Beliefs & Demand for Mental Health Services Among University Students*

Alisher Batmanov

UC San Diego

Idaliya Grigoryeva

UC San Diego

Bruno Calderón-Hernández

ITAM

Roberto González-Téllez

ITAM

Alejandro Guardiola

Tec de Monterrey

March 24, 2025

Abstract

This paper investigates the role of beliefs and stigma in shaping students' use of professional mental health services at a large private university in Mexico, where supply-side barriers are minimal and services are readily accessible. In an online experiment with 680 students, we estimate a large treatment gap with nearly 50% of students in distress not receiving professional mental health support despite a high level of awareness and perceived effectiveness. We document stigmatized beliefs and misconceptions correlated with the treatment gap. For example, three-quarters of students incorrectly believe that those in distress perform worse academically, and many underestimate how common therapy use is among their peers. To correct inaccurate beliefs, we implement an information intervention and find that it increases students' willingness to share on-campus mental health resources with peers and encourages them to recommend these resources when advising a friend in distress. However, we also find that it lowers their willingness to pay for external services, suggesting a potential substitution effect from private therapy to free on-campus resources.

1   Introduction

Student mental health and wellbeing are issues of growing concern, with suicide being the 3rd leading cause of death among 15–29 years-old's and rates of depression and anxiety continuously rising (WHO 2021). At the same time, among over 100,000 adults surveyed across 30 countries in the World Mental Health Surveys, more than 80% of those struggling with depression, anxiety, or substance use disorders report not receiving any professional support, contributing to the "treatment gap" (Patel et al. 2018). That is despite widely recognized treatments to reduce depression and anxiety, such as cognitive behavioral therapy (CBT) (Cuijpers et al. 2013, 2016). This treatment gap exceeds 90% in most developing countries, with a staggering 95% of people in distress lacking professional help in countries like Mexico. (Wang et al. 2007). Importantly, such a wide gap is present even in settings where treatments are available (mitigating supply-side constraints) as the low take-up of such interventions is attributed to cognitive and behavioral biases, as well as low perceived effectiveness or need (Ridley et al. 2020; Patel et al. 2018; Andrade et al. 2014; Thapar et al. 2022).

Mental distress often carries far-reaching consequences for educational attainment and other long-term economic outcomes (Ridley et al. 2020), with depression and anxiety — two of the most prevalent mood disorders[1] — disrupting students' educational trajectories and potentially constraining future employment and socio-economic mobility (Oreopoulos 2007; Cornaglia et al. 2015; Fletcher 2008). Facing the pressure to perform academically while becoming independent adults, college students stand to benefit substantially from getting professional help, which can help prevent more severe depression and anxiety. However, previous studies among students who have access to university-provided counseling still show that the majority of students in distress do not receive professional mental health support (Acampora et al. 2022)[2], which raises a question about the role of demand-side factors related to student beliefs and attitudes toward therapy. Taking the setting with free on-campus counseling and high perceived effectiveness of therapy, we designed our study to determine whether incorrect beliefs drive the remaining treatment gap and assess the potential of a light-touch belief-correction intervention to reduce it.

We conduct a survey experiment with a representative sample of 680 students from a large private university in Mexico to document the size of the treatment gap, examine students'

---

[1]In this paper, we focus on mood disorders, and more specifically depression and anxiety, where CBT and other talk therapy treatments have been demonstrated to be effective and often are provided by the university to students. We will not address more severe mental illnesses, such as schizophrenia or bipolar disorder, which typically necessitate psychiatric interventions combined with medications.

[2]Studies of non-representative or not college-specific student samples limit the analysis of the demand factors without data on the supply. Acampora et al. (2022) conducted the only comparable study focusing on a single institution measuring demand for university-specific and outside mental health services, which was conducted in a large university in the Netherlands.

beliefs about mental health and therapy use, and correct potential misconceptions in an information treatment. To our knowledge, this is the first study in a developing country that leverages a broadly representative sample of the student population at the university level to document both the prevalence of psychological distress and the factors shaping demand for support, while also assessing how inaccurate beliefs contribute to the treatment gap.[3]. Focusing on a private university where mental health services are readily available and widely known allows us to examine barriers to treatment in a setting where the supply of services is not a limiting factor, providing a suggestive lower bound on the treatment gap compared to students at public institutions, who face greater financial and access barriers. Given that perceived effectiveness of therapy is high in this context, addressing inaccurate beliefs about who seeks therapy and the stigmatized relation between mental health support and academic performance may be more effective in reducing barriers to care.

We find that there is a significant mental health treatment gap among university students in our sample, despite the availability of free on-campus counseling services. In our sample, nearly 1 in 4 students exhibit moderate to severe symptoms of depression or anxiety, and among them nearly half do not receive professional mental health support: we estimate a treatment gap of nearly 50% of students in distress not having used any professional mental health support services in the last 12 months. Notably, this gap is present even though over 90% of students in distress agree that therapy can improve their mental wellbeing substantially, and 80% of them believe the university provides a good support system for mental or emotional health. The treatment gap is significantly larger among male students, as well as students not open to sharing their mental health struggles with classmates, potentially hinting at some negative beliefs about therapy and seeking help. Interestingly, while financial stress is highly positively correlated with mental distress, there is no significant association between financial stress and treatment gap.[4]

Further analysis indicates that this gap is associated with stigmatized beliefs and prevalent negative stereotypes related to mental distress and help-seeking. We identify a particularly pervasive misconception as 3 out of 4 respondents believe that students in mental distress academically perform *worse or much worse* compared to students not in distress — despite no observed correlation between GPA and mental distress score across students in our sample. This highlights

---

[3]Most existing studies on university students' mental health and treatment use come from developed countries. For instance, an empirical study in the Netherlands examines an intervention targeting student mental health and therapy use (Acampora et al. 2022), while survey-based studies have documented related measures in Norway (Sæther et al. 2021) and among college students in the World Mental Health Surveys across 21 countries (Auerbach et al. 2016). A systematic review by Mortier et al. (2018) provides further references to studies using college-student data.

[4]Surprisingly, we even observe that students with a stressful financial situation are marginally *more likely* to seek help when in distress, in particular by being much more likely to seek professional help on campus compared to students not reporting struggling with finances, although these differences are not statistically significant.

a prevalent stereotype of associating mental health struggles with low academic achievement, which may discourage students from sharing their mental health struggles or revealing going to therapy, as these could be construed as signals of lower performance. Among students in distress who do not seek help, 81% guess a negative correlation in an incentivized question, relative to 74% among the rest of the students.

We document that many students underestimate how many of their peers seek professional mental health help and are open to discussing mental health struggles while overestimating the prevalence of self-stigma, resulting in a more pessimistic view of public perceptions of stigma and their peers' attitudes toward mental distress.[5] Our results broadly echo the findings from recent online and field experimental studies that identify misconceptions around willingness to discuss mental health issues and the prevalence of mental-health-related beliefs among others as potential evidence of stigma (Roth et al. 2024a; Ridley 2025; Ronak & Khandelwal 2024; Acampora et al. 2022).

Having identified that the treatment gap is partially driven by inaccurate beliefs and associated perceived stigma, we design an information intervention to correct misperceptions about mental health in three ways: (1) conveying that psychotherapy has long-term (4–5 years) benefits in reducing instances of depression, (2) normalizing therapy by noting that most students at their university who seek it do not have severe symptoms, reinforcing that therapy isn't just for those in crisis, and (3) countering the negative bias linking distress to academic performance by informing students that GPA and mental distress are uncorrelated. While 97% of subjects had a correct prior on the long-term effectiveness of psychotherapy (prior 1), we find that nearly half held incorrect priors about the proportion of students in therapy with mild or no symptoms (prior 2), and 75% incorrectly believed there was a negative correlation between GPA and mental distress (prior 3).[6]

To evaluate the effectiveness of interventions aimed at addressing these misperceptions, respondents were randomized into one of three treatment conditions: *Information + Reflection* (T1), *Information Only* (T2), and *Control*. Both treatment groups received an information-based intervention, while T1 included an additional reflection component prompting participants to engage with a vignette and consider strategies for managing mental health challenges. The Control

---

[5]Public/social stigma refers to societal disapproval of individuals perceived as deviating from norms. Self-stigma, in contrast, occurs when individuals internalize these negative societal views, leading to feelings of shame or diminished self-worth. Experiencing mental distress can be associated with both forms of stigma.

[6]Note that while we document wide-spread belief in effectiveness (and corresponding to it correct guesses on prior 1 about long-term positive effects), we did not assume this at the design stage and included this treatment based on well-documented role of low perceived effectiveness as one of the main barriers to seeking help and drivers of the treatment gap (Andrade et al. 2014).

group did not receive any mental health-related information but instead answered questions about general campus services to ensure comparable cognitive engagement across conditions.[7]

The information intervention yields three main insights. First, participants in the treatment groups were more likely to engage with and share the link to the campus psychological counseling services, with a click-through rate nearly twice that of the control group. Since they were encouraged to share the link, this hints at potentially greater dissemination of the resource among their peers. This effect was primarily driven by the *Information Only* condition (T2), suggesting that providing information alone was sufficient to encourage sharing resources with others. When asked to provide incentivized advice to a hypothetical friend in distress, treated participants were slightly more likely to mention on-campus counselling services and significantly less likely to give passive advice.[8] Lastly, somewhat unexpectedly, participants in the treatment groups exhibited a lower willingness to pay (WTP) for a one-month online therapy subscription for themselves and for a friend. This result may be due to a substitution effect, as the intervention emphasized campus-provided services rather than external options, such as the online therapy, highlighting a potential limitation of this outcome measure.

This paper contributes to the literature on mental health economics, behavioral frictions in help-seeking behavior, and the role of information interventions in addressing misperceptions and treatment gaps, particularly in developing countries. We provide new evidence on demand-side constraints in a setting where professional mental health services are available on campus, allowing us to isolate attitudinal and informational barriers from structural supply-side constraints. While previous work has examined the role of affordability and availability (Patel et al. 2017; Barker et al. 2022; Haushofer et al. 2021; Bhat et al. 2022), we contribute by documenting how belief distortions and stigma inhibit take-up despite widespread recognition of therapy's benefits. This extends the literature on behavioral constraints affecting mental health decisions (Schilbach et al. 2016; Shreekumar & Vautrey 2023) and connects to broader discussions on the psychology of poverty and its implications for economic decision-making (Schilbach et al. 2016; Rao et al. 2021).

Second, we contribute to the literature on mental health stigma and misconceptions by systematically documenting belief distortions among students regarding therapy use and academic performance. We find that students systematically overestimate the negative relationship between mental distress and GPA—a belief that may contribute to stigma and discourage help-seeking behavior, complementing an earlier result on productivity and mental distress in a stylized online setting (Ridley 2025) with a measurable and relevant productivity measure in

---

[7]Average time to complete the survey is not statistically different between joint treatment and control groups ($p = 0.909$). We also fail to reject the null hypothesis that survey duration is equal between the two treatment groups ($p = 0.559$).

[8]Treated participants were more likely to validate their friend's feelings and use language that signaled attentiveness in their responses.

an academic setting. We further complement existing online experiments with US adults (Roth et al. 2024a,b) with a more real-life setting and an interpersonally connected student sample from a single university in a developing country, with additional insights capturing behaviors around promoting help-seeking among students via link sharing and giving advice to a friend. Extending on the two main field experiments related to mental health, stigma and treatment take-up in India and Nepal (Ronak & Khandelwal 2024; Lacey et al. 2024), we leverage the setting where supply is reasonably available to zoom in on demand-side factors and beliefs. While prior work has explored information provision as a tool for reducing stigma and increasing take-up (Osman et al. 2022; Acampora et al. 2022; Ronak & Khandelwal 2024), our study provides suggestive evidence that correcting misperceptions may shift help-seeking behavior toward free, on-campus resources rather than increasing overall demand.

The rest of the paper proceeds as follows. Section 2 describes the setting, survey design and data collection. Section 3 documents the prevalence of mental distress, professional help utilization, and the magnitude of the treatment gap. Section 4 presents evidence on student beliefs related to mental health, the effectiveness and prevalence of professional mental health support, and relationship between miscalibrated beliefs, stigma, and the demand for therapy. Section 5 presents the design and the results of our information intervention. Finally, Section 6 discusses the implications of our findings for policy and future research.

2   Student Survey

2.1   Background and Setting

Mental health is an issue of rising importance and concern, with around 280 million people around the world diagnosed with some form of depression (World Health Organization 2021), accounting for about 5% of all adults suffering from this disorder. Based on several recent surveys (Healthy Minds Survey 2022), university students are experiencing even higher rates of depression and anxiety, drawing further attention to this population in research and supporting an unmet need for support (Abrams 2022). In Mexico specifically, mental health issues have gained increasing attention as a recent report by the OECD (OECD Report 2022) places Mexico among the top-3 OECD countries with the highest prevalence of depression post-pandemic (See Figure B1), indicating a concerning rise in the prevalence of mental health conditions in recent years. While there are no systematic representative surveys of college students, one of the largest student surveys on mental health and wellbeing by coverage in the US identifies 44% and 37% of students struggling with depression and anxiety, respectively (Eisenberg et al. 2022). Furthermore, while over 80% of students report needing help, only 37% receive counseling, indicating a large potential treatment gap (Eisenberg et al. 2022).

In Mexico, there is limited data on mental health and wellbeing among young people and students in particular. A mental health survey—conducted in 2005—of a large representative sample of adolescents (12–17 year-olds, a sample of over 3,000 children) living in Mexico City reveals the prevalence in the past 12 months of any anxiety disorder at almost 30% and any mood disorder (including depression) at 7.2% (Benjet et al. 2009). To the best of our knowledge, the only systematically collected source of information related to mental health is the Mexican Health and Nutrition Survey (ENSANUT). In Figure B2, we show the distribution of a depression screening questionnaire[9]—consisting of questions such as "During last week, did you feel sad/depressed?"—for Mexicans in 2023 using the ENSANUT data. According to this screening, about 12% of the Mexican population and 10% of those aged 17–28 scored above the half-score cutoff, consistent with experiencing such symptoms most or almost all of the days. However, while ENSANUT is representative at the national and regional levels,[10] it is limited in the representativeness for subgroups of the population, particularly our population of interest: university students.

As university enrollments rise, growing attention is drawn to mental health issues among students, a demographic going through critical life transitions and often being in a vulnerable emotional state. In particular, there has been a growing concern over suicides in major schools, including important ones in Mexico (Salud Mental 2022; Velazquez Hernandez 2017). As the number of university enrollments in Mexico surged by almost 50% from 2008 to 2022, reaching over 4 million students (Ministry of Education, 2023), an expanded demographic may be at risk, confounded by low availability of mental health services that are both affordable and effective. The mental health crisis is then further exacerbated by existing stigma and prejudice against recognizing mental distress and seeking treatment (Lagunes-Cordoba et al. 2021; Mascayano et al. 2016; Brewer et al. 2023).

For this project, we partnered with a large private university in Mexico with about 20,000 students[11] on the site of our study to conduct our survey. While not as low-resource as an average public university in Mexico or any developing country, this institutional setting allows us to focus on the demand for MH services when supply is largely available through on-campus therapy. Additionally, the quality of the support system for students in need of professional help offered by the university is deemed good by 85% of our survey respondents, suggesting quality concerns are not an issue in our setting. This implies that we would get a lower-bound estimate of the treatment gap and other potential miscalibrated beliefs compared to students

---

[9]While these are not standard questions used in measures such as PHQ-9 (Kroenke et al. 2001), they are similar and allow us to make suggestive conclusions about distress prevalence in Mexico.

[10]Regions are defined as a partition of the set of Mexican Federal States.

[11]In 2021, the university was home to more than 16% of the State's university students, with similar figures in the recent past (INEGI Statistics 2000-2023).

from public institutions who both face more financial constraints, lack on-campus services in accessing therapy, and are more likely to be exposed to mental-health-related stigma.

2.2 Survey Structure

Our survey begins with consent and ID collection, followed by screening questions about mental health (PHQ-4, GAD-4) and demographics. We then gather data on therapy usage, barriers and students' beliefs about therapy effectiveness, before assessing stigma-related questions and knowledge about on-campus services. We follow by splitting survey respondents into three experimental conditions—one control and two treatment groups. After a light-touch intervention and a placebo one for treatment and control groups, respectively, we elicit students' (incentive-compatible) WTP for an online therapy service. Finally, we present students with links for sharing mental health- and therapy-related information and we ask them to provide advice to a (hypothetical) friend in distress.[12]

We conducted our survey in November 2024, prior to the period scheduled for final exams. Our survey was advertised as a "Student Experience Survey" through multiple channels, including departmental communications, faculty outreach and social media. To maintain privacy while enabling payment processing, participants created a Unique ID[13] and verified their university affiliation through institutional e-mail addresses in a separate form. The final sample consists of 680 responses after filtering out those who do not pass the attention check, and the median completion time was 21 minutes. Participants were incentivized through a combination of guaranteed payments, random lottery draws and performance-based bonuses.[14]

Descriptive Statistics

Prior to the main discussion of depression and anxiety prevalence among students, we describe the main characteristics of our student sample and highlight its representativeness of the population we are studying in Table 1. The majority of our subjects are undergraduate students, the mean respondent's age is 20 years, approximately half are female, 75% identify as heterosexual, and 69% receive some amount of scholarship. Both parents of more than 70% of the respondents have at least a Bachelor's degree.

---

[12]The study was pre-registered on AEA RCT Registry website under AEARCTR-0014804. The study pre-analysis plan can be accessed on the Open Science Framework (OSF) platform via https://osf.io/6ng8x/. For further detail on the survey flow see Appendix Figure B10.

[13]The Unique ID combined the respondent mother's name initials, respondent's birth day, last name initials and last two phone number digits.

[14]We guaranteed $200 MXN ($10 USD) to the first 100 respondents, we randomly drew twenty respondents, each of which won a $2,000 MXN gift card, and we give $50 MXN for correctly answering one randomly selected bonus question out of eight.

Table 1: Survey Participant Characteristics

Characteristic	Mean	SD
Female (%)	51	50
Age (Years)	20.2	1.9
Heterosexual (%)	74.9	43.4
Pursuing Bachelor's (%)	91.2	28.4
Full scholarship (%)	7.9	27.1
Partial scholarship (%)	69.1	46.2
Both parents w/ college degree (%)	71.3	45.3

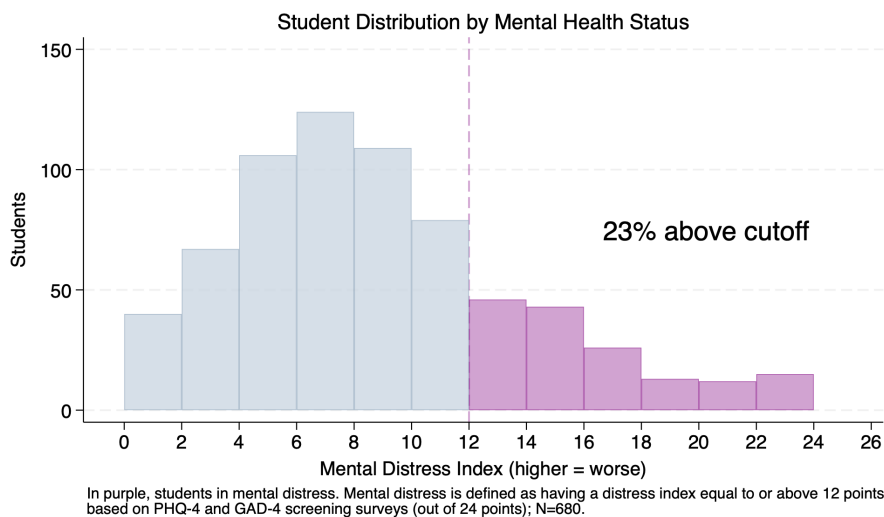| Fields of Study	Population	Sample
STEM (%)	42	46
Business (%)	25	18
Medicine & Health (%)	10	20
Law, Econ, Government (%)	8	11
Creative Studies (%)	8	3
Architecture & Environment (%)	7	2

Notes: This table shows, on the left, sample means and standard deviations of student participants' characteristics and, on the right, the distribution of the University population and student sample across fields of study.

3 Student Mental Health

In this section, we document the mental health treatment gap in the student population of a Mexican university. First, we illustrate the prevalence of depression and anxiety in our sample based on mental health questionnaires (PHQ and GAD). To document the size of the treatment gap, we then show the incidence of professional help utilization by university students, distinguishing it from informal sources of support. Finally, we look at student observables that correlate with mental distress and examine heterogeneity in the magnitude of the treatment gap between groups, highlighting the role of stigma and larger gaps among men.

3.1 Mental Distress and Professional Help

Figure 1: Mental Distress Index Distribution

Student Distribution by Mental Health Status

23% above cutoff

In purple, students in mental distress. Mental distress is defined as having a distress index equal to or above 12 points based on PHQ-4 and GAD-4 screening surveys (out of 24 points); N=680.

Notes: This figure shows the distribution of the mental distress index across students in our sample. Blue bars represent observations for students with a mental distress score below the cutoff of 12 points, while purple bars denote observations for students with mental distress scores above the cutoff.

We measure students' mental health using eight diagnostic questions from shortened versions of two widely used questionnaires. The PHQ-4 assesses the incidence of four symptoms of *major depressive disorder*, while the GAD-4 measures the incidence of four symptoms of *generalized anxiety disorder* over the past two weeks (Kroenke et al. 2001; Spitzer et al. 2006). Students' responses are quantified on a 0–3 scale for each question and aggregated into a distress index ranging from 0 to 24. A score of 12 or higher indicates moderate to severe symptoms, and we classify students meeting this threshold as being in distress.[15] Recent applications of PHQ and GAD in economic research include their use as tools to screen depression and anxiety among graduate students in economics departments in the US and their utilization to analyze the bi-directional relationship between poverty and depression in India (Bolotnyy et al. 2022; Ridley et al. 2020).

For our sample of 680 students, Figure 1 depicts the distribution of mental distress index values with higher values indicating poorer mental health. Given the right-skewed shape of the distribution, the mean distress index is around 8.4 out of 24 possible points, which is slightly above the median value of 8. In our sample, 155 students are at or above the 12 point cutoff for distress, constituting 22.8% of all students with a 95% confidence interval of $[19.6\%, 26\%]$.[16] This shows that the prevalence of poor mental health in our sample of Mexican university students is substantial. To compare, during 2013—2016, 8.1% of American adults aged 20 and over experienced depression in a given two-week period, according to the Centers for Disease Control and Prevention (CDC) (Brody et al. 2018). In a large meta-analysis involving 44,503 participants aged 18 or older from 100 eligible studies, the prevalence of major depression was 10% (Negeri et al. 2021).

