlling for demographics (race/ethnicity and parental level of education), $n = 7,263$. For the model controlling for demographics (race/ethnicity and parental level of education) and sex assigned at birth, $n = 7,254$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each sexual orientation are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (straight/heterosexual).

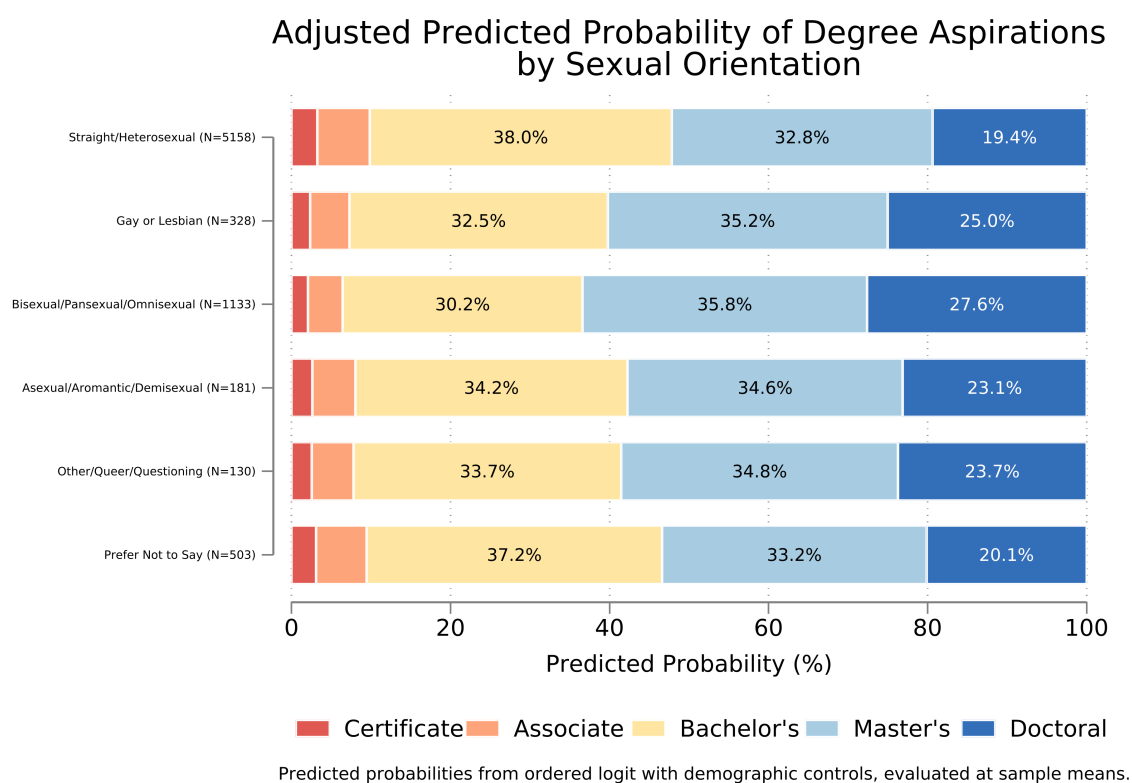
3.17. Predicted Probabilities of Degree Aspirations

FIGURE 3.17. Predicted Probability of Degree Aspirations by Gender Identity



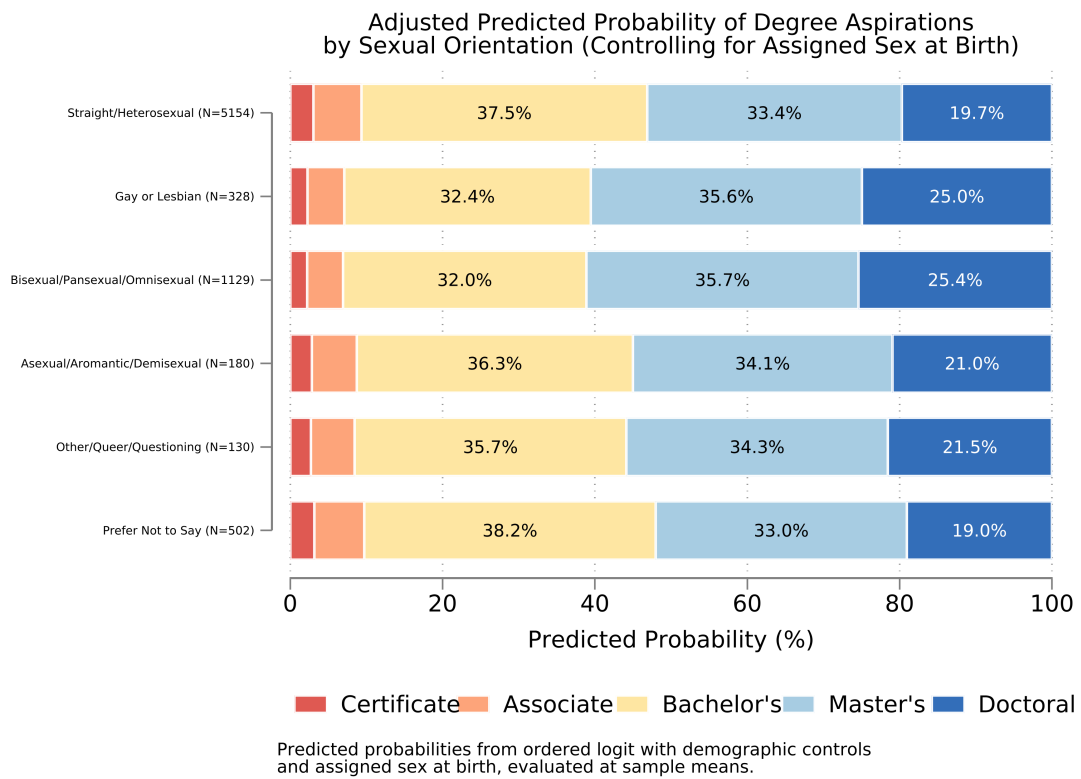
Notes: Controlling for demographics (race/ethnicity and parent education).

FIGURE 3.18. Predicted Probability of Degree Aspirations by Sexual Orientation



Notes: Controlling for demographics (race/ethnicity and parent education).

