# Trade-offs Between Interpersonal and Informational Privacy Concerns in Mental Health Service Mode Choice

Zhichen Chen , Gordon Burtch, Chrysanthos Dellarocas and Bin Gu

Questrom School of Business, Boston University

**Abstract**

Many individuals experiencing mental health challenges do not seek professional help due to various barriers. One particularly important yet often overlooked factor is interpersonal privacy concerns, e.g., fear of social stigma or judgment. Digital technologies offer avenues to mitigate these concerns in mental health support. In particular, telehealth and large language model (LLM)-based conversational AI can provide individuals with access to mental health services while reducing the potential of social exposure. However, digital technologies may *increase* other privacy concerns, namely those related to the collection and preservation of personal information. We explore that tension here through a series of experiments. We begin with an online experiment at Prolific involving 650 participants who reported dealing with moderate to severe mental health challenges. Participants' concerns about social exposure were manipulated exogenously, and each participant completed a standard survey scale evaluating information privacy concerns. Participants were then asked to provide a relative evaluation of three alternative mental health service modes: in-person treatment, telehealth, and an AI-enabled chatbot. We show that participants assigned to higher social exposure concerns exhibited a significantly increased preference for AI and telehealth modes relative to in-person services. Further, however, we show that participants' information privacy concerns interact with concerns about social exposure (interpersonal privacy concerns), such that information privacy plays a systematically weaker role in service mode preference when interpersonal privacy concerns are greater. Simply put, we show that concerns about social stigma or judgment crowd out concerns about personal data. Subsequent experiments demonstrate that although the inclusion of telehealth and AI service offerings as part of a larger bundle can effectively increase participants' willingness to engage with mental health services, the extent of that increase again depends on the interaction between information and interpersonal privacy concerns. Finally, we demonstrate the external validity of our findings through a field experiment conducted in collaboration with a major Chinese university, communicating available mental health service options to its students.

# 1 Introduction

The global rise in mental health issues underscores the need for effective treatment, yet adoption remains modest (Bolotnyy et al., 2022). In 2021, 57.8 million U.S. adults were suffering from mental illness, but only 47.2% received mental health services in the past year [1]. This is due to several factors, among which stigma-related privacy concerns are particularly important yet often overlooked. These concerns arise because stigmatized individuals often manage their identities through concealment, masking, or selective disclosure to avoid negative social consequences (Goffman, 1963), a significant barrier that often discourages individuals from seeking support (Thornicroft et al., 2016; Bharadwaj et al., 2017; Corredor-Waldron and Currie, 2024). Digital technologies offer new solutions for reducing stigma in mental health support. In particular, telehealth and conversational artificial intelligence (AI) chatbots have great potential to mitigate stigma concerns by removing the need for in-person visits and minimizing the risk of face-to-face exposure (Organization et al., 2013; Jackson et al., 2014; Lucas et al., 2014; Stiefel, 2018; Palmer and Schwan, 2022; Blease and Torous, 2023; Erol and Wu, 2025).[2]

While AI offers promising solutions to mitigate stigma-related barriers in mental health support, it simultaneously raises a distinct source of privacy concern—namely, data security and confidentiality risks associated with sharing personal information with digital technologies. A growing body of research highlights the risks associated with AI in healthcare, including data leakage (Rahman, 2023; Chen and Esmaeilzadeh, 2024), unauthorized sharing with third parties (Wang et al., 2023), data misuse (Jin, 2018), and concerns over data retention (Tucker, 2018; Wang et al., 2024). These risks are particularly relevant in the context of AI-driven mental health services, where sensitive user information is collected and processed. Unlike traditional healthcare settings, AI chatbots often lack compliance with established data privacy frameworks, such as the Health Insurance Portability and Accountability Act (HIPAA), raising concerns about the security and confidentiality of personal health data (Marks and Haupt, 2023). Given

[1] https://www.nimh.nih.gov/health/statistics/mental-illness

[2] We collected all posts and comments from the r/ChatGPT subreddit on Reddit and observed a heated debate about using ChatGPT as a therapist. This debate highlights ChatGPT's effectiveness, its non-judgmental nature, and various data privacy concerns. As an example, see `https://www.reddit.com/r/ChatGPT/comments/1gloehd/will_chatgpt_replace_your_therapist_check_these/`. We also present excerpts from Reddit posts and Google Play reviews discussing mental health AI chatbots in Appendix M.

these challenges, individuals may be hesitant to engage with AI-based mental health services unless strong data privacy assurances are in place.

In this paper, we highlight a tension in AI-enabled mental health care seeking: while AI can reduce stigma-related privacy concerns, it can simultaneously exacerbate concerns about data security and confidentiality. Formally, we distinguish between (i) *interpersonal privacy concerns*, which arise in social and face-to-face interactions with other human beings, and (ii) *informational privacy concerns*, which stem from the collection, storage, and potential use of personal data by digital technologies. In particular, people prioritize interpersonal privacy concerns tied to stigma and social exposure over informational privacy risks when deciding whether to use conversational generative AI chatbots for mental health support; when those interpersonal privacy concerns are less salient, informational privacy risks tied to data security and confidentiality in digital technologies become more consequential.

Building on the above discussion, we aim to answer the main research questions as follows:

- RQ1. How do individuals evaluate telehealth and generative AI chatbots relative to in-person services for mental health support? And how does offering these digital channels alongside an in-person option (i.e., a multi-channel bundle) change willingness to engage?
- RQ2. To what extent do interpersonal privacy concerns shift preferences across mental health support modalities and multi-channel bundles?
- RQ3. To what extent do informational privacy concerns shift preferences toward generative AI chatbots?
- RQ4. Does there exist a trade-off between interpersonal and informational privacy concerns in individuals' decisions to seek mental health care via generative AI chatbots?

To answer our research questions, we first conducted a pre-registered online experiment on Prolific where each participant with moderate and severe mental health issues evaluated all three service options (in-person services, telehealth, and an AI chatbot) and provided individual rankings reflecting their relative willingness to employ each. The experiment features two treatment arms: a high interpersonal privacy

concerns condition (in which no corrective information is provided to shift participants' beliefs) and a low interpersonal privacy concerns condition (in which participants are provided with truthful information about the relatively low prevalence of stigma in the population). Hence, our manipulation mirrors the design implemented by Bursztyn et al. (2020). We then measure participants' baseline informational privacy concerns using the Internet Users' Information Privacy Concerns (IUIPC) scale (Malhotra et al., 2004). We report formal manipulation checks demonstrating that the manipulation of interpersonal privacy concerns operates as intended. To address additional confounds, we eliminate price and insurance considerations by stating that all services are free, and we elicit ex post perceptions of convenience, effectiveness, and data privacy for each service, which are included as controls. Our findings show that, first, AI is less attractive while telehealth is more attractive than in-person services, however, higher interpersonal privacy concerns make both AI and telehealth more attractive. Second, informational privacy concerns constrain AI's attractiveness only when interpersonal privacy concerns are low, but not when interpersonal privacy concerns are high.

We then conducted another online experiment on Prolific to investigate how offering these digital channels alongside an in-person option (i.e., a multi-channel bundle) change willingness to engage. In this experiment, participants were randomly assigned to two independent "treatments": one manipulated their stigma perceptions using an intervention of disclosing new public information, while the other presented them with different bundles of mental health service offerings, including combinations of offline services, telehealth services, and a conversational AI chatbot with varying policies emphasizing informational privacy assurances or personalization at the expenses of information privacy. Our findings show that, first, individuals are more likely to engage with mental health services when AI options were provided alongside with telehealth and in-person services, compared to in-person service alone, especially when interpersonal privacy concerns are higher. Second, AI personalization at the expense of informational privacy plays a bigger role than strong informational privacy safeguard when interpersonal privacy concerns are more pronounced.

To enhance the generalizability of our findings, especially the tradeoff between interpersonal and informational privacy concerns, we then conducted a field experiment in collaboration with a university

in China. In this field experiment, we sent messages to students encouraging them to seek mental health services. Similar to the second online experiment, this field experiment involved two independent interventions. First, we varied the framing of messages to influence students' stigma perceptions regarding mental health support. Second, these messages presented two bundles of mental health services that included all three modalities (in-person counseling, telehealth, and a conversational AI chatbot), differing only in how the AI privacy policies were framed—either emphasizing informational privacy assurances or personalization at the expense of information privacy. These messages included links to access the services, and we tracked whether students clicked on these links within a two-week period after delivery. The findings from this field experiment align closely with those from the two online experiments, reinforcing the interplay between interpersonal and informational privacy concerns in shaping engagement with mental health services.

This paper contributes to the literature in the following directions. First, it is one of the first studies to experimentally examine to what extent digital technologies encourage mental health services seeking, especially when interpersonal privacy concerns are pronounced. Second, while most existing literature focuses on technology adoption in healthcare from the provider's perspective (Jussupow et al., 2021; Agarwal et al., 2023; Wang et al., 2023), our study contributes and expands the scarce literature investigating individuals' decisions to adopt digital technologies, including telehealth and AI, from the consumer's perspective (Longoni et al., 2019; Cadario et al., 2021). Third, our paper extends the privacy and AI literature (Golda et al., 2024; Leschanowsky et al., 2024; Varghese et al., 2024) by experimentally examining how different sources of privacy concern, namely interpersonal and informational privacy concerns, shift and interact during generative AI adoption, in the context of a widely relevant and high-stakes application: generative AI chatbots for mental health services. Lastly, our study complements and extends prior human–computer interaction research showing greater willingness to disclose sensitive information to machines than to humans by focusing on a policy-relevant population with mental health needs, examining healthcare service-seeking decisions rather than general disclosure, and experimentally manipulating interpersonal privacy concerns to causally assess when technology-based options—particularly AI chatbots—become

more attractive [3].

The rest of the paper proceeds as follows. Section 2 describes literature review and theoretical foundations. Section 4 introduces the experimental design and data. Section 4.3 describes our experimental results. Section 5 expands on the main analysis, evaluating the generalizability of our results. Finally, Section 6 concludes and discusses implications.

# 2 Literature Review

## 2.1 Barriers and Solutions to Seeking Mental Health Care: Interpersonal Privacy Concerns and Digital Interventions

Stigmatization operates through heightened sensitivity to social exposure, judgment, and visibility. According to Goffman Goffman (1963), when stigmatizing attributes become perceptible in interaction, individuals experience this as an invasion of privacy, prompting strategies of concealment, masking, or selective disclosure. This view is consistent with Altman (1975)'s definition of privacy as the selective control of access to the self, which highlights privacy as an active, dynamic, and fundamentally social process rather than a purely informational one. We synthesize these perspectives to define interpersonal privacy concerns as concerns about being seen, known, judged, or socially evaluated by others in direct interaction.

Seeking in-person therapy often entails face-to-face disclosure to a human therapist, potential encounters with family members or acquaintances, and physical presence in clinics or waiting rooms, all of which heighten concerns about social exposure and stigma. Prior research suggests that such threats of social exposure can trigger strong and immediate reactions, particularly in the short term, because they are concrete, salient, and emotionally charged (Henderson et al., 2014; Thornicroft et al., 2016; Bharadwaj et al., 2017; Corredor-Waldron and Currie, 2024; Cronin et al., 2024). As a result, these interpersonal privacy concerns can loom especially large when individuals contemplate in-person mental health care.

[3] We include a comparative table (Table A1) that summarizes (i) the main findings of representative HCI studies on disclosure to computers, (ii) their methods and focal mechanisms, and (iii) how our study differs in terms of population, mechanism, outcome, and design.

Digital technologies offer new solutions for reducing stigma in mental health support. In particular, telehealth and conversational artificial intelligence (AI) chatbots have great potential to mitigate stigma concerns by removing the need for in-person visits and minimizing the risk of face-to-face exposure.

A large body of literature documents that teletherapy (i.e., the delivery of therapeutic services through digital technologies) can substantially reduce stigma associated with mental health treatment. By allowing individuals to receive care without physically attending clinics, telehealth eliminates visible actions such as scheduling appointments, waiting in public spaces, or being observed by others (Goh et al., 2021; Erol and Wu, 2025).

Recent rapid progress in large language models (LLMs) has made AI-based mental health support technically viable. LLM-based AI has demonstrated expert-level medical knowledge and empathetic communication (Ayers et al., 2023; Rao et al., 2023), in particular, LLM-enabled mental health tools are increasingly used for assessment, consultation, and therapy, with growing evidence that they can improve mental health symptoms (Fitzpatrick et al., 2017; Abd-Alrazaq et al., 2020; He et al., 2023). As a result, generative AI–based mental health chatbots are gaining widespread adoption and represent a promising, stigma-free alternative for individuals reluctant to seek traditional care, potentially extending the stigma-reducing benefits of telehealth by further reducing direct human interaction (Stiefel, 2018; Palmer and Schwan, 2022; Blease and Torous, 2023).

## 2.2 Barriers to using Generative AI in Mental Health Support: Informational Privacy Concerns

Informational privacy refers to individuals' claims to control when, how, and to what extent information about them is communicated to others (Westin, 1968). The literature highlights significant privacy concerns associated with the adoption of health information technologies (HIT) and AI in healthcare. For example, electronic medical records and health information exchanges have been shown to improve patient outcomes and enhance care coordination; at the same time, privacy concerns remain a significant barrier to their adoption (Miller and Tucker, 2011; Adjerid et al., 2016). In particular, generative AI introduces privacy risks, including data leakage (Rahman, 2023; Chen and Esmaeilzadeh, 2024), unauthorized sharing with

third parties (Wang et al., 2023), data misuse (Jin, 2018), and concerns over data retention (Tucker, 2018; Wang et al., 2024).

These risks are particularly concerning in mental health applications, where confidentiality is paramount (Varghese et al., 2024). Specifically, in the context of AI-based mental health chatbots, informational privacy concerns primarily relate to data collection, storage, algorithmic processing, and potential downstream uses such as personalization or targeted advertising, based on patients' mental heath data as well as conversation transcripts. While these concerns are consequential, prior work shows that they are often perceived as probabilistic, distant, and malleable (Acquisti et al., 2015).

## 2.3 Interpersonal Privacy Concerns Overshadow Informational Privacy Concerns

Related work underscores that privacy violations in social contexts are often experienced as visceral and immediate, involving mental suffering, judgment, and loss of dignity (Ernst and Schwartz, 1977; Link and Phelan, 2001; Pryor et al., 2004). On the other hand, literature in economics and information systems emphasizes that informational privacy concerns are often abstract, uncertain, and context-dependent, shaped by perceptions of data use, secondary use, and institutional safeguards (Malhotra et al., 2004; John et al., 2011; Acquisti et al., 2016; Buckman et al., 2019; Kummer and Schulte, 2019; Al-Natour et al., 2020; Xu and Zhang, 2022).

Bringing these two streams together, we now explicitly theorize the trade-off between interpersonal and informational privacy concerns in mental health care decisions. Drawing on research on risk perception and privacy calculus (Culnan and Bies, 2003; Dinev and Hart, 2006; Slovic, 2016), we argue that individuals weigh multiple, potentially competing privacy risks, and that some risks may overshadow others depending on their salience, immediacy, and perceived controllability. In particular, people prioritize interpersonal privacy concerns tied to stigma and social exposure over informational privacy risks when deciding whether to use conversational generative AI chatbots for mental health support; when those interpersonal privacy concerns are less salient, informational privacy risks tied to data security and confidentiality in digital technologies become more consequential.