The global treatment gap for mental health is significant, with over 80% of people with common mental health disorders – rising to more than 90% in poorer countries — not receiving treatment despite the availability of cost-effective solutions (Chisholm et al. 2016). Given a steady supply of counseling services in the university environment we are studying, it is not obvious ex-ante what the size of the treatment gap would be. The availability of and knowledge about services could, in principle, close the gap, but factors such as a lack of mental health literacy, stigma, and shame could, on the other hand, reduce demand.

---

[15]Four PHQ questions ask respondents how often over the past two weeks they have been bothered by 'little interest or pleasure in doing things', 'feeling down, depressed or hopeless', 'feeling tired or having little energy' and 'feeling bad about yourself - or that you are a failure or have let yourself or your family down'. Four GAD questions ask respondents how often over the past two weeks they have been bothered by 'worrying too much about different things', 'becoming easily annoyed or irritable', 'being so restless that it is hard to sit still', and 'feeling nervous, anxious or on edge.'

[16]Using a more lenient cutoff of 10 points yields a hefty 34.4% of students in distress, with a 95% confidence interval $[30.8\%, 38\%]$.

Table 2: Professional Mental Health Help Use by Mental Distress

	Used Prof. Help	No Prof. Help	Total
In Distress	80	75	155 (23%)
Not in Distress	190	335	525 (77%)
Total	270 (40%)	410 (60%)	680 (100%)

Notes: This table shows the cross-tabulation of students who have used professional mental health in the last 12 months and those who are in mental distress. We consider a student to be in distress if their mental health distress score is above or equal to 12.

We asked students in our survey about their use of professional mental health help in the last 12 months and, by splitting their responses based on whether they are in distress or not, categorized them in Table 2 into one of the four groups.[17] Out of 680 respondents, 270 report using professional help either on-campus or off-campus, meaning 2 out of 5 students in our sample receive some form of support from a mental health professional. Notably, when focusing only on those in distress, we observe that 80 out of 155 students (52%) with moderate or severe symptoms of depression or anxiety have received professional treatment in the last year.[18] Therefore, the estimate of the *treatment gap* in our sample of university students in Mexico is 48%. This indicates that roughly a half of students experiencing mental or emotional challenges are not receiving the psychological help they could benefit from, even though 80% of these students agree there is a good support system on campus for students who need professional help for their mental or emotional health.[19]

3.2 Predictors of Treatment Gap

In this section, we discuss the observable characteristics in our sample that are correlated with mental distress. Several individual covariates exhibit notable differences between students in distress and those not in distress, as shown in Table 3. Students in distress are significantly less likely to identify as heterosexual (64.5% vs. 77.9%, $p < 0.001$) and are more likely to be in their third year or above (63.9% vs. 50.5%, $p = 0.003$). They also report higher financial stress (70.3% vs. 51.4%, $p < 0.001$), suggesting that economic concerns may contribute to mental health disparities. Additionally, students in distress are slightly older on average (20.4 vs. 20.1

---

[17]Specifically, we asked students whom they had turned to for help with mental health challenges in the past 12 months and recorded the share who selected either the '*mental health professionals at my university*' option, the '*mental health professionals outside of my university*' option, or both.

[18]When splitting the components of distress, we find that around 47% of those exhibiting symptoms of depression and 47% of those exhibiting symptoms of anxiety have received professional help.

[19]One could argue that a person in distress might not realize this, so even if they are aware that the campus provides support, they might not seek it. In our sample, 94% of those in distress report experiencing mental health challenges in the last 12 months (e.g., frequent stress, feeling anxious or down), which indicates a high level of awareness of their own mental distress.

years, $p = 0.042$), and the fraction of students identifying as female is higher among those in distress (56.8% vs. 49.3%, $p = 0.104$). Other factors, such as GPA, scholarship status, and parental education, do not show statistically significant differences between the two groups.

In addition to the use of professional help, we also elicited respondents' perceived sufficiency of received support by asking them if in the past 12 months they received (i) *as much support* (counseling, therapy, or medication) as wanted, (ii) *less support* than wanted, or if they (iii) *were not seeking support* for their mental or emotional health. As Figure B5 in Appendix B portrays, 187 (133) students out of 680 report having received as much support as (less support than) desired. Strikingly, out of the remaining 360 people (53%) who reported not seeking support, 62 (17%) are in distress according to our mental health questionnaires.

Table 3: Comparison of Individual Covariates By Mental Distress

	In Distress (N=155)	Not in Distress (N=525)	p-value
Female (%)	56.8	49.3	0.104
Age (years)	20.4	20.1	0.042
Heterosexual (%)	64.5	77.9	<0.001
Year 3 or above (%)	63.9	50.5	0.003
GPA (0–100 scale)	90.5	91.1	0.155
Full scholarship (%)	9.7	7.4	0.364
Partial scholarship (%)	67.7	69.5	0.674
Financially stressed (%)	70.3	51.4	<0.001
Moved from hometown (%)	58.1	61.9	0.390
Both parents with college degree (%)	51.7	48.1	0.445

Notes: This table shows the means and p-value of the difference in means for covariates among students in distress and not in distress.

Students experiencing mental distress exhibit notable heterogeneity based on financial stress and perceived need for help. Financially stressed students report significantly higher distress levels, suggesting that economic constraints may play a crucial role in mental health challenges. Additionally, the perceived need for help strongly correlates with distress—many students recognize their struggles but refrain from seeking professional support due to financial concerns, time limitations, or skepticism about therapy's effectiveness. This self-awareness yet lack of action contributes to the overall treatment gap in mental health service utilization.

The treatment gap is particularly pronounced among specific student subgroups. Students who are not open to discussing mental health issues with classmates exhibit an 18 percentage point higher gap, indicating that stigma or discomfort with vulnerability may deter them from seeking help. Similarly, students who self-report as unlikely to seek mental health services experience a 9 percentage point larger treatment gap, reinforcing the impact of personal attitudes

on service utilization. Male students also show a significantly greater treatment gap of 12 percentage points compared to female students, aligning with broader patterns of lower help-seeking behavior among men. Figure B6 shows that students with mental health challenges in the past 12 months primarily relied on informal support networks, with over 40% turning to family and friends at university.

In contrast, formal professional help, whether at or outside the university, was sought by a much smaller fraction, while some students did not seek help at all. This highlights the importance of social support networks but also underscores a significant treatment gap, driven by stigma, reluctance, or barriers like financial stress and time constraints. Addressing these barriers is essential to improving access to professional mental health services. While informal networks provide some relief, they may not always offer adequate or evidence-based care. Additionally, students who misjudge their academic performance—either overestimating or underestimating their GPA—are more likely to experience a higher treatment gap, possibly due to the influence of academic self-perception on mental health help-seeking behavior.

4 Beliefs and Demand for Mental Health Help

Previous research attributes the treatment gap primarily to attitudinal barriers, such e.g., perceptions of low need or low effectiveness of existing treatments (Andrade et al. 2014) and stigma (Schnyder et al. 2017), despite rigorous and overwhelming evidence of their demonstrated effectiveness (Cuijpers et al. 2013), including in developing-country settings (Patel et al. 2017; Barker et al. 2022; Lacey et al. 2024) and specifically for college students (Cuijpers et al. 2016). In this section, we document evidence from our student sample in Mexico on the prevalence of incorrect beliefs or misconceptions related to key aspects that might affect their help-seeking: perceived effectiveness, perceived need (partially covered in section 3), beliefs about mental health and academic performance and other potentially stigmatized beliefs.
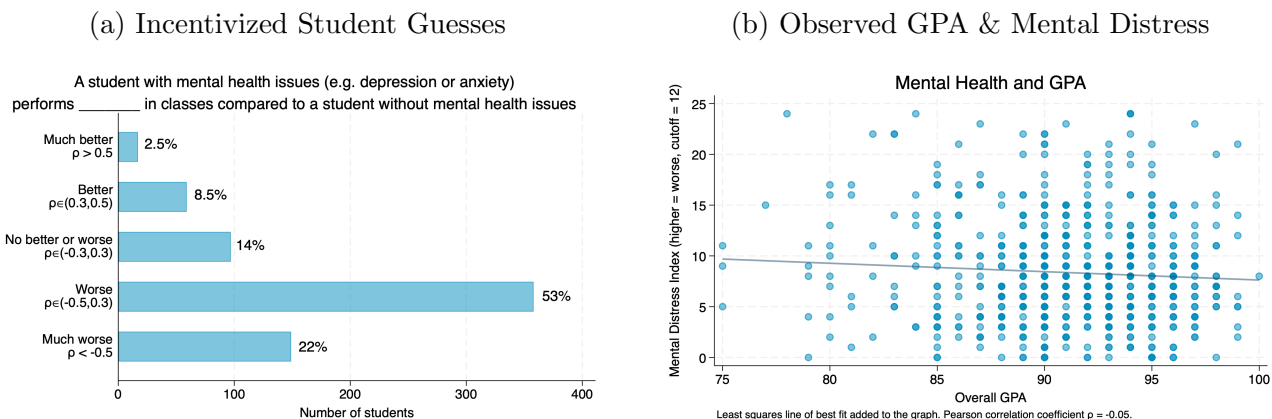
4.1 GPA & Mental Distress

Stereotypes and misconceptions about mental health often shape beliefs about productivity and performance, which in turn influence individuals' willingness to disclose their mental health status. Furthermore, as seeking therapy may be perceived as a signal of poor mental well-being, some might feel discouraged from talking about their mental health struggles and accessing professional help. Prior research by Ridley (2022) found that people strongly believe workers experiencing mental distress perform worse on a communication-related task in an online experimental setting, yet his results demonstrate no actual difference in performance. Our exploratory field visits revealed similar patterns in personal anecdotes and in-depth interviews, constituting a prevalent stereotype that we document in our student sample: 75% respondents believe that

the students in mental distress perform worse or much worse in classes compared to students not in distress – despite the actual data showing no such relationship.[20] This perception may contribute stigma, discouraging students from sharing their mental health struggles or seeking therapy, potentially contributing to the treatment gap.

We identify a particularly pervasive misconception related to mental distress and academic performance with 3 in 4 respondents guessing that students in mental distress perform worse or much worse in classes (have a lower GPA), despite the actual data showing no such relationship. This is one of the stereotypes that may discourage students from revealing their struggles or seeking professional help, mirroring prior findings by Ridley (2022), who demonstrated that individuals similarly overestimate the impact of mental health on workplace performance.

Figure 2 illustrates this discrepancy. The left panel shows that most students believe there is a strong negative relationship between mental distress and GPA. However, the right panel, based on actual data, reveals no meaningful correlation ($\rho = -0.05$). This systematic overestimation of the negative correlation of mental distress across students may reinforce stigma and discourage some from seeking help. Correcting this belief as a component of our information intervention is particularly promising since there is a relatively small fraction of subjects who would update their belief in the opposite direction (those who thought there was a positive correlation, around 11% of all subjects).

Figure 2: Correlation between Mental Distress and GPA

(a) Incentivized Student Guesses

(b) Observed GPA & Mental Distress

Notes: Panel (a) shows that most students (75%) guess that the relationship between GPA and mental distress across students is negative. We elicit their beliefs in an incentivized question, clarifying that the correct answer will be calculated across the participants based on their GPA and answers to the MH questionnaire. Panel (b) shows that there is no significant relationship between mental distress and GPA, with the correlation coefficient of $\rho = -0.05$. We also test this relationship using a binary distress measure (in distress if score above 12), and equivalently find no significant relationship.

---

[20]Specifically, we find that 75% of respondents predicted a strong negative correlation between mental distress and GPA in an incentivized question

Interestingly, the treatment gap is substantially larger for students who guess a negative relationship (51%) compared to those who guess the relationship correctly (33%)[21], suggesting that correcting this belief might contribute to reducing the treatment gap.

4.2   Perceived Therapy Effectiveness & Therapy Use

Next, we explore student beliefs about the effectiveness of mental health treatments to assess the potential of low perceived effectiveness to contribute to the treatment gap. We find (almost surprisingly) overwhelming evidence that the perceived effectiveness of therapy is high, with over 90% agreeing that it can improve both their own and other people's mental well-being. Additionally, there is strong social support for seeking therapy as over 91% believe their friends would support them seeking therapy, and 87% report the same for parents. These findings suggest a generally positive attitude toward mental health treatments, both in terms of personal experience and broader societal acceptance. This is a striking finding for a developing-country setting where typically there might be more stigma and conservative views around seeking help (Bhat et al. 2022; Ronak & Khandelwal 2024), and it can be explained by the fact that our field site and the students in this school are relatively better off. Yet, at the same time, this still underscores the importance of understanding other contributors to the treatment gap where the basis supply and effectiveness are already closer to desired levels.

In our sample, the majority (66%) of respondents have received professional mental health help at some point, and 87% report having a friend who has received help (a high share highlighting the potential for more information transmission and personal experience sharing among students) (See Table 4). In the past 12 months, 40% of students have received professional help for mental health issues, with about 20% receiving assistance at the university and 26% outside the university (there is a small overlap between two groups who have utilized both support resources).

The key patterns in professional help-seeking behavior and perceived effectiveness of therapy across distress levels are highlighted in Table 5. Notably, perceived effectiveness of therapy does not differ significantly between students in distress and those not in distress, suggesting that skepticism about therapy's efficacy is unlikely to be a primary driver of the treatment gap. Instead, the gap appears to be shaped by differences in actual help-seeking behavior. While distressed students are more likely to have sought professional help compared to their non-distressed peers, they are also more likely to avoid seeking help altogether. This polarization—where some distressed students engage with therapy while others abstain despite likely need—indicates that barriers beyond perceived effectiveness, such as stigma or other beliefs or constraints, may play a crucial role in shaping mental health treatment decisions.

---

[21]The magnitude of the difference is large, however, it is not statistically significant ($p = 0.13$)

Figure 3: Student Guesses of the Prevalence of Professional Help-Seeking

(a) All Students

(b) By prior therapy usage
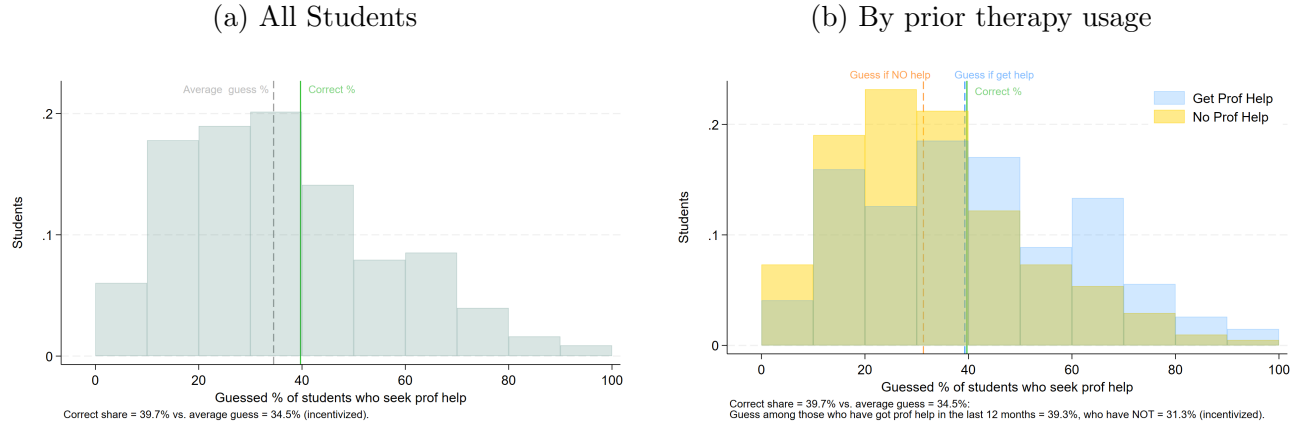
Notes: This figure shows the distribution of guesses of the percentage of students who seek professional help among University students. In Panel (a) we show all respondents while on Panel (b) we split the sample by an indicator of whether the respondent got professional help in the last 12 months (self-reported). See the CDFs of guesses by prior therapy use in Appendix Figure B7.

Table 4: Summary Statistics: Perceived Effectiveness, Support, and Therapy Use

	Mean
Perceived Effectiveness of Therapy	
Agree: Therapy improves **my own** well-being	0.904
Agree: Therapy improves **people's** well-being	0.924
Agree with both	0.897
Perceived Support for Therapy	
Agree: Friends would support me going to therapy	0.913
Agree: Parents would support me going to therapy	0.872
Agree that both friends and parents would support	0.843
Professional Help Received	
Have ever received professional MH help	0.662
Have a friend who received professional MH help	0.876
Have a friend who would benefit from therapy	0.894
(Last 12 Months)	
Sought help from mental health professionals (last 12m)	0.397
→ help from mental health professionals **at the university**	0.203
→ help from mental health professionals **outside the university**	0.260
Sample size	680

Notes: This table shows means for questions on perceived effectiveness, support and therapy use. For items under the Perceived Effectiveness of Therapy and Perceived Support for Therapy panels we ask *How much do you agree or disagree with the following statements?* (1) *Going to therapy can improve my own mental health* (2) *In general, going to therapy can improve people's mental wellbeing* (4) *My friends would show support if I told them I am going to therapy* (5) *My parents would show support if I told them I am going to therapy*; we code as "agree" responses which state Somewhat Agree, Agree or Strongly Agree. For items under the Professional Help Received panel we ask the following Yes/No questions: (i) Have you ever received professional mental help? (ii) Do you have a friend who is currently receiving or has previously received professional mental health?, and (iii) Do you have a friend or someone you know closely who you think would benefit from therapy? Finally, we ask *If you experienced mental health challenges in the last 12 months, [...], to who did you turn for help? Select ALL that apply* for items under the (Last 12 Months) panel.

16

Table 5: Perceived Effectiveness & Help-Seeking by Distress

	(1) Not in Distress Mean (SD)	(2) Distress Mean (SD)	(2)-(1) Pairwise t-test Mean difference
A. Perceived Effectiveness: & Support			
Guess # studies ↓ depression (correct 22)	17.02 (4.39)	17.33 (4.32)	0.31
Agree: Therapy can improve my own well-being	0.90 (0.31)	0.94 (0.25)	0.04
Agree: Therapy can improve people's own well-being	0.92 (0.27)	0.94 (0.25)	0.02
Agree: Friends would support me going to therapy	0.91 (0.28)	0.92 (0.28)	0.00
Agree: Parents would support me going to therapy	0.88 (0.32)	0.83 (0.37)	−0.05*
B. Use of Professional Mental Health Help:			
Have a friend who received professional MH help	0.88 (0.33)	0.87 (0.34)	−0.01
Have a friend who would benefit from therapy	0.88 (0.33)	0.95 (0.22)	0.07**
Have ever received professional MH help	0.63 (0.48)	0.77 (0.42)	0.14***
Sought professional mental health help (last 12 months)	0.36 (0.48)	0.52 (0.50)	0.15***
→ professional MH help **at the university**	0.19 (0.39)	0.26 (0.44)	0.07*
→ professional MH help **outside university**	0.23 (0.42)	0.37 (0.49)	0.15***
Unlikely to seek help when struggling with mental health issues	0.15 (0.36)	0.28 (0.45)	0.13***
Sample size	525	155	680

Notes: This table shows difference in mean responses across students who are/are not in distress for questions related to perceived effectiveness and support as well on use of professional mental health. Difference = Distress - No distress. Sample size (680). ***, **, * indicate 1, 5, 10% significance.

From Figure 3, one can see a wide distribution of student guesses about the prevalence of professional help-seeking, with most students underestimating this share. In our sample, 39.7% used professional mental health services, and the students who have received professional help before provide, on average, estimates that are close to the actual rate, with an average guess of 39.3%. In contrast, students who have not sought therapy tend to underestimate its prevalence more, with an average guess of 31.3% and with over 70% of respondents underestimating the share. This pattern suggests that students' own experiences shape their perceptions of therapy-seeking norms, with those outside the system systematically underestimating how common it is, reinforcing potential misconceptions about the accessibility or acceptability of seeking professional mental health support.

4.3 Prior Beliefs and Misconceptions

We further explore students' personal beliefs and public perceptions related to stigma, seeking help and discussing mental health struggles. We document that many students underestimate how many of their peers seek professional mental health help and are open to discussing mental health struggles while overestimating the prevalence of self-stigma, resulting in a more pessimistic view of public perceptions of stigma and their peers' attitudes toward mental distress. Our results broadly echo the findings from recent online and field experimental studies that identify misconceptions around willingness to discuss mental health issues and the prevalence of mental-health-related beliefs among others as potential evidence of stigma (Roth et al. 2024a; Ridley 2022; Ronak & Khandelwal 2024; Acampora et al. 2022).
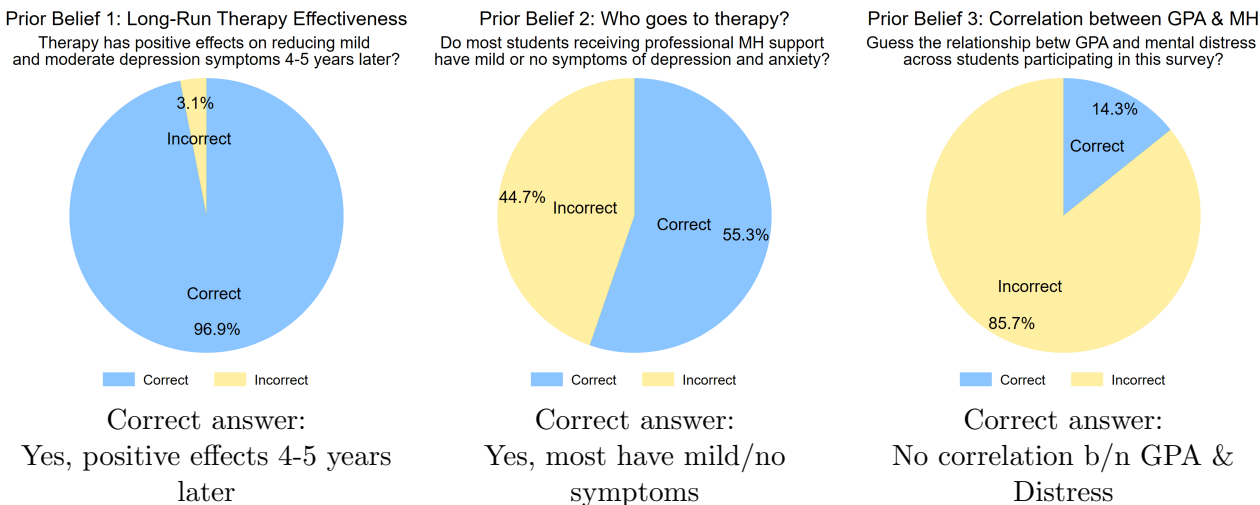
Our data reveal a notable projection bias among respondents, with individuals' personal experiences and attitudes influencing their perceptions of their peers' behaviors (Figure 4). Addressing these biases can lead to more accurate perceptions, fostering a supportive environment that encourages help-seeking behaviors among students. This aligns with Bushong & Gagnon-Bartsch (2024), who demonstrated that individuals often project their own behaviors onto others, leading to skewed perceptions of social norms.