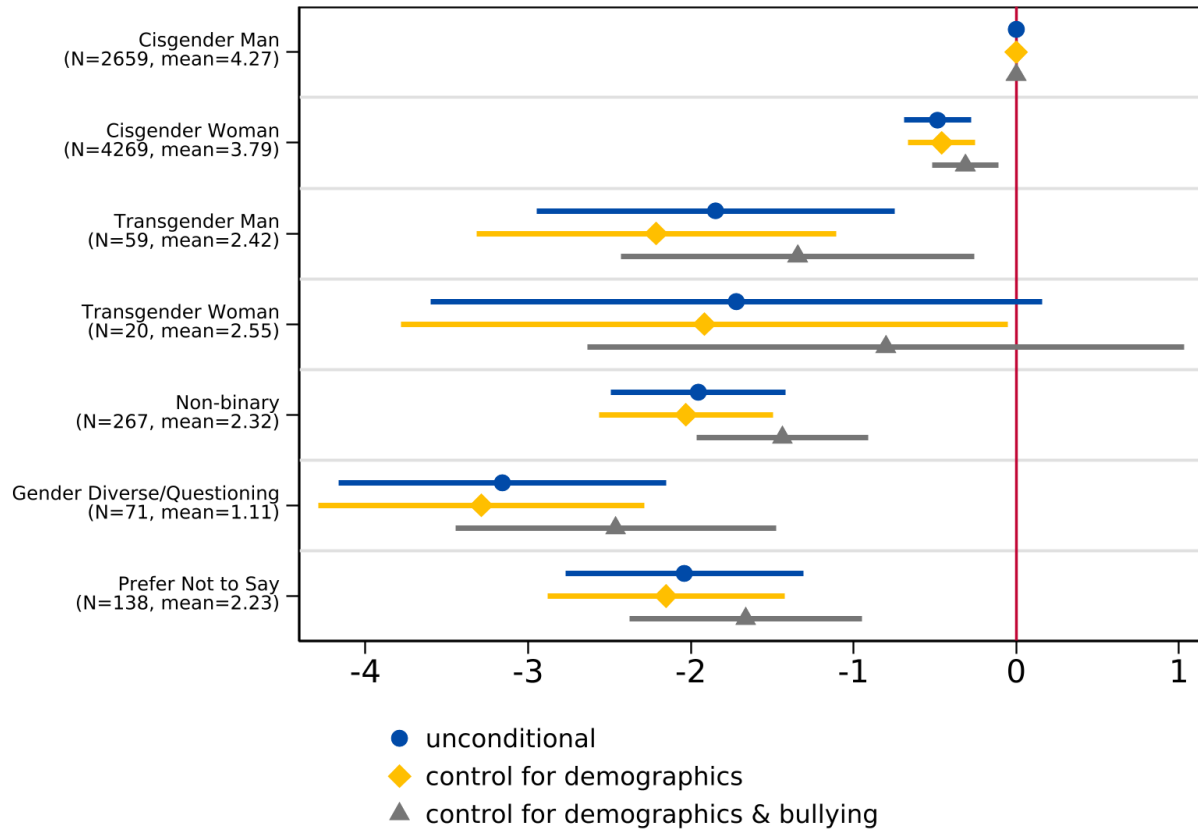
FIGURE 3.19. Predicted Probability of Degree Aspirations by Sexual Orientation (Controlling for Assigned Sex at Birth)



Notes: Controlling for demographics (race/ethnicity and parent education) and assigned sex at birth.

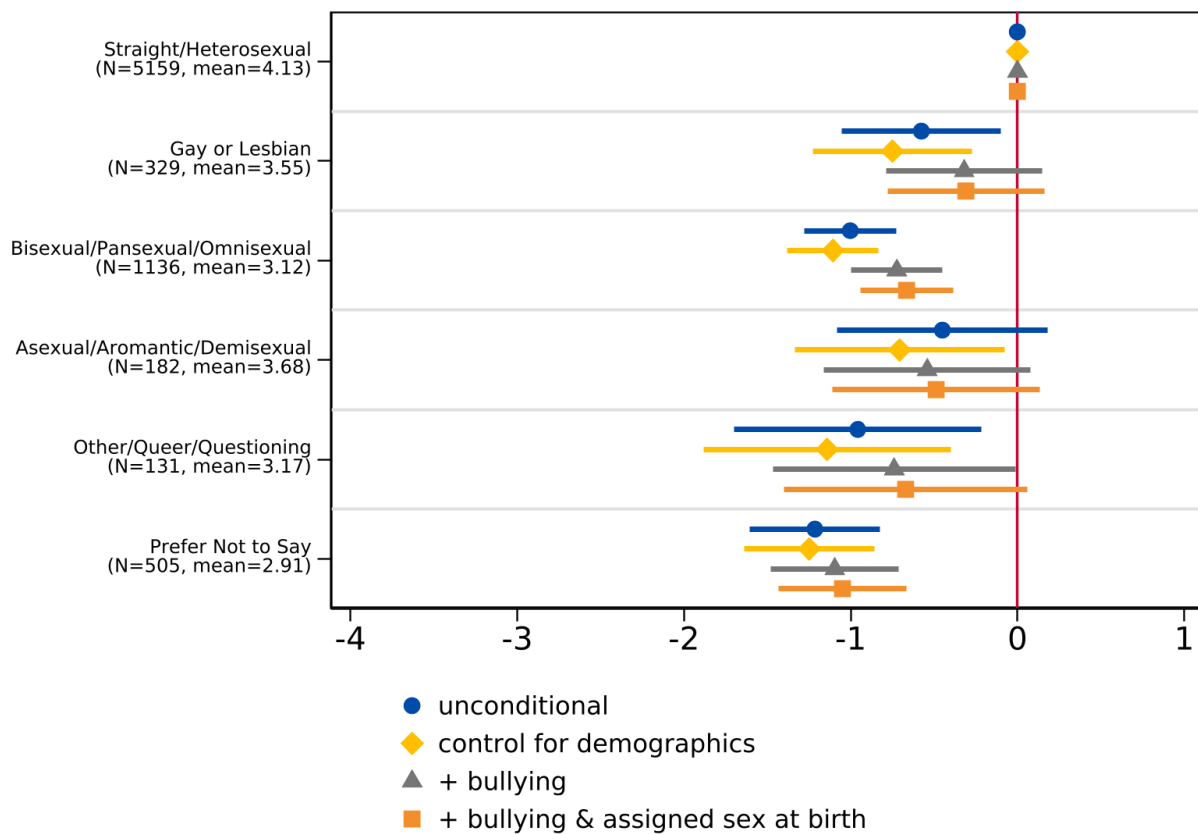
3.18. Bullying as Mediator for High School Experience

FIGURE 3.20. Regression of HS Experience Index on Gender Identity (Controlling for Bullying)



Notes: For the unconditional model, $n = 7,483$. For the model controlling for demographics (race/ethnicity and parental level of education), $n = 7,464$. For the model controlling for demographics and frequency of bullying, $n = 7,456$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each gender identity are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (cisgender man).