# 3 Main Experiment: Online Ranking Experiment

To answer our research questions, we first conducted a pre-registered online experiment on Prolific involving 650 participants with moderate and severe mental health issues based on their PHQ-4 score [4]. Participants were informed that the survey would take no more than 10 minutes to complete and that they would be paid $2 for their time, translating to a wage of $12/hour. We undertake an online experiment at Prolific for two key reasons. First, effectively manipulating interpersonal privacy concerns requires participants to answer survey questions that elicit their private beliefs and their perceptions of others' beliefs, which necessitates their awareness of being part of a research study. Second, Prolific allows us to screen for a non-standard subject pool based on participants' mental health conditions, enhancing the credibility of their behaviors and responses. Participants' behavior in such an experiment—framed as an artefactual field experiment by Levitt and List (2007), a conventional lab experiment with a non-standard subject pool —can explain behavior in the field (Harrison and List, 2004).

The experiment features two treatment arms: a high interpersonal privacy concerns condition (in which no corrective information is provided to shift participants' beliefs) and a low interpersonal privacy concerns condition condition (in which participants are provided with truthful information about the relatively low prevalence of stigma in the population). Hence, our manipulation mirrors the design implemented by Bursztyn et al. (2020). Each participant evaluated all three service options (in-person services, telehealth, and an AI chatbot) and provided individual rankings reflecting their relative willingness to employ each (see Appendix Appendix D for full texts of the service offerings). We next elicit each participant's perceptions of multiple potentially relevant attributes of each service. We then measure participants' informational privacy concerns using the Internet Users' Information Privacy Concerns (IUIPC) scale (Malhotra et al., 2004) (See Appendix Appendix E for the full texts of the scale items). Finally, participants answer a short list of questions about their demographic characteristics, whether they have ever used or sought the services

[4]We initially recruited 1,500 participants. Although participants were pre-screened for having experienced mental health issues, the average PHQ-4 score was 5.17, corresponding to mild mental distress. It is motivated by prior literature showing that individuals with moderate to severe conditions are both more in need of care and more sensitive to stigma (DeRubeis et al., 2005; Cronin et al., 2024). Note that we were not able to prescreen based on PHQ-4 for a Prolific study, thus we initially recruited 1,500 participants and restricted the focal sample to those PHQ-4 $\geq$ 6 based on their survey answers, resulting the sample size for analysis being 650.

mentioned in the email, and their usage of AI tools on a monthly basis.

## 3.1 Manipulation of Interpersonal Privacy Concerns

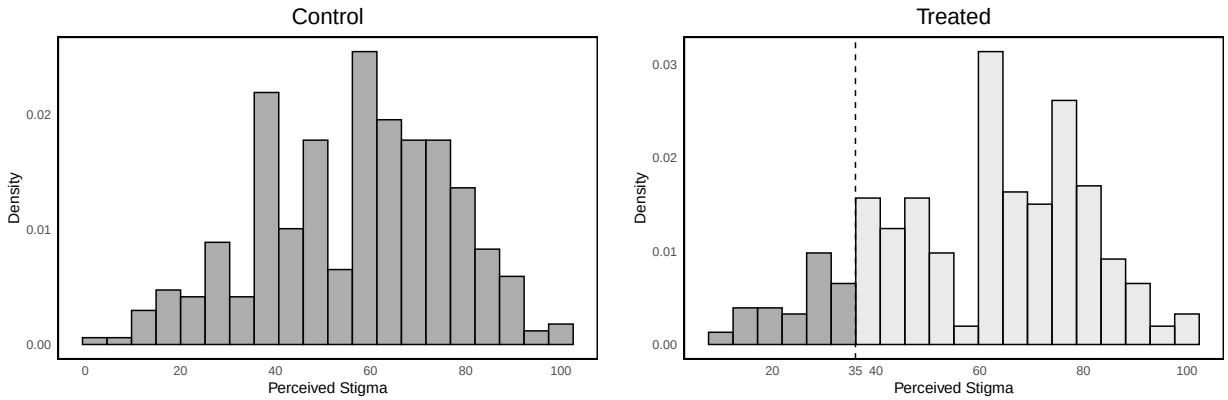
In society, individuals often update their perceptions based on new public information about what others think. When people receive aggregated information about private opinions, it can shift their perceptions of social norms and influence behavior (Bursztyn et al., 2020,?). In this study, we manipulate perceptions of stigma by providing participants with fabricated public information about how others perceive mental health.

To implement this manipulation, each participant is first asked to indicate the extent to which they agree with the following statement: *Others would view me negatively if they believed that I experienced a mental health issue.* They are then asked to estimate the percentage of other participants who agree with this statement. Afterwards, participants are randomly assigned to either a treatment or control condition. In the treatment condition, participants are shown their own response alongside a fabricated percentage—specifically, that 35% of others agree with the statement. If a participant's estimate differs from this fabricated statistic, they are informed of the discrepancy. For example, if a participant estimates that 60% of others agree with the statement, they are told that they have overestimated by 25%. In the control condition, participants receive no additional information after making their estimates.[5]

Figure 1 identifies the participants who are intended to be treated. Notably, the majority of participants in the treatment group report *ex ante* perceived stigma levels exceeding this stated level (the lighter-shaded bars in the right panel of Figure 1). Therefore, the information treatment *can be used* to correct misperceptions (Bursztyn et al., 2020): Treated participants who were informed that only 35% of others held stigmatizing views, and whose *ex ante* perceptions exceeded 35%, should (in theory) update their beliefs downward relative to control participants, who did not receive any information about others' opinions. For example, a participant in the treatment group who initially guesses that 60% of other participants hold stigmatizing views should in theory (Bursztyn et al., 2020) revise this belief downward upon learning that the true figure is 35%.

[5]See Appendix Appendix B for more details of this manipulation, and the full texts of questions asked in the survey.

Figure 1: Distribution of participants' perceived IPR in the control (high IPR) and treatment (low IPR) groups prior to treatment




*Notes:* We manipulate (perceived) interpersonal privacy concerns by providing new (fabricated) information about actual stigma only to participants in the treatment group. In the control (high interpersonal privacy concerns) condition, participants receive no information about actual stigma. In the treatment (low interpersonal privacy concerns) condition, participants are presented with (fabricated) information about actual stigma stating that 35% of other Prolific participants in this study hold stigmatizing views. Notably, the majority of participants in the treatment group report perceived stigma levels exceeding this stated level (the lighter-shaded bars).

## 3.2 Manipulation Checks

We conduct a manipulation check by examining (1) the treatment effect of manipulation of interpersonal privacy concerns on *ex post* perceptions of mental health stigma (elicited by a new question included in the new experiment asking the extent to which respondents believe in-person mental health services are free from stigma, with scores ranging from 1 to 100, see Appendix Appendix C for full texts of question asked in the survey), and (2) the treatment effect on preferences for in-person services (i.e., reversed ranking of in-person services, where 1 indicates the lowest preference and 3 indicates the highest preference; see columns (5) and (6) in Table 1). A negative and statistically significant coefficient on the experimentally high IPR (vs. low IPR) condition in both outcomes indicates that the manipulation worked as intended: respondents assigned to the high IPR condition are less likely to perceive in-person services as stigma-free and less likely to rank in-person services highly (note that rankings are reverse-coded, such that a lower number indicates a lower preference).

Table 1: Manipulation Checks

| DV: | convenience (1) | effectiveness (2) | privacy (3) | iuipc (4) | stigma (5) | r.ranking (6) |
|---|---|---|---|---|---|---|
| IPR | 0.529 | 1.265 | 1.128 | 0.073 | -4.309** | -0.097* |
| | (2.052) | (1.380) | (1.571) | (0.058) | (2.147) | (0.058) |
| convenience | | | | | 0.198*** | 0.009*** |
| | | | | | (0.043) | (0.001) |
| effectiveness | | | | | 0.264*** | 0.004* |
| | | | | | (0.076) | (0.002) |
| privacy | | | | | 0.239*** | 0.000 |
| | | | | | (0.065) | (0.002) |
| Constant | 47.040*** | 84.316*** | 84.034*** | 5.981*** | 64.012*** | 2.132*** |
| | (1.455) | (0.979) | (1.114) | (0.041) | (1.522) | (0.041) |
| Observations | 650 | 650 | 650 | 650 | 650 | 650 |
| $R^2$ | 0.0001 | 0.001 | 0.0008 | 0.002 | 0.142 | 0.119 |

*Note:* The dependent variables convenience, effectiveness, privacy and stigma measure respondents' perceptions about the extent to which in-person services are convenient, of high quality, of high data privacy and free from stigma, measured on a 1–100 scale (higher values indicate higher convenience, effectiveness, data privacy and lower perceived stigma). iuipc measures informational privacy concern using the Internet Users' Information Privacy Concerns (IUIPC) scale, which comprises three dimensions—collection, control, and awareness—and is constructed by averaging responses to 10 items measured on a 1 (strongly disagree)–7 (strongly agree) scale. r.ranking measures preference for in-person services using a reversed ranking, where 1 indicates the lowest preference and 3 indicates the highest preference. The independent variable IPR is an indicator equal to 1 for interpersonal privacy concerns and 0 otherwise. In columns (5) and (6), convenience, effectiveness, and privacy are demeaned to facilitate interpretation of the regression intercepts and to ensure consistency with subsequent regression tables. ***, ** and * represent statistical significance at the 1%, 5%, and 10% levels, respectively. IPR refers to the experimental arm with high interpersonal privacy concerns.

## 3.3 Econometric Specification

**How do individuals evaluate telehealth and AI chatbots relative to in-person services & the Effects of Interpersonal Privacy Concerns**

Each participant evaluated and provided rankings for all three service options (in-person, telehealth, and AI chatbot), and we transform the data from wide form to long form, where each participant ID has three observations corresponding to their rankings for each service.

Service-specific perceived attributes (e.g., perceived convenience, effectiveness, and informational privacy) are included as controls to account for incomparability across service options that may affect rankings (see Appendix Appendix C for full texts of question asked in the survey), thereby improving the precision of the estimated treatment effect. When the data are structured in long form, we include respondent fixed effects. Respondent-invariant characteristics are not included in these specifications because they are absorbed by respondent fixed effects.

Specifically, we estimate a regression of the form reflected by Equation 1,

$$reversed_rank_{ij} \;=\; \alpha_i \;+\; \beta_1\, 1\{j = \text{AI}\} \;+\; \beta_2\, 1\{j = \text{Online}\} \;+\; \mathbf{X}_{ij}^{\top}\boldsymbol{\gamma} \;+\; \varepsilon_{ij}, \tag{1}$$

where $i$ indexes respondents and $j$ indexes service modality. $reversed_rank_{ij}$ is the reversed ranking on the three-point scale, assigned by respondent $i$ to modality $j$, with the omitted (reference) modality being in-person ($j$ = In-person). $\alpha_i$ are respondent fixed effects. $\mathbf{X}_i$ are perceived attributes of each service as covariates: convenience, effectiveness and privacy.

To examine whether interpersonal privacy concerns systematically shift preferences across modalities, we first estimate Equation 1 separately for respondents with high and low IPR. We then estimate Equation 2, which interacts the modality indicators with IPR, to formally test whether these differences are statistically significant.

$$
\begin{aligned}
reversed_rank_{ij} = \alpha_i &+ \beta_1 \, 1\{j = \text{AI}\} + \beta_2 \, 1\{j = \text{Online}\} \\
&+ \theta_1 (IPR_i \times 1\{j = \text{AI}\}) + \theta_2 (IPR_i \times 1\{j = \text{Online}\}) \\
&+ \mathbf{X}_{ij}^{\top} \boldsymbol{\gamma} + \varepsilon_{ij}.
\end{aligned} \tag{2}
$$

where $\theta_1$ captures how interpersonal privacy concerns modify the relative ranking of AI (compared to in-person services), while $\theta_2$ captures how interpersonal privacy concerns modify the relative ranking of telehealth.

**The Effects of Informational Privacy Concerns Towards AI Chatbots**

To investigate the role of informational privacy concerns, we focus on preferences for AI chatbots. We begin by estimating Equation 3, which examines the direct association between informational privacy concerns (IUIPC) and the ranking of AI.

$$
reversed_rank_ai_i = \alpha + \beta_1 \, \text{IUIPC}_i + \mathbf{X}_i^{\top} \boldsymbol{\gamma} + \varepsilon_i, \tag{3}
$$

where $i$ indexes individuals, and $reversed_rank_ai_i$ denotes the reversed ranking of AI for individual $i$, where higher values indicate a stronger preference for AI. $\text{IUIPC}_i$ measures individual $i$'s *informational privacy concern*. $\mathbf{X}_i$ are covariates. $\beta_1$ captures how an increase in informational privacy concern shifts an individual's preference for AI, holding perceived attributes constant.

To test whether this informational privacy effect depends on interpersonal privacy concerns, we first estimate Equation 3 separately for respondents with high and low IPR. We then estimate Equation 4, which interacts IUIPC with IPR:

$$
reversed_rank_ai_i = \alpha + \beta_1 \, \text{IUIPC}_i + \beta_2 \, \text{IPR}_i + \beta_3 \big(\text{IUIPC}_i \times \text{IPR}_i\big) + \mathbf{X}_i^{\top} \boldsymbol{\gamma} + \varepsilon_i, \tag{4}
$$

where $\text{IPR}_i$ denotes *interpersonal privacy concerns*. $\beta_1$ represents the effect of informational privacy concerns when interpersonal privacy concerns are low (i.e., $\text{IPR}_i = 0$). The interaction coefficient $\beta_3$ captures how this effect changes when interpersonal privacy concerns are high.

Table F2 display descriptive statistics for several pre-treatment respondent characteristics as well as dependent variables.

## 3.4 Results

**How do individuals evaluate telehealth and AI chatbots relative to in-person services & the Effects of Interpersonal Privacy Concerns**

Table 2 reports results from the ranking experiment, where higher values of r.ranking indicate a stronger preference for in-person services. Column (1) shows that, relative to in-person services, AI chatbots receive significantly lower rankings ($\beta = -0.568$, $p < 0.01$), indicating that AI-based mental health support may not yet be viewed as a primary alternative. In contrast, telehealth (Online) receives significantly higher rankings than in-person services ($\beta = 0.318$, $p < 0.01$), suggesting that respondents generally view telehealth as more attractive than traditional face-to-face services.

Columns (2) and (3) examine heterogeneity by interpersonal privacy concerns (IPR). Although AI remains less preferred than in-person services in both subsamples, the magnitude of the relative preference gap is smaller in the high-IPR group (Column (2): $\beta = -0.479$) than in the low-IPR group (Column (3): $\beta = -0.651$). A similar pattern emerges for telehealth: the positive coefficient is larger under high IPR ($\beta = 0.384$) than under low IPR ($\beta = 0.255$).