Figure 4: Perceptions: Self-Stigma and Openness to Share

(a) Perceived Self-Stigma

(b) Perceived Willingness to Share Mental Health Challenges
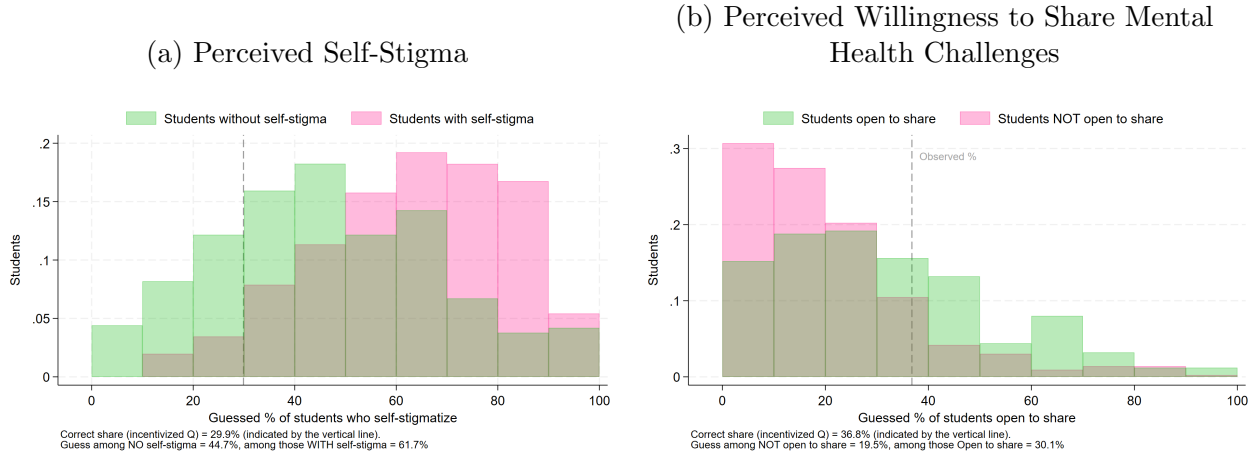
Notes: This figure shows, in Panel (a), the distribution of guesses of the percentage of students who would be disappointed in themselves if they had a mental health issue, and in Panel (b), the distribution of guesses of the percentage of students who would be open to share their mental health challenges with classmates who are not necessarily their friends. We show the distributions by respondents who do/do not self stigmatize, and by respondents who would/would not be open to share.

Figure 5: Prior Beliefs about Therapy and Mental Health

Prior Belief 1: Long-Run Therapy Effectiveness
Therapy has positive effects on reducing mild
and moderate depression symptoms 4-5 years later?

Correct answer:
Yes, positive effects 4-5 years
later

Prior Belief 2: Who goes to therapy?
Do most students receiving professional MH support
have mild or no symptoms of depression and anxiety?

Correct answer:
Yes, most have mild/no
symptoms

Prior Belief 3: Correlation between GPA & MH
Guess the relationship betw GPA and mental distress
across students participating in this survey?

Correct answer:
No correlation b/n GPA &
Distress

Notes: This figure shows the share of students who answered each of our three "prior beliefs" questions correctly or incorrectly.

4.4 Beliefs about Stigma

Stigma is a multidimensional construct shaped by individual beliefs, social perceptions, and structural norms, making it challenging to measure directly (Link & Phelan 2001; Brouwers 2020). Given established misperceptions and projection biases in our data, these beliefs might reinforce stigma and influence help-seeking behaviors. To capture these complexities, we construct a stigma index using Principal Component Analysis (PCA), incorporating perceived public stigma (negative judgments from peers and authority figures), perceived prevalence of self-stigma, and personal stigma (students' biases against other students with mental health issues), allowing us to examine how stigma affects willingness to pay for therapy.

With the established dimensions of stigma, we run descriptive statistics and analyze the correlation between the constructed stigma indices—stigma index 1 (PCA1) and stigma index 2 (PCA2), derived from the first and second principal components (See Table 7), respectively—and their corresponding component variables. Table 7 shows that stigma index 1 (PCA1) primarily captures variations in perceived public stigma, with strong associations with stigma from students, professors, and parents. Self-stigma also contributes to PCA1 but plays a secondary role, while personal stigma variables, such as preferences for low GPA over mental health symptoms or discussions, show minimal influence. This suggests PCA1 is a reliable measure of public stigma perceptions. Stigma index 2 (PCA2), on the other hand, reflects variations in personal stigma, as indicated by its strong correlations with preferences prioritizing academic performance over addressing mental health issues. It is largely independent of perceived public stigma but

shows a slight relationship with self-stigma. This highlights PCA2 as capturing an internalized, individual dimension of stigma distinct from external perceptions[22].

Table 6: Mental Health Stigma Variables

	Definition
Perceived Public Stigma	
From students	Percentage of students that the respondent believes would view a student negatively for experiencing mental health issues like anxiety or depression.
From professors	Percentage of professors that the respondent believes would view a student negatively for experiencing mental health issues like anxiety or depression.
From parents	Percentage of student parents that the respondent believes would view a student negatively for experiencing mental health issues like anxiety or depression.
Self-Stigma	
Self-stigma	Respondent's estimate of how many out of 100 students would feel disappointed in themselves if they had a mental health issue.
Personal Stigma	
Low GPA over MH symptoms	Dummy variable where it has a value of 1 if the respondent ranked a student with a low GPA as preferred as a class project teammate rather than a student experiencing mental health distress; 0 otherwise.
Low GPA over MH talk	Dummy variable where it has a value of 1 if the respondent ranked a student with a low GPA as preferred as a class project teammate rather than a student openly talking about mental health issues; 0 otherwise.

Notes: This table shows the definition of variables used as inputs for constructing our stigma index.

---

[22]For more technical details on the construction and interpretation of the stigma indices, please refer to the appendix. See full correlation table in Table B8

Table 7: Correlation of Stigma Index PCA1 and PCA2 with Components

Variable	PCA1	PCA2
PCA1	1.00	
PCA2	0.00	1.00
Stigma Students	0.83***	0.04
Stigma Professors	0.87***	0.04
Stigma Parents	0.83***	-0.03
Guess Self-Stigma	0.50***	0.11**
Low GPA Symptoms	-0.09*	0.84***
Low GPA Talk	-0.04	0.85***

Notes: This table shows the correlation coefficient of our stigma indices with the components used for constructing them. Significance levels: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure 6: Stigma Measures By Distress
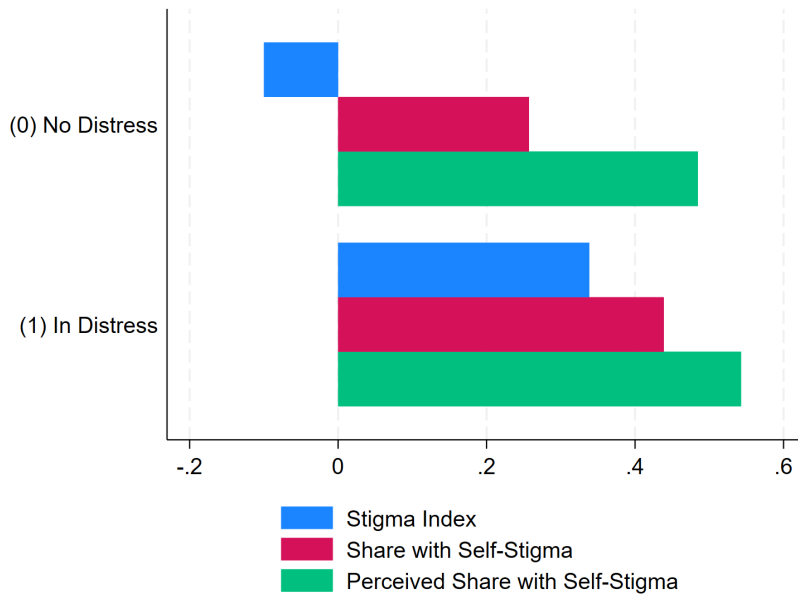
Notes: This figure shows (i) the average of the stigma index (measured in standard deviations), (ii) the share of students who have self-stigma (percent share on the axis), and (iii) the average guess of students from University who would be disappointed if themselves if they had a mental issue (percent share on the axis). We show averages and shares by an indicator of whether the student is in distress or not according to the mental distress index.
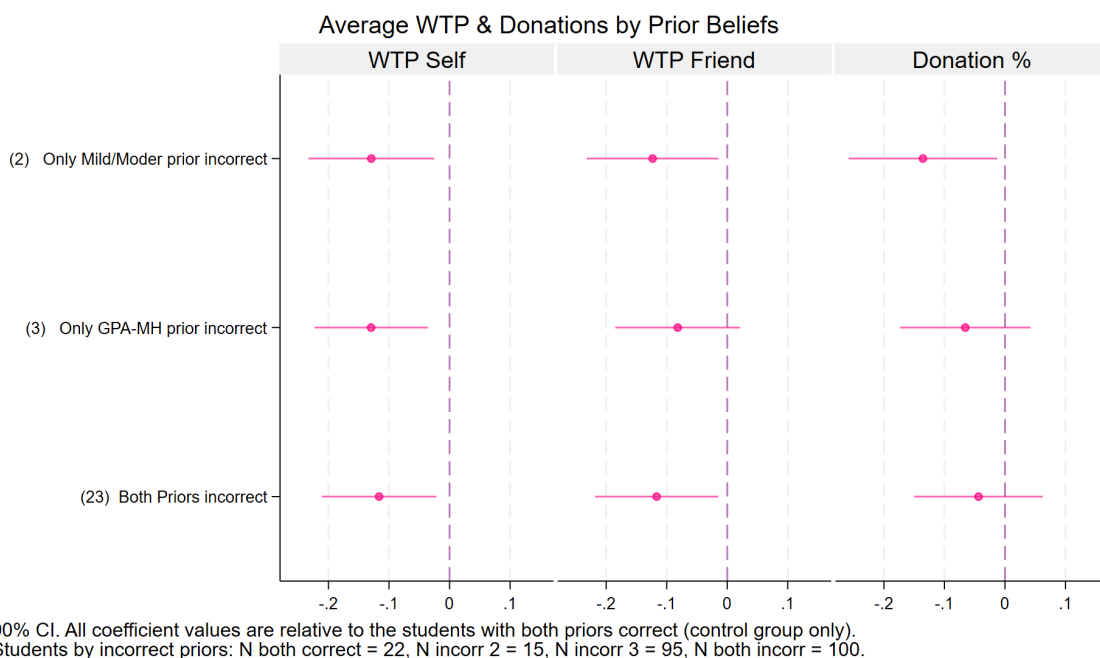
Figure 6 illustrates variation in stigma measures across students by mental distress, with more prevalent stigmatized beliefs among students in distress, suggesting that stigmatized beliefs might be related to the treatment gap. While interesting, this correlation, however, does not shed light on the directionality of the connection, whether having stigmatized beliefs leads to more distress (potentially by not seeking therapy) or whether being in distress in itself affects your perception of the outside world.

The stigma section reveals that stigma perceptions are shaped by distress levels, professional help usage, and prior beliefs about mental health. Stigma index 1 (PCA1), which captures perceived public stigma from peers, professors, and parents, provides a robust measure of how external societal attitudes influence mental health-related decisions. In contrast, stigma index 2 (PCA2) reflects personal stigma and internalized biases, such as prioritizing academic performance over mental health, but has a narrower focus and limited explanatory power.

4.5 Demand for Therapy and Beliefs

The willingness to pay (WTP) for therapy and the propensity to donate to mental health initiatives provide important insights into how students' beliefs influence their demand for mental health services. The graph examines the average WTP and donations by the different classifications of incorrect prior beliefs about mental health. Students with both prior beliefs incorrect—concerning the prevalence of mild or moderate mental health issues and the relationship between GPA and mental health—demonstrated the negative WTP for therapy, both for themselves and for their friends. This suggests that the presence of misconceptions is strongly associated with diminished students' perceived value of private therapy, for themselves and other students[23].

Figure 7: Willingness To Pay for Private Therapy and Donation

90% CI. All coefficient values are relative to the students with both priors correct (control group only).
Students by incorrect priors: N both correct = 22, N incorr 2 = 15, N incorr 3 = 95, N both incorr = 100.

Notes: This figure shows the difference in mean estimates on willingness to pay and donation outcomes. We estimate $Y_i = \alpha + \beta_M \texttt{MildIncorrect}_i + \beta_G \texttt{GPAIncorrect}_i + \beta_B \texttt{BothIncorrect}_i + \varepsilon_i$, where Y_i is the outcome of interest, $\texttt{MildIncorrect}$ is an indicator equal to 1 if the respondent only answered the "Mild/Moderate"-prior question incorrectly, $\texttt{GPAIncorrect}$ is an indicator equal to 1 if the respondent only answered the "GPA-MH"-prior question incorrectly and $\texttt{BothIncorrect}$ is an indicator equal to 1 if the respondent answered both the "Mild/Moderate"- and the "GPA-MH"-questions incorrectly. The reference group is the group of respondents who answered all priors' questions correctly. Horizontal lines represent 90% confidence intervals.

---

[23]Appendix subsubsection C.4.3 shows some demand measures by stigma, with no notable differences.

5 Information Intervention

5.1 Treatment Design

To gauge the effectiveness of potential interventions aimed at reducing stigma and prejudice, as well as promoting supportive attitudes towards people suffering from mental health issues, we design a light-touch intervention, which combines two components: belief correction and a reflection activity, focusing on potential attitudinal barriers to seeking help: perceived effectiveness (which we found is already high as a result of the intervention), identifying that therapy is not just for students in severe distress, and debunking the stereotype of low GPA being correlated with higher mental distress across students. We chose these facts based on prior work highlighting the role of attitudinal barriers in the treatment gap (Andrade et al. 2014) and early-stage fieldwork insights highlighting the above-mentioned stereotypes.

We randomly assign students into one of three possible groups: a control and two different treatment groups.

- **Treatment 1 (T1): Information + Reflection** ($n = 227$) Students in this group were shown three different sets of information in the form of infographics. The first infographic shows a recent study found that offering psychotherapy leads to an 11% drop in mild depression and an 8% drop in moderate depression four to five years later. The second infographic showed information disclosing that "Among [University] students who are receiving professional mental help, 2 out of 3 have only mild or no symptoms of depression and anxiety." The third and last infographic showed that "Among 53 [University] students, 3 out of 4 respondents believe that a student with mental health issues performs **worse** or **much worse** academically than a student without mental health issues. But our survey data show **no relationship** between students' GPA and mental distress."

  In addition to the infographics, students in this group were prompted with the following message: "*Many university students sometimes struggle with feelings of being overwhelmed, anxious, or depressed. Based on your experience, what are some effective ways students can manage these types of mental health challenges? Please explain your thoughts.*" Furthermore, we showed students in this group one of two vignettes[24] with an image of a fictitious student from their university and describing a hypothetical situation in which this student seeks help from a therapist after suffering a panic attack.

---

[24]The only differences across vignettes are the sex of the student appearing in the images and the name of the student. We did this to rule out treatment effects being driven by the sex of the student in the hypothetical situation.

- **Treatment 2 (T2): Information Only** ($n = 221$) Students in this group were shown the same infographics as the ones shown to students in the Information + Reflection treatment with the difference that no reflection activities or vignette components were part of the treatment for this group.

- **Control (C):** ($n = 232$) Students in the control condition were not shown the infographics nor any of the vignettes. They answered additional questions about various university services to keep the overall survey time closer to that in the treatment groups.

Figure 8: Treatment Assignment
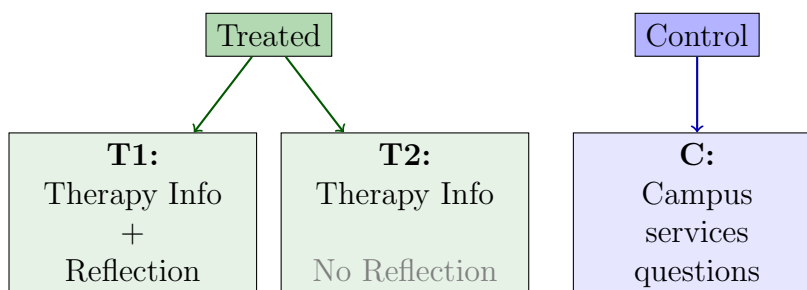
Notes: This figure depicts the assignment of students in our sample to the three experimental groups.

Importantly, students randomized into the Control condition or the Information Only treatment were also asked to tell us about their experience using on-campus services including sports facilities, academic counseling and career services in an attempt to hold mental effort during the survey completion comparable to that of the Information + Reflection treatment group. To show balance, in Table 8 we present evidence that treatment and control groups were not statistically different on pre-determined covariates.[25]

5.2 Empirical Strategy

Information Sharing

We consider two groups: a treatment group (T) with n_T individuals and a control group (C) with n_C individuals. Let k_T and k_C be the total observed clicks from the treatment and control groups, respectively. We wish to test whether the underlying click rates in the two groups differ. Since each participant in our study could generate an unbounded number of link clicks, we modeled the click counts using a Poisson process. Denote by λ_T the (unknown) rate of clicks per person in the treatment group and by λ_C the (unknown) rate in the control group. The null hypothesis asserts that both groups share the same click rate, i.e. $H_0 : \lambda_T = \lambda_C$, whereas the alternative is $H_1 : \lambda_T \neq \lambda_C$. In practice, this is often expressed as testing whether the *rate ratio* λ_T/λ_C equals 1.[26]

---

[25]See appendix Table B1 for a balance table comparing all three experimental groups.

[26]One could in principle model this environment as a comparison of two binomial random variables, where each observation can either result in success or failure. Thus, a binomial framework assumes a fixed upper limit on the

Table 8: Covariate Balance

Variable	N	(1) Control Mean/(SD)	N	(2) Treated Mean/(SD)	N	(1)-(2) Pairwise t-test Mean difference
Age	232	20.159 (1.848)	448	20.145 (2.031)	680	0.014
Female	232	0.461 (0.500)	448	0.536 (0.499)	680	-0.075*
Financially Stressed	232	0.530 (0.500)	448	0.571 (0.495)	680	-0.041
Has Scholarship	232	0.651 (0.478)	448	0.712 (0.453)	680	-0.061
Receives a full scholarship	232	0.082 (0.275)	448	0.078 (0.269)	680	0.004
Moved Residence	232	0.591 (0.493)	448	0.621 (0.486)	680	-0.030
GPA	232	90.897 (4.659)	448	91.007 (4.727)	680	-0.110
MH Score	232	8.569 (5.132)	448	8.237 (5.054)	680	0.332
Used Therapy L12 Months	232	0.233 (0.424)	448	0.234 (0.424)	680	-0.002
Open to Share MH Challenges	232	0.392 (0.489)	448	0.355 (0.479)	680	0.037
Self-stigmatize	232	0.323 (0.469)	448	0.286 (0.452)	680	0.038

Notes: We pool T1 and T2 into a "Treated" group. This table shows balance on covariates across treatment groups. For each covariate we show each experimental group's sample mean and standard deviation, as well as the difference in means across both groups. Age measures the respondent's age in years, female is an indicator equal to one if the respondent is female-born, financially stressed is an indicator equal to one if the respondent described her financial situation as "Always", "Often" or "Sometimes" stressful and equal to 0 if she reported it as "Rarely" or "Never" stressful, Has scholarship is an indicator equal to one if the respondent has at least some amount of scholarship, receives a full scholarship is an indicator equal to one if the respondent's scholarship covers 100% of tuition, moved residence is an indicator equal to one if the respondent moved her residence city to pursue her current studies, GPA measures the respondent's current overall GPA on a scale from 0–100, MH score measures the student's mental health score as described in section 2, used therapy in L12 months is an indicator equal to one if the respondent states having used therapy in the last 12 months, open to share MH challenges is an indicator equal to one if the respondent states she would be willing to share about her own personal MH challenges with others and self-stigmatize is an indicator equal to one if the respondent states she would be disappointed in herself if she suffered from mental distress. Standard errors for the difference in means test are heteroskedasticity robust. Significance levels: * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$

---

number of "successes" each participant can contribute (e.g., at most 1 click per person). In our study, however, each participant could potentially produce multiple clicks, so there is no obvious upper bound. We, therefore, model such unbounded count data using Poisson distribution, with each group's total number of events (clicks) assumed to be Poisson($\lambda_T n_T$) or Poisson($\lambda_C n_C$), respectively.

Given that our click counts are relatively small, in addition to running a test relying on large-sample approximations (Wald test, in our case), we also employed an *exact test* for two-sample Poisson comparisons (in the spirit of Fisher's exact p-value test on binomial data). Under H_0, the total number of clicks $k_T + k_C$ is fixed, and the conditional distribution of k_T (the count in the treatment group) is binomial with parameter

$$p = \frac{n_T}{n_T + n_C}.$$

Thus, the test assesses whether the observed k_T is unreasonably large or small relative to this binomial distribution, thereby providing an exact p-value for the hypothesis $H_0 : \lambda_T = \lambda_C$.[27]

In addition to the joint treatment (T1&T2) vs. control comparison, we separately tested other pairwise differences (e.g., T1 vs. control, T2 vs. control, and T1 vs. T2). For each comparison, the method returns (i) a *rate ratio*, $\hat{\lambda}_T/\hat{\lambda}_C$, estimated by the ratio of observed click rates, (ii) an *exact* two-sided p-value, and (iii) an indicator of whether we reject H_0 at 5% level. Unlike approximate Poisson methods, the exact approach remains valid even when k_T and k_C are small. However, it does not provide a confidence interval for the rate ratio in the current implementation; we therefore focus on p-values and the estimated ratio to interpret group differences in click rates.

Main Regression Specifications

To estimate treatment effects on our primary outcomes, we use two main regression specifications. These models evaluate the effectiveness of our interventions while accounting for potential confounders. Our primary specification examines the pooled effect of any intervention (T1 or T2) compared to the control group. The regression model is specified as follows:

$$Y_i = \alpha + \beta \, Info\,Treatment_i + X_i^{'}\gamma + \epsilon_i,$$

In this specification, Y_i represents the outcome of interest for individual i, such as willingness to pay for therapy, self-reported stigma, or the advice quality index. The variable $Info\,Treatment_i$ is an indicator equal to 1 if individual i received any of the treatment conditions (T1 or T2), and 0 otherwise. Additionally, X_i denotes a vector of baseline covariates included to improve precision, such as gender, GPA, or mental health score. Finally, ε_i represents the error term. The coefficient β captures the average treatment effect of the pooled intervention on the specified outcome.