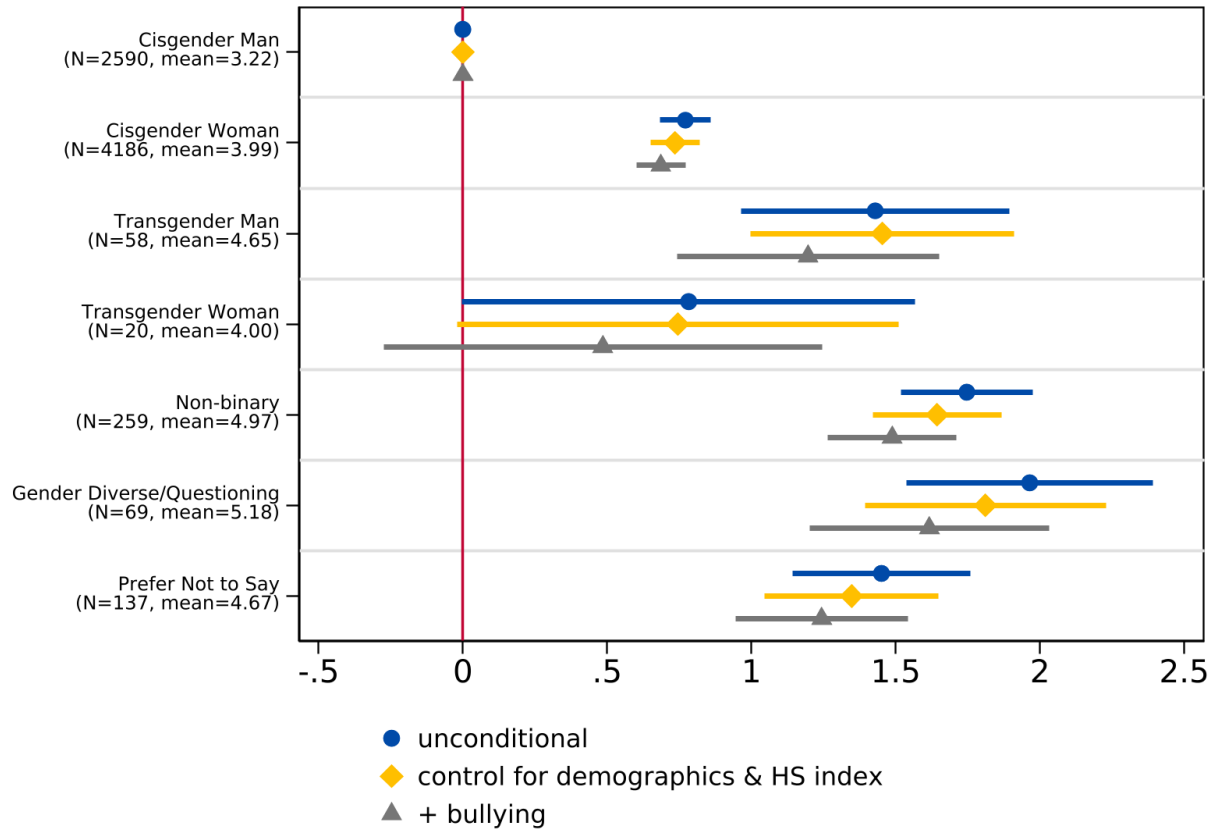
FIGURE 3.21. Regression of HS Experience Index on Sexual Orientation (Controlling for Bullying)



Notes: For the unconditional model, $n = 7,442$. For the model controlling for demographics (race/ethnicity and parental level of education), $n = 7,422$. For the model controlling for demographics and frequency of bullying, $n = 7,414$. For the model additionally controlling for assigned sex at birth, $n = 7,404$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each sexual orientation are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (straight/heterosexual).

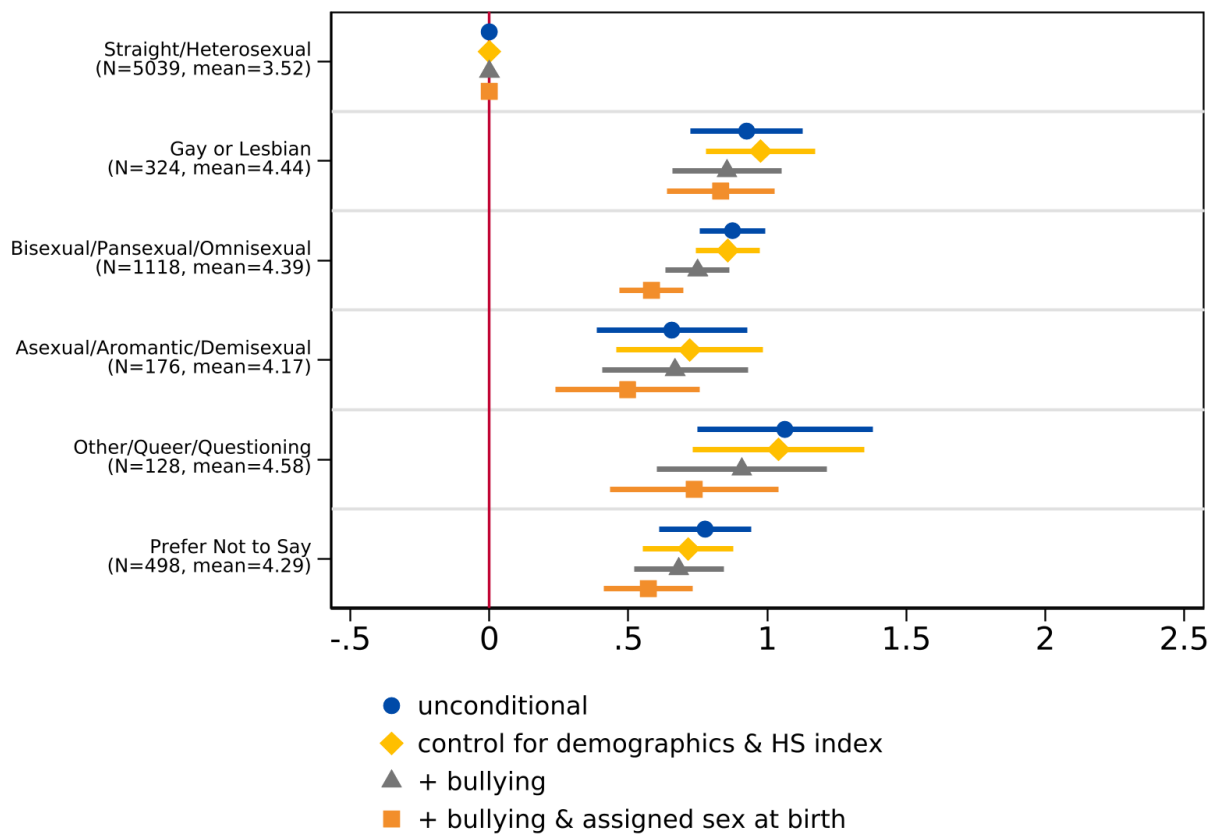
3.19. Bullying as Mediator for College Worries

FIGURE 3.22. Regression of General College Worries on Gender Identity (Controlling for Bullying)