Column (4) formally tests these differences using interaction terms. The positive and marginally significant interaction between AI and IPR ($\beta = 0.171$, $p < 0.10$) indicates that stronger interpersonal privacy concerns narrow the relative preference gap between AI and in-person services. Likewise, the interaction between Online and IPR is positive and significant ($\beta = 0.128$, $p < 0.10$), suggesting that stronger interpersonal privacy concerns further increase the relative attractiveness of telehealth relative to in-person services. Substantively, under high (vs. low) interpersonal privacy concerns, AI is ranked 0.171 positions higher, and telehealth is ranked 0.128 positions higher, both relative to in-person services. Given the baseline preference gap of 0.568 for AI relative to in-person services, the interaction effect of 0.171 reduces this gap by approximately 30%, indicating that interpersonal privacy concerns meaningfully narrow the relative difference in preferences between AI and in-person services. We also report the

rank-ordered logit results in Appendix Appendix H, which allow for a utility-based interpretation of the coefficients.

The coefficients on respondents' perceptions about the three attributes of each service also have meaningful implications. Focusing on column (5) of Table 2, within a respondent, a one SD increase in (normalized) convenience (mean: 0; SD: 26.07) leads to a $0.006 \times 26.07$ (0.16) unit increase in ranking positions on the three-point scale; a one SD increase in effectiveness (mean: 0; SD: 23.00) leads to a $0.018 \times 23.00$ (0.41) unit increase in ranking positions on the three-point scale; and a one SD increase in perceived informational privacy[6] (mean: 0; SD: 25.79) leads to a $0.005 \times 25.79$ (0.13) unit increase in ranking positions on the three-point scale—all statistically significant.

**The Effects of Informational Privacy Concerns Towards AI Chatbots**

Table 3 examines how informational privacy concerns (IUIPC) shape individuals' preference for generative AI chatbots. Column (1) shows that IUIPC is negatively and significantly associated with AI ranking ($\beta = -0.094$, $p < 0.01$). Because higher values of r.ranking indicate stronger preference for AI, this coefficient implies that individuals who are more concerned about information collection, control, and awareness systematically rank AI chatbots lower.

The magnitude is economically meaningful: a one SD increase in (normalized) IUIPC (mean: 0; SD: 0.74) leads to a $0.094 \times 0.74$ (0.07) unit decrease in ranking positions of AI on the three-point scale, even controlling for perceptions of other attributes of AI chatbots, such as convenience and effectiveness. This finding suggests that informational privacy concerns constitute a distinct barrier to AI adoption in mental health contexts, beyond perceptions of service quality or usability.

Columns (2) and (3) explore heterogeneity by interpersonal privacy concerns (IPR). The negative association between IUIPC and AI preference is substantially stronger in the low-IPR group (Column (3): $\beta = -0.156$, $p < 0.01$) than in the high-IPR group (Column (2): $\beta = -0.041$, $p > 0.1$). A one SD increase in (normalized) IUIPC (mean: 0; SD: 0.74) leads to a $0.156 \times 0.74$ (0.12) unit decrease in ranking positions of AI on the three-point scale, when interpersonal privacy concerns are lower. These findings are formally validated by the positive interaction term between $iuipc$ and $IPR$ in column (4). Substantively, this pattern

[6]Different from IUIPC, this is a perceived construct elicited by asking respondents the extent to which they believe a service guarantees data privacy.

Table 2: The Role of Interpersonal Privacy Concerns in Preferences for Individual Services (Online Ranking Experiment)

| Sample: | Full Sample | IPR (High) | IPR (Low) | Full Sample |
|---|---|---|---|---|
| DV: r.ranking | (1) | (2) | (3) | (4) |
| AI | -0.568*** | -0.479*** | -0.651*** | -0.654*** |
| | (0.044) | (0.063) | (0.062) | (0.062) |
| Online | 0.318*** | 0.384*** | 0.255*** | 0.254*** |
| | (0.039) | (0.054) | (0.055) | (0.055) |
| convenience | 0.006*** | 0.007*** | 0.005*** | 0.006*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| effectiveness | 0.018*** | 0.018*** | 0.019*** | 0.018*** |
| | (0.001) | (0.002) | (0.002) | (0.001) |
| privacy | 0.004*** | 0.006*** | 0.003* | 0.005*** |
| | (0.001) | (0.002) | (0.002) | (0.001) |
| AI × IPR | | | | 0.171* |
| | | | | (0.089) |
| Online × IPR | | | | 0.128* |
| | | | | (0.077) |
| # respondents | 650 | 327 | 323 | 650 |
| # Observations | 1,950 | 981 | 969 | 1,950 |
| $R^2$ | 0.483 | 0.481 | 0.492 | 0.485 |
| Respondent fixed effects | ✓ | ✓ | ✓ | ✓ |

*Note:* The reference group for AI and Online is the in-person service. The dependent variables for all columns are r.ranking, measuring preference for each service using a reversed ranking, where 1 indicates the lowest preference and 3 indicates the highest preference. The independent variable IPR refers to the experimental arm with high interpersonal privacy concerns, and is an indicator equal to 1 for interpersonal privacy concerns and 0 otherwise, and IPR is omitted when perfectly collinear with the respondent fixed effects. convenience, effectiveness and privacy measure respondents' perceptions about the extent to which each service is convenient, of high quality, of high data privacy and free from stigma, measured on a 1–100 scale (higher values indicate higher convenience, effectiveness, data privacy and lower perceived stigma). convenience, effectiveness, privacy are demeaned to facilitate interpretation of the regression intercepts. ***, ** and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

suggests a trade-off between two forms of privacy risk: when concerns about social exposure or stigma are salient (high IPR), individuals appear more willing to tolerate informational privacy risks associated with AI. Conversely, when interpersonal concerns are low, informational privacy concerns exert a stronger deterrent effect. Such a tradeoff is not observed in Table 4, where the dependent variable is the ranking of in-person services. In this setting, the IUIPC scale is not designed to capture informational privacy concerns related to seeking in-person services.

Table 3: The Role of Information Privacy Concerns in Preferences for AI Services (Online Ranking Experiment)

| Sample: | Full Sample | IPR (High) | IPR (Low) | Full Sample |
|---|---|---|---|---|
| DV: r.ranking of AI | (1) | (2) | (3) | (4) |
| iuipc | -0.094*** | -0.041 | -0.156*** | -0.154*** |
| | (0.034) | (0.048) | (0.047) | (0.048) |
| convenience | 0.002** | 0.002 | 0.002* | 0.002** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| effectiveness | 0.013*** | 0.013*** | 0.013*** | 0.013*** |
| | (0.001) | (0.002) | (0.002) | (0.001) |
| privacy | 0.002** | 0.002 | 0.002 | 0.002* |
| | (0.001) | (0.002) | (0.001) | (0.001) |
| IPR | | | | 0.053 |
| | | | | (0.049) |
| iuipc × IPR | | | | 0.111* |
| | | | | (0.066) |
| Constant | 1.515*** | 1.540*** | 1.486*** | 1.487*** |
| | (0.024) | (0.036) | (0.033) | (0.034) |
| # Observations / Respondents | 650 | 327 | 323 | 650 |
| $R^2$ | 0.383 | 0.368 | 0.405 | 0.386 |

*Note:* The dependent variables for all columns are r.ranking, measuring preference for AI chatbots using a reversed ranking, where 1 indicates the lowest preference and 3 indicates the highest preference. The independent variable IPR refers to the experimental arm with high interpersonal privacy concerns, and is an indicator equal to 1 for interpersonal privacy concerns and 0 otherwise. iuipc measures informational privacy concern using the Internet Users' Information Privacy Concerns (IUIPC) scale, which comprises three dimensions—collection, control, and awareness—and is constructed by averaging responses to 10 items measured on a 1 (strongly disagree)–7 (strongly agree) scale. convenience, effectiveness and privacy measure respondents' perceptions about the extent to which AI is convenient, of high quality, of high data privacy and free from stigma, measured on a 1–100 scale (higher values indicate higher convenience, effectiveness, data privacy and lower perceived stigma). convenience, effectiveness, privacy and iuipc are demeaned to facilitate interpretation of the coefficients. ***, ** and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

In Appendix Appendix G, we explore heterogeneity across the IUIPC dimensions—awareness (Ta-

Table 4: Falsification Test: The Role of Information Privacy Concerns in Preferences for In-person Services (Online Ranking Experiment)

| Sample: | Full Sample | IPR (High) | IPR (Low) | Full Sample |
|---|---|---|---|---|
| DV: r.rankings of in-person | (1) | (2) | (3) | (4) |
| iuipc | 0.170*** | 0.185*** | 0.158*** | 0.172*** |
| | (0.040) | (0.057) | (0.056) | (0.056) |
| convenience | 0.010*** | 0.010*** | 0.009*** | 0.010*** |
| | (0.001) | (0.002) | (0.002) | (0.001) |
| effectiveness | 0.003 | -0.001 | 0.006** | 0.003 |
| | (0.002) | (0.003) | (0.003) | (0.002) |
| privacy | -0.001 | 0.000 | -0.002 | -0.001 |
| | (0.002) | (0.003) | (0.002) | (0.002) |
| IPR | | | | -0.107* |
| | | | | (0.058) |
| iuipc × IPR | | | | 0.002 |
| | | | | (0.078) |
| Constant | 2.083*** | 2.031*** | 2.138*** | 2.137*** |
| | (0.029) | (0.041) | (0.040) | (0.041) |
| # Observations / Respondents | 650 | 327 | 323 | 650 |
| $R^2$ | 0.139 | 0.139 | 0.152 | 0.144 |

*Note:* The dependent variables for all columns are r.ranking, measuring preference for in-person services using a reversed ranking, where 1 indicates the lowest preference and 3 indicates the highest preference. The independent variable IPR refers to the experimental arm with high interpersonal privacy concerns, and is an indicator equal to 1 for interpersonal privacy concerns and 0 otherwise. iuipc measures informational privacy concern using the Internet Users' Information Privacy Concerns (IUIPC) scale, which comprises three dimensions—collection, control, and awareness—and is constructed by averaging responses to 10 items measured on a 1 (strongly disagree)–7 (strongly agree) scale. convenience, effectiveness and privacy measure respondents' perceptions about the extent to which an in-person service is convenient, of high quality, of high data privacy and free from stigma, measured on a 1–100 scale (higher values indicate higher convenience, effectiveness, data privacy and lower perceived stigma). convenience, effectiveness, privacy and iuipc are demeaned to facilitate interpretation of the coefficients. ***, ** and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

ble G3), control (Table G4), and collection (Table G5)—and find that the moderation effect is primarily driven by concerns about data collection. We also report the ordered logit results in Appendix Appendix H, which allow for a utility-based interpretation of the coefficients.

# 4 Complementarity: Online Bundled Experiment

We then conducted another online experiment on Prolific using a revised design. To complement the main experiment, this study serves three goals. First, it examines how offering digital channels alongside an in-person option (i.e., a multi-channel bundle) affects willingness to engage, and the extent to which interpersonal privacy concerns shift preferences across such bundles. Second, it experimentally manipulates informational privacy rather than measuring it. Third, the bundled design provides suggestive evidence on substitution and complementarity across service modalities.

This online experiment involving 438 participants. Consistent with the ranking experiments, these participants were screened as having moderate to severe mental health issues based on their PHQ-4 scores [7]. Participants were informed that the survey would take no more than 10 minutes to complete and that they would be paid $2 for their time, translating to a wage of $12/hour.

Participants were randomly assigned to two independent "treatments": one manipulated their stigma perceptions using an intervention of new public information[8], while the other presented them with different bundles of mental health service offerings, including combinations of offline services, telehealth services, and a conversational AI chatbot with varying privacy assurances [9]. We then ask participants about their intentions to seek any services offered in the bundles.

## 4.1 Experimental Design

**Mental health service offerings and Information Privacy Policies**

To examine how offering digital channels alongside an in-person option affects willingness to engage, and how informational privacy influences preferences for AI-involved bundles, we manipulate service

[7] We initially recruited 1,024 participants. Although participants were pre-screened for having experienced mental health issues, the average PHQ-4 score was 5.17, corresponding to mild mental distress. It is motivated by prior literature showing that individuals with moderate to severe conditions are both more in need of care and more sensitive to stigma (DeRubeis et al., 2005; Cronin et al., 2024). Note that we were not able to prescreen based on PHQ-4 for a Prolific study, thus we initially recruited 1,024 participants and restricted the focal sample to those PHQ-4 $\geq$ 6 based on their survey answers, resulting the sample size for analysis being 438.

[8] See Appendix Appendix I for the details of the manipulation, which follows the same approach as the ranking experiment, with a minor modification.

[9] A multi-level randomized experiment is widely implemented in the literature (Campos et al., 2017).

offerings as well as AI-related information privacy policies. Participants were instructed to imagine that they had recently experienced some mental health issues, and they were randomly assigned to receive one of five hypothetical email scenarios. These emails outlined various service options, including (see Appendix Appendix C for full texts of the manipulation of service offerings):

1. *In-person services* only (*in-person* bundle)
2. Virtual (telehealth) consultations in addition to in-person services (*online* bundle)
3. In-person and telehealth services, plus a conversational AI chatbot with no privacy information provided (AI NoP)
4. In-person and telehealth services, plus a conversational AI chatbot with privacy policies emphasizing informational privacy assurances (*AI Privacy* bundle)
5. In-person and telehealth services, plus a conversational AI chatbot with privacy policies emphasizing personalization at the expenses of information privacy (*AI Personalization* bundle)

Importantly, real-world AI privacy policy positioning typically reflects a trade-off between personalization and data protection. Accordingly, the AI Privacy and AI Personalization manipulations are designed to capture this practical tension rather than convey a purely abstract informational privacy signal. In practice, emphasizing data collection or reduced information privacy is usually accompanied by an explanation of its purpose—most commonly, enabling improved AI personalization [10].

After reading the hypothetical emails, participants are then asked about their intention to use any of the aforementioned services in the email they just read. Finally, participants answer a short list of questions about their demographic characteristics, whether they have ever used or sought the services mentioned in the email, and their usage of AI tools on a monthly basis.

[10] See, for example, `https://openai.com/policies/row-privacy-policy/`, Section 2, "How we use Personal Data".

## 4.2 Econometric Specification

For the empirical analysis, the focal sample of $N = 438$ participants is divided into two subsamples corresponding to the two distinct research objectives.

First, to estimate the effects of offering digital channels alongside an in-person option and the moderating role of interpersonal privacy concerns, we focus on participants assigned to the *in-person*, *online*, and *AI NoP* bundles ($N = 237$). This subsample isolates variation in service modality while holding AI informational privacy framing absent (i.e., no additional data privacy information is provided), allowing us to identify how adding telehealth and/or AI channels shifts willingness to engage relative to in-person services alone.

Second, to examine the causal effects of AI informational privacy policies, we restrict attention to participants assigned to the *AI Privacy* and *AI Personalization* bundles ($N = 165$). This subsample holds the service modality bundle fixed (in-person + telehealth + AI) and varies only the informational privacy framing of the AI chatbot, enabling a clean test of how privacy assurances versus personalization-oriented policies affect engagement intentions and how these effects interact with interpersonal privacy concerns.

**The Effects of Offering Digital Channels and the Role of Interpersonal Privacy Concerns**

We first examine whether adding digital channels (telehealth or AI) to an in-person option shifts participants' willingness to engage. To do so, we estimate Equation 5, which compares different service bundles relative to the in-person-only baseline:

$$\text{Intent_to_Use}_i = \alpha + \sum_{g=2}^{3} \alpha_g \cdot \mathbf{1}(\text{Email_group}_i = g) + X_i\gamma + \epsilon_i, \tag{5}$$

where $\text{Intent_to_Use}_i$ is a binary indicator equal to 1 if respondent $i$ reports being likely to use any service in the bundle. The indicators $\mathbf{1}(\text{Email_group}_i = g)$ capture assignment to the online bundle or the AI NoP bundle, with the in-person-only bundle as the omitted reference category. Thus, each $\alpha_g$ measures how adding telehealth or AI changes engagement relative to in-person service alone, holding covariates constant.