---

[27]We carried out the exact Poisson test using `statsmodels` in Python with the `method="exact-cond"` option.

To disentangle the effects of each intervention (T1 and T2), we also estimate the following model:

$$Y_i = \alpha + \kappa_1 T_{1i} + \kappa_2 T_{2i} + X_i'\gamma + \varepsilon_i,$$

Where T_{1i} is an indicator variable equal to 1 if individual i was assigned to the Information + Reflection treatment, and 0 otherwise, and T_{2i} is an indicator variable equal to 1 if individual i was assigned to the Information Only treatment, and 0 otherwise. All other terms are defined as in the previous model. In this specification, κ_1 and κ_2 measure the treatment effects for T1 and T2, respectively, compared to the control group.

Both models allow us to identify the impact of our interventions on key outcomes. The first specification focuses on the pooled treatment effect, providing a simple and intuitive estimate of the intervention's overall effectiveness. The second specification, on the other hand, enables us to compare the effects of the two distinct treatments.

5.3 Treatment Effects

On-Campus Resources Sharing

As part of our survey, students were presented with an opportunity to share a link to on-campus counseling services with their peers. We tracked both the total number of human clicks and the number of unique users who clicked the link across three experimental conditions: *Information+Reflection* (T1), *Information Only* (T2), and a control group (C). In total, we recorded 44 clicks in T1, 92 in T2, and 35 in the control group. Of these, 35 were unique clicks in T1, 59 in T2, and 24 in the control group. Given the number of respondents in each condition (T1: $n = 227$, T2: $n = 221$, C: $n = 232$), the unique click-through rates were approximately 15.4% in T1, 26.7% in T2, and 10.3% in the control group (pooling two treatment groups yields a rate of 21%). These results indicate that the *Information Only* condition led to the highest engagement with the counseling resources, while the control group had the lowest level of participation.

A simple comparison of click-through rates indicates that the information treatment led to a higher propensity to share or engage with on-campus resources—the unique click-through rate is more than double that of the control group. We, however, cannot directly run statistical tests for the comparison of means as the data is aggregated at the treatment level instead of providing us with individual-level observations. The alternative is to leverage observed click-through rates and run a Poisson test to compare whether the click-through rates in treatment and control groups, captured by parameters $\widehat{\lambda}_T = 0.209$ and $\widehat{\lambda}_C = 0.103$ are statistically distinguishable or not.

The results based on both the approximated and exact Poisson versions of the statistical test with the null $H_0 : \lambda_T = \lambda_C$ are presented in Table 9. Panel A shows the estimated click-

Table 9: Poisson Test Results for Link Clicks

Test	Rate Ratio	Approx. Poisson		Exact Poisson	
		p-value	Reject H_0	p-value	Reject H_0
Panel A: Total Clicks					
T1 vs C	1.32	0.221	False	0.260	False
T2 vs C	2.69	<0.001	True	<0.001	True
T1 vs T2	0.49	<0.001	True	<0.001	True
T1&T2 vs C	2.01	<0.001	True	<0.001	True
Panel B: Unique Clicks					
T1 vs C	1.53	0.107	False	0.118	False
T2 vs C	2.51	<0.001	True	<0.001	True
T1 vs T2	0.61	0.019	True	0.023	True
T1&T2 vs C	2.03	<0.001	True	0.001	True

Notes: This table shows estimates for the rate ratio of treatment group link clicks to those of the control groups. In Panel A we focus on total clicks, while on Panel B we focus on "unique" clicks. For each comparison we show the p-value associated to the null hypothesis of rate ratios equal to 1, both using Wald-type and permutation-based tests.

through rate ratios, p-values and test conclusions for total number of clicks, while Panel B shows the estimates for unique human clicks. It can be clearly seen that regardless of the type of the test or the type of link clicks, it is always the case that we reject the null of equality of two rates between joint treatment (T1&T2) and control groups ($p \leq 0.001$). The effect is driven by a high click-through rate in *Information Only* treatment group as the rate ratio is the highest (and p-value is the smallest) when we compare T1 and C. The click rates in *Information + Reflection* and control groups are not statistically different from each other and we cannot reject the equality of two means ($p = 0.260$ for total clicks and $p = 0.118$ for unique clicks).

Peer Advice and Donation

We observe a sizable increase of 3.8 *pp* (73% of the control group mean) in the share of students who mention some form of on-campus therapy for the treatment groups relative to the control. The share of words or phrases related to passive advice is 2.9 *pp* lower for students in treatment conditions.[28] This decrease is mostly driven by students in the Treatment + Reflection group, whose share of words or phrases mentioned decreases by 3.9 *pp*, whereas that of students in the Information Only group also decreases but only by a mild—and insignificant—1.8 *pp*. Turning to suggestive advice we do not observe any differences between the share of words/phrases mentioned by students in the control group and those in either of the treatment groups. These results, along with the documented high level of knowledge about on-campus services, suggest students who are provided with information about therapy effectiveness change the composition of the advice

---

[28]See appendix section C.5 for a detailed explanation of advice processing.

they give to friends in distress. In particular, they substitute advice in which they state they, e.g., "are there for them" or "are there if they want to talk", in favor of more targeted advice where they prompt their friend about available on-campus services.

Table 10: ATE on Advice

|	(1) Campus Help	(2) Passive Advice	(3) Suggestive Advice	(4) Campus Help	(5) Passive Advice	(6) Suggestive Advice
Any Treatment	0.038* (0.020)	-0.029* (0.017)	-0.004 (0.015)			
Info + Reflection				0.036 (0.024)	-0.039** (0.018)	0.001 (0.018)
Info Only				0.039 (0.024)	-0.018 (0.020)	-0.009 (0.017)
Observations	680	680	680	680	680	680
R2	0.004	0.005	0.000	0.005	0.007	0.001
Control Mean	0.052	0.216	0.184	0.052	0.216	0.184

Notes: This table shows treatment effect estimates on advice given by survey respondents.

Willingness to Pay and Donation

As a proxy for participants' demand for therapy, we use incentive-compatible BDM-style willingness to pay (WTP) measure (Becker et al. 1964). Specifically, we measured the maximum amount participants were willing to pay for a one-month therapy subscription from *BetterHelp*, both for themselves and for a friend (two separate incentivized questions). The average WTP in the control group was USD $135 for self and USD $134 for a friend (the original price for this service is around USD $320). In addition, participants could donate a percentage of their total survey earnings to fund a therapy subscription for a financially constrained student at their university – the baseline donation rate in the control group is roughly 20%. The results are presented in Figure 9.

The treatment effects reveal a negative average treatment effect (ATE) on WTP for therapy for oneself, as indicated in the Figure 9. This suggests that the interventions may have reduced participants' perceived personal value of therapy. Similarly, the effects on WTP for a friend are comparable (even if marginally insignificant). These findings indicate that our interventions have a small negative effect on individuals' valuation of therapy for themselves and a friend, with the negative effects primarily identified from the lower-WTP groups as indicated by the quantile regression results (See Figure 10). The donation behavior shows limited responsiveness to the treatments, with no significant shifts in the percentage of earnings donated. Taken

Figure 9: ATE for therapy WTP and Donations
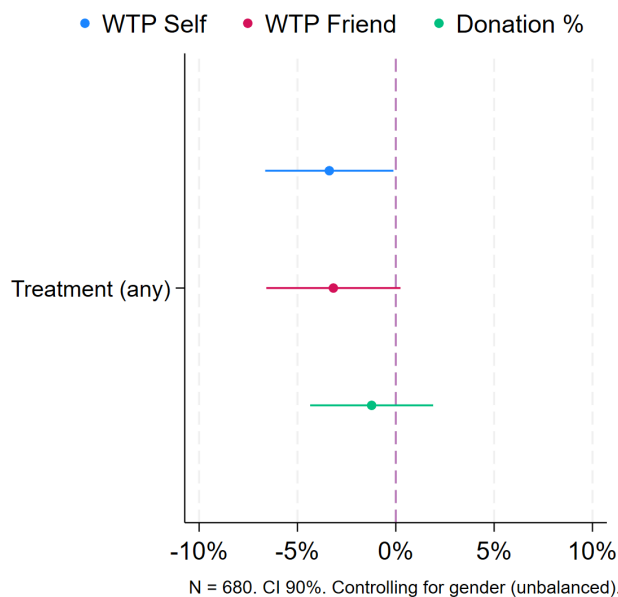
Notes: This figure shows treatment effect estimates and 90% confidence intervals of quantile regressions for WTP for oneself, a friend, and the percentage of earnings donated for another participant's therapy subscription.

together, these results suggest that the interventions modestly alters personal valuation of private therapy, reducing WTP by slightly more than 3 p.p.

5.4 Treatment Heterogeneity Analysis

To explore heterogeneity in treatment effects, we interact the pooled treatment indicator with several key variables (focusing on the pre-registered specifications). We explore heterogeneity by groups defined by combinations of mental distress and professional help use: (0) No distress + no professional help (reference group), (1) No distress + professional help, (2) Distress + no professional help, (3) Distress + professional help.

In Figure 11, we observe that for the group that might be the main target of potential interventions (Distress + No Professional Help), we observe a significant negative effect on own WTP and a smaller not significant negative effect on the WTP for a friend. While we can not robustly show the main driver of this negative updating, we conjecture that this may be aligned with the substitution effect of switching to free on-campus therapy we have discussed above. At the same time, this also highlights some limitations of using WTP measures for a private service analogue in a setting where free services are provided, limiting out discussion of this measure as an imperfect proxy for the overall demand.

In Figure 12, we test whether the treatment effects differ based on the accuracy of students' prior beliefs. Specifically, we examine the following groups: (0) Both priors correct (reference

Figure 10: Quantile regressions
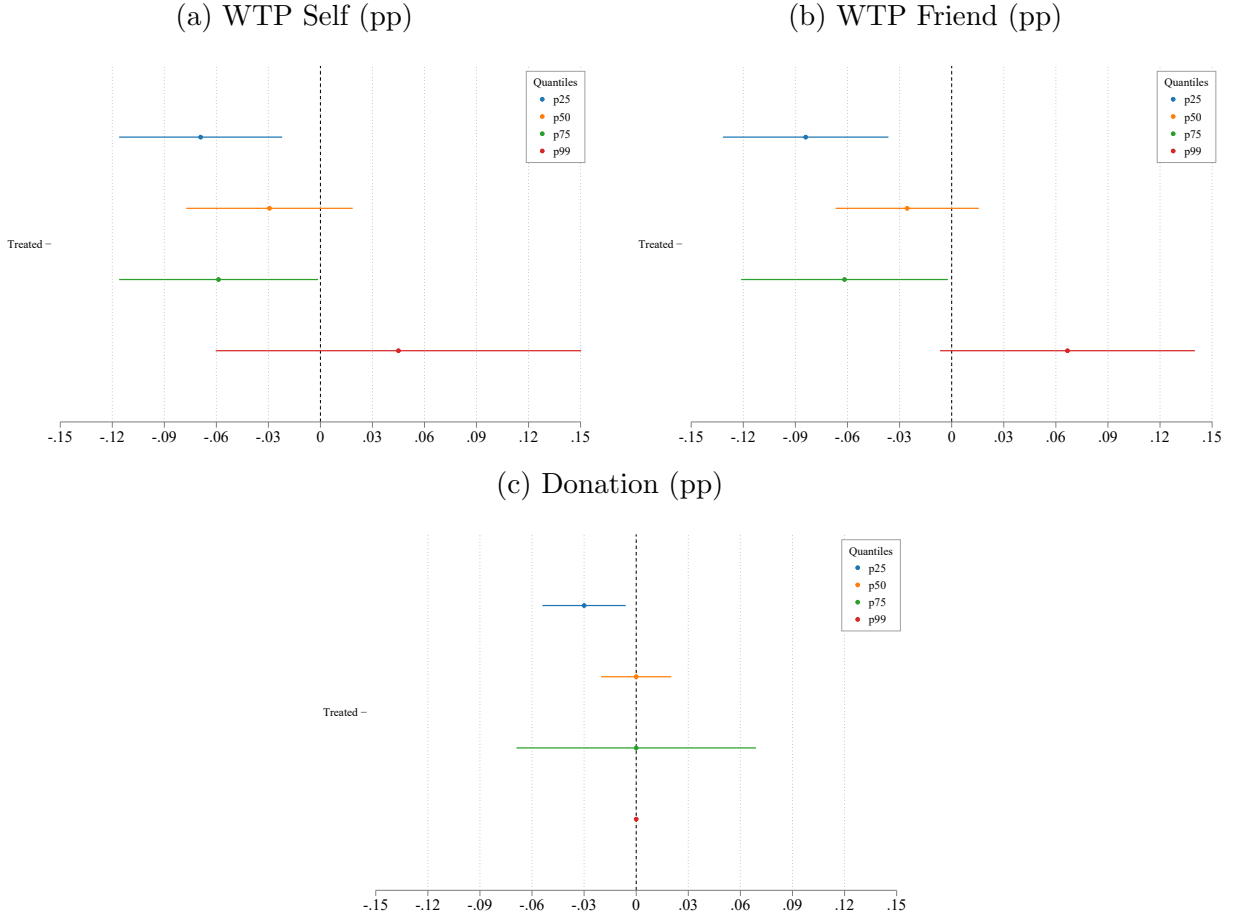
(a) WTP Self (pp)

(b) WTP Friend (pp)

(c) Donation (pp)

Notes: This figure shows treatment effect estimates and 90% confidence intervals of quantile regressions for WTP for oneself, a friend, and the percentage of earnings donated for another participant's therapy subscription. We denote the first quartile by p25, the median by p50, the third quartile by p75, and the 99th percentile by p99. Quantile results are not the upper quartile for donation given the bunching of chosen percentages for donations in subject responses.

group), (1) Prior 2 incorrect only, (2) Prior 3 incorrect only, (3) Both priors incorrect. We observe no heterogeneity highlighting that our information intervention might have also carried the salience effect promoting subjects to think more about the shared facts beyond just updating on specific facts or statements. This is one of the limitations we face in having treated subjects with multiple facts concurrently due to constraints on power, but in future studies it might be worth exploring the differential effects of individual statements and different means of delivering them.

We also examine how treatment effects vary by levels of the stigma index (Figure 13). Marginally insignificant negative coefficients on the interaction term provide suggestive evidence that the effects are suggestively stronger for more stigmatized individuals in the willingness-to-pay outcomes.

Figure 11: Heterogeneity by Distress and Professional Help Usage
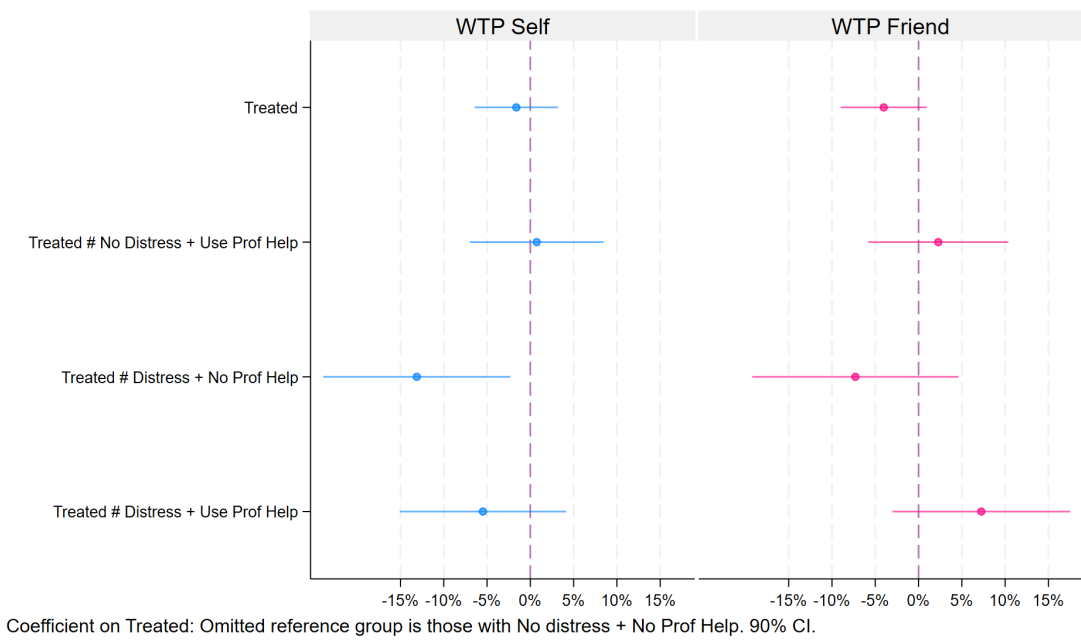
Coefficient on Treated: Omitted reference group is those with No distress + No Prof Help. 90% CI.

Notes: This figure shows estimates and 90% confidence intervals for willingness to pay. The top-most estimates report the ATE, while the three bottom-most coefficients show heterogeneous effects across subgroups of the population of interest. In particular, we estimate heterogeneity on those who are not in distress but have used professional mental health help, those who are in distress but have not used professional mental health help and those who are in distress and have used professional mental health help.

Figure 12: Heterogeneity by Incorrect Priors
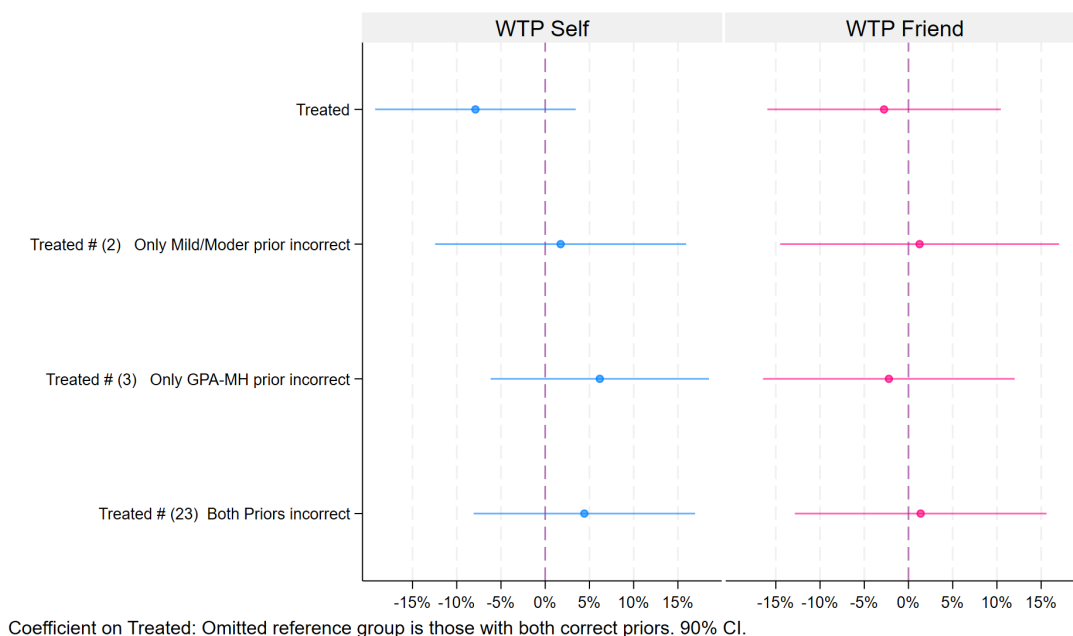
Notes: This figure shows estimates and 90% confidence intervals for willingness to pay. The top-most estimates report the ATE, while the three bottom-most coefficients show heterogeneous effects across subgroups of the population of interest. In particular, we estimate heterogeneity on those who only answered the "Mild/Moderate"-prior question incorrectly, those who only answered the "GPA-MH"-prior question incorrectly and those who answered both these questions incorrectly.

Figure 13: Treatment Effect Heterogeneity by Stigma
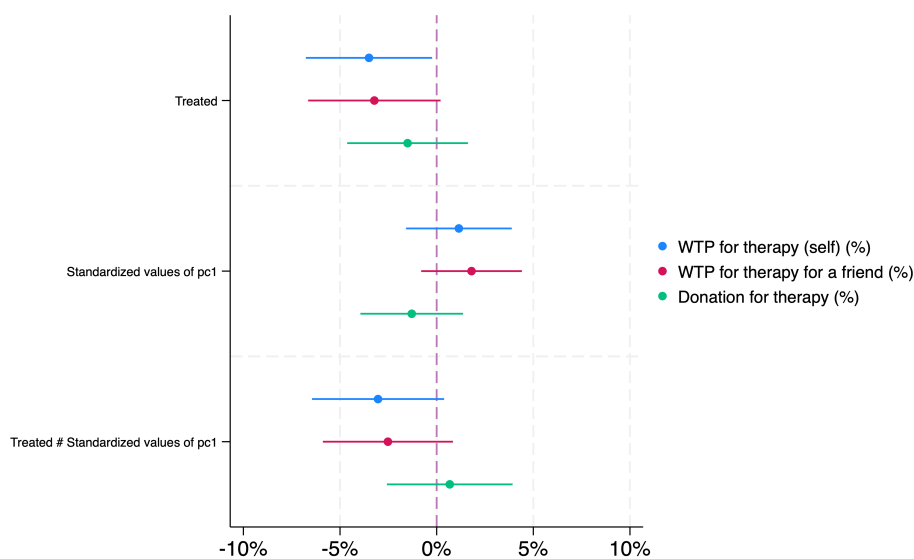
Notes: This figure shows estimated treatment effect and their interactions with underlying stigma measure and 90% confidence intervals of quantile regressions for WTP for oneself, a friend, and the percentage of earnings donated for another participant's therapy subscription. Standardized values of the PCA-derived stigma index (component 1, measured in st. dev.) are included separately and interacted with the main treatment variable.

6 Discussion

Our findings highlight the role of misperceptions and stigma as contributors to the mental health treatment gap among students, especially in settings where financial and structural barriers to therapy access are minimal with free therapy provided to students and higher perceived effectiveness of therapy. While previous research has emphasized attitudinal constraints such as low perceived need or skepticism about effectiveness among adults, we show that misperceptions about who seeks therapy and the perceived relationship between distress and grades correlate with students' underutilization of professional help. The incorrect belief that therapy is primarily for those with severe distress and the inaccurate perception that mental distress is highly correlated with poor academic performance across students may deter them from engaging with available support services.

By correcting these misperceptions, our light-touch intervention increases students' willingness to share campus mental health resources and provide more proactive advice to peers. However, we also find that it lowers individual demand for private therapy measured by students' willingness to pay (WTP) for online therapy subscription for themselves and for a friend, suggesting a potential substitution effect toward free on-campus services. This highlights a previously overlooked dimension of information interventions: while they can reduce stigma and increase awareness, they may also shift help-seeking behavior in ways that interact with existing institutional constraints and resource allocation. This echoes the results in Roth et al. (2024a) who also find that reducing public stigma perceptions lowers individual WTP for online therapy.