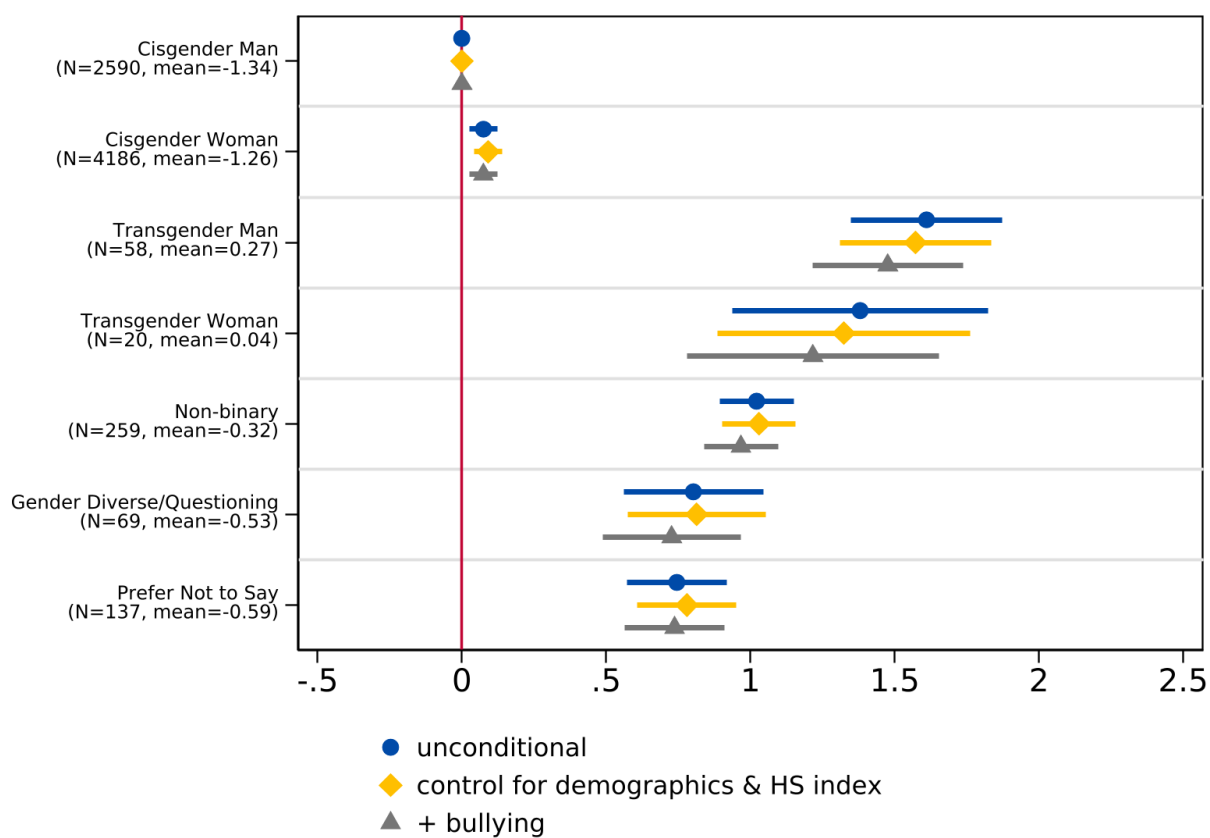
Notes: For the unconditional model, $n = 7,319$. For the model controlling for demographics (race/ethnicity and parental level of education) and the high school experience index, $n = 7,276$. For the model additionally controlling for frequency of bullying, $n = 7,269$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each gender identity are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (cisgender man).

FIGURE 3.23. Regression of General College Worries on Sexual Orientation (Controlling for Bullying)



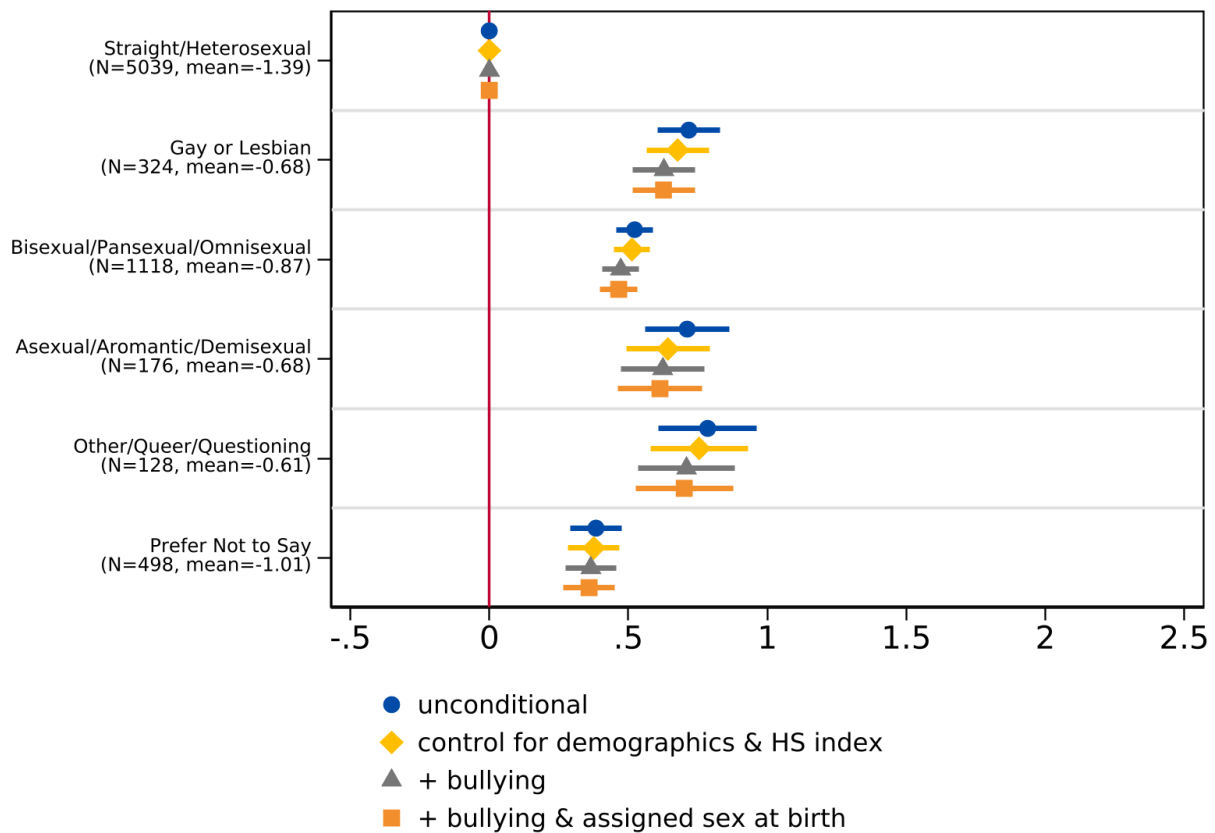
Notes: For the unconditional model, $n = 7,283$. For the model controlling for demographics (race/ethnicity and parental level of education) and the high school experience index, $n = 7,239$. For the model additionally controlling for frequency of bullying, $n = 7,232$. For the model additionally controlling for assigned sex at birth, $n = 7,223$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each sexual orientation are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (straight/heterosexual).

FIGURE 3.24. Regression of College Worries About Discrimination on Gender Identity (Controlling for Bullying)