To examine whether these effects differ under high and low interpersonal privacy concerns, we first

estimate Equation 5 separately for respondents with high and low IPR. We then estimate Equation 6, which interacts the bundle indicators with IPR:

$$\begin{aligned}\text{Intent_to_Use}_i = \alpha + \sum_{g=2}^{3} \alpha_g \cdot \mathbf{1}(\text{Email_group}_i = g) + \delta \cdot \text{IPR}_i \\ + \sum_{g=2}^{3} \theta_g \Big[\mathbf{1}(\text{Email_group}_i = g) \times \text{IPR}_i\Big] + X_i\gamma + \epsilon_i.\end{aligned} \tag{6}$$

where $\theta_g$ captures how interpersonal privacy concerns modify the incremental effect of each digital bundle relative to in-person services.

**The Effects of AI Informational Privacy Policies**

To isolate the effect of AI informational privacy framing, we restrict the sample to respondents assigned to either the AI Privacy bundle or the AI Personalization bundle. We estimate:

$$\text{Intent_to_Use}_i = \alpha + \beta \cdot \mathbf{1}(\text{Email_group}_i = \text{AI Personalization}) + X_i\gamma + \epsilon_i, \tag{7}$$

where the omitted category is the AI Privacy bundle. Here, $\beta$ measures how emphasizing personalization at the expense of informational privacy changes willingness to engage relative to emphasizing informational privacy assurances.

To test whether this informational privacy effect depends on interpersonal privacy concerns, we first estimate Equation 7 separately for respondents with high and low IPR. We then estimate Equation 8, which interacts the information privacy policy indicator with IPR:

$$\begin{aligned}\text{Intent_to_Use}_i = \alpha + \beta \cdot \mathbf{1}(\text{Email_group}_i = \text{AI Personalization}) + \delta \cdot \text{IPR}_i \\ + \theta\Big[\mathbf{1}(\text{Email_group}_i = \text{AI Personalization}) \times \text{IPR}_i\Big] + X_i\gamma + \epsilon_i,\end{aligned} \tag{8}$$

where $\theta$ captures how interpersonal privacy concerns moderate the impact of AI informational privacy framing.

Table J9 in Appendix display descriptive statistics for several pre-treatment covariates as well as dependent variables by experiment arms.

## 4.3 Results

**The Effects of offering digital channels alongside an in-person option & the Effects of Interpersonal Privacy Concerns**

Table 5 examines how individuals evaluate multi-channel bundles relative to in-person services alone. Column (1) shows that adding an online channel to in-person service increases intention to use (ITU) by 9.8 percentage points, although the estimate is not statistically significant. In contrast, adding both telehealth and an AI chatbot significantly increases ITU by 14.4 percentage points ($p < 0.05$). These results suggest that incorporating generative AI into the service offering meaningfully enhances engagement intentions.

Columns (2) and (3) explore heterogeneity by interpersonal privacy concerns (IPR). Among participants in the high-IPR condition (Column (2)), the AI bundle increases ITU by 17.0 percentage points (marginally significant), whereas the effect is substantially smaller and not statistically significant in the low-IPR group (Column (3)). The telehealth-only bundle follows a similar directional pattern, though estimates remain imprecise.

Column (4) directly interacts IPR with bundle offerings. The coefficient on (+Online + AI NoP) $\times$ IPR is positive (0.175), indicating that higher interpersonal privacy concerns amplify the relative attractiveness of AI-involved bundles. However, the interaction term is not statistically significant. Given the moderate sample size ($N = 273$), this imprecision is likely due to limited statistical power. Table 6 examines the interaction between bundle offerings and IPR across subsamples defined by different PHQ-4 thresholds. While the interaction term between the AI bundle and IPR is positive but statistically insignificant in the broader (i.e., less distress-targeted) subsamples (e.g., PHQ-4 $\geq 5$ and PHQ-4 $\geq 6$), it becomes statistically significant when we restrict attention to respondents with more severe mental distress (PHQ-4 $\geq 7$). Notably, for the PHQ-4 $\geq 8$ subsample, the $p$-value of the interaction term is 0.110, indicating borderline statistical significance, given the smaller sample size ($N = 164$).

Substantively, this pattern suggests that when individuals are more concerned about interpersonal

Table 5: The Role of Interpersonal Privacy Concerns in Preferences for Bundled Services (Online Bundled Experiment)

| Ref. Group: In-person<br>DV: ITU Any Services | Full Sample<br>(1) | IPR (High)<br>(2) | IPR (Low)<br>(3) | Full Sample<br>(4) |
|---|---|---|---|---|
| + Online | 0.098 | 0.146 | 0.096 | 0.075 |
| | (0.066) | (0.097) | (0.098) | (0.091) |
| + Online + AI NoP | 0.144** | 0.170* | 0.057 | 0.058 |
| | (0.065) | (0.095) | (0.097) | (0.091) |
| IPR | | | | -0.022 |
| | | | | (0.093) |
| (+ Online) * IPR | | | | 0.051 |
| | | | | (0.134) |
| (+ Online + AI NoP) * IPR | | | | 0.175 |
| | | | | (0.131) |
| Constant | 0.177 | -0.038 | 0.482 | 0.209 |
| | (0.287) | (0.429) | (0.415) | (0.294) |
| Control Variables | ✓ | ✓ | ✓ | ✓ |
| # Observations | 273 | 132 | 141 | 273 |
| $R^2$ | 0.173 | 0.284 | 0.216 | 0.182 |

*Note:* The independent variable IPR refers to the experimental arm with high interpersonal privacy concerns, and is an indicator equal to 1 for interpersonal privacy concerns and 0 otherwise. Standard errors in parentheses. ***, ** and * represent statistical significance at the 1%, 5%, and 10% levels, respectively. Stigma beliefs, perceptions of others, mental status, previous mental health services utilization, previous AI utilization, and other demographic characteristics (age, gender, education, household income, race, marital status, occupation) are incorporated as control variables. NoP stands for no AI privacy information.

privacy concerns, they respond more positively to combinations of modalities potentially reduce social exposure and stigma. Notably, in Table 2, where we compare preferences for each individual service modality, AI is ranked lower than in-person services, and this preference gap is mitigated when interpersonal privacy concerns are high. To complement this finding, here we shift the focus from relative rankings of standalone services to overall willingness to engage when digital channels are offered alongside in-person services in a multi-channel bundle. We find that adding digital options—particularly bundles that include an AI chatbot—increases overall willingness to engage with mental health support, especially when IPR is high. Thus, although AI is not necessarily preferred as a standalone modality, its inclusion as an additional channel expands engagement in bundled settings, with the effect being stronger among individuals concerned about interpersonal privacy.

Table 6: Interaction Terms (Online Bundled Experiment)

| Ref. Group: In-person<br>DV: ITU Any Services | PHQ-4 ≥ 5<br>(1) | PHQ-4 ≥ 6<br>(2) | PHQ-4 ≥ 7<br>(3) | PHQ-4 ≥ 8<br>(4) |
|---|---|---|---|---|
| + Online | 0.073 | 0.075 | 0.165* | 0.223* |
| | (0.085) | (0.091) | (0.098) | (0.116) |
| + Online + AI NoP | 0.084 | 0.058 | 0.023 | 0.018 |
| | (0.083) | (0.091) | (0.098) | (0.115) |
| IPR | -0.081 | -0.022 | -0.069 | 0.013 |
| | (0.086) | (0.093) | (0.108) | (0.122) |
| (+ Online) × IPR | 0.165 | 0.051 | 0.009 | -0.072 |
| | (0.120) | (0.134) | (0.154) | (0.184) |
| (+ Online + AI NoP) × IPR | 0.170 | 0.175 | 0.329** | 0.278 |
| | (0.121) | (0.131) | (0.148) | (0.172) |
| Constant | 0.354 | 0.209 | 0.002 | 0.044 |
| | (0.257) | (0.294) | (0.334) | (0.414) |
| p-value of (+ Online + AI NoP) × IPR | 0.159 | 0.183 | 0.028 | 0.110 |
| Control Variables | Yes | Yes | Yes | Yes |
| # Observations | 335 | 273 | 210 | 164 |
| $R^2$ | 0.153 | 0.182 | 0.270 | 0.258 |

*Note:* We test the significance of the interaction term between service treatments with stigma treatment across different levels of PHQ-4. Standard errors in parentheses. ***, ** and * represent statistical significance at the 1%, 5%, and 10% levels, respectively. Stigma beliefs, perceptions of others, mental status, previous mental health services utilization, previous AI utilization, and other demographic characteristics (age, gender, education, household income, race, marital status, occupation) are incorporated as control variables. NoP stands for no AI privacy information. IPR refers to the experimental arm with high interpersonal privacy concerns.

The bundled design also allows us to provide suggestive evidence on substitution and complementarity

across service modalities. We observe that adding a telehealth as an option crowds out the preference for in-person services, but further adding AI does not. Moreover, adding AI as an option does not crowd out the preference for telehealth. More intuitively, increased engagement with the online bundle relative to in-person service (though not significant, but marginal) can be attributed to switching from in-person to telehealth services; increased engagement with the AI bundle relative to in-person service can not be suggestively explained by the crowding-out, but istead, the inclusion of AI chatbots indicates a complementary relationship between other modalities [11]. This substitution-complementarity analysis requires measuring intentions to use each service individually, which we also collected but were not used in the main analysis, and we report the regression results on substitution and complementarity patterns in the Appendix K, as this is not the main focus of this paper.

The bundled design also allows us to provide suggestive evidence on substitution and complementarity across service modalities. We observe that adding telehealth as an option tends to crowd out preference for in-person services, whereas further adding AI does not generate additional crowding-out of in-person services. Moreover, introducing AI does not crowd out telehealth intended usage. More intuitively, the (marginally) higher engagement with the online bundle relative to the in-person-only bundle appears to reflect substitution from in-person to telehealth. In contrast, the increased engagement observed in the AI bundle relative to in-person services cannot be fully explained by substitution across channels. Instead, the inclusion of AI chatbots suggests a complementary relationship with other modalities, expanding overall engagement rather than merely reallocating it across channels.[12] This substitution–complementarity analysis relies on participants' reported intentions to use each service individually, which we collected but do not use in the main specifications. For completeness, we report the corresponding regression results in Appendix Appendix K, as this mechanism analysis is not the central focus of the paper.

**The Effects of AI Informational Privacy Policies**

Table 7 examines whether emphasizing personalization at the expense of informational privacy shifts ITU for bundled services that include AI chatbots. The reference category is the AI Privacy condition,

[11] We also find anecdotal evidence of complementarity (rather than substitution) between AI chatbots and human-based mental health services in Reddit discussions. See Appendix M.

[12] We also find anecdotal evidence of complementarity (rather than substitution) between AI chatbots and human-based mental health services in Reddit discussions. See Appendix Appendix M.

which highlights stronger informational privacy assurances.

Column (1) shows that, in the full sample, compared to the AI Privacy framing, the AI Personalization framing reduces ITU by 1.4 percentage points, although the estimate is not statistically significant. Columns (2) and (3) reveal substantial heterogeneity by interpersonal privacy concerns (IPR). Among individuals in the low-IPR group (Column (3)), emphasizing personalization at the expense of informational privacy significantly reduces ITU ($\beta = -0.152$, $p < 0.10$). In contrast, among individuals in the high-IPR group (Column (2)), the coefficient becomes positive ($\beta = 0.149$, $p > 0.10$), and it is not statistically significant.

Column (4) formally tests this moderation effect. The interaction between AI Personalization and IPR is positive and statistically significant ($\beta = 0.233$, $p < 0.10$), indicating that, when interpersonal privacy concerns are high, the negative effect of reduced informational privacy protection is partially offset by the perceived benefits of improved personalization. This trade-off between interpersonal and informational privacy concerns is consistent with the ranking experiment.

# 5 Generalizability: Field Experiment

To assess the generalizability of our findings on informational privacy and its interaction with interpersonal privacy concerns from our ranking and bundled online experiments conducted via Prolific,[13] we partnered with a university in China to conduct a randomized field experiment. We employed a cluster randomization, where we conducted the randomization at the student cohort level (i.e., "class" in Chinese universities). Within each cohort, all students were assigned to the same experimental condition. This approach has two key advantages: first, it minimizes contamination between experimental arms [14]. Second, students were not required to identify themselves, respond to any questions, or explicitly consent to participate in the experiment. Instead, their participation was tracked using student cohort IDs. This approach strengthens the external validity of the experiment. Additionally, the treatment units (student cohorts) are stratified by school (college), with randomization conducted within each college to increase statistical power. To

[13] As discussed below, statistical power constraints prevented us from including an AI NoP arm in our initial field experiment design; consequently, we cannot replicate the first two research questions in the field experiment.

[14] A student level randomization is not desirable in this experiment in that students within the same student cohort might share information about the messages they received from the university.

Table 7: The Role of Information Privacy Concerns in Preferences for Bundled Services (Online Bundled Experiment)

| Ref. Group: AI Privacy<br>DV: ITU Any Services | Full Sample<br>(1) | IPR (High)<br>(2) | IPR (Low)<br>(3) | Full Sample<br>(4) |
|---|---|---|---|---|
| AI Personalization | -0.014 | 0.149 | -0.152* | -0.111 |
| | (0.057) | (0.114) | (0.078) | (0.076) |
| IPR | | | | -0.096 |
| | | | | (0.084) |
| AI Personalization × IPR | | | | 0.233* |
| | | | | (0.118) |
| Constant | 0.535* | 0.829 | -0.099 | 0.628** |
| | (0.302) | (0.575) | (0.482) | (0.305) |
| Control Variables | ✓ | ✓ | ✓ | ✓ |
| # Observations | 165 | 71 | 94 | 165 |
| $R^2$ | 0.241 | 0.272 | 0.365 | 0.262 |

*Note: AI Personalization* includes three services: In-person and telehealth services, plus a conversational AI chatbot emphasizing informational privacy assurances. *AI Privacy* includes three services: In-person and telehealth services, plus a conversational AI chatbot emphasizing personalization at the expenses of information privacy (see appendix for full texts of privacy policy manipulation). The independent variable IPR refers to the experimental arm with high interpersonal privacy concerns, and is an indicator equal to 1 for interpersonal privacy concerns and 0 otherwise. Standard errors in parentheses. ***, ** and * represent statistical significance at the 1%, 5%, and 10% levels, respectively. Stigma beliefs, perceptions of others, mental status, previous mental health services utilization, previous AI utilization, and other demographic characteristics (age, gender, education, household income, race, marital status, occupation) are incorporated as control variables.

encourage greater engagement, the experiment was conducted two weeks before midterm exam week—a period when students typically experience heightened academic pressure.