These findings have important implications for mental health policy in university settings and beyond, particularly in developing countries where mental health stigma remains high. Our results suggest that addressing psychological frictions through belief correction can be a cost-effective strategy to improve engagement with available resources. However, the substitution effect we document raises questions about whether similar interventions would be effective in settings where free or subsidized therapy is not widely available. More broadly, the findings contribute to the development economics literature on behavioral barriers to human capital investment, aligning with recent work on how cognitive frictions shape education, labor, and health decisions (Schilbach et al. 2016; Rao et al. 2021). Future research could explore whether such interventions can increase treatment take-up in contexts where financial and logistical barriers are more binding and whether belief correction can generate longer-term shifts in mental health norms and behaviors. As universities and policymakers seek to expand mental health support, understanding the mechanisms underlying help-seeking decisions will be critical in designing interventions that effectively close the treatment gap.

References

Abrams, Z. (2022). Student mental health is in crisis. Campuses are rethinking their approach. *APA Monitor in Psychology*, 53(7), 60.

Acampora, M., Capozza, F., & Moghani, V. (2022). Mental Health Literacy, Beliefs and Demand for Mental Health Support among University Students. *Working Paper.*

Andrade, L. H., Alonso, J., Mneimneh, Z., Wells, J. E., Al-Hamzawi, A., Borges, G., Bromet, E., Bruffaerts, R., de Girolamo, G., de Graaf, R., Florescu, S., Gureje, O., Hinkov, H. R., Hu, C., Huang, Y., Hwang, I., Jin, R., Karam, E. G., Kovess-Masfety, V., Levinson, D., Matschinger, H., O'Neill, S., Posada-Villa, J., Sagar, R., Sampson, N. A., Sasu, C., Stein, D. J., Takeshima, T., Viana, M. C., Xavier, M., & Kessler, R. C. (2014). Barriers to mental health treatment: results from the WHO World Mental Health surveys. *Psychological medicine*, 44(6), 1303–17.

Auerbach, R. P., Alonso, J., Axinn, W. G., Cuijpers, P., Ebert, D. D., Green, J. G., Hwang, I., Kessler, R. C., Liu, H., Mortier, P., Nock, M. K., Pinder-Amaker, S., Sampson, N. A., Aguilar-Gaxiola, S., Al-Hamzawi, A., Andrade, L. H., Benjet, C., Caldas-De-Almeida, J. M., Demyttenaere, K., Florescu, S., De Girolamo, G., Gureje, O., Haro, J. M., Karam, E. G., Kiejna, A., Kovess-Masfety, V., Lee, S., McGrath, J. J., O'Neill, S., Pennell, B. E., Scott, K., Ten Have, M., Torres, Y., Zaslavsky, A. M., Zarkov, Z., & Bruffaerts, R. (2016). Mental disorders among college students in the World Health Organization World Mental Health Surveys. *Psychological Medicine*, 46(14), 2955–2970.

Barker, N., Bryan, G., Karlan, D., Ofori-Atta, A., & Udry, C. (2022). Cognitive Behavioral Therapy among Ghana's Rural Poor Is Effective Regardless of Baseline Mental Distress. *American Economic Review: Insights*, 4(4), 527–45.

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral science*, 9(3).

Benjet, C., Borges, G., Medina-Mora, M. E., Zambrano, J., & Aguilar-Gaxiola, S. (2009). Youth mental health in a populous city of the developing world: Results from the mexican adolescent mental health survey. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 50(4), 386–395.

Bhat, B., de Quidt, J., Haushofer, J., Patel, V., Rao, G., Schilbach, F., & Vautrey, P.-L. (2022). The Long-Run Effects of Psychotherapy on Depression, Beliefs, and Economic Outcomes. *SSRN Electronic Journal.*

Bolotnyy, V., Basilico, M., & Barreira, P. (2022). Graduate Student Mental Health: Lessons from American Economics Departments. *Journal of Economic Literature*, 60(4), 1188–1222.

Brewer, K. B., Gibson, R., Tomar, N., Washburn, M., Giraldo-Santiago, N., Hostos-Torres, L. R., & Gearing, R. E. (2023). Why Culture and Context Matters: Examining Differences in Mental Health Stigma and Social Distance Between Latino Individuals in the United States and Mexico. *Journal of Immigrant and Minority Health.*

Brody, D. J., Pratt, L. A., & Hughes, J. P. (2018). *Prevalence of Depression Among Adults Aged 20 and Over: United States, 2013–2016 Key findings Data from the National Health and Nutrition Examination Survey.* Technical Report 303.

Brouwers, E. P. (2020). Social stigma is an underestimated contributing factor to unemployment in people with mental illness or mental health issues: Position paper and future directions. *BMC Psychology*, 8(1).

Bushong, B. & Gagnon-Bartsch, T. (2024). An Experiment on Interpersonal Projection Bias. *Working Paper.*

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: A global return on investment analysis. *The Lancet Psychiatry*, 3(5).

Cornaglia, F., Crivellaro, E., & McNally, S. (2015). Mental health and education decisions. *Labour Economics*, 33, 1–12.

Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Canadian Journal of Psychiatry*, 58(7), 376–85.

Cuijpers, P., Cristea, I. A., Ebert, D. D., Koot, H. M., Auerbach, R. P., Bruffaerts, R., & Kessler, R. C. (2016). Psychological Treatment of Depression in College Students: A Metaanalysis. *Depression and Anxiety*, 33(5), 400–414.

Eisenberg, D., Lipson, S. K., Heinze, J., & Zhou, S. (2022). *Healthy Minds Study – US National Report 2021-2022.* Technical report.

Fletcher, J. M. (2008). Adolescent depression: Diagnosis, treatment, and educational attainment. *Health Economics*, 17(11), 1215–1235.

Haushofer, J., Mudida, R., & Shapiro, J. P. (2021). The Comparative Impact of Cash Transfers and a Psychotherapy Program on Psychological and Economic Well-being. *SSRN Electronic*

Journal.

Healthy Minds Survey (2022). Data for Research - Healthy Minds Network.

Jaadi, Z. & Whitfield, B. (2024). Principal Component Analysis (PCA): A Step-by-Step Explanation.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9).

Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: The PHQ-4. *Psychosomatics*, 50(6).

Lacey, L., Mishra, N., Mukherjee, P., Prakash, N., Prakash, N., Quinn, D., Sabarwal, S., & Saraswat, D. (2024). Can destigmatizing mental health increase willingness to seek help? Experimental evidence from Nepal. *Journal of Policy Analysis and Management*, 44(1), 97–124.

Lagunes-Cordoba, E., Davalos, A., Fresan-Orellana, A., Jarrett, M., Gonzalez-Olvera, J., Thornicroft, G., & Henderson, C. (2021). Mental Health Service Users' Perceptions of Stigma, From the General Population and From Mental Health Professionals in Mexico: A Qualitative Study. *Community Mental Health Journal*, 57(5), 985–993.

Link, B. G. & Phelan, J. C. (2001). Conceptualizing stigma. *Annual Review of Sociology*, 27, 363–385.

Mascayano, F., Tapia, T., Schilling, S., Alvarado, R., Tapia, E., Lips, W., & Yang, L. H. (2016). Stigma toward mental illness in Latin America and the Caribbean: a systematic review. *Brazilian Journal of Psychiatry*, 38(1), 73–85.

Mortier, P., Cuijpers, P., Kiekens, G., Auerbach, R. P., Demyttenaere, K., Green, J. G., Kessler, R. C., Nock, M. K., & Bruffaerts, R. (2018). The prevalence of suicidal thoughts and behaviours among college students: a meta-analysis. *Psychological medicine*, 48(4), 554–565.

Negeri, Z. F., Levis, B., Sun, Y., He, C., Krishnan, A., Wu, Y., Bhandari, P. M., Neupane, D., Brehaut, E., Benedetti, A., & Thombs, B. D. (2021). Accuracy of the Patient Health Questionnaire-9 for screening to detect major depression: updated systematic review and individual participant data meta-analysis.

OECD Report (2022). *Tackling the mental health impact of the COVID-19 crisis: An integrated, whole-of-society response - OECD.* Technical report.

Oreopoulos, P. (2007). Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics*, 91(11-12), 2213–2229.

Osman, N., Michel, C., Schimmelmann, B. G., Schilbach, L., Meisenzahl, E., & Schultze-Lutter, F. (2022). Influence of mental health literacy on help-seeking behaviour for mental health problems in the Swiss young adult community: a cohort and longitudinal case–control study. *European Archives of Psychiatry and Clinical Neuroscience*.

Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., Chisholm, D., Collins, P. Y., Cooper, J. L., Eaton, J., Herrman, H., Herzallah, M. M., Huang, Y., Jordans, M. J., Kleinman, A., Medina-Mora, M. E., Morgan, E., Niaz, U., Omigbodun, O., Prince, M., Rahman, A., Saraceno, B., Sarkar, B. K., De Silva, M., Singh, I., Stein, D. J., Sunkel, C., & UnÜtzer, J. (2018). The Lancet Commission on global mental health and sustainable development.

Patel, V., Weobong, B., Weiss, H. A., Anand, A., Bhat, B., Katti, B., Dimidjian, S., Araya, R., Hollon, S. D., King, M., Vijayakumar, L., Park, A. L., McDaid, D., Wilson, T., Velleman, R., Kirkwood, B. R., & Fairburn, C. G. (2017). The Healthy Activity Program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: a randomised controlled trial. *The Lancet*, 389(10065), 176–185.

Rao, G., Redline, S., Schilbach, F., Schofield, H., & Toma, M. (2021). Informing sleep policy through field experiments. *Science*, 374(6567), 530–533.

Ridley, M. (2022). Essays on the economics of health and education.

Ridley, M. (2025). Mental Illness Discrimination. *Working Paper (Former JPM)*.

Ridley, M., Rao, G., Schilbach, F., & Patel, V. (2020). Poverty, depression, and anxiety: Causal evidence and mechanisms. *Science*, 370(6522).

Ronak, J. & Khandelwal, V. (2024). Silent Networks : The Role of Inaccurate Beliefs. *JMP*.

Roth, C., Schwardmann, P., & Tripodi, E. (2024a). Depression Stigma. *CESifo Working Paper no. 11012*.

Roth, C., Schwardmann, P., & Tripodi, E. (2024b). Misperceived Effectiveness and the Demand for Psychotherapy Misperceived Effectiveness and the Demand for Psychotherapy *. *Journal of Public Economics*.

Sæther, M. H., Sivertsen, B., & Bjerkeset, O. (2021). Mental Distress, Help Seeking, and Use of Health Services Among University Students. The SHoT-Study 2018, Norway. *Frontiers in*

Psychiatry, 12, 727237.

Salud Mental (2022). Que hay detras de los suicidios de estudiantes de Medicina? Esto explica la UNAM – El Financiero. *El Financiero (Mexico)*.

Schilbach, F., Schofield, H., & Mullainathan, S. (2016). The psychological lives of the poor. *American Economic Review*, 106(5), 435–440.

Schnyder, N., Panczak, R., Groth, N., & Schultze-Lutter, F. (2017). Association between mental health-related stigma and active help-seeking: Systematic review and meta-analysis. *British Journal of Psychiatry*, 210(4), 261–268.

Shreekumar, A. & Vautrey, P.-L. (2023). Managing Emotions: The Effects of Online Mindfulness Meditation on Mental Health and Economic Behavior. *Working Paper*.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, 166(10).

Thapar, A., Eyre, O., Patel, V., & Brent, D. (2022). Depression in young people. *The Lancet*, 400(10352), 617–631.

Velazquez Hernandez, A. (2017). Suicidio y depresion en estudiantes y residentes de Medicina en Mexico. *Salud Mental, Vice.com*.

Wang, P. S., Aguilar-Gaxiola, S., Alonso, J., Angermeyer, M. C., Borges, G., Bromet, E. J., Bruffaerts, R., de Girolamo, G., de Graaf, R., Gureje, O., Haro, J. M., Karam, E. G., Kessler, R. C., Kovess, V., Lane, M. C., Lee, S., Levinson, D., Ono, Y., Petukhova, M., Posada-Villa, J., Seedat, S., & Wells, J. E. (2007). Worldwide Use of Mental Health Services for Anxiety, Mood, and Substance Disorders: Results from 17 Countries in the WHO World Mental Health (WMH) Surveys. *Lancet*, 370(9590), 841.

WHO, T. W. H. O. (2021). Mental health of adolescents.

World Health Organization (2021). Depression.

A   Appendix: Description of Outcomes

Mental Distress

Apart from the clinical screening procedures, the most common approach to assess depression or anxiety symptoms in a sample is to utilize standardized screening questionnaires. We elicit information on student respondents' mental health in our study using PHQ-4 and GAD-4 short surveys, each containing 4 questions with a total of 8 questions in the questionnaire. These are the shortened versions of PHQ-9 and GAD-7, self-administered diagnostic instruments for reliable and valid measures of depression and anxiety severity (Kroenke et al. 2001; Spitzer et al. 2006). These screening surveys, widely used for clinical screening in physicians' offices and hospitals as well as for epidemiological measurement, are characterized by high sensitivity (the probability of testing positive when the disorder is present) and high specificity (the probability of testing negative when the disorder is absent).[29]

Each question has four possible answers: 'not at all' scores 0 points, 'several days' scores 1 point, 'more than half the days' scores 2 points and 'nearly every day' scores 3 points. This means that participants' scores on each survey (PHQ-4 and GAD-4) range from 0 to 12, so for most of the analysis in the paper we use a combined distress measure relying on all eight questions with a distress index for each person ranging between 0-24 (higher = worse mental health).

Practitioners sometimes use an ultra-brief screening scale that combines four most diagnostic symptoms – a score of 3 or above (out of 6) is typically used as the cutoff point for concluding that major depressive disorder (for the PHQ) or generalized anxiety disorder (for the GAD) is likely (Kroenke et al. 2009). Consistent with that, we use a threshold of 12 out of 24 points to indicate that a respondent in our survey exhibits symptoms associated with being in mental distress (distress, hereafter) consistent with moderate or severe depression/anxiety.[30]

Professional Help Use

We measure usage of professional mental health services by directly asking respondents whether they are currently receiving professional mental health help and whether they have ever received

---

[29]Major depressive disorder screening questionnaire PHQ-9 exhibits 88% sensitivity and 88% specificity. Generalized anxiety disorder screening questionnaire exhibits 89% sensitivity and 82% specificity (Kroenke et al. 2001; Spitzer et al. 2006).

[30]The midpoint cutoff of 12 points is a direct generalization from the cutoff of 6 points for PHQ-2 & GAD-2 instrument - we double the threshold given our scale doubles the range. Given the commonly used threshold of 10 for both PHQ-9 (27 points) and GAD-7 (21 points), one could in principle use the same threshold of 10 for the joint PHQ-4& GAD-4 questionnaire (24 points) we have in our survey – from this end, the threshold of 12 can be viewed as more conservative yielding in a lower bound of student share in distress. Throughout the paper, we focus on a binary split of students into those in distress and not in distress, based on the 12 points cutoff.

such help in the past. About 24% of participants report currently receiving professional help, while 66% have received professional help at some point in their lives.

To understand the social context of mental health, we ask participants if they have friends who are currently or have previously received professional help. Moreover, participants are asked whether they believe a friend or someone they know closely would benefit from therapy. A significant majority of participants—87.6% and 89.4%, respectively—reported knowing a friend who has used professional help and believing that a friend would benefit from therapy.

Beliefs About Mental Health and Help-Seeking

Our survey contains a number of questions asking students about their beliefs surrounding mental distress as well as therapy use by them and their peers, in an incentivized way. For example, we ask respondents to guess the share of students at their university who have received professional help in the last 12 months and if their guess is within 5% of the true number, they earn a bonus. We also elicit students' beliefs about mental health stigma - in one of the questions they are asked to predict the percentage of students who would be open to sharing their mental health challenges with a classmate – the incentivization scheme is the same for this type of questions with a 5% margin of error. As a proxy for perceived self-stigmatization, we elicit respondents' guesses about share of students who would feel disappointed in themselves if they had a mental health issue.

We elicit incentivized beliefs in this manner on several other points: guess about the long-run effectiveness of cognitive behavioral therapy (yes/no), guess about the number of studies from a meta-study that showed the effectiveness of therapy in reducing depression (count of studies out of 22, bonus given is picked the correct number), guess about whether professional help users are mostly students with more severe symptoms of mental distress (yes/no), guess about the correlation between GPA and the mental distress score across students (categorical). Finally, we have several questions without a bonus incentive including questions about the share of other students, their parents, and university professors who would view a student with symptoms of mental distress negatively.

Peer Advice and Donation

This outcome examines participants' ability to provide advice to peers struggling with emotional challenges and their willingness to donate a portion of their earnings to support access to therapy for others. Participants were asked to imagine a scenario where a friend approaches them for emotional support due to personal struggles (advice). They were then prompted to provide open-ended advice, which was evaluated across multiple dimensions to assess the thoughtfulness, relevance, and overall quality of their responses. These evaluations help gauge participants' em-

pathy, problem-solving abilities, and understanding of mental health support, providing insight into their capacity to engage in peer-to-peer support effectively.

In addition, participants were asked about their willingness to donate a percentage of their compensation to help fund therapy for others. Specifically, they were informed that their donations would be directed toward covering the cost of a therapy session for a fellow [University A] student who reported that financial constraints were preventing them from seeking therapy. Participants were notified that any donation they pledged would be automatically deducted from their payment and allocated toward this funded therapy session. This approach allowed us to capture participants' altruism and their support for increasing access to mental health services in a tangible, actionable way.

Together, these variables – open-ended advice and willingness to donate – highlight the interconnected roles of peer-to-peer emotional support and financial contributions in fostering a supportive mental health environment. They provide insight into students' attitudes toward helping others, both through direct advice and through reducing barriers to accessing professional mental health care.

Demand for Private Therapy

To assess whether our interventions have an effect on demand for therapy, we briefly introduce BetterHelp – an online therapy service that matches users with licensed therapists – before eliciting students' (incentivized) willingness to pay (WTP) for this service. The first step involves measuring respondents' willingness to pay for therapy for themselves. Students are asked to indicate the maximum amount they would spend on four weeks of therapy with BetterHelp, which is typically priced at approximately MXN $6,500. To ensure responses are incentive-compatible, we inform participants that a subset of respondents will be randomly selected, and their choices will be implemented. In this implementation, a computer will bid an amount between MXN $0 and MXN $7,000 against the student's bid. If the student's WTP exceeds the computer's bid, they will receive four weeks of therapy for free. If the computer's bid is higher, the participant will receive an Amazon Gift Card valued at the computer's bid amount.

After eliciting respondents' willingness to pay for therapy for themselves, we extend the question to their willingness to pay for therapy for a friend. This question is structured similarly to the first but asks for the maximum amount participants would be willing to spend on four weeks of therapy with BetterHelp for their friend. This dual approach provides insights into both personal and altruistic demand for therapy services.

Additionally, if participants choose a WTP of zero, they are prompted with a follow-up question that elicits their willingness to accept (WTA) – how much they would need to be

compensated to use this mental health tool. This mechanism addresses potential concerns about a high prevalence of zero WTP responses, which, while informative to some extent, could limit understanding of the spectrum of demand for mental health services and the barriers to utilizing certain treatments. Overall, this provides a nuanced measure of student demand for mental health resources. It also allows us to explore how actual and perceived stigma and discrimination against people seeking mental health treatment influence their willingness to seek therapy. By capturing both WTP and WTA, we aim to deepen our understanding of the psychological and social factors that drive or hinder engagement with mental health services.

B Appendix: Figures and Tables
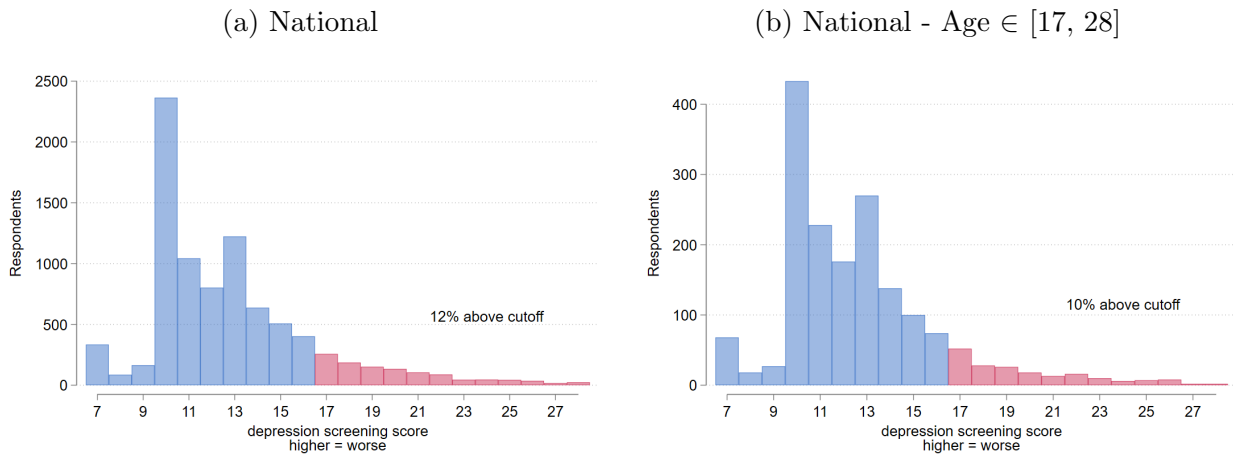
B.1 National Statistics

Figure B1: National Estimates of Anxiety Prevalence (OECD 2022)
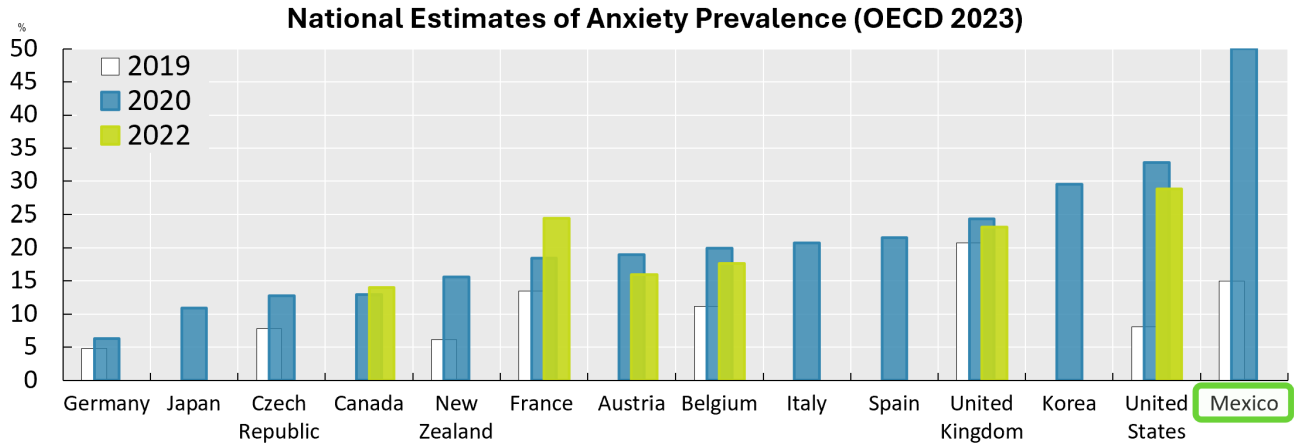
Notes: This figure shows national estimates of anxiety prevalence across OECD countries over time.