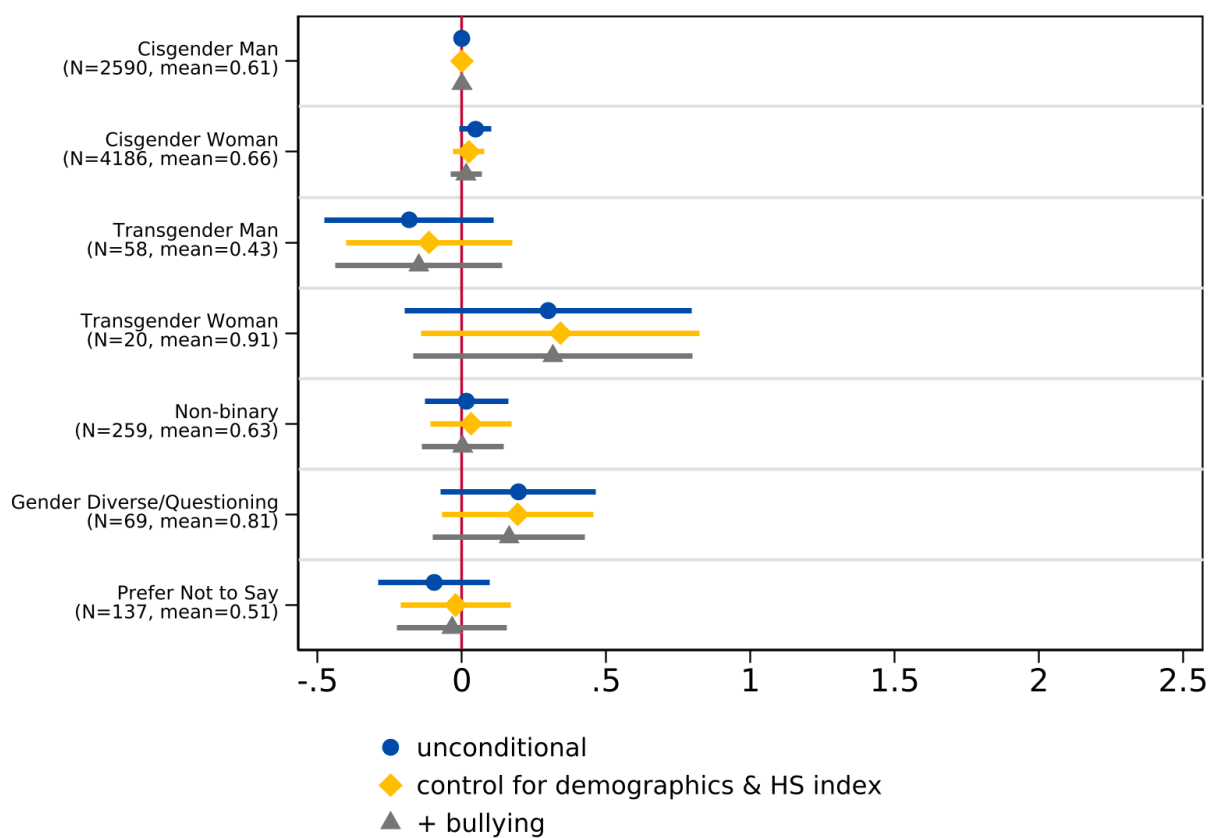
Notes: For the unconditional model, $n = 7,319$. For the model controlling for demographics (race/ethnicity and parental level of education) and the high school experience index, $n = 7,276$. For the model additionally controlling for frequency of bullying, $n = 7,269$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each gender identity are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (cisgender man).

FIGURE 3.25. Regression of College Worries About Discrimination on Sexual Orientation (Controlling for Bullying)



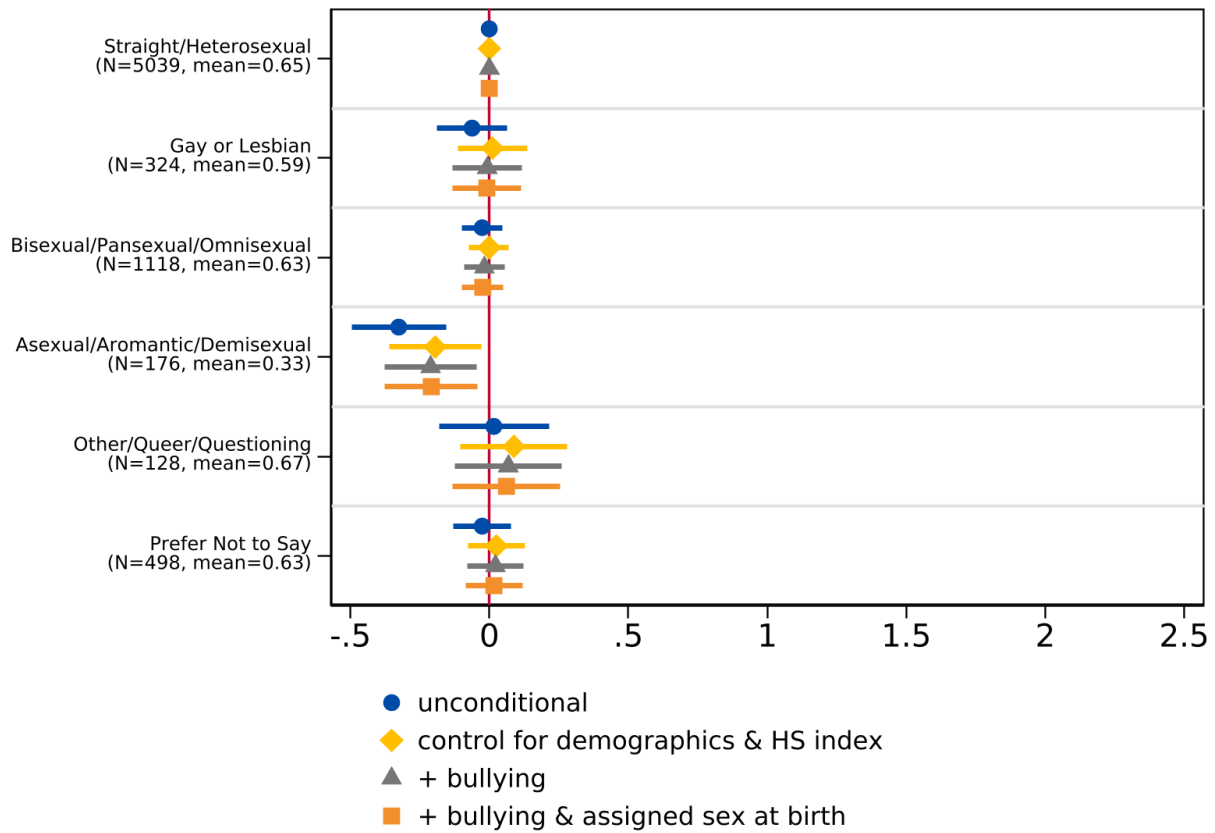
Notes: For the unconditional model, $n = 7,283$. For the model controlling for demographics (race/ethnicity and parental level of education) and the high school experience index, $n = 7,239$. For the model additionally controlling for frequency of bullying, $n = 7,232$. For the model additionally controlling for assigned sex at birth, $n = 7,223$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each sexual orientation are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (straight/heterosexual).

FIGURE 3.26. Regression of College Worries About Financial Burdens on Gender Identity (Controlling for Bullying)



Notes: For the unconditional model, $n = 7,319$. For the model controlling for demographics (race/ethnicity and parental level of education) and the high school experience index, $n = 7,276$. For the model additionally controlling for frequency of bullying, $n = 7,269$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each gender identity are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (cisgender man).

FIGURE 3.27. Regression of College Worries About Financial Burdens on Sexual Orientation (Controlling for Bullying)



Notes: For the unconditional model, $n = 7,283$. For the model controlling for demographics (race/ethnicity and parental level of education) and the high school experience index, $n = 7,239$. For the model additionally controlling for frequency of bullying, $n = 7,232$. For the model additionally controlling for assigned sex at birth, $n = 7,223$. The lines indicate 95% confidence intervals while the sample frequency and mean index of each sexual orientation are reported in parentheses. The vertical line at 0 represents the coefficient for the reference group (straight/heterosexual).