However, this design has notable limitations. First, the field experiment targeted a different population, regular college students, whereas the online experiment focused on individuals who had reported experiencing mental health challenges. Unlike the latter group, college students may be less responsive to information about mental health resources, even when messages are delivered at a potentially high-stress period. Second, student cohort-level randomization yielded only 333 observations, which limited our statistical power. Third, due to privacy constraints and our collaborator's requirements, we could only track click-through rates (CTR) rather than actual service usage. Consequently, nearly all participants clicked on only one link [15], even when multiple options were provided. We suspect that once a participant clicked on a link, they focused on the subsequent website content (e.g., scheduling a telehealth visit or learning about the AI chatbot) and did not return to explore other services. A related limitation is that, considerations of statistical power prevented us from including an AI NoP arm in our initial field experiment design. As a result, we are unable to replicate the first part of the bundled experiment in the field—namely, examining how offering multi-channel bundles affects willingness to engage and how interpersonal privacy concerns shift preferences across such bundles. Therefore, our final field analysis focuses on the two arms manipulating informational privacy policies and on replicating the effects of informational privacy policies and their interaction with interpersonal privacy concerns, both of which are experimentally manipulated in the field setting. Fourth, for the same reason as mentioned above, the manipulation of interpersonal privacy concerns was conducted differently from the online experiment. Instead of asking students to express their own stigma beliefs, perceptions of other students' beliefs, or estimates of others' stigma perceptions, they were presented with messages that framed mental health differently based on the treatment condition—one suggesting potential social judgment and the other conveying an environment free from such judgment.

[15] Except for one individual who clicked on two service links.

## 5.1 Experimental Design

In this field experiment, we sent messages to students encouraging them to seek mental health services by offering different bundles of services, similar to those mentioned in the online experiment. These messages nudged students with various interpersonal privacy concerns associated with seeking help and included links for accessing the services. We tracked whether students clicked on the provided links and calculated the student cohort-level click-through rates within a two-week period following the message delivery.

In this field experiment, we sent messages to students encouraging them to seek mental health services. Similar to the second online experiment, this field experiment involved two independent interventions. First, we varied the framing of messages to influence students' stigma perceptions regarding mental health support. Second, these messages presented two bundles of mental health services that included all three modalities (in-person counseling, telehealth, and a conversational AI chatbot), differing only in how the AI privacy policies were framed—either emphasizing informational privacy assurances or personalization at the expense of information privacy. These messages included links to access the services, and we tracked whether students clicked on these links within a two-week period after delivery

First, we varied the framing of messages to influence students' stigma perceptions regarding mental health support. Second, we presented different bundles of mental health services, including combinations of offline counseling, telehealth, and a conversational AI chatbot with varying levels of privacy assurances. These messages included links to access the services, and we tracked whether students clicked on these links within a two-week period after delivery.

The focal university is located in one of the biggest cities in southwest China and has over 26,000 undergraduate students in 16 schools. Students chose their majors in the college entrance examination, and students in the same major consist of one to three student cohorts, where each student cohort elected one cohort leader when they matriculated, totaling up to 790 student cohorts, with most of the student cohorts having between 30 and 40 students.

All student cohort leaders received a message from the Student Union through an official QQ group chat two weeks prior to midterm exam week. QQ is a popular platform in China similar to WhatsApp,

and a QQ group chat can accommodate between 500 and 5,000 participants. The message first instructed student cohort leaders to retrieve their student cohort names from a designated link. Once they completed this step, they were shown a message from the university encouraging students to utilize available mental health support services and were asked to copy the message and post it in the QQ group chat notice, where each student in that student cohort would be notified.

Each student cohort was randomly assigned into four groups ($2 \times 2$)[16]. At the top of each message, a general description is provided on how students can utilize different options for mental health support. This message is framed differently depending on the treatment condition (high vs. low interpersonal privacy concerns). For the high interpersonal privacy concerns group, the message emphasizes concerns about social judgment, highlighting that many students may feel reluctant to seek mental health support due to fear of being viewed negatively by others. It reinforces the idea that discussing mental health issues is uncommon and that stigma around mental health persists in the student community. For the low interpersonal privacy concerns group, the message focuses on normalizing mental health care, framing it as a common and accepted practice. It emphasizes that many students are using these resources and that there is no social judgment attached to seeking help. By presenting contrasting public perceptions of stigma, the messages manipulates students' social views on mental health.

Next, the message is displayed in two different ways emphasizing information privacy or personalization at the expense of information privacy, similar to what we have for the experiment on Prolific:

In-person, telehealth services plus AI chatbot with privacy policies emphasizing informational privacy assurances (*AI Privacy* bundle) —This message presents students with the option to access support via in-person or telehealth services, as well as a conversational AI chatbot. The chatbot, a leading mental health AI in China, is available for free. In this AI Privacy group, students are informed that their data will be stored with strong privacy protections, including end-to-end encryption, and that their information will remain confidential and not be shared with any third parties. The message includes a link to access the chatbot and make appointments for other services.

[16]We initially designed four experimental groups manipulating service offerings. However, due to the limitations discussed above, we restricted the field experiment to service arms offering all three services—in-person services, telehealth, and AI chatbots—with privacy policies emphasizing either informational privacy assurances (AI Privacy bundle) or personalization at the expense of informational privacy (AI Personalization bundle).

In-person, telehealth services plus AI chatbot with privacy policies emphasizing personalization at the expense of informational privacy (*AI Personalization* bundle) —This version also offers students the same three service options. However, students are informed that while the chatbot is free to use, the company may collect and store their data with fewer privacy protections. They are told that some data might be shared with third parties for research or service improvement purposes, such as training their generative AI model. The message includes a link for accessing services and making appointments.

Each channel is presented with clear instructions for accessing the services or making appointments, ensuring that students understand their options. The full content of the messages is provided in the D1.

## 5.2 Econometric Specification

**The Effects of Informational Privacy Concerns Towards AI Chatbots**

**Replication for Ranking Experiment: The Effects of AI Informational Privacy Policies**

In the field experiment, the messages to students presented two bundles of mental health services that included all three modalities (in-person counseling, telehealth, and a conversational AI chatbot), differing only in how the AI privacy policies were framed (AI Privacy and AI Personalization). These messages included links to access the services, and we tracked whether students clicked on each individual link when all three were shown. As noted above, nearly all participants clicked on only one link, even when multiple options were provided. As a result, the field experiment effectively mimics the main ranking experiment, in which individuals ranked individual services. Therefore, we estimate a regression of the form reflected by Equation 9 and 10,

$$\text{CTR_AI_link}_{cs} = \alpha + \beta \cdot \text{AI_Personalization}_c + \lambda_s + \epsilon_{cs}, \tag{9}$$

where $c$ indexes student cohorts and $\lambda_s$ denotes college fixed effects. $CTR_AI_link_{cs}$ denotes the student cohort–level click-through rate on the AI chatbot link, calculated as the number of students who clicked on the AI link divided by cohort size. $AI_Personalization_c$ is an indicator for assignment to the AI Personalization bundle, with the AI Privacy bundle serving as the reference group. Standard errors are

clustered at the college level.

To examine whether these effects differ under high and low interpersonal privacy concerns, we first estimate Equation 9 separately for respondents with high and low IPR. We then estimate Equation 10, which interacts the information privacy policy indicator with IPR:

$$\text{CTR_AI_link}_{cs} = \alpha + \beta \cdot \text{AI_Personalization}_c + \delta \cdot \text{IPR}_c + \theta \left(\text{AI_Personalization}_c \times \text{IPR}_c\right) + \lambda_s + \epsilon_{cs}, \quad (10)$$

where $IPR_c$ denotes interpersonal privacy concerns.

**Replication for Bundled Experiment: The Effects of AI Informational Privacy Policies**

To replicate the experiment design of the online bundled experiment, we construct a dependent variable, $CTR_any_link_{cs}$, denoting the student cohort–level click-through rate on any links presented in the message, calculated as the number of students who clicked on any service link divided by cohort size. Therefore, we estimate a regression of the form reflected by Equation 11 and 12,

$$\text{CTR_any_link}_{cs} = \alpha + \beta\, \text{AI_Personalization}_c + \lambda_s + \epsilon_{cs}, \quad (11)$$

where $c$ indexes student cohorts and $\lambda_s$ denotes college fixed effects. $CTR_any_link_{cs}$ denotes the student cohort–level click-through rate, calculated as the number of students who clicked on any service link divided by cohort size. $AI_Personalization_c$ is an indicator for assignment to the AI Personalization bundle, with the AI Privacy bundle serving as the reference group. Standard errors are clustered at the college level.

To examine whether these effects differ under high and low interpersonal privacy concerns, we first estimate Equation 11 separately for respondents with high and low IPR. We then estimate Equation 12, which interacts the information privacy policy indicator with IPR:

$$\text{CTR_any_link}_{cs} = \alpha + \beta\, \text{AI_Personalization}_c + \delta\, \text{IPR}_c + \theta \left(\text{AI_Personalization}_c \times \text{IPR}_c\right) + \lambda_s + \epsilon_{cs}, \quad (12)$$

where $IPR_c$ denotes interpersonal privacy concerns.

Section K1 in Appendix display descriptive statistics for several pre-treatment covariates as well as dependent variables by experiment arms.

## 5.3 Results

**Generalizability for Ranking Experiment: The Effects of AI Informational Privacy Policies**

Table 8 examines to what extent informational privacy affects engagement with AI services in the field setting. The omitted (reference) category is the AI Privacy condition, which emphasizes informational privacy assurances; coefficients therefore capture the effect of switching to the AI Personalization condition, which emphasizes personalization while downplaying privacy assurances. Column (1) shows that switching from AI Privacy to an AI Personalization framing reduces the click-through rate (CTR) to AI services by 0.1 percentage points, although the estimate is not statistically significant.

Columns (2) and (3) explore heterogeneity by interpersonal privacy concerns (IPR). The negative effect of the AI Personalization framing is concentrated in the low-IPR subsample (Column (3): $-0.004$, $p < 0.10$), while it is economically smaller and statistically insignificant in the high-IPR group (Column (2): 0.002, $p > 0.10$)).

Column (4) formally tests this moderation effect. The interaction term between AI Personalization and IPR is positive and marginally significant ($\beta = 0.006$, $p < 0.10$), indicating that the negative impact of weakened informational privacy assurances is attenuated when interpersonal privacy concerns are high. In other words, students who are more concerned about social exposure or stigma are less deterred by reductions in informational privacy protection. Given that the baseline CTR is approximately 0.5%, this coefficient represents a non-trivial proportional change. As a falsification test, such a tradeoff is not observed in Table 9, where the dependent variable is the CTR to in-person services.

The evidence from the field is consistent with the ranking experiment: when interpersonal privacy concerns are low, students are more sensitive to informational privacy framing, and emphasizing personalization at the expense of data protection reduces AI engagement. However, when interpersonal privacy concerns are high, the deterrent effect of reduced informational privacy assurances weakens, suggesting

that concerns about social judgment partially offset concerns about data security.

Table 8: The Role of Information Privacy Concerns in Preferences for AI Services (Field Experiment)

| Ref. Group: AI Privacy<br>DV: CTR AI Services | Full Sample<br>(1) | IPR (High)<br>(2) | IPR (Low)<br>(3) | Full Sample<br>(4) |
|---|---|---|---|---|
| AI Personalization | -0.001 | 0.002 | -0.004* | -0.004* |
| | (0.002) | (0.003) | (0.002) | (0.002) |
| IPR | | | | -0.000 |
| | | | | (0.002) |
| AI Personalization × IPR | | | | 0.006* |
| | | | | (0.003) |
| Constant | 0.005*** | 0.005*** | 0.005*** | 0.005*** |
| | (0.001) | (0.001) | (0.001) | (0.002) |
| Control Variables | ✓ | ✓ | ✓ | ✓ |
| # Student Cohorts | 162 | 84 | 76 | 162 |
| $R^2$ | 0.035 | 0.088 | 0.154 | 0.068 |

*Note:* *AI Personalization* includes three services: In-person and telehealth services, plus a conversational AI chatbot emphasizing informational privacy assurances. *AI Privacy* includes three services: In-person and telehealth services, plus a conversational AI chatbot emphasizing personalization at the expenses of information privacy (see appendix for full texts of privacy policy manipulation). This regression is conducted at the student cohort level. Dependent variable CTR is the student cohort-level click-through rate, calculated as number of students who click divided by cohort size. Standard errors in parentheses, are clustered at the school (college) level. ***, ** and * represent statistical significance at the 1%, 5% and 10% levels, respectively. Control variables include school (college) fixed effects. 1, 1, 2 and 1 singleton observations are dropped for column (1), (2), (3) and (4) respectively due to adding fixed effects. IPR refers to the experimental arm with high interpersonal privacy concerns.

The field results examining the effects of informational privacy and its interaction with interpersonal privacy on overall engagement with mental health service bundles are reported in the Appendix Appendix L.

Table 9: Falsification Test: The Role of Information Privacy Concerns in Preferences for In-person Services (Field Experiment)

| Ref. Group: AI Privacy<br>DV: CTR In-person Services | Full Sample<br>(1) | IPR (High)<br>(2) | IPR (Low)<br>(3) | Full Sample<br>(4) |
|---|---|---|---|---|
| AI Personalization | -0.000 | -0.002 | -0.001 | 0.000 |
| | (0.004) | (0.007) | (0.002) | (0.002) |
| IPR | | | | 0.003 |
| | | | | (0.005) |
| AI Personalization × IPR | | | | -0.001 |
| | | | | (0.006) |
| Constant | 0.004** | 0.007 | 0.003*** | 0.003 |
| | (0.002) | (0.004) | (0.001) | (0.002) |
| Control Variables | ✓ | ✓ | ✓ | ✓ |
| # Student Cohorts | 162 | 84 | 76 | 162 |
| $R^2$ | 0.145 | 0.203 | 0.192 | 0.148 |

*Note: AI Personalization* includes three services: In-person and telehealth services, plus a conversational AI chatbot emphasizing informational privacy assurances. *AI Privacy* includes three services: In-person and telehealth services, plus a conversational AI chatbot emphasizing personalization at the expenses of information privacy (see appendix for full texts of privacy policy manipulation). This regression is conducted at the student cohort level. Dependent variable CTR is the student cohort-level click-through rate, calculated as number of students who click divided by cohort size. Standard errors in parentheses, are clustered at the school (college) level. ***, ** and * represent statistical significance at the 1%, 5% and 10% levels, respectively. Control variables include school (college) fixed effects. 1, 1, 2 and 1 singleton observations are dropped for column (1), (2), (3) and (4) respectively due to adding fixed effects. IPR refers to the experimental arm with high interpersonal privacy concerns.

# 6 Discussion

Our findings show that telehealth and AI chatbots are effective in encouraging the use of mental health services, and both interpersonal and informational privacy concerns shape the willingness of individuals to engage with mental health services. In contexts where interpersonal privacy concerns are low, individuals with weaker informational privacy concerns exhibit higher engagement with AI chatbots, and similarly, AI chatbots that emphasize stronger informational privacy protections enhance engagement. In contexts where interpersonal privacy is a major concern, users appear to be willing to overlook informational privacy risks to access AI-based support, thereby sustaining high levels of engagement and preserving its complementary role alongside telehealth.