## Depression Screening Scores – ENSANUT

Figure B2: Mental Distress in Mexico - 2023

(a) National (b) National - Age $\in [17, 28]$

Notes: This figure shows the distribution of depression screening scores using data from the 2023 Mexican Health and Nutrition Survey (ENSANUT). The survey is representative of the national population. In panel (a) we show the distribution among ENSANUT respondents aged 10 years old or older (sample size = 8,696). In panel (b) we subset respondents to those between 17 and 28 years old to more closely approximate the population of university students (sample size = 1,720).
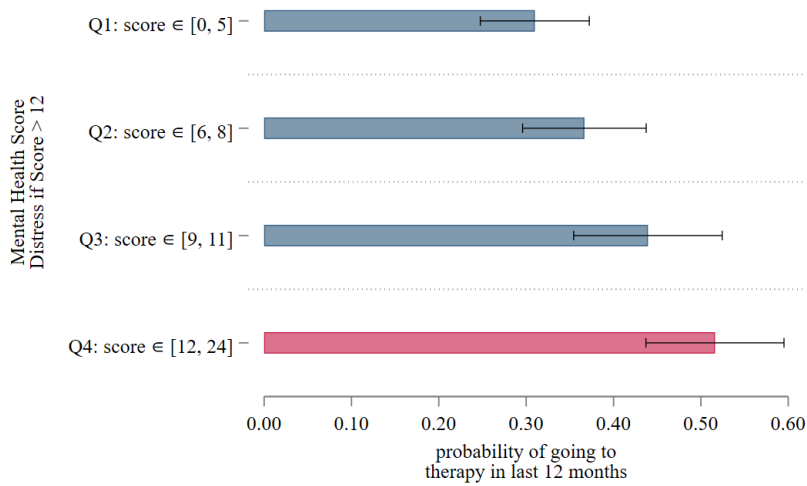
B.2   Mental Health Index and Professional Help Use

Figure B3: Mental Distress Share by Financial Stress
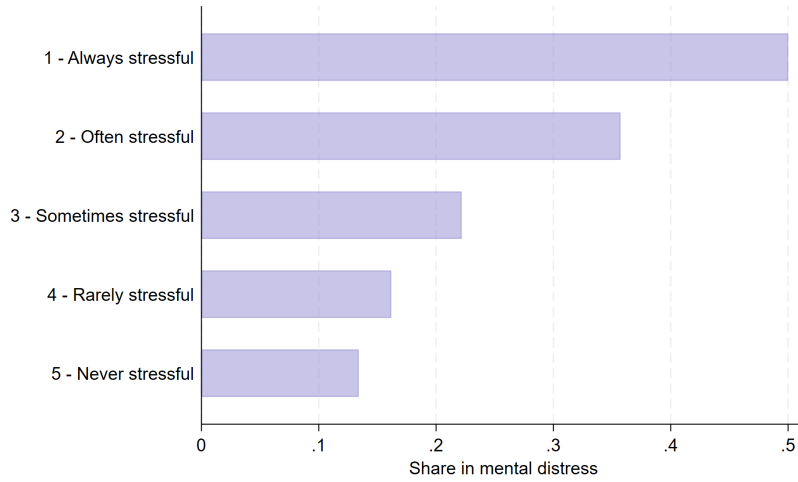
Notes: This figure shows the share of respondents in mental distress within each possible answer option to the question of *How would you describe your financial situation?*

Figure B4: Therapy Use Share by Mental Distress Quartile

Notes: This figure shows estimates of the probability of having used therapy in the last 12 months conditional on being in a given quartile of the mental health score distribution. We show 95% confidence intervals in black capped spikes.

Figure B5: Mental Distress Index Distribution by Perceived Support Level

Student Distribution by Perceived Support Level
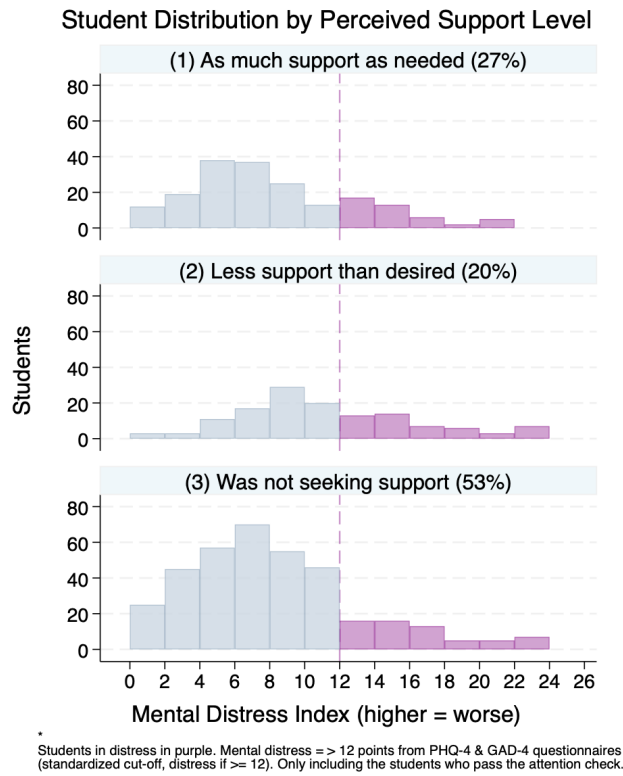
Notes: This figure shows the distribution of mental health scores across responses to the question of whether *In the last 12 months I received...* More help than I needed/Less help than I needed/I was not seeking support.

Figure B6: Who Did You Turn For Help?
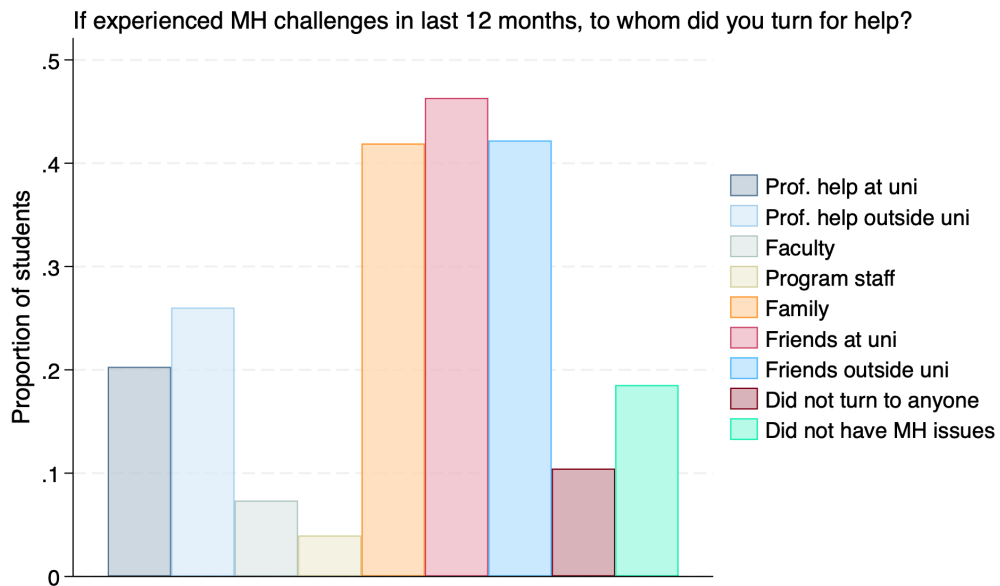
Notes: This figure shows the share of students who chose each of the options to the question of whom did the respondent turn for help in case she experienced mental health challenges in the past 12 months.

Figure B7: CDFs of guesses of therapy use by prior own use.
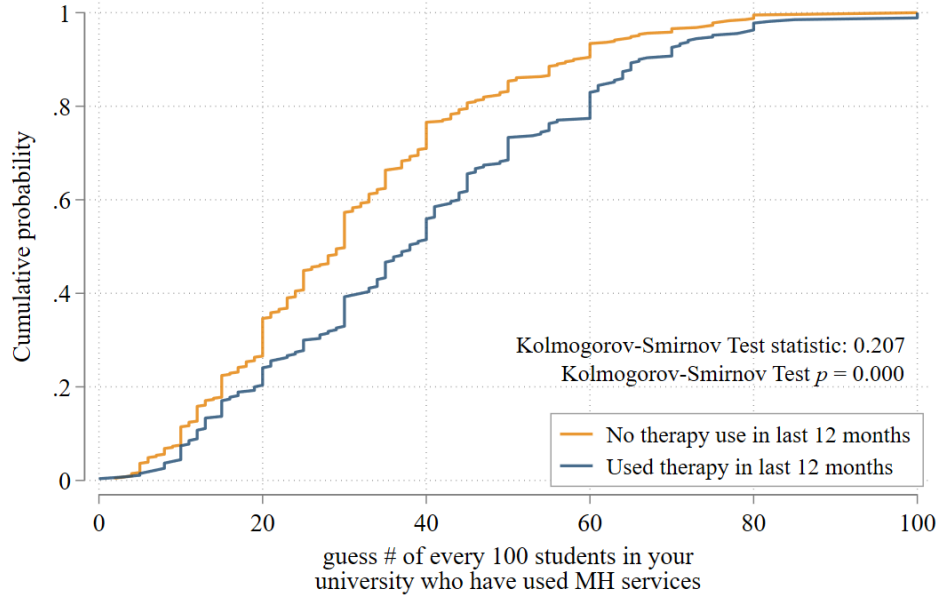
Notes: This figure shows the cumulative distribution function of the guesses of the number of students out of every 100 students from their university who have used professional mental health services in the last 12 months, splitting the sample by those who have/have not used professional mental health services in the last 12 months.

Figure B8: Resource link sharing

Notes: This figure shows the infographic containing the QR code for sharing campus services information.

Figure B9: Self-Stigma and Openness to Share

(a) Self-Stigma

(b) Openness to Sharing Mental Health Challenges
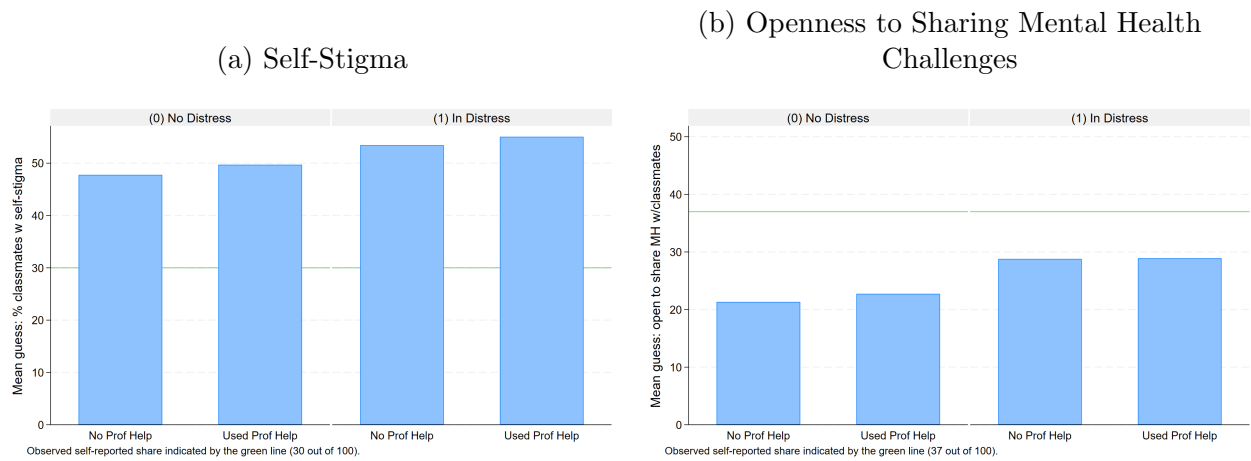
Notes: This figure shows, in Panel (a) the mean guess of students who would be disappointed in themselves if they had a mental health issue, splitting the sample on whether the respondent is in mental distress or not, and on whether the respondent has/has not used professional mental health services in the last 12 months. In Panel (b) we show the mean guess of the number of students out of 100 who would be open to sharing if they had mental health issues with students who are not necessarily their friends. We also split the sample on whether the student is in mental distress or not, and on whether the student has/has not used professional mental health services in the last 12 months.

B.3 Balance on observables
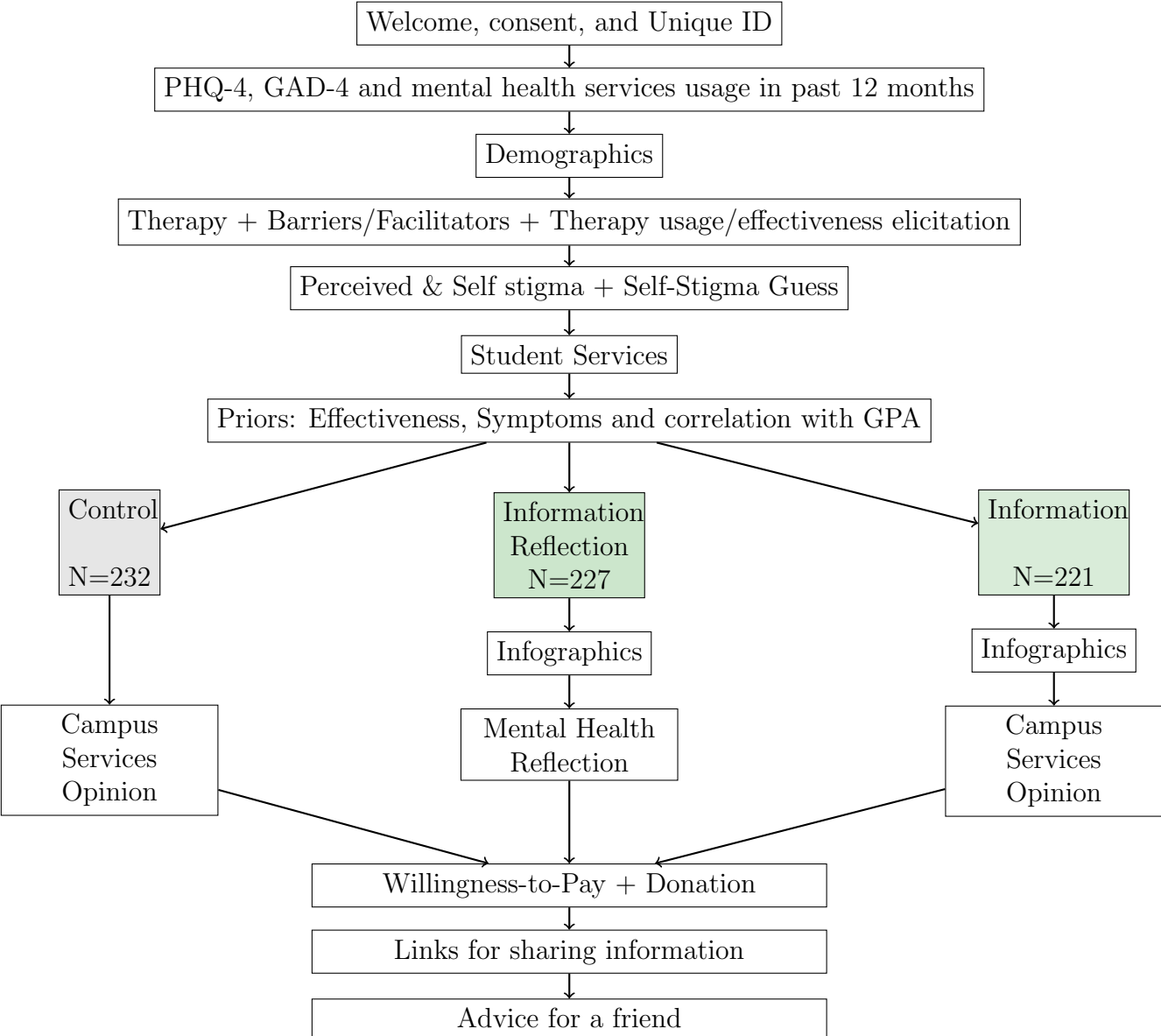
Table B1: Covariate Balance

Variable	N	(1) (0) Control Mean/(SD)	N	(2) (1) T1: Info + Reflection Mean/(SD)	N	(3) (2) T2: Info only Mean/(SD)	N	(1)-(2) Mean difference	N	(1)-(3) Pairwise t-test Mean difference	N	(2)-(3) Mean difference
Age	232	20.159 (1.848)	227	20.084 (2.218)	221	20.208 (1.822)	459	0.076	453	-0.049	448	-0.124
Female	232	0.461 (0.500)	227	0.533 (0.500)	221	0.538 (0.500)	459	-0.072	453	-0.077	448	-0.005
Financially Stressed	232	0.530 (0.500)	227	0.599 (0.491)	221	0.543 (0.499)	459	-0.069	453	-0.013	448	0.056
Has Scholarship	232	0.651 (0.478)	227	0.718 (0.451)	221	0.706 (0.457)	459	-0.067	453	-0.055	448	0.012
Receives a full scholarship	232	0.082 (0.275)	227	0.084 (0.278)	221	0.072 (0.260)	459	-0.002	453	0.009	448	0.011
Moved Residence	232	0.591 (0.493)	227	0.626 (0.485)	221	0.615 (0.488)	459	-0.035	453	-0.025	448	0.010
GPA	232	90.897 (4.659)	227	90.784 (5.394)	221	91.235 (3.925)	459	0.112	453	-0.339	448	-0.451
MH Score	232	8.569 (5.132)	227	8.048 (5.003)	221	8.430 (5.110)	459	0.521	453	0.139	448	-0.381
Used Therapy L12 Months	232	0.233 (0.424)	227	0.181 (0.386)	221	0.290 (0.455)	459	0.052	453	-0.057	448	-0.109***
Open to Share MH Challenges	232	0.392 (0.489)	227	0.339 (0.474)	221	0.371 (0.484)	459	0.053	453	0.021	448	-0.032
Self-stigmatize	232	0.323 (0.469)	227	0.295 (0.457)	221	0.276 (0.448)	459	0.028	453	0.047	448	0.019

Notes: This table shows balance on covariates across treatment groups. For each covariate we show each experimental group's sample mean and standard deviation, as well the difference in means across pairs of groups. Age measures the respondent's age in years, female is an indicator equal to one if the respondent is female-born, financially stressed is an indicator equal to one if the respondent described her financial situation as "Always", "Often" or "Sometimes" stressful and equal to 0 if she reported it as "Rarely" or "Never" stressful, Has scholarship is an indicator equal to one if the respondent has at least some amount of scholarship, receives a full scholarship is an indicator equal to one if the respondent's scholarship covers 100% of tuition, moved residence is an indicator equal to one if the respondent moved her residence city to pursue her current studies, GPA measures the respondent's current overall GPA on a scale from 0–100, MH score measures the student's mental health score as described in section 2, used therapy in L12 months is an indicator equal to one if the respondent states having used therapy in the last 12 months, open to share MH challenges is an indicator equal to one if the respondent states she would be willing to share about her own personal MH challenges with others and self-stigmatize is an indicator equal to one if the respondent states she would be disappointed in herself if she suffered from mental distress. Standard errors for the difference in means test are heteroskedasticity robust. Significance levels: * $p < 0.1$, ** $p < 0.05$ and *** $p < 0.01$

C Appendix: Further Analyses

C.1 Survey flowchart

Figure B10: Survey Flow

```
┌─────────────────────────────────────────────┐
│        Welcome, consent, and Unique ID        │
└─────────────────────────────────────────────┘
                      ↓
┌─────────────────────────────────────────────────────────┐
│ PHQ-4, GAD-4 and mental health services usage in past 12 months │
└─────────────────────────────────────────────────────────┘
                      ↓
            ┌──────────────────┐
            │   Demographics   │
            └──────────────────┘
                      ↓
┌──────────────────────────────────────────────────────────────────┐
│ Therapy + Barriers/Facilitators + Therapy usage/effectiveness elicitation │
└──────────────────────────────────────────────────────────────────┘
                      ↓
    ┌──────────────────────────────────────────┐
    │ Perceived & Self stigma + Self-Stigma Guess │
    └──────────────────────────────────────────┘
                      ↓
            ┌──────────────────┐
            │ Student Services │
            └──────────────────┘
                      ↓
┌──────────────────────────────────────────────────────┐
│ Priors: Effectiveness, Symptoms and correlation with GPA │
└──────────────────────────────────────────────────────┘
```

Control	Information Reflection	Information
N=232	N=227	N=221

Information Reflection → Infographics → Mental Health Reflection

Information → Infographics → Campus Services Opinion

Control → Campus Services Opinion

```
┌─────────────────────────────────┐
│   Willingness-to-Pay + Donation  │
└─────────────────────────────────┘
                ↓
┌─────────────────────────────────┐
│    Links for sharing information │
└─────────────────────────────────┘
                ↓
┌─────────────────────────────────┐
│        Advice for a friend       │
└─────────────────────────────────┘
```

Notes: This figure depicts the survey flow.

C.2 Heterogeneity Analysis Specifications

Heterogeneity by Incorrect Beliefs

We test whether the treatment effects differ based on the accuracy of students' prior beliefs. Specifically, we examine the following groups: (0) Both priors correct (reference group), (1) Prior 2 incorrect only, (2) Prior 3 incorrect only, (3) Both priors incorrect.

The regression model is:

$$Y_i = \alpha + \sum_{j=1}^{3} \delta_j BeliefGroup_{ij} + \sum_{j=1}^{3} \phi_j (InfoTreatment_i \cdot BeliefGroup_{ij}) + X_i'\gamma + \varepsilon_i,$$

where $BeliefGroup_{ij}$ is an indicator variable for individual i being in belief group j (e.g., "Prior 2 incorrect only"), $InfoTreatment_i \cdot BeliefGroup_{ij}$ is the interaction term capturing the differential treatment effect for each belief group j, ϕ_j represents the difference in treatment effects for each group relative to the baseline (both priors correct).