Bibliography

- Aisch, G., Gebeloff, R., & Quealy, K. (2014). Where We Came From and Where We Went, State by State. New York Times, 19 August.
- Alesina, A., Carlana, M., Ferrara, E. L., & Pinotti, P. (2024). Revealing Stereotypes: Evidence from Immigrants in Schools (No. 7), National Bureau of Economic Research. <https://doi.org/10.1257/aer.20191184>
- Alesina, A., Miano, A., & Stantcheva, S. (2023). Immigration and Redistribution. Review of Economic Studies, 90(1), 1–39. <https://doi.org/10.1093/restud/rdac011>
- Angrist, J. D., & Lang, K. (2004). Does school integration generate peer effects? evidence from boston’s metco program. American Economic Review, 94(5), 1613–1634.
- Aragon, S., Poteat, V., Espelage, D., & Koenig, B. (2014). The influence of peer victimization on educational outcomes for lgbtq and non-lgbtq high school students. Journal of LGBT Youth, 11(1), 1–19. <https://doi.org/10.1080/>
- Arrow, K. J. (1973). The Theory of Dscrimination. In O. Ashenfelter & A. Rees (Eds.), Discrimination in labor markets (pp. 3–33). Princeton University Press. <http://www.jstor.com/stable/j.ctt13x10hs.5%0AJSTOR>
- Austin, G., Hanson, T., Bala, N., & Zheng, C. (2023). Student engagement and well-being in california, 2019-21: Results of the eighteenth biennial state california healthy kids survey, grades 7, 9, and 11. https://data.calschls.org/resources/18th_
- Azrieli, Y., Chambers, C. P., & Healy, P. J. (2018). Incentives in experiments: A theoretical analysis. Journal of Political Economy, 126(4), 1472–1503. <https://doi.org/10.1086/698136>
- Badgett, M., Goldberg, N., Conron, K., & Gates, G. (2009). Best practices for asking questions about sexual orientation on surveys. The Williams Institute. <https://williamsinstitute.law.ucla.edu/wp-content/uploads/Best-Practices-SO-Surveys-Nov-2009.pdf>
- Becker, G. S. (1971). The Economics of Discrimination (2nd Editio). The University of Chicago Press. <https://doi.org/10.7208/chicago/9780226041049.001.0001>

- Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. Behavioral Science, 9(3), 226–232.
- Beemyn, G. (2022). The changing nature of gender in the 21st century: How trans and nonbinary students applying to college today self-identify. Campus Pride. https://www.campuspride.org/wp-content/uploads/CampusPride_ChangingNature
- Bergerson, A. (2009). College choice and access to college: Moving policy, research, and practice to the 21st century. ASHE Higher Education Report, 35(4), 1–141. <https://doi.org/10.1002/aehe.3504>
- Betts, J., & Fairlie, R. (2003). Does immigration induce ‘native flight’ from public schools into private schools. Journal of Public Economics, (87), 987–1012.
- Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). The more the merrier? the effect of family size and birth order on children’s education*. The Quarterly Journal of Economics, 120(2), 669–700.
- Bohren, J. A., Haggag, K., Imas, A., & Pope, D. G. (2020). Inaccurate Statistical Discrimination: An Identification Problem. <https://doi.org/10.2139/ssrn.3406060>
- Bohren, J. A., Hascher, J., Imas, A., Ungeheuer, M., & Weber, M. (2024). A Cognitive Foundation for Perceiving Uncertainty (tech. rep.). <https://ssrn.com/abstract=4706147>
- Bohren, J. A., Imas, A., & Rosenberg, M. (2019). The dynamics of discrimination: Theory and evidence. American Economic Review, 109(10), 3395–3436. <https://doi.org/10.1257/aer.20171829>
- Boje-Kovacs, B., Mulalic, I., Saiz, A., Sant’Anna, V., & Schultz-Nielsen, M. L. (2024). Immigrants and native flight: Geographic extent and heterogeneous preferences. Rockwool Foundation Research Unit.
- Breining, S., Doyle, J., Figlio, D. N., Karbownik, K., & Roth, J. (2020). Birth order and delinquency: Evidence from denmark and florida. Journal of Labor Economics, 38(1), 95–142.
- Brenøe, A. A., & Zölitz, U. (2020). Exposure to more female peers widens the gender gap in stem participation. Journal of Labor Economics, 38(4), 1009–1054.
- California Legislative Information. (2023). Ab 1955, support academic futures and educators for today’s youth (safety) act. https://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=202320240AB1955

- Callis, A. (2014). Bisexual, pansexual, queer: Non-binary identities and the sexual borderlands. *Sexualities*, *17*(1-2), 63–80. <https://doi.org/10.1177/1363460713511094>
- Campos-Mercade & Mengel Friederike. (2022). Non-Bayesian Statistical Discrimination. *134*. <https://ssrn.com/abstract=3843579>
- Carrell, S. E., Hoekstra, M., & Kuka, E. (2018). The long-run effects of disruptive peers. *American Economic Review*, *108*(11), 3377–3415. <https://doi.org/10.1257/aer.20160763>
- Cascio, E. U., & Lewis, E. G. (2012). Cracks in the melting pot: Immigration, school choice, and segregation. *American Economic Journal: Economic Policy*, *4*(3), 91–117. <https://doi.org/10.1257/pol.4.3.91>
- Center, P. R. (2019). Public views of immigration. Pew Research Center.
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, *104*(9), 2633–2679.
- Chetty, R., Hendren, N., & Katz, L. F. (2016). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, *106*(4), 855–902.
- Choi, K., Kelly Raley, R., Muller, C., & Riegle-Crumb, C. (2008). Class composition: Socioeconomic characteristics of coursemates and college enrollment. *Social Science Quarterly*, *89*(4), 846–866. <https://doi.org/10.1111/j.1540-6237.2008.00587.x>
- Clemens, M. A. (2021). Violence, development, and migration waves: Evidence from central american child migrant apprehensions. *Journal of Urban Economics*, *124*, 103355. <https://doi.org/https://doi.org/10.1016/j.jue.2021.103355>
- Coffman, K. B., Exley, C. L., & Niederle, M. (2019). The role of beliefs in driving gender discrimination (No. 18-054). https://www.hbs.edu/faculty/Publication%20Files/18-054_1668ace1-4d48-4133-bf7d-22a5146fee10.pdf
- Coffman, K. B., Kostyshak, S., & Saygin, P. O. (2024). Choosing and Using Information in Evaluation Decisions (tech. rep.). <https://ssrn.com/abstract=4783704>
- Cools, A., & Patacchini, E. (2021). Peer effects in education (J. Dronkers, E. Velez, & G. Fink, Eds.), —. <https://doi.org/10.4324/9780429202520-11>