Our findings suggest that the value of digital technology-driven mental health care is highly contingent on the interplay between interpersonal and informational privacy concerns. Although one might expect that high interpersonal privacy concerns would deter usage due to information privacy risks, our results reveal the opposite: in the context of high interpersonal privacy concerns, individuals appear more willing to overlook information privacy issues—despite their prevalence in digital health technologies—in order to access the support these AI chatbots offer. This implies that digital mental health technologies are generally accepted by users despite security and confidentiality risks. In the age of AI, where data reuse is inevitable, users are often willing to tolerate these risks when the primary benefit is the mitigation of social judgment. If this were not the case, AI-driven solutions might fail to support those who need them most, reinforcing rather than alleviating stigma-related barriers.

The results have important managerial and policy implications. First, the success of telehealth and conversational AI in encouraging the use of mental health services suggests that digital interventions are effective tools in combating stigma, despite data security risks. For managers and policymakers, this highlights the need to expand access to these services, particularly in environments with high risks of interpersonal exposures, where individuals may be reluctant to seek traditional face-to-face help. By offering telehealth or AI-based options, organizations can make mental health services more accessible to a wider population.

Second, while we show that interpersonal privacy concerns appears to be a stronger barrier than informational privacy concerns in mental health settings, the role of information securities and confidentiality cannot be overlooked, particularly as mental health stigma may decrease over time or be less relevant for certain groups of people. Managers and policymakers must ensure that robust privacy protections are in place, particularly for geenrative AI-based mental health services.

Finally, it is essential to address interpersonal and informational concerns separately, especially in the design and marketing of AI-based mental health services. To maximize user engagement, companies should emphasize reduction of interpersonal exposure in stigma-sensitive environments and focus on data security assurances in less stigma-sensitive contexts. Differentiating between these concerns will help enhance user experience and ensure broader adoption of AI-based mental health services.

# References


Abd-Alrazaq, A. A., A. Rababeh, M. Alajlani, B. M. Bewick, and M. Househ (2020). Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. *Journal of medical Internet research 22*(7), e16021.

Acquisti, A., L. Brandimarte, and G. Loewenstein (2015). Privacy and human behavior in the age of information. *Science 347*(6221), 509–514.

Acquisti, A., C. Taylor, and L. Wagman (2016). The economics of privacy. *Journal of economic Literature 54*(2), 442–492.

Adjerid, I., A. Acquisti, R. Telang, R. Padman, and J. Adler-Milstein (2016). The impact of privacy regulation and technology incentives: The case of health information exchanges. *Management Science 62*(4), 1042–1063.

Agarwal, N., A. Moehring, P. Rajpurkar, and T. Salz (2023). Combining human expertise with artificial intelligence: Experimental evidence from radiology. Technical report, National Bureau of Economic Research.

Al-Natour, S., H. Cavusoglu, I. Benbasat, and U. Aleem (2020). An empirical investigation of the antecedents and consequences of privacy uncertainty in the context of mobile apps. *Information Systems Research 31*(4), 1037–1063.

Altman, I. (1975). The environment and social behavior: privacy, personal space, territory, and crowding.

Ayers, J. W., A. Poliak, M. Dredze, E. C. Leas, Z. Zhu, J. B. Kelley, D. J. Faix, A. M. Goodman, C. A. Longhurst, M. Hogarth, et al. (2023). Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA internal medicine*.

Beggs, S., S. Cardell, and J. Hausman (1981). Assessing the potential demand for electric cars. *Journal of econometrics 17*(1), 1–19.

Bharadwaj, P., M. M. Pai, and A. Suziedelyte (2017). Mental health stigma. *Economics Letters 159*, 57–60.

Blease, C. and J. Torous (2023). Chatgpt and mental healthcare: balancing benefits with risks of harms. *BMJ Ment Health 26*(1).

Bolotnyy, V., M. Basilico, and P. Barreira (2022). Graduate student mental health: lessons from american economics departments. *Journal of Economic Literature 60*(4), 1188–1222.

Buckman, J. R., J. C. Bockstedt, and M. J. Hashim (2019). Relative privacy valuations under varying disclosure characteristics. *Information Systems Research 30*(2), 375–388.

Bursztyn, L., G. Egorov, and S. Fiorin (2020). From extreme to mainstream: The erosion of social norms. *American economic review 110*(11), 3522–3548.

Bursztyn, L., A. L. González, and D. Yanagizawa-Drott (2020). Misperceived social norms: Women working outside the home in saudi arabia. *American economic review 110*(10), 2997–3029.

Bursztyn, L. and D. Y. Yang (2022). Misperceptions about others. *Annual Review of Economics 14*, 425–452.

Cadario, R., C. Longoni, and C. K. Morewedge (2021). Understanding, explaining, and utilizing medical artificial intelligence. *Nature human behaviour 5*(12), 1636–1642.

Campos, F., M. Frese, M. Goldstein, L. Iacovone, H. C. Johnson, D. McKenzie, and M. Mensmann (2017). Teaching personal initiative beats traditional training in boosting small business in west africa. *Science 357*(6357), 1287–1290.

Chen, Y. and P. Esmaeilzadeh (2024). Generative ai in medical practice: in-depth exploration of privacy and security challenges. *Journal of Medical Internet Research 26*, e53008.

Corredor-Waldron, A. and J. Currie (2024). To what extent are trends in teen mental health driven by changes in reporting?: The example of suicide-related hospital visits. *Journal of Human Resources 59*(S), S14–S40.

Cronin, C. J., M. P. Forsstrom, and N. W. Papageorge (2024). What good are treatment effects without treatment? mental health and the reluctance to use talk therapy. *Review of Economic Studies*, rdae061.

Culnan, M. J. and R. J. Bies (2003). Consumer privacy: Balancing economic and justice considerations. *Journal of social issues 59*(2), 323–342.

DeRubeis, R. J., S. D. Hollon, J. D. Amsterdam, R. C. Shelton, P. R. Young, R. M. Salomon, J. P. O'Reardon, M. L. Lovett, M. M. Gladis, L. L. Brown, et al. (2005). Cognitive therapy vs medications in the treatment of moderate to severe depression. *Archives of general psychiatry 62*(4), 409–416.

Dinev, T. and P. Hart (2006). An extended privacy calculus model for e-commerce transactions. *Information systems research 17*(1), 61–80.

Ernst, M. L. and A. U. Schwartz (1977). Privacy: The right to be let alone.

Erol, A. C. and L. Wu (2025). Offline privacy concerns in mental health care delivery through telehealth. *Available at SSRN 5368267*.

Fitzpatrick, K. K., A. Darcy, and M. Vierhile (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): a randomized controlled trial. *JMIR mental health 4*(2), e7785.

Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. Simon and schuster.

Goh, Y.-S., Q. Y. J. Ow Yong, and W.-S. W. Tam (2021). Effects of online stigma-reduction programme for people experiencing mental health conditions: A systematic review and meta-analysis. *International journal of mental health nursing 30*(5), 1040–1056.

Golda, A., K. Mekonen, A. Pandey, A. Singh, V. Hassija, V. Chamola, and B. Sikdar (2024). Privacy and security concerns in generative ai: a comprehensive survey. *IEEE Access 12*, 48126–48144.

Hair, J. F., W. C. Black, B. J. Babin, R. E. Anderson, R. L. Tatham, et al. (2006). Multivariate data analysis (vol. 6).

Harrison, G. W. and J. A. List (2004). Field experiments. *Journal of Economic literature 42*(4), 1009–1055.

He, T., G. Fu, Y. Yu, F. Wang, J. Li, Q. Zhao, C. Song, H. Qi, D. Luo, H. Zou, et al. (2023). Towards a psychological generalist ai: A survey of current applications of large language models and future prospects. *arXiv preprint arXiv:2312.04578*.

Henderson, C., J. Noblett, H. Parke, S. Clement, A. Caffrey, O. Gale-Grant, B. Schulze, B. Druss, and G. Thornicroft (2014). Mental health-related stigma in health care and mental health-care settings. *The Lancet Psychiatry 1*(6), 467–482.

Ho, A., J. Hancock, and A. S. Miner (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication 68*(4), 712–733.

Jackson, S. L., R. Mejilla, J. D. Darer, N. V. Oster, J. D. Ralston, S. G. Leveille, J. Walker, T. Delbanco, and J. G. Elmore (2014). Patients who share transparent visit notes with others: characteristics, risks, and benefits. *Journal of medical Internet research 16*(11), e247.

Jin, G. Z. (2018). Artificial intelligence and consumer privacy. In *The economics of artificial intelligence: An agenda*, pp. 439–462. University of Chicago Press.

John, L. K., A. Acquisti, and G. Loewenstein (2011). Strangers on a plane: Context-dependent willingness to divulge sensitive information. *Journal of consumer research 37*(5), 858–873.

Jussupow, E., K. Spohrer, A. Heinzl, and J. Gawlitza (2021). Augmenting medical diagnosis decisions? an investigation into physicians' decision-making process with artificial intelligence. *Information Systems Research 32*(3), 713–735.

Kummer, M. and P. Schulte (2019). When private information settles the bill: Money and privacy in google's market for smartphone applications. *Management Science 65*(8), 3470–3494.

Leschanowsky, A., S. Rech, B. Popp, and T. Bäckström (2024). Evaluating privacy, security, and trust perceptions in conversational ai: A systematic review. *arXiv preprint arXiv:2406.09037*.

Levitt, S. D. and J. A. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic perspectives 21*(2), 153–174.

Link, B. G. and J. C. Phelan (2001). Conceptualizing stigma. *Annual review of Sociology 27*(1), 363–385.

Longoni, C., A. Bonezzi, and C. K. Morewedge (2019). Resistance to medical artificial intelligence. *Journal of Consumer Research 46*(4), 629–650.

Lucas, G. M., J. Gratch, A. King, and L.-P. Morency (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior 37*, 94–100.

Malhotra, N. K., S. S. Kim, and J. Agarwal (2004). Internet users' information privacy concerns (iuipc): The construct, the scale, and a causal model. *Information systems research 15*(4), 336–355.

Marden, J. I. (1996). *Analyzing and modeling rank data*. CRC Press.

Marks, M. and C. E. Haupt (2023). Ai chatbots, health privacy, and challenges to hipaa compliance. *Jama*.

Miller, A. R. and C. E. Tucker (2011). Can health care information technology save babies? *Journal of Political Economy 119*(2), 289–324.

Moon, Y. (2000). Intimate exchanges: Using computers to elicit self-disclosure from consumers. *Journal of consumer research 26*(4), 323–339.

Nass, C., Y. Moon, and P. Carney (1999). Are people polite to computers? responses to computer-based interviewing systems 1. *Journal of applied social psychology 29*(5), 1093–1109.

Organization, W. H. et al. (2013). Comprehensive mental health action plan 2013–2020. *Geneva: World Health Organization 48*.

Palmer, A. and D. Schwan (2022). Beneficent dehumanization: Employing artificial intelligence and carebots to mitigate shame-induced barriers to medical care. *Bioethics 36*(2), 187–193.

Pryor, J. B., G. D. Reeder, C. Yeadon, and M. Hesson-McInnis (2004). A dual-process model of reactions to perceived stigma. *Journal of personality and social psychology 87*(4), 436.

Punj, G. N. and R. Staelin (1978). The choice process for graduate business schools. *Journal of Marketing research 15*(4), 588–598.

Rahman, M. A. (2023). A survey on security and privacy of multimodal llms-connected healthcare perspective. In *2023 IEEE Globecom Workshops (GC Wkshps)*, pp. 1807–1812. IEEE.

Rao, A., M. Pang, J. Kim, M. Kamineni, W. Lie, A. K. Prasad, A. Landman, K. Dreyer, and M. D. Succi (2023). Assessing the utility of chatgpt throughout the entire clinical workflow: development and usability study. *Journal of Medical Internet Research 25*, e48659.

Slovic, P. (2016). Perception of risk. In *The perception of risk*, pp. 220–231. Routledge.

Stiefel, S. (2018). The chatbot will see you now: protecting mental health confidentiality in software applications. *Colum. Sci. & Tech. L. Rev. 20*, 333.

Thornicroft, G., N. Mehta, S. Clement, S. Evans-Lacko, M. Doherty, D. Rose, M. Koschorke, R. Shidhaye, C. O'Reilly, and C. Henderson (2016). Evidence for effective interventions to reduce mental-health-related stigma and discrimination. *The Lancet 387*(10023), 1123–1132.

Tucker, C. (2018). Privacy, algorithms, and artificial intelligence. In *The economics of artificial intelligence: An agenda*, pp. 423–437. University of Chicago Press.

Varghese, M. A., P. Sharma, and M. Patwardhan (2024). Public perception on artificial intelligence–driven mental health interventions: Survey research. *JMIR Formative Research 8*, e64380.

Wang, C., S. Liu, H. Yang, J. Guo, Y. Wu, and J. Liu (2023). Ethical considerations of using chatgpt in health care. *Journal of Medical Internet Research 25*, e48009.

Wang, W., G. Gao, and R. Agarwal (2023). Friend or foe? teaming between artificial intelligence and workers with variation in experience. *Management Science*.

Wang, Y., C. Liu, K. Zhou, T. Zhu, and X. Han (2024). Towards regulatory generative ai in ophthalmology healthcare: a security and privacy perspective. *British Journal of Ophthalmology*.

Weisband, S. and S. Kiesler (1996). Self disclosure on computer forms: Meta-analysis and implications. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 3–10.

Westin, A. F. (1968). Privacy and freedom. *Washington and Lee Law Review 25*(1), 166.

Xu, H. and N. Zhang (2022). From contextualizing to context theorizing: assessing context effects in privacy research. *Management Science 68*(10), 7383–7401.