Heterogeneity by Distress and Professional Help Use

We explore heterogeneity by groups defined by combinations of mental distress and professional help use: (0) No distress + no professional help (reference group), (1) No distress + professional help, (2) Distress + no professional help, (3) Distress + professional help.

The regression model is:

$$Y_i = \alpha + \sum_{j=1}^{3} \delta_j DistressGroup_{ij} + \sum_{j=1}^{3} \phi_j (InfoTreatment_i \cdot DistressGroup_{ij}) + X_i'\gamma + \varepsilon_i,$$

where $DistressGroup_{ij}$ is an indicator variable for individual i being in distress group j (e.g., "Distress + no professional help"), $InfoTreatment_i \cdot DistressGroup_{ij}$ is the interaction term capturing the differential treatment effect for each distress/help group j, ϕ_j represents the difference in treatment effects for each group relative to the baseline (no distress + no professional help).

Heterogeneity by Stigma Index

We also examine how treatment effects vary by levels of the stigma index. The model is specified as:

$$Y_i = \alpha + \beta_1 InfoTreatment_i + \delta StigmaIndex_i + \phi (InfoTreatment_i \cdot StigmaIndex_i) + X_i'\gamma + \varepsilon_i,$$

where $StigmaIndex_i$ is a continuous variable representing individual i's stigma index score, $InfoTreatment_i \cdot StigmaIndex_i$ is the interaction term capturing how treatment effects vary with levels of stigma, ϕ measures the marginal change in treatment effect per unit increase in the stigma index.

Each specification allows us to analyze differential treatment effects. In the first specification, ϕ_j quantifies whether treatment effects vary based on prior beliefs, relative to those with

both priors correct. In the second one, ϕ_j captures how treatment effects differ for combinations of mental distress and professional help use, relative to the baseline group (no distress + no professional help). In the third specification, ϕ indicates whether treatment effects are stronger or weaker depending on the level of stigma.

These models provide insights into whether the intervention's effectiveness is moderated by key characteristics of participants.

C.3 Incentivized bonus questions

The eight bonus questions included: (1) guessing the percentage of "Yes" responses to the question regarding therapy usage in the past 12 months which was compared to the actual calculated percentage; (2) guessing the percentage of "Yes" responses to the question on willingness to share therapy information, which was similarly compared to the actual percentage; (3) responding "22" to a specific survey question, which earned the bonus if correct; (4) guessing the percentage of "Agree" responses (including "Strongly Agree," "Agree," and "Somewhat Agree") to a question on self-stigma, validated against the computed percentage; (5) answering "Yes" to a question about therapy effectiveness, which directly earned the bonus; (6) answering "Yes" to a question about therapy effectiveness for mild-to-moderate conditions , which similarly earned the bonus; (7) categorizing the correlation between mental health scores and grade point averages into predefined categories such as "Better" or "Much Worse", with correctness determined by the computed correlation; and (8) providing open-ended advice on a specific topic, where responses deemed "Very useful" during review earned the bonus.

C.4 Stigma Index

In the context of the study we seek to create a unified measure of stigma taking into account three distinct dimensions.

- **Perceived Public Stigma**: This dimension is defined by three variables that measure the perception fo stigma of other students, professors, and parents.

- **Self-Stigma**: This dimension corresponds to a variable that measures the number of people out of 100 that would feel disappointment for experiencing any mental health issues.

- **Personal Stigma**: This third dimension corresponds to two dummy variables measuring preference of a lower GPA over experiencing mental health symptoms and talking about mental healthy issues.

The following tables provides a comprehensive description of the variables present across the 3 dimensions. From these classifications we aim to implement not only a PCA to generated an index, but also a weighted average.

Table B2: Mental Health Stigma Variables

	Definition
Perceived Public Stigma	
From students	Percentage of students that the respondent believes would view a student negatively for experiencing mental health issues like anxiety or depression.
From professors	Percentage of professors that the respondent believes would view a student negatively for experiencing mental health issues like anxiety or depression.
From parents	Percentage of student parents that the respondent believes would view a student negatively for experiencing mental health issues like anxiety or depression.
Self-Stigma	
Self-stigma	Respondent's estimate of how many out of 100 students would feel disappointed in themselves if they had a mental health issue.
Personal Stigma	
Low GPA over MH symptoms	Dummy variable where it has a value of 1 if the respondent ranked a student with a low GPA as preferred as a class project teammate rather than a student experiencing mental health distress; 0 otherwise.
Low GPA over MH talk	Dummy variable where it has a value of 1 if the respondent ranked a student with a low GPA as preferred as a class project teammate rather than a student openly talking about mental health issues; 0 otherwise.

Notes: This table shows the definition of variables used as inputs for constructing our stigma index.

C.4.1 Weighted Average

To create a unified measure of mental health stigma, we developed indices that account for three distinct dimensions of stigma: *Perceived Public Stigma*, *Self-Stigma*, and *Personal Stigma*. Each dimension was represented by relevant variables described in the table above, and the methodology for index construction is outlined below.

Table B3: Summary Statistics for Stigma Dimensions (Mean and Median Thresholds)

Dimension	Threshold Type	Mean	Std. Dev.	Min–Max
Perceived Public Stigma	Mean-based	0.428	0.396	0–1
	Median-based	0.492	0.397	0–1
Self-Stigma	Mean-based	0.513	0.500	0–1
	Median-based	0.469	0.499	0–1
Personal Stigma	Mean-based	0.204	0.337	0–1
	Median-based	0.204	0.337	0–1

Notes: Perceived Public Stigma reflects stigma perceptions from students, professors, and parents. Self-Stigma measures internalized stigma based on perceived social disappointment in experiencing mental health issues. Personal Stigma captures preferences for GPA trade-offs over experiencing or discussing mental health concerns.

Table B4: Summary Statistics for Composite Stigma Indices

Index	Standardization	Mean	Std. Dev.	Min–Max
Composite Index (Mean-based)	Raw	-0.107	1.847	-2.787 − 4.573
	Standardized	-0.055	0.948	-1.431 − 2.348
Composite Index (Median-based)	Raw	-0.117	1.836	-2.879 − 4.498
	Standardized	-0.061	0.958	-1.502 − 2.347

Notes: Composite indices classify stigma levels using mean- and median-based thresholds. Perceived public stigma is assessed separately for students, professors, and parents. Self-stigma is binarized based on a defined threshold, while personal stigma captures preferences for GPA trade-offs over experiencing or discussing mental health issues.

For each variable within the stigma dimensions, we created binary indicators based on whether the value exceeded the dimension-specific mean or median. Perceived public stigma was assessed separately for students, professors, and parents by comparing their reported percentages against predefined thresholds. Specifically, for students, values greater than 26.35 (mean) or 20 (median) indicated perceived public stigma, while for professors, the corresponding thresholds were greater than 26.61 (mean) or 20 (median). For parents, the thresholds were set at greater than 40.11 (mean) or 39.5 (median). Self-stigma was binarized using a threshold of greater than 49.79 (mean) or 50 (median). Finally, personal stigma was represented by two binary variables: the preference for a lower GPA over experiencing mental health symptoms and the preference for a lower GPA over talking about mental health issues.

Aggregation Within Dimensions

For each stigma dimension, aggregated measures were computed based on the proportion of satisfied binary indicators. Perceived public stigma was calculated as the mean of the three binary indicators corresponding to students, professors, and parents. Self-stigma, being a single variable, was directly represented by its binary indicator. Personal stigma was aggregated as

the mean of the two binary indicators reflecting preferences related to GPA and mental health concerns.

To account for potential variation across treatment groups, the aggregated shares for each dimension were standardized. This was achieved by centering the values around the control group's mean and dividing by the standard deviation.

Table B5: Correlations Between Stigma Dimensions and Composite Indices

Dimension	Perceived Public Stigma	Self-Stigma	Personal Stigma	Composite Index
Mean-Based Composite Index				
Perceived Public Stigma	1.0000	0.2455	-0.0177	0.6612
Self-Stigma	0.2455	1.0000	0.0255	0.6863
Personal Stigma	-0.0177	0.0255	1.0000	0.5248
Composite Index	0.6612	0.6863	0.5248	1.0000
Median-Based Composite Index				
Perceived Public Stigma	1.0000	0.2550	-0.0293	0.6695
Self-Stigma	0.2550	1.0000	-0.0041	0.6805
Personal Stigma	-0.0293	-0.0041	1.0000	0.5053
Composite Index	0.6695	0.6805	0.5053	1.0000

Notes: Perceived Public Stigma reflects beliefs about how students, professors, or parents view mental health issues. Self-Stigma captures internalized negative attitudes toward one's own mental health. Personal Stigma represents preferences related to GPA trade-offs over experiencing or discussing mental health concerns. The Composite Index combines these stigma dimensions into standardized measures.

Weighted Average Stigma Index Amongst Distressed and Non-Distressed Individuals

Table B6: Summary of Composite Stigma Indices by Mental Distress Groups

Mental Distress Group	Mean (Mean-Based Index)	Mean (Median-Based Index)	SD (Mean-Based Index)	SD (Median-Based Index)
No Distress (0)	-0.1023	-0.1075	0.9784	0.9855
In Distress (1)	0.1058	0.0952	0.8192	0.8414
Total	-0.0549	-0.0613	0.9479	0.9578

Notes: This table summarizes composite stigma indices by distress groups. The mean-based and median-based indices classify stigma levels using different statistical cutoffs. Individuals in distress show higher stigma levels compared to those without distress.

Table B7: T-Test Results for Composite Stigma Indices by Distress Groups

Index	Group	Obs	Mean	Std. Err.	95% CI	p-value (two-tailed)
Mean-Based Index	No Distress	525	-0.1023	0.0427	[-0.1862, -0.0184]	0.0162
	In Distress	155	0.1058	0.0658	[-0.0242, 0.2358]	
	Difference		-0.2081	0.0863	[-0.3776, -0.0386]	
Median-Based Index	No Distress	525	-0.1075	0.0430	[-0.1920, -0.0230]	0.0205
	In Distress	155	0.0952	0.0676	[-0.0383, 0.2287]	
	Difference		-0.2027	0.0873	[-0.3741, -0.0314]	

Notes: This table shows two-sample t-test results comparing composite stigma indices between individuals with and without distress. The mean-based and median-based indices classify stigma perceptions using different statistical thresholds. Results show significant differences, with distressed individuals exhibiting higher stigma levels.

The results of the two-sample t-tests indicate a significant difference in stigma indices (mean and median) between individuals with "No Distress" (No D) and those "In Distress" (In D). For the stigma index based on the mean, individuals in the "No D" group had a significantly lower stigma index (Mean = -0.102, Std. Dev = 0.978) compared to those in the "In D" group (Mean = 0.106, Std. Dev = 0.819), with a mean difference of -0.208 (95% CI: -0.378 to -0.039; $t = -2.4101$, $p = 0.0162$ for the two-tailed test). Similarly, for the stigma index based on the median, the "No D" group had a lower stigma index (Mean = -0.108, Std. Dev = 0.986) compared to the "In D" group (Mean = 0.095, Std. Dev = 0.841), with a mean difference of -0.203 (95% CI: -0.374 to -0.031; $t = -2.3228$, $p = 0.0205$ for the two-tailed test).

These findings suggest that individuals in distress experience higher levels of stigma compared to those not in distress. The statistical significance ($p < 0.05$) and confidence intervals that exclude zero provide strong evidence that these differences are unlikely due to random chance. Although the effect sizes (mean differences of -0.208 and -0.203) are relatively small, the results underscore the need for targeted interventions to address stigma among distressed individuals.
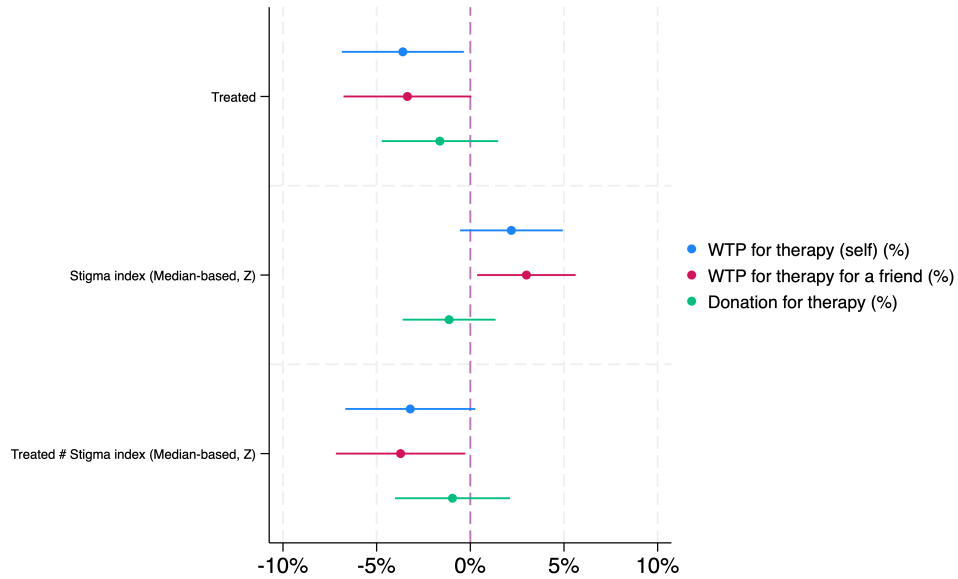
Index from Weighted Average

Figure B11: Main Effects by Mean Stigma Index
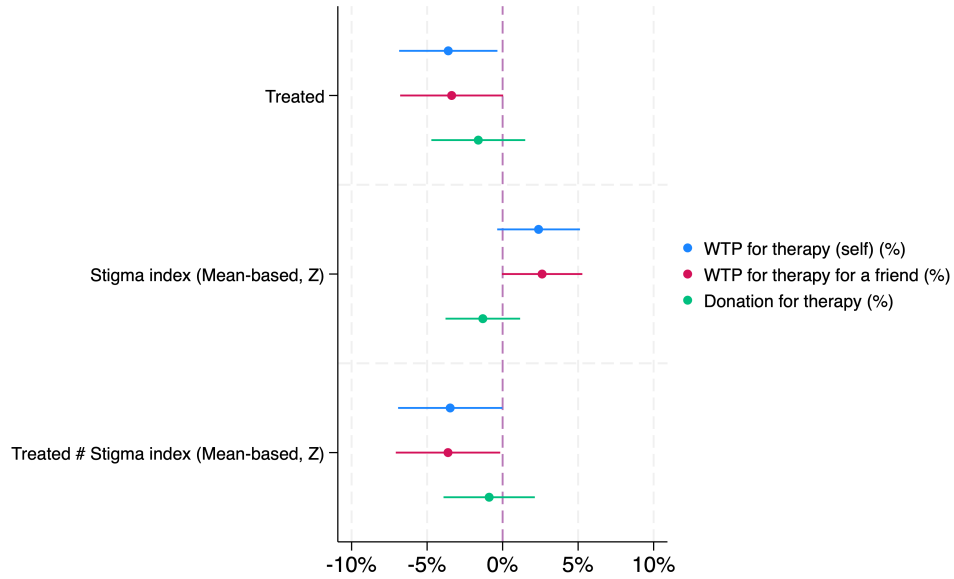
Notes: This figure shows treatment effects by mean stigma index, illustrating its relationship with willingness to pay (WTP) for therapy for oneself, for a friend, and donation amounts. The mean stigma index reflects perceived public stigma, self-stigma, and personal stigma, capturing overall attitudes toward mental health.

Figure B12: Main Effects by Median Stigma Index

Notes: This figure shows treatment effects by median stigma index, showing its relationship with willingness to pay (WTP) for therapy for oneself, for a friend, and donation amounts. The median stigma index captures perceived public stigma, self-stigma, and personal stigma, summarizing overall attitudes toward mental health.

C.4.2 PCA

Principal Component Analysis (PCA) is employed as a dimensionality reduction technique to distill key insights from a dataset with multiple variables while minimizing the loss of critical information (Jaadi & Whitfield (2024)), the context of our research on mental health stigma among university students, PCA enables us to synthesize a complex set of variables—such as perceptions of therapy, barriers to seeking help, and beliefs about peer behavior—into a smaller number of components. These components capture the majority of the variance within the original dataset, providing a simplified yet meaningful representation of the underlying patterns.

In this analysis, PCA helps identify the primary dimensions of mental health stigma, which we use to construct an index reflecting the most significant factors influencing students' attitudes and behaviors. Initially, all components and loadings are considered, but subsequent iterations focus on those with the highest explained variance and loadings of 0.3 or above. This filtering ensures that we emphasize the most informative relationships between variables. Prior to applying PCA, all variables are standardized to ensure comparability and to give equal weight to each variable, regardless of its original scale.

This approach not only simplifies our data analysis but also provides a robust foundation for understanding the most influential factors shaping students' mental health perceptions and their decision-making regarding therapy.

Table B8: Correlation of Stigma Index PCA1 and PCA2 with Components

Variable	PCA1	PCA2	Stigma Students	Stigma Professors	Stigma Parents	Guess Self-Stigma	Low GPA Symptoms	Low GPA Talk
PCA1	1.00							
PCA2	0.00	1.00						
Stigma Students	0.83***	0.04	1.00					
Stigma Professors	0.87***	0.04	0.64***	1.00				
Stigma Parents	0.83***	-0.03	0.53***	0.66***	1.00			
Guess Self-Stigma	0.50***	0.11**	0.30***	0.27***	0.25***	1.00		
Low GPA Symptoms	-0.09*	0.84***	-0.02	-0.03	-0.06	-0.00	1.00	
Low GPA Talk	-0.04	0.85***	-0.01	0.00	-0.04	0.04	0.45***	1.00

Notes: This table shows the correlations between the two principal components (PCA1 and PCA2) and the key stigma-related variables. PCA1 primarily captures perceived public stigma from students, professors, and parents, while PCA2 reflects attitudes related to academic performance and mental health. The stigma variables represent perceived stigma from different groups, Guess Self-Stigma measures internalized stigma, and Low GPA Symptoms and Low GPA Talk capture preferences for avoiding mental health symptoms or discussions even at the cost of lower academic performance.

Table B9: Summary Statistics for Stigma Variables and PCA Indexes

Variable	Observations	Mean	Std. Dev.	Min–Max
Stigma Index PCA1	680	0.000	1.00	-1.93 − 3.47
Stigma Index PCA2	680	0.000	1.00	-0.80 − 2.66
Stigma Students (Std.)	680	0.000	1.00	-1.22 − 3.40
Stigma Professors (Std.)	680	0.000	1.00	-1.20 − 3.31
Stigma Parents (Std.)	680	0.000	1.00	-1.57 − 2.34
Guess Self-Stigma (Std.)	680	0.000	1.00	-2.20 − 2.21
Low GPA Symptoms (Std.)	680	0.000	1.00	-0.61 − 1.65
Low GPA Talk (Std.)	680	0.000	1.00	-0.40 − 2.50

Notes: PCA1 captures public stigma perceptions from students, professors, and parents, while PCA2 reflects attitudes toward academic performance and mental health. Stigma variables measure perceived stigma from different groups, self-stigma represents internalized beliefs, and low GPA variables capture preferences for academic performance over mental health concerns.

PCA Results

Table B10: Principal Components Analysis Summary

Component	Eigenvalue	Difference	Proportion	Cumulative
Component 1	2.3835	0.9320	0.3972	0.3972
Component 2	1.4515	0.6114	0.2419	0.6392
Component 3	0.8401	0.2902	0.1400	0.7792
Component 4	0.5499	0.0839	0.0917	0.8708
Component 5	0.4660	0.1569	0.0777	0.9485
Component 6	0.3090	–	0.0515	1.0000

Summary Statistics:	
Number of observations	680
Number of components	2
Trace	2
Rotation (unrotated)	Principal
Rho	0.6392

Notes: The eigenvalues measure the variance explained by each principal component. According to the Kaiser criterion (Jaadi and Whitfield, 2024), only components with eigenvalues above 1 should be retained. In this case, only the first two components meet this criterion, capturing 63.92% of the total variance. These components will be used for further analysis, such as examining the loadings.

Figure B13: Screeplot
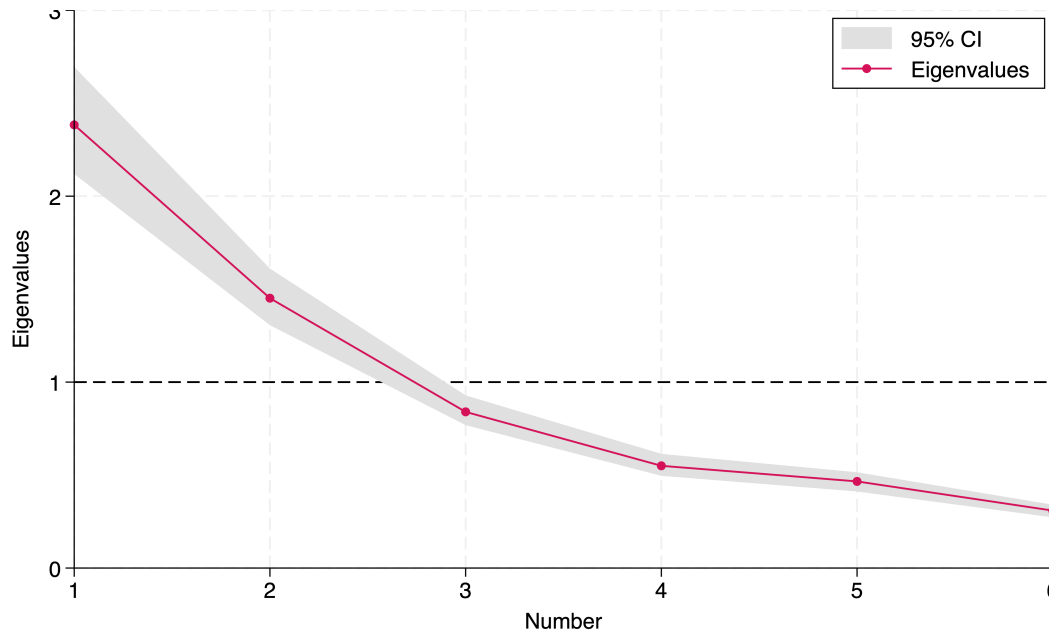
Notes: This figure shows the variance explained by each principal component. We keep components 1 and 2, which exceed the threshold of 1 for being kept in further analyses.

Eigenvalues are the measure of how much variance (information) each principal component explains in the dataset. Larger eigenvalues indicate components that explain more variance Jaadi & Whitfield (2024). From the initial PCA results in the table above and from the screeplot we can observe that only the first two components have eigenvalues of 1 and above - meaning they each explain greater variance than the rest of the components - which will be the ones we shall be keeping, and the only components we shall be considering when looking at the loadings.