- Cortes, K. E. (2004). Are refugees different from economic immigrants? some empirical evidence on the heterogeneity of immigrant groups in the united states. The Review of Economics and Statistics, 86(2), 465–480.
- Coutts, A. (2019). Good news and bad news are still news: experimental evidence on belief updating. Experimental Economics, 22(2), 369–395. <https://doi.org/10.1007/s10683-018-9572-5>
- Day, J., Perez-Brumer, A., & Russell, S. (2018). Safe schools? transgender youth’s school experiences and perceptions of school climate. Journal of Youth and Adolescence, 47(8), 1731–1742. <https://doi.org/10.1007/s10964-018-0866-x>
- De Quidt, J., Haushofer, J., & Roth, C. (2018). Measuring and bounding experimenter demand. American Economic Review, 108(11), 3266–3302. <https://doi.org/10.1257/aer.20171330>
- Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from head start. American Economic Journal: Applied Economics, 1(3), 111–34. <https://doi.org/10.1257/app.1.3.111>
- Denning, J. T., Murphy, R., & Weinhardt, F. (2023). Class rank and long-run outcomes. The Review of Economics and Statistics, 105(6), 1426–1441.
- DesJardins, S., Toutkoushian, R., Hossler, D., & Chen, J. (2019). Time may change me: Examining how aspirations for college evolve during high school. The Review of Higher Education, 43(1), 263–294. <https://doi.org/10.1353/rhe.2019.0096>
- Dugan, J., Kusel, M., & Simounet, D. (2012). Transgender college students: An exploratory study of perceptions, engagement, and educational outcomes. Journal of College Student Development, 53(5), 719–736. <https://doi.org/10.1353/csd.2012.0067>
- Dustmann, C., Frattini, T., & Preston, I. P. (2013). The effect of immigration along the distribution of wages. Review of Economic Studies, 80(1), 145–173.
- Equality California. (2025). 2025 legislative priorities. <https://www.eqca.org/>
- Eyting, M. (2022). Why do we Discriminate? The Role of Motivated Reasoning (No. 356). <https://doi.org/10.2139/ssrn.4210315>
- Feldman, S., Watson, R., & Gallik, C. (2022). College aspirations, gender sexuality alliances, and teacher support among diverse lgbtq youth. Educational Review, 74(2), 281–297. <https://doi.org/10.1080/00131911.2020.1816907>

- Fernandes, T., Alves, B., & Gato, J. (2023). Between resilience and agency: A systematic review of protective factors and positive experiences of lgbtq+ students. *Healthcare*, *11*(14). <https://doi.org/10.3390/healthcare11142098>
- Figlio, D., & Ozek, U. (2019). Unwelcome guests? the effects of refugees on the educational outcomes of incumbent students. *Journal of Labor Economics*, *4*(37), 1061–1096.
- Figlio, D., Giuliano, P., Marchingiglio, R., Ozek, U., & Sapienza, P. (2021). Diversity in schools: Immigrants and the educational performance of u.s. born students. *IZA Working Paper #14196*.
- Fletcher, J. (2015). Social interactions and college enrollment: A combined school fixed effects/instrumental variables approach. *Social Science Research*, *52*, 494–507. <https://doi.org/10.1016/j.ssresearch.2015.03.004>
- Fotopoulou, A. (2012). Intersectionality queer studies and hybridity: Methodological frameworks for social research. *Journal of International Women's Studies*, *13*(2), 19–32. <https://vc.bridgew.edu/jiws/vol13/iss2/3>
- Frydman, C., & Jin, L. J. (2023). *On the Source and Instability of Probability Weighting* (tech. rep.). SSRN Working Paper No. 4493313. <https://ssrn.com/abstract=4493313>
- Galupo, M., Mitchell, R., & Davis, K. (2015). Sexual minority self-identification: Multiple identities and complexity. *Psychology of Sexual Orientation and Gender Diversity*, *2*(4), 355–364. <https://doi.org/10.1037/sgd0000131>
- Gershenson, S., Hart, C. M. D., Hyman, J., Lindsay, C. A., & Papageorge, N. W. (2022). The long-run impacts of same-race teachers. *American Economic Journal: Economic Policy*, *14*(4), 300–342. <https://doi.org/10.1257/pol.20190573>
- Getik, D., Sjögren, A., & Sundberg, A. (2024). *Migration inflow and the school performance of incumbent students* (tech. rep.). Institute for Evaluation of Labour Market and Education Policy (IFAU).
- Gould, E. D., Lavy, V., & Paserman, M. D. (2009). Does immigration affect the long-term educational outcomes of natives? quasi-experimental evidence. *Economic Journal*, (540), 1243–1269.
- Gupta, N., Rigotti, L., & Wilson, A. (2021). *The Experimenters' Dilemma: Inferential Preferences over Populations* (July). <http://arxiv.org/abs/2107.05064>

- Guryan, J. (2004). Desegregation and black dropout rates. American Economic Review, 94(4), 919–943.
- Haaland, I., & Roth, C. (2019). Beliefs About Racial Discrimination and Support for Pro-Black Policies (August). <https://doi.org/10.2139/ssrn.3334961>
- Haaland, I., & Roth, C. (2020). Labor market concerns and support for immigration. Journal of Public Economics, 191. <https://doi.org/10.1016/j.jpubeco.2020.104256>
- Hammack, P., Hughes, S., Atwood, J., Cohen, E., & Clark, R. (2022). Gender and sexual identity in adolescence: A mixed-methods study of labeling in diverse community settings. Journal of Adolescent Research, 37(2), 167–220. <https://doi.org/10.1177/>
- Hanson, G. H. (2010). The economics and policy of illegal immigration in the united states (Report). Migration Policy Institute.
- Hatchel, T., Valido, A., Pedro, K., Huang, Y., & Espelage, D. (2018). Minority stress among transgender adolescents: The role of peer victimization, school belonging, and ethnicity. Journal of Child and Family Studies, 28(9), 2467–2476. <https://doi.org/10.1007/s10826-018-1168-3>
- Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Corwin.
- Heck, N., Lindquist, L., Machek, G., & Cochran, B. (2014). School belonging, school victimization, and the mental health of lgbt young adults: Implications for school psychologists. School Psychology Forum, 8(1), 28–37.
- Hertwig, R., & Erev, I. (2009, December). The description-experience gap in risky choice. <https://doi.org/10.1016/j.tics.2009.09.004>
- Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. (7867). <https://doi.org/10.3386/w7867>
- Hunter, J., Butler, C., & Cooper, K. (2021). Gender minority stress in trans and gender diverse adolescents and young people. Clinical Child Psychology and Psychiatry, 26(4), 1182–1195. <https://doi.org/10.1177/13591045211033187>

- Johnson, R. C. (2011, January). Long-run impacts of school desegregation & school quality on adult attainments (NBER Working Papers No. 16664). National Bureau of Economic Research, Inc.
- Khaw, M. W., Li, Z., & Woodford, M. (2021). Cognitive Imprecision and Small-Stakes Risk Aversion. Review of Economic Studies, 88(4), 1979–2013. <https://doi.org/10.1093/restud/rdaa044>
- Klasik, D., Holzman, B., & Jones, A. (2025). The many paths to college enrollment: Re-conceptualizing the transition to college [EdWorkingPaper: 25-1269]. Annenberg Institute at Brown University.]. <https://doi.org/10.26300/54c6-2316>
- Kosciw, J., Clark, C., & Menard, L. (2022). The 2021 national school climate survey: The experiences of lgbtq+ youth in our nation's schools. GLSEN. <https://www.glsen.org/sites/default/files/2022-10/NSCS-2021-Full-Report.pdf>
- Krueger, A. B. (1999). Experimental estimates of education production functions. Quarterly Journal of Economics, 114(2), 497–532.
- Kurlaender, M., Reed, S., & Hurtt, A. (2019, August). Improving college readiness: A research summary and implications for practice. Policy Analysis for California Education. https://edpolicyinca.org/sites/default/files/R_Kurlaender_Aug19_01.pdf
- Kuziemko, I., Norton, M. I., Saez, E., & Stantcheva, S. (2015, April). How elastic are preferences for redistribution? Evidence from randomized survey experiments. <https://doi.org/10.1257/aer.20130360>
- Linley, J., Remm, K., & Woodford, M. (2018). Examining the ecological systems of lgbtq stem majors. Journal of Women and Minorities in Science and Engineering, 24(1), 1–16. <https://doi.org/10.1615/jwomenminorscieng.2017018836>
- Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. The Review of Economic Studies, 60(3), 531–542.
- Massey, D. S., Durand, J., & Malone, N. J. (2002). Beyond smoke and mirrors: Mexican immigration in an era of economic integration. Russell Sage Foundation.
- Mayda, A. M., Peri, G., & Steingress, W. (2022). The political impact of immigration: Evidence from the united states. American Economic Journal: Applied Economics, 14(1), 358–389.