# Appendix A Contextualizing our contribution against disclosure-to-computers studies

Table A1: Contextualizing our contribution against disclosure-to-computers studies

| Paper | (i) Focal outcome (measurement) | (ii) Setting / population | (iii) Method + manipulation | (iv) Main findings | (v) Focal mechanism(s) | (vi) Comparison to our paper |
|---|---|---|---|---|---|---|
| Nass et al. (1999) | Politeness and evaluation of the system, measured by positivity and variance of ratings about the computer following task completion. | General lab participants interacting with text- or voice-based computers in a computer-based interviewing context. | Laboratory experiments manipulating the locus of evaluation: the same computer evaluates itself versus a different computer or a paper-and-pencil questionnaire. | Participants provided more positive and less varied evaluations when the same computer asked about its own performance, consistent with social responses predicted by CASA. | CASA and politeness norms: people apply social rules to computers even when the interaction is purely technological; privacy concerns are not theorized. | We study service adoption rather than politeness or evaluation. Our mechanism is explicitly interpersonal privacy concerns (stigma), experimentally manipulated, rather than general social-norm compliance toward computers. |

*Continued on next page*

| Paper | (i) Focal outcome (measurement) | (ii) Setting / population | (iii) Method + manipulation | (iv) Main findings | (v) Focal mechanism(s) | (vi) Comparison to our paper |
|---|---|---|---|---|---|---|
| Moon (2000) | Self-disclosure of intimate personal information to a computer, along with downstream behavioral effects in subsequent interactions. | Consumer participants in computer-mediated marketing-research style disclosure tasks. | Laboratory experiments manipulating reciprocity (computer discloses vs. not) and sequence (order of disclosure exchanges). | Reciprocity and appropriate sequencing increase willingness to disclose intimate information to computers; disclosure also affects later consumer behavior. | Social exchange dynamics (reciprocity and sequencing norms) governing disclosure to computers; no stigma or privacy-risk framework. | Our paper does not rely on reciprocity or sequence. Instead, we manipulate interpersonal privacy pressure and service-option bundles, focusing on mental health help-seeking rather than eliciting disclosure per se. |

*Continued on next page*

| Paper | (i) Focal outcome (measurement) | (ii) Setting / population | (iii) Method + manipulation | (iv) Main findings | (v) Focal mechanism(s) | (vi) Comparison to our paper |
|---|---|---|---|---|---|---|
| Ho et al. (2018) | Psychological, relational, and emotional outcomes following emotional versus factual disclosure in conversations. | General participants interacting with a supposed chatbot or a human conversational partner. | Experiment manipulating perceived partner type (chatbot vs. human) and disclosure type (emotional vs. factual). | Emotional disclosure produced equivalent downstream outcomes whether participants believed they were interacting with a chatbot or a human, supporting media equivalency. | Media equivalency: the consequences of disclosure are similar across human and machine partners when conversational cues are comparable. | We focus on adoption decisions under privacy tradeoffs, not on the emotional consequences of disclosure. Our design directly manipulates interpersonal privacy concerns rather than only partner identity beliefs. |

*Continued on next page*

| Paper | (i) Focal outcome (measurement) | (ii) Setting / population | (iii) Method + manipulation | (iv) Main findings | (v) Focal mechanism(s) | (vi) Comparison to our paper |
|---|---|---|---|---|---|---|
| Lucas et al. (2014) | Willingness to disclose, fear of self-disclosure, impression management, emotional expression intensity, and observer-rated disclosure behavior. | Participants in a health-screening interview with a virtual human interviewer. | Experiment manipulating belief about agency: whether the virtual human is controlled by a human operator or by automation. | Believing the virtual human was automated reduced fear of disclosure and impression management, and increased emotional expression and perceived willingness to disclose. | Implied reduction in social-evaluative threat when human judgment is absent, though privacy is not explicitly conceptualized or decomposed. | We make the mechanism explicit by framing social-evaluative threat as interpersonal privacy (stigma), manipulate it directly, and examine service choice among AI and human options rather than disclosure during an interview. |

*Continued on next page*

| Paper | (i) Focal outcome (measurement) | (ii) Setting / population | (iii) Method + manipulation | (iv) Main findings | (v) Focal mechanism(s) | (vi) Comparison to our paper |
|---|---|---|---|---|---|---|
| Weisband and Kiesler (1996) | Self-disclosure levels aggregated across studies comparing computer, face-to-face, and paper-based administration. | Meta-analysis of 39 studies (1969–1994), including sensitive topics and some medical or psychiatric contexts. | Meta-analysis; moderators include administration mode, sensitivity of information, subject population, and interface characteristics. | Across studies, computer administration increased self-disclosure, especially relative to face-to-face interviews and for sensitive information; mechanisms remain unclear. | Documents systematic patterns and moderators but cannot isolate a single causal mechanism for increased disclosure to computers. | We extend this literature by providing a causal test of a specific mechanism—interpersonal privacy concerns—and by studying contemporary AI-based mental health services using adoption decisions rather than self-reported disclosure. |

*Continued on next page*

| Paper | (i) Focal outcome (measurement) | (ii) Setting / population | (iii) Method + manipulation | (iv) Main findings | (v) Focal mechanism(s) | (vi) Comparison to our paper |
|---|---|---|---|---|---|---|
| **Our paper** | Service adoption intentions and decisions among mental health service offerings, interpreted as revealed preferences under privacy tradeoffs. | Individuals with mental health concerns making help-seeking decisions in a mental health context. | Experiment manipulating interpersonal privacy concerns (stigma) and combinations of service offerings that vary in recipient type and modality. | When interpersonal privacy concerns are more pressing, participants are more likely to choose AI-based services, despite salient informational privacy risks. | Interpersonal privacy concerns (stigma and fear of social judgment) as the focal mechanism, explicitly contrasted with informational privacy concerns. | n/a |

# Appendix B Manipulation of Interpersonal Privacy Concerns

To implement this manipulation, each participant is first asked to indicate the extent to which they agree with the following statement: *Others would view me negatively if they believed that I experienced a mental health issue.* They are then asked to estimate the percentage of others who agree with this statement. Afterwards, participants are randomly assigned to either a treatment or control condition. In the treatment condition, participants are shown their own response alongside a fabricated percentage—specifically, that 35% of others agree with the statement. If a participant's estimate differs from this fabricated statistic, they are informed of the discrepancy.

Specifically, we elicit two *ex ante* stigma beliefs from each participant. First, we elicit their *private belief* about mental health stigma—that is, the extent to which *they themselves* believe mental health is stigmatized (i.e., that people hold stigmatizing views toward individuals with mental health issues). Second, we elicit their *perceptions about others' beliefs*—that is, the percentage of *other participants they think* believe that people hold stigmatizing views toward individuals with mental health issues.

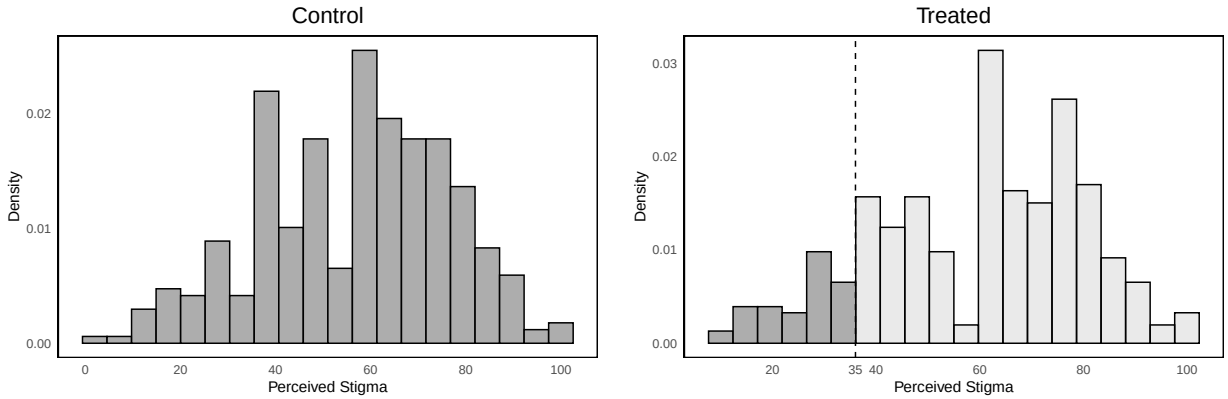
Notably, eliciting both *ex ante* (baseline) stigma beliefs, i.e., participants' own beliefs about stigma and their perceptions of others' beliefs, is necessary for our experimental design. Without baseline private beliefs, we could not credibly frame the information shown to each participant as an aggregate of other participants' beliefs, which requires eliciting each participant's beliefs in the first place. Without baseline perceptions of others' beliefs, there would be no misperception to "correct.". In this paper, our focus is on *perceived* stigma (i.e., we correct misperceptions about others' beliefs, not participants' own beliefs), which can shape mental health help-seeking behavior and can be shifted when updated aggregate information arrives (Bursztyn and Yang, 2022).

The majority of participants in the treatment group report perceived stigma levels exceeding this stated level (the lighter-shaded bars in the Figure B1). Therefore, the information treatment *can be used* to correct misperceptions (Bursztyn et al., 2020): Treated participants who were informed that only 35% of others held stigmatizing views, and whose *ex ante* perceptions exceeded 35%, should (in theory) update their beliefs downward relative to control participants, who did not receive any information about others'

opinions (in the right panel of Figure B1, participants whose ex ante perceptions exceeded 35% are those intended to be treated, shown in the lighter-shaded bars). For example, a participant in the treatment group who initially guesses that 60% of other participants hold stigmatizing views should (in theory (Bursztyn et al., 2020)) revise this belief downward upon learning that the true figure is 35%.

Figure B1 identifies the participants who are intended to be treated. We emphasize that our treatment is a type of information intervention where we manipulate perceived stigma by providing new (fabricated) information about actual stigma only to participants in the treatment group. In the control (high-stigma) condition, participants receive no information about actual stigma. In the treatment (low-stigma) condition, participants are presented with (fabricated) information about actual stigma stating that 35% of other Prolific participants in this study hold stigmatizing views.

Figure B1: Distribution of participants' perceived stigma in the control (high stigma) and treatment (low stigma) groups prior to treatment




*Notes:* We manipulate perceived stigma by providing new (fabricated) information about actual stigma only to participants in the treatment group. In the control (high-stigma) condition, participants receive no information about actual stigma. In the treatment (low-stigma) condition, participants are presented with (fabricated) information about actual stigma stating that 35% of other Prolific participants in this study hold stigmatizing views. Notably, the majority of participants in the treatment group report perceived stigma levels exceeding this stated level (the lighter-shaded bars).

## B1 Full Texts of Survey

Do you agree with this statement: **"Others would view me negatively if they believed that I experienced a mental health issue."**



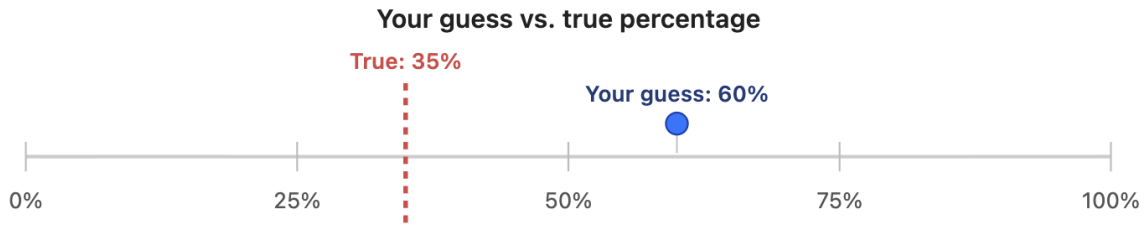
- ○ Strongly disagree
- ○ Somewhat disagree
- ○ Neither agree nor disagree
- ○ Somewhat agree
- ○ Strongly agree

Based on several hundred responses from other US-based Prolific workers so far, **what percentage** do you think agreed (either **Somewhat agree** or **Strongly agree**) with the statement we just asked you. Please enter a number between 0 and 100 below.

Others would view me negatively if they believed that I experienced a mental health issue.

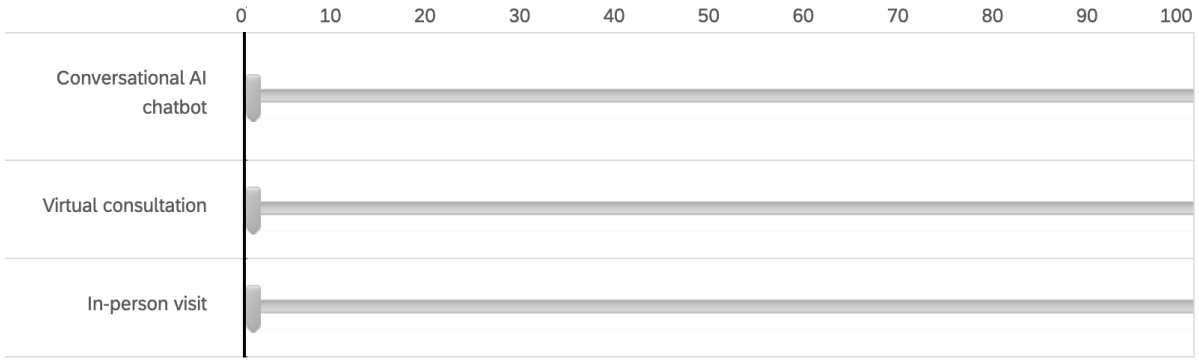
You estimated that **60%** of other participants agreed with the statement: **"Others would view me negatively if they believed that I experienced a mental health issue."** In reality, **35%** of other respondents agreed with the statement. Your estimate was **25% higher** than the estimates of other participants. Others were more accepting of mental health issues than you thought.

Estimate what share agreed with: *"Others would view me negatively if they believed I experienced a mental health issue."*

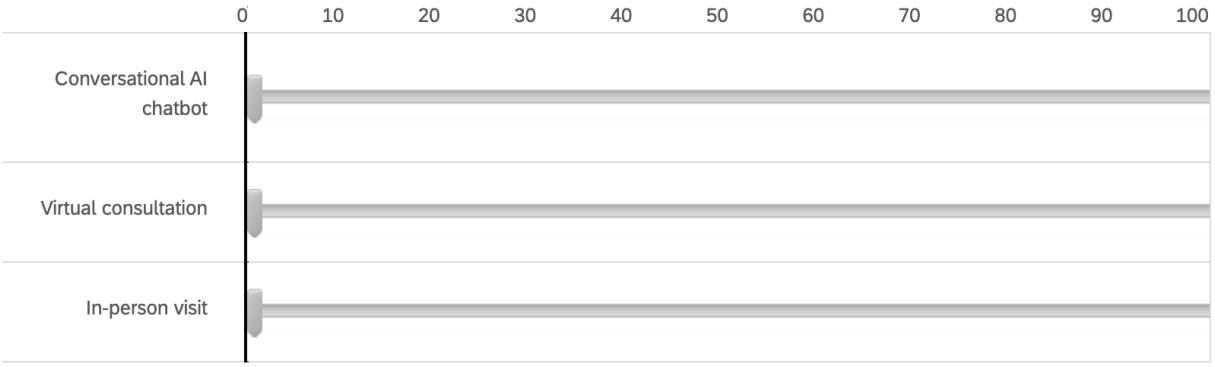


# Appendix C Full Texts of Survey: Perceived Attributes of Each Service

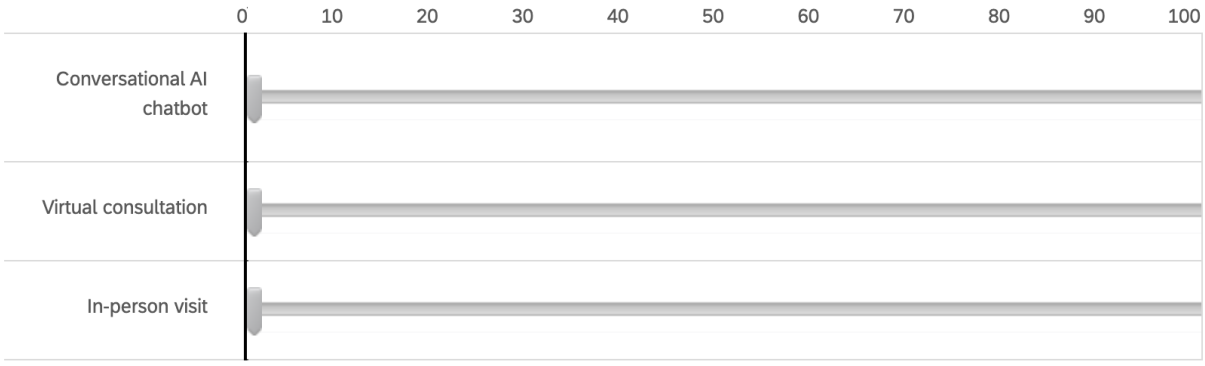
How convenient would each of the following services be for you to use (e.g. ease of access, time savings, and flexibility)? (0 = Not at all convenient, 100 = Extremely convenient)