Table B11: Principal Components (Eigenvectors)

Variable	Component 1	Component 2	Component 3	Component 4	Component 5	Component 6
stigma_students	0.5358	0.0351	-0.1130	0.0557	-0.7420	0.3810
stigma_professors	0.5649	0.0359	-0.2256	-0.0323	0.0352	-0.7915
stigma_parents	0.5359	-0.0216	-0.2206	-0.0214	0.6617	0.4747
self_stigma	0.3206	0.0934	0.9365	0.0610	0.0811	-0.0327
low_GPA_over_sympt	-0.0269	0.7059	-0.0124	-0.7062	-0.0142	0.0446
low_GPA_over_talk	-0.0551	0.7000	-0.1020	0.7021	0.0593	-0.0045

Notes: This table shows the loadings (coefficients) from the principal component analysis (PCA), representing how much each variable contributes to a given component. In the next step, only loadings greater than 0.3 will be considered to improve interpretability.

The loadings - the coefficients, or weights - from the Principal Component Eigenvectors table above represent the contribution of each variable to a given principal component. In the next iteration only loadings above .3 will be considered in order to better interpret components. Subsequently, we make sure the first two components we have focused on are not correlated amongst each other.

60

Table B12: Correlation Matrix of Principal Components

	pc1	pc2
pc1	1.0	–
pc2	0.0	1.0

Notes: This table shows the correlation matrix between both of our Principal Components. PC1 captures public stigma perceptions from students, professors, and parents, while PC2 reflects attitudes toward academic performance and mental health.

PCA Interpretation

Component 1 primarily captures perceptions of stigma from various groups (students, professors, parents), while Component 2 reflects preferences related to mental health versus academic performance (low GPA acceptance). The following loading plot showcases the previous loadings and how related they are to each component.

Figure B14: PCA Loading Plot
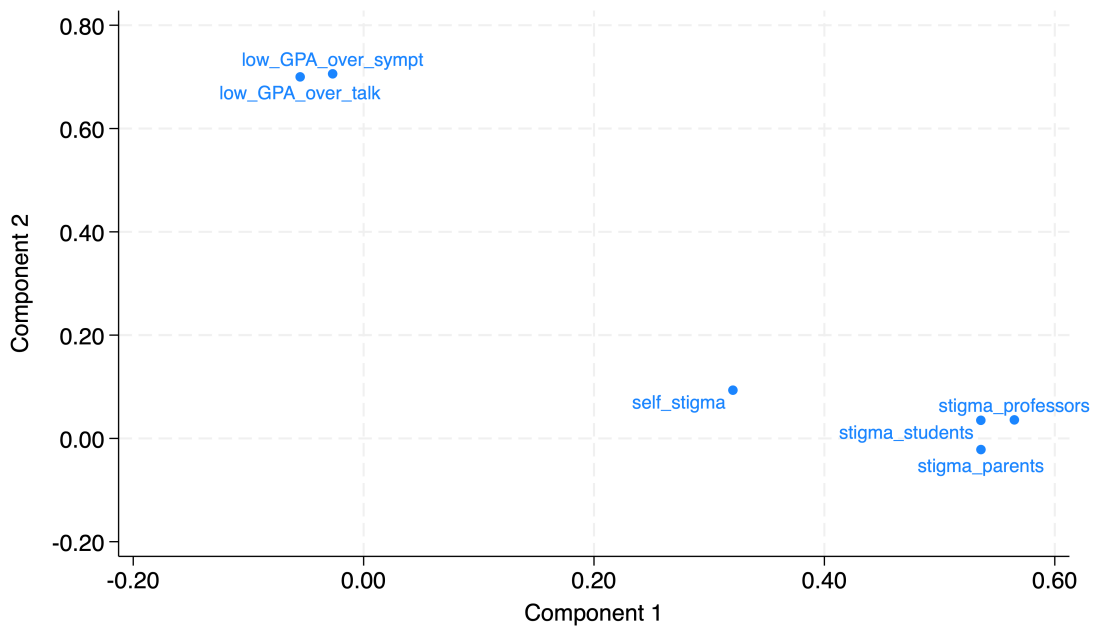
Notes: This figure shows the scatter plot of PCA loadings for each of our Principal Components. PC1 is primarily explained by perceived public stigma from students, professors, and parents, while PC2 reflects attitudes toward academic performance and mental health, particularly preferences related to GPA trade-offs.

Index from PCA

After having done and examined the 2 components of the PCA, we proceed to construct two indexes from component 1 and component 2, we then proceed to interact of our treatment groups with the our standardized stigma index with outcome variables being WTP for therapy, for self, for a friend and lastly the donation amount an individual is willing to give to help someone access mental health services.

Figure B15: Main Effects by Component 1 Stigma Index
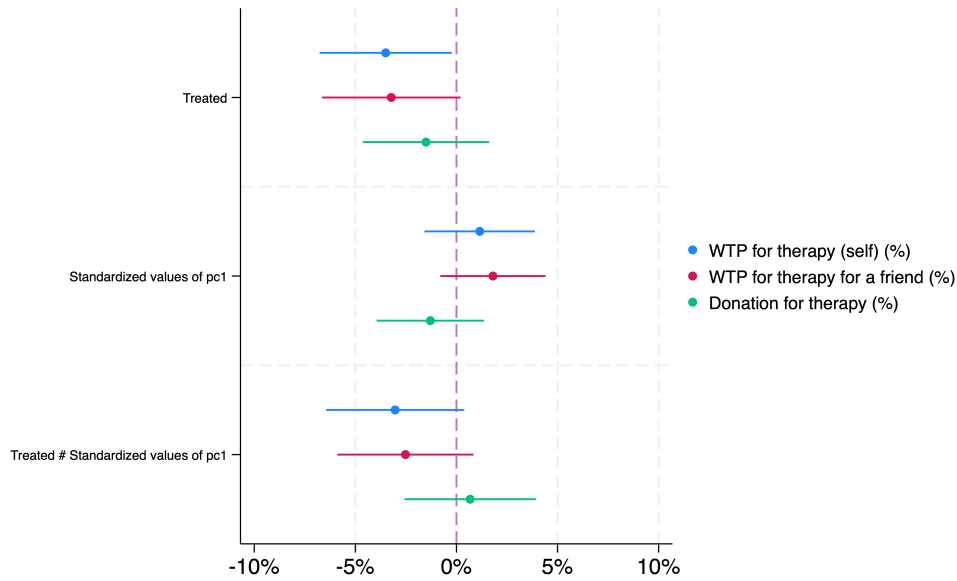
Notes: This figure shows treatment effects by Component 1 (PC1) of the stigma index, illustrating its relationship with willingness to pay (WTP) for therapy for oneself, for a friend, and donation amounts. PC1 primarily captures perceived public stigma from students, professors, and parents, summarizing external attitudes toward mental health.

Figure B16: Main Effects by Component 2 Stigma Index
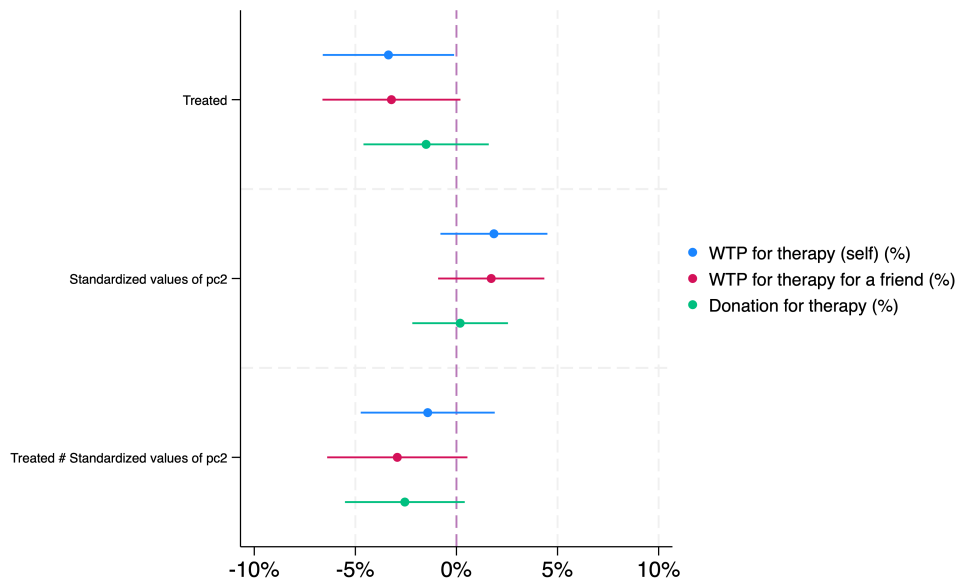
Notes: This figure shows treatment effects by Component 2 (PC2) of the stigma index, illustrating its relationship with willingness to pay (WTP) for therapy for oneself, for a friend, and donation amounts. PC2 primarily reflects attitudes toward academic performance and mental health, particularly preferences related to GPA trade-offs.

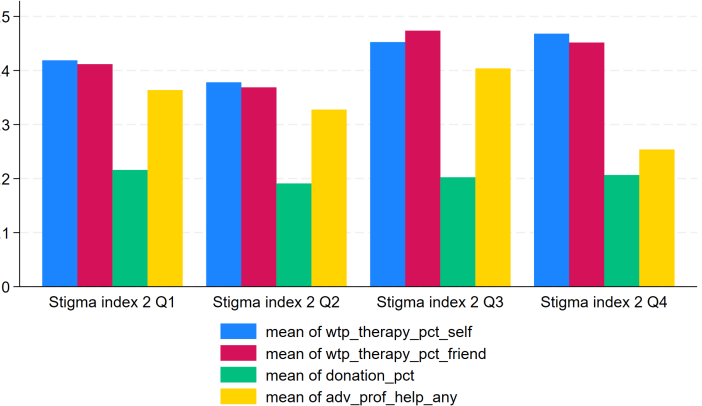### C.4.3 Stigma & Demand for Mental Health Services

Figure B17: Outcome Means by Stigma Index 1
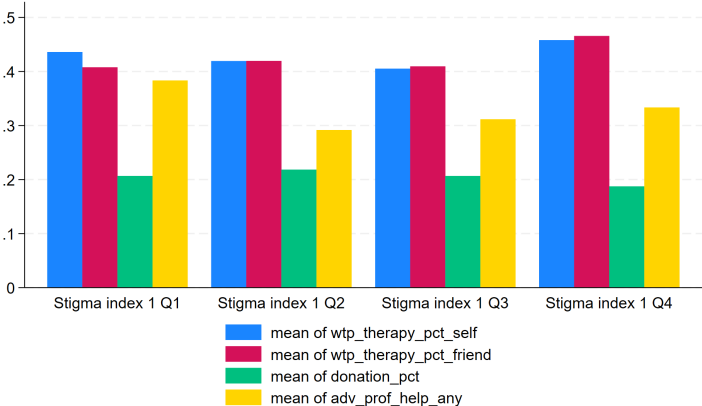
Notes: This figure illustrates how outcome measures vary across quartiles of stigma index 1, capturing differences in willingness to pay (WTP) for therapy, advocacy for professional mental health services, and donation behavior. Higher stigma levels are associated with increased personal investment in therapy but reduced advocacy for professional help.

Figure B17 shows how outcome measures vary across the quartiles of stigma index 1. In Q1 (lowest stigma), the WTP for therapy for oneself and a friend is high, alongside higher rates of advocating for professional mental health services. In Q4 (highest stigma) there is an increase in WTP for therapy for oneself and a friend, while a stark decrease in advocacy for professional help; donations seem relatively uniform across all quartiles. The differences across quartiles reveal that higher stigma is associated with increased personal investment in therapy but decreased engagement with broader supportive actions, such as advocating for professional help or donating.

Figure B18: Outcome Means by Stigma Index 2

Notes: This figure illustrates how outcome measures vary across quartiles of stigma index 2, capturing differences in willingness to pay (WTP) for therapy, advocacy for professional mental health services, and donation behavior. Higher stigma levels are associated with increased personal investment in therapy but reduced advocacy for professional help.

When looking at outcome means by stigma index 2, which primarily focuses on personal stigma—a dimension that mainly measures the preferences of academic performance over addressing mental health issues—one can observe that there is an increasing trend in WTP for self and friend from Q1 to Q4, and a decrease in advocacy for mental health services. Similarly to outcome means by stigma index 1, donations seem relatively stable across quartiles. These parallel trends between stigma index 1 and stigma index 2 suggest that while the two indices capture different dimensions of stigma (public and personal, respectively), their influence on behavioral outcomes, such as WTP, advocacy, and donations, is aligned. This alignment reinforces the robustness of stigma index 1 in explaining how stigma—whether public or personal—affects mental health-related decisions and highlights the consistency of stigma's negative impact on broader support for professional mental health resources.

Table B13: Correlation of Demand Variables with Stigma Indices

Variable	WTP Therapy (Self)	WTP Therapy (Friend)	Donation (%)	Stigma Index PCA1	Stigma Index PCA2
Stigma Index PCA1	-0.03	0.01	-0.04	1.00	
Stigma Index PCA2	0.04	-0.00	-0.06	0.00	1.00

Notes: WTP Therapy (Self) and WTP Therapy (Friend) measure willingness to pay for therapy for oneself and for a friend, respectively. Donation (%) represents the percentage of income participants are willing to donate to mental health causes. Stigma Index PCA1 captures public stigma perceptions, while Stigma Index PCA2 reflects attitudes toward academic performance and mental health. The table shows correlations between these variables.

The Table B13 shows that willingness to pay (WTP) for therapy for oneself and for a friend are strongly and positively correlated, suggesting that individuals who value therapy for themselves also value it for others. Donations, while positively correlated with both WTP measures, exhibit weaker associations, indicating a different motivational factor driving altruistic behavior. Neither stigma index (PCA1 or PCA2) shows significant correlations with WTP or donations, highlighting that perceived public stigma (PCA1) and personal stigma (PCA2) are not directly linked to demand for therapy or altruistic behavior in this context.

The coefficient plots demonstrate that neither Stigma Index 1 (PCA1), representing perceived public stigma, nor Stigma Index 2 (PCA2), capturing personal stigma, has a significant impact on the demand variables. For PCA1, the effects on willingness to pay (WTP) for therapy (self and friend) and donations for therapy are minimal, indicating that public stigma perceptions do not strongly influence these behaviors. Similarly, PCA2 shows near-zero effects across the same variables, suggesting that individual attitudes and personal stigma are not major drivers of therapy demand or donation behavior.

The stigma section reveals that stigma perceptions are shaped by distress levels, professional help usage, and prior beliefs about mental health. Stigma index 1 (PCA1), which captures perceived public stigma from peers, professors, and parents, provides a robust measure of how external societal attitudes influence mental health-related decisions. In contrast, stigma index 2 (PCA2) reflects personal stigma and internalized biases, such as prioritizing academic performance over mental health, but has a narrower focus and limited explanatory power.

Figure B19: Stigma Indices and WTP/Donate

(a) Impact of Stigma Index 1 on Demand
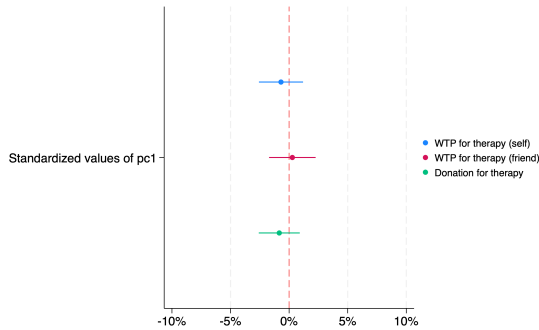
(b) Impact of Stigma Index 2 on Demand
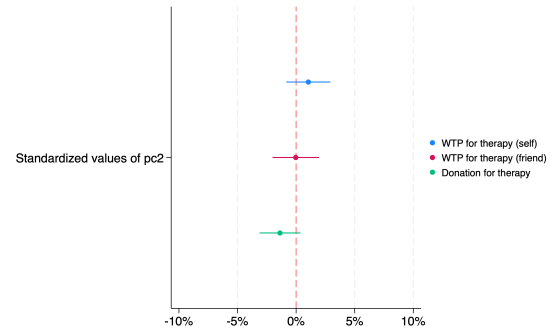
Notes: This figure shows the impact of stigma indices on demand, measured by willingness to pay (WTP) for therapy for oneself and a friend, as well as donations. Stigma Index 1 (PC1) captures public stigma perceptions from students, professors, and parents, while Stigma Index 2 (PC2) reflects personal stigma, particularly preferences for academic performance over addressing mental health issues.

The lack of significant correlations between the stigma indices and demand variables (WTP for therapy and donations) suggests that neither public nor personal stigma directly drives these behaviors. Instead, the data implies that stigma impacts broader societal norms and individual perceptions rather than immediate willingness to invest in therapy. Stigma index 1's comprehensive design offers valuable insights into the broader societal dynamics of stigma, but a more expansive experimental framework could better capture its multifaceted effects on mental health outcomes and decision-making. This underscores the need for future research to refine stigma measures and incorporate additional dimensions for a more complete understanding.

Furthermore, Figure B20 highlights a positive correlation across measures of stigma and financial stress, indicating a potential connection to demand as financial stress may act as a proxy for financial constraints.

Figure B20: Stigma & Mental Distress by Financial Stress
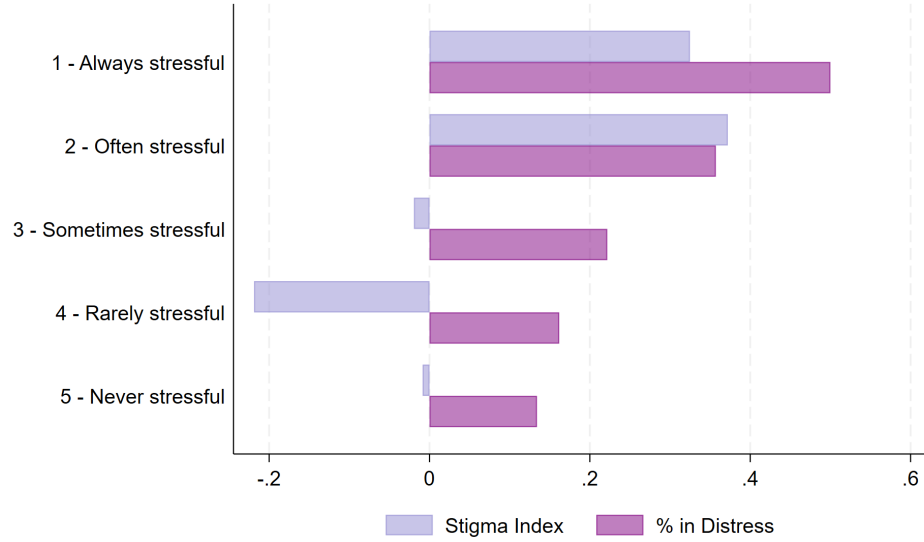
Notes: This figure highlights a positive correlation across measures of stigma and financial stress, indicating a potential connection to demand as financial stress may act as a proxy for financial constraints.

Figure B21: Relationship Between Stigma Index and Perceived Use of Professional Mental Health Services
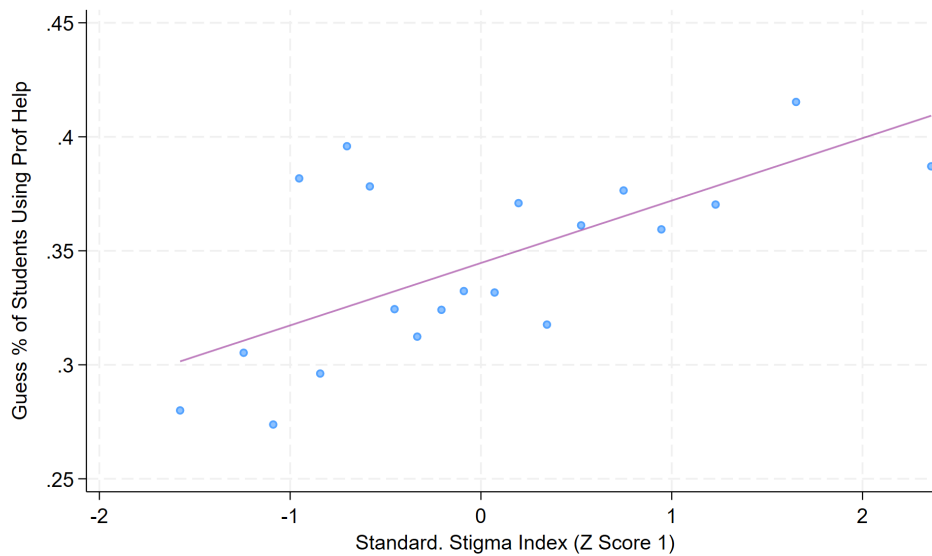
Notes: This figure shows a strong positive correlation between stigma index 1 and predicted therapy use. Individuals with higher stigma levels may assume greater concealment of therapy use among peers, influencing their broader perceptions of mental health treatment.

Furthermore, in the given context, there appears to be a strong positive correlation between stigma index 1 and predicted therapy use, as shown in Figure B21. This trend aligns with the notion that individuals rationalize their assumptions about therapy use based on their own beliefs and perceived social norms. Specifically, individuals with higher stigma levels may assume higher rates of concealing therapy use among peers, which in turn influences their predictions of broader usage patterns. This rationalization mechanism underscores the role of stigma in shaping perceptions of mental health treatment, particularly through the lens of assumed societal concealment behaviors.

C.5   Advice Indicators for mentioning words or phrases

We generate indicator variables based on the inclusion of specific words or phrases in incentivized advice provided by subjects.

We build the "Passive advice" variable as an indicator equal to one if any of the following variables are mentioned by the respondent: Listen is an indicator equal to one if the respondent mentions "listen." Be attentive is an indicator equal to one if the respondent mentions "I am here for you/him/her/them", "I am there for you/him/her/them." Empathy is an indicator equal to one if the respondent mentions "empathy" or "understood." Validate feelings is an indicator equal to one if the respondent mentions it is "not bad to feel bad", "it is normal not to feel ok" or "it is completely normal." Show support is an indicator equal to one if the respondent mentions "I support you", "you have my support" or "I love you."

We build the "Suggestive advice" variable as an indicator equal to one if any of the following variables are mentioned by the respondent: Give opinion is an indicator equal to one if the respondent mentions "opinion", "advice", "what to do", "recommend", "you/he/she should", "my experience." Seek help is an indicator equal to one if the respondent mentions "seek help/support", "refer to professional", "find resources." Mention therapy is an indicator equal to one if the respondent mentions "therapy", "psychologist" or "counseling." Do stuff you enjoy is an indicator equal to one if the respondent mentions "do something you enjoy", "activities you like" or "do stuff you like."