- McBride, R. (2021). A literature review of the secondary school experiences of trans youth. Journal of LGBT Youth, 18(2), 103–134. <https://doi.org/10.1080/19361653>.
- Means, B., Wang, H., Young, V., Peters, V., & Lynch, S. (2016). Stem-focused high schools as a strategy for enhancing readiness for postsecondary stem programs. Journal of Research in Science Teaching, 53(5), 709–736. <https://doi.org/10.1002/tea.21313>
- Mobius, M. M., Niederle, M., Niehaus, P., & Rosenblat, T. S. (2014). Managing Self-Confidence.
- Morales, C. (2022). Do refugee students affect the academic achievement of peers? evidence from a large urban school district. Economics of Education Review, 89, 102283.
- Morgan, E. (2013). Contemporary issues in sexual orientation and identity development in emerging adulthood. Emerging Adulthood, 1(1), 52–66. <https://doi.org/10.1177/2167696812469187>
- Movement Advancement Project. (2025). Equality maps: Safe schools laws. https://www.lgbtmap.org/equality-maps/safe_school_laws
- Mummolo, J., & Peterson, E. (2019). Demand effects in survey experiments: An empirical assessment. American Political Science Review, 113(2), 517–529. <https://doi.org/10.1017/S0003055418000837>
- Oprea, R., & Vieider, F. M. (2024). Minding the Gap: On the Origins of Probability Weighting and the Description-Experience Gap * (tech. rep.).
- Ottaviano, G. I., & Peri, G. (2012). Rethinking the effect of immigration on wages. Journal of the European economic association, 10(1), 152–197.
- Pedersen, M. J., & Nielsen, V. L. (2024). Understanding Discrimination: Outcome-Relevant Information Does Not Mitigate Discrimination. Social Problems, 71(1), 77–105. <https://doi.org/10.1093/socpro/spac006>
- Pedraza, S. (2007). Political disaffection in cuba’s revolution and exodus. Cambridge University Press.
- Peer, E., Rothschild, D., Gordon, A., Evernden, Z., & Damer, E. (2022). Data quality of platforms and panels for online behavioral research. Behavior Research Methods, 54(4), 1643–1662. <https://doi.org/10.3758/s13428-021-01694-3>
- Pennell, S. (2016). Queer cultural capital: Implications for education. Race Ethnicity and Education, 19(2), 324–338. <https://doi.org/10.1080/13613324.2015.1013462>

- Peri, G. (2012). The effect of immigration on productivity: Evidence from us states. Review of Economics and Statistics, 94(1), 348–358.
- Phelps, E. S. (1972). The Statistical Theory of Racism and Sexism. American Economic Review, 62(4), 659–661.
- Reed, S., Hurtt, A., Kurlaender, M., Luu, J., & Merritt, C. (2023). Inequality in academic preparation for college ([Report].). Policy Analysis for California Education. https://edpolicyinca.org/sites/default/files/2023-07/r_reed-july2023.pdf
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. American Economic Review, 94(2), 247–252. <https://doi.org/10.1257/0002828041302244>
- Sacerdote, B. (2011). Chapter 4 - peer effects in education: How might they work, how big are they and how much do we know thus far? (E. A. Hanushek, S. Machin, & L. Woessmann, Eds.). 3, 249–277. <https://doi.org/https://doi.org/10.1016/B978-0-444-53429-3.00004-1>
- Sansone, D. (2019). Lgbt students: New evidence on demographics and educational outcomes. Economics of Education Review, 73, 101933. <https://doi.org/10.1016/>
- Schultz, B. (2025). President trump's early actions undo Biden efforts to protect LGBTQ+ students. Education Week. <https://www.edweek.org/policy-politics/president-trumps-early-actions-undo-biden-efforts-to-protect-lgbtq-students/2025/01>
- Székel, G. J., & Rizzo, M. L. (2004). Testing for equal distributions in high dimension. InterStat, 5(16.10), 1249–1272.
- Texas Education Agency. (2017). Attendance, admission, enrollment records, and tuition [Section IV summarizes enrollment verification requirements]. <https://tea.texas.gov/about-tea/news-and-multimedia/correspondence/taa-letters/attendance-admission-enrollment-records-and-tuition-august-2017>
- Toutkoushian, R., Stollberg, R., & Slaton, K. (2018). Talking 'bout my generation: Defining "first-generation college students" in higher education research. Teachers College Record, 120(4), 1–38. <https://doi.org/10.1177/016146811812000407>
- Tuttle, C. (2019). The long-run economic effects of school desegregation. Available at SSRN, 3460993.
- Ungar, M. (2011). The social ecology of resilience: Addressing contextual and cultural ambiguity of a nascent construct. American Journal of Orthopsychiatry, 81(1), 1–17. <https://doi.org/10.1111/j.1939-0025.2010.01067.x>

- Uwah, C., McMahon, H., & Furlow, C. (2008). School belonging, educational aspirations, and academic self-efficacy among african american male high school students: Implications for school counselors. Professional School Counseling, 11(5), 296–305. <https://doi.org/10.1177/2156759x0801100503>
- Van der Werf, C. (2021). The impact of refugees on native students' academic achievement. Louisiana State University.
- Watson, R., Wheldon, C., & Puhl, R. (2020). Evidence of diverse identities in a large national sample of sexual and gender minority adolescents. Journal of Research on Adolescence, 30(S2), 431–442. <https://doi.org/10.1111/jora.12488>
- White, A., Moeller, J., Ivcevic, Z., & Brackett, M. (2018). Gender identity and sexual identity labels used by us high school students: A co-occurrence network analysis. Psychology of Sexual Orientation and Gender Diversity, 5(2), 243–252. <https://doi.org/10.1037/08912432211038689>
- White, S., Potter, L., You, H., Valencia, L., Jordan, J., & Pecotte, B. (2016). Introduction to Texas Domestic Migration. Office of the State Demographer.
- Wilkinson, L., Shifrer, D., & Pearson, J. (2021). Educational outcomes of gender-diverse youth: A national population-based study. Gender & Society, 35(5), 806–837. <https://doi.org/10.1177/08912432211038689>
- Woodford, M. (2012). Prospect theory as efficient perceptual distortion. American Economic Review, 102, 41–46. <https://doi.org/10.1257/aer.102.3.41>
- Wozniak, D., & Macneill, T. (2018). Ethnic Discrimination in the Lab: Evidence of Statistical and Taste-Based Discrimination 1 (tech. rep.). <https://ssrn.com/abstract=2613075>
- Zimmermann, F. (2020). The dynamics of motivated beliefs. American Economic Review, 110(2), 337–363. <https://doi.org/10.1257/aer.20180728>