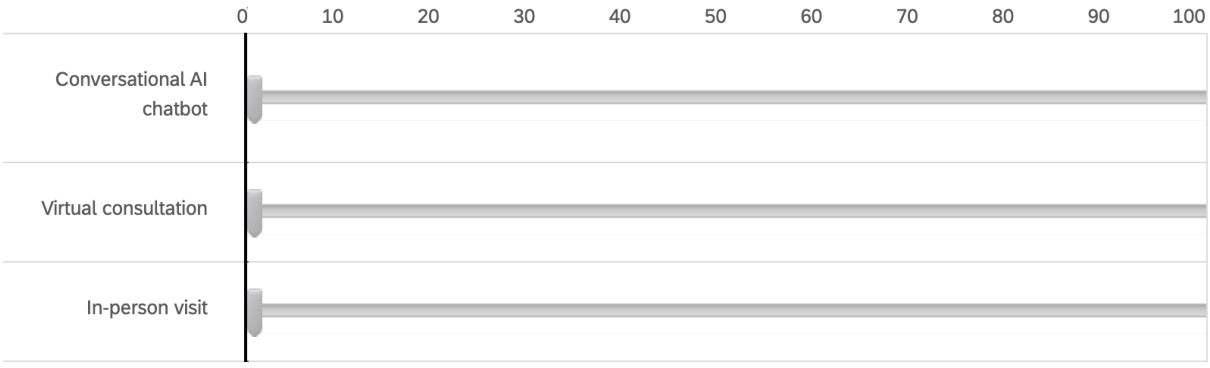

What is your expected quality of the care for each of the following services? (0 = Extremely low quality, 100 = Extremely high quality)



How would you rate the data privacy of each of the following services (e.g., security of information, and control over personal data)? (0 = Very poor privacy, 100 = Excellent privacy)




To what extent would using each of the following services make you feel *free from* stigma or judgment from others (e.g., friends, family, coworkers, therapists, or the public)? (0 = Not at all free from stigma or judgement, 100 = Completely free from stigma or judgement)



# Appendix D Full Text of the Hypothetical Emails

## Online Ranking Experiment

### Explore Your Mental Health Support Options: Conversational AI Chatbot, Virtual Consultations, or In-Person Visits

We know taking the first step after a difficult time can feel heavy. We're here to make it easier. Below is a quick overview of our three **board-certified** support options so you can decide what feels right today. **No out-of-pocket cost. No insurance required.**

#### Conversational AI: Have a conversation with an AI chatbot

- **Main Features:** Chat in real time with a board-certified conversational AI chatbot for mental health.



### Virtual Consultations: Schedule an online meeting with a mental health professional

- **Main Features:** Connect with a board-certified mental health professional from your home using video platforms (e.g., Zoom).




### In-Person Visits: Directly meet with a mental health professional

- **Main Features:** Meet with a board-certified professional face-to-face at one of our local offices.




### What to Expect

- Whether chatting with an AI chatbot, meeting virtually, or sitting down in person — you'll be welcomed with care, respect, and support tailored to your needs.

If you're in immediate distress or thinking about harming yourself, call or text **988** in the U.S. for 24/7 support, or contact your local emergency number.

## Online Bundled Experiment

## In-person + Telehealth + AI Privacy

# Explore Your Mental Health Support Options: ChatGPT-Like AI, Virtual Consultations, or In-Person Visits

We are excited to offer you three distinct avenues to support your mental health journey. **Each option offers a different form of engagement.** Read on to discover which service might be the best fit for you:

**1. ChatGPT-Like AI: Have a conversation with an AI chatbot that ensures high-level data privacy protection**

- **Main Features:** Engage in a completely private, text-based conversation with our advanced ChatGPT-Like AI. The AI chatbot service is available 24/7 and can be accessed from anywhere. It allows you to seek support without fear of being seen by strangers or overheard by friends or family members.
- **High Data Privacy:**
  - No personal login required; use a nickname only.
  - Any conversation data voluntarily provided by you.
  - Data encryption ensures your information remains secure.
  - No third-party access to your data.

**2. Virtual Consultations: Schedule an online meeting with a mental health professional**

- **Main Features:** Connect with our mental health professionals from the comfort of your home using video conferencing platforms (e.g., Zoom). This allows you to avoid the commute to local offices and ensures your privacy without any risk of being seen by the public.

**3. In-Person Visits: Directly meet with a mental health professional**

- **Main Features:** Our In-Person Visits allow you to meet with a mental health professional face-to-face at one of our local offices.

Select the option that best fits your needs!

**In-person + Telehealth + AI Personalization**

## Explore Your Mental Health Support Options: ChatGPT-Like AI, Virtual Consultations, or In-Person Visits

We are excited to offer you three distinct avenues to support your mental health journey. **Each option offers a different form of engagement.** Read on to discover which service might be the best fit for you:

**1. ChatGPT-Like AI: Have a conversation with an AI chatbot**

- **Main Features:** Engage in a completely private, text-based conversation everywhere with our advanced ChatGPT-Like AI. The AI chatbot service is available 24/7 and can be accessed from anywhere. It allows you to seek support without fear of being seen by strangers or overheard by friends or family members.
- **Privacy Policy Highlights:**
  - **Data Collection:** We collect user content, including chat communications.
  - **Usage Purposes:**
    - Personalize services to better meet your needs.
    - Analyze, maintain, improve, modify, and measure services.
    - Train our AI and machine learning models.
    - Develop new programs and services for added value.

**2. Virtual Consultations: Schedule an online meeting with a mental health professional**

- **Main Features:** Connect with our mental health professionals from the comfort of your home using video conferencing platforms (e.g., Zoom). This allows you to avoid the commute to local offices and ensures your privacy without any risk of being seen by the public.

**3. In-Person Visits: Directly meet with a mental health professional**

- **Main Features:** Our In-Person Visits allow you to meet with a mental health professional face-to-face at one of our local offices.

Select the option that best fits your needs!

## In-person + Telehealth + AI (no privacy information)

# Explore Your Mental Health Support Options: ChatGPT-Like AI, Virtual Consultations, or In-Person Visits

We are excited to offer you three distinct avenues to support your mental health journey. **Each option offers a different form of engagement.** Read on to discover which service might be the best fit for you:

**1. ChatGPT-Like AI: Have a conversation with an AI chatbot**

- **Main Features:** Engage in a completely private, text-based conversation everywhere with our advanced ChatGPT-Like AI. The AI chatbot service is available 24/7 and can be accessed from anywhere. It allows you to seek support without fear of being seen by strangers or overheard by friends or family members.

**2. Virtual Consultations: Schedule an online meeting with a mental health professional**

- **Main Features:** Connect with our mental health professionals from the comfort of your home using video conferencing platforms (e.g., Zoom). This allows you to avoid the commute to local offices and ensures your privacy without any risk of being seen by the public.

**3. In-Person Visits: Directly meet with a mental health professional**

- **Main Features:** Our In-Person Visits allow you to meet with a mental health professional face-to-face at one of our local offices.

Select the option that best fits your needs!

**In-person + Telehealth**

## Explore Your Mental Health Support Options: Virtual Consultations, or In-Person Visits

We are excited to offer you two distinct avenues to support your mental health journey. **Each option offers a different form of engagement.** Read on to discover which service might be the best fit for you:

**1. Virtual Consultations: Schedule an online meeting with a mental health professional**

- **Main Features:** Connect with our mental health professionals from the comfort of your home using video conferencing platforms (e.g., Zoom). This allows you to avoid the commute to local offices and ensures your privacy without any risk of being seen by the public.

**2. In-Person Visits: Directly meet with a mental health professional**

- **Main Features:** Our In-Person Visits allow you to meet with a mental health professional face-to-face at one of our local offices.

Select the option that best fits your needs!

# In-person

## Explore Your Mental Health Support through In-Person Visits

We are excited to offer you a valuable avenue to support your mental health journey: In-Person Visits.

**In-Person Visits: Directly meet with a mental health professional**

Our In-Person Visits provide a traditional setting where you meet directly with health professionals:

- **Main Features:** Our In-Person Visits allow you to meet with a mental health professional face-to-face at one of our local offices.

# D1 Field Experiment

## In-person + Telehealth + AI Privacy

Conversational Artificial Intelligence (AI) Chatbot: Text-based interaction with an AI chatbot

- **Key Features**: Students can access AI-based psychological counseling services powered by large language models. The AI counseling service enables fully private, text-based conversations with a chatbot and is available 24 hours a day, 7 days a week, from any location. Using AI counseling can alleviate concerns about being seen by faculty, peers, or others when visiting in-person counseling centers, and can also reduce the risk of being overheard during online counseling sessions.
- **Enhanced Privacy Protection**:
  - You have the right to voluntarily provide, access, correct, or supplement your personal information.
  - You may request account deactivation and deletion of all associated personal data.
  - Collected data are encrypted and stored on [CLOUD SERVICE PROVIDER].
  - Your personal information and chat data are not shared with third-party organizations.
- **How to Access**
  - Click [WEBSITE URL] , enter your mobile phone number to receive a verification code, and log in using the code

Online Video Counseling: Remote counseling sessions conducted via video

- **Key Features**: Online counseling allows students to communicate with professional counselors from home or their dormitory. This option reduces the likelihood of being seen by faculty, peers, or others while traveling to in-person counseling centers.
- **Appointment Process**:
  - Click [WEBSITE URL] to download and complete the counseling appointment form and informed consent form, and send them via email to [EMAIL ADDRESS] (please indicate "Online Counseling" in the email subject line).
  - Online counseling is conducted via video; please add the counseling contact in advance: [CONTACT ID].
  - You may also schedule an appointment by phone: [APPOINTMENT PHONE NUMBER].

In-Person Counseling: Face-to-face counseling sessions with a professional counselor

- **Key Features**: In-person counseling enables direct, face-to-face communication with professional counselors at the university's counseling service center.
- **Counseling Location**: [COUNSELING LOCATION].
- **Appointment Process**:
  - Click [WEBSITE URL] to make an online appointment.
  - You may also schedule an appointment by phone: [APPOINTMENT PHONE NUMBER].

# D2 Field Experiment

## In-person + Telehealth + AI Personalization

Conversational Artificial Intelligence (AI) Chatbot: Text-based interaction with an AI chatbot

- **Key Features**: Students can access AI-based psychological counseling services powered by large language models. The AI counseling service enables fully private, text-based conversations with a chatbot and is available 24 hours a day, 7 days a week, from any location. Using AI counseling can alleviate concerns about being seen by faculty, peers, or others when visiting in-person counseling centers, and can also reduce the risk of being overheard during online counseling sessions.
- **Privacy Policy Information**:
  - Data collection: Conversation data are collected for product improvement purposes.
  - Data usage:
    - To provide personalized services that better address mental health needs.
    - To analyze, maintain, and improve services and assess service quality.
    - To train artificial intelligence models.
    - To develop new programs and services.
- **How to Access**
  - Click [WEBSITE URL] , enter your mobile phone number to receive a verification code, and log in using the code

Online Video Counseling: Remote counseling sessions conducted via video

- **Key Features**: Online counseling allows students to communicate with professional counselors from home or their dormitory. This option reduces the likelihood of being seen by faculty, peers, or others while traveling to in-person counseling centers.
- **Appointment Process**:
  - Click [WEBSITE URL] to download and complete the counseling appointment form and informed consent form, and send them via email to [EMAIL ADDRESS] (please indicate "Online Counseling" in the email subject line).
  - Online counseling is conducted via video; please add the counseling contact in advance: [CONTACT ID].
  - You may also schedule an appointment by phone: [APPOINTMENT PHONE NUMBER].

In-Person Counseling: Face-to-face counseling sessions with a professional counselor

- **Key Features**: In-person counseling enables direct, face-to-face communication with professional counselors at the university's counseling service center.
- **Counseling Location**: [COUNSELING LOCATION].
- **Appointment Process**:
  - Click [WEBSITE URL] to make an online appointment.
  - You may also schedule an appointment by phone: [APPOINTMENT PHONE NUMBER].

# Appendix E IUIPC Scale

Seven-point scales anchored with “strongly disagree” and “strongly agree” for each question.

**Control:**

- Consumer online privacy is really a matter of consumers’ right to exercise control and autonomy over decisions about how their information is collected, used, and shared.
- Consumer control of personal information lies at the heart of consumer privacy.
- I believe that online privacy is invaded when control is lost or unwillingly reduced as a result of a marketing transaction.

**Awareness:**

- Companies seeking information online should disclose the way the data are collected, processed, and used.
- A good consumer online privacy policy should have a clear and conspicuous disclosure.
- It is very important to me that I am aware and knowledgeable about how my personal information will be used.

**Collection:**

- It usually bothers me when online companies ask me for personal information.
- When online companies ask me for personal information, I sometimes think twice before providing it.
- It bothers me to give personal information to so many online companies.
- I’m concerned that online companies are collecting too much personal information about me.

# Appendix F Summary Statistics for Online Ranking Experiment

Table F2: Descriptive Statistics (Online Ranking Experiment)

| | All | | Low IPR | | High IPR | |
|---|---|---|---|---|---|---|
| Variables: | Mean | SD | Mean | SD | Mean | SD |
| % perceived stigma | 58.565 | 19.568 | 59.579 | 19.584 | 57.563 | 19.531 |
| GAD | 4.537 | 1.324 | 4.399 | 1.316 | 4.673 | 1.320 |
| PHQ | 4.028 | 1.537 | 3.950 | 1.540 | 4.104 | 1.533 |
| Ever used services | 0.929 | 0.257 | 0.916 | 0.277 | 0.942 | 0.234 |
| AI Utilization | 2.308 | 1.315 | 2.300 | 1.297 | 2.315 | 1.334 |
| age groups | 2.046 | 1.234 | 1.963 | 1.202 | 2.128 | 1.261 |
| income groups | 1.772 | 1.555 | 1.876 | 1.515 | 1.670 | 1.590 |
| gender ==Male | 0.295 | 0.457 | 0.319 | 0.467 | 0.272 | 0.446 |
| gender ==Female | 0.655 | 0.476 | 0.644 | 0.480 | 0.667 | 0.472 |
| gender ==Others | 0.049 | 0.217 | 0.037 | 0.189 | 0.061 | 0.240 |
| races ==White | 0.814 | 0.390 | 0.802 | 0.399 | 0.826 | 0.380 |
| races ==Black | 0.068 | 0.251 | 0.068 | 0.252 | 0.067 | 0.251 |
| races ==Asian | 0.028 | 0.164 | 0.031 | 0.173 | 0.024 | 0.155 |
| races ==Others | 0.091 | 0.288 | 0.099 | 0.299 | 0.083 | 0.276 |
| educ == Associate degree | 0.391 | 0.488 | 0.350 | 0.478 | 0.431 | 0.496 |
| educ ==Bachelor or above | 0.405 | 0.491 | 0.440 | 0.497 | 0.370 | 0.484 |
| educ == Others | 0.205 | 0.404 | 0.211 | 0.408 | 0.199 | 0.400 |
| marital ==Never Married | 0.292 | 0.455 | 0.316 | 0.466 | 0.269 | 0.444 |
| marital ==Married | 0.695 | 0.461 | 0.675 | 0.469 | 0.716 | 0.452 |
| marital ==Others | 0.012 | 0.110 | 0.009 | 0.096 | 0.015 | 0.123 |
| occupation ==Working | 0.640 | 0.480 | 0.659 | 0.475 | 0.621 | 0.486 |
| occupation ==Unemployed | 0.123 | 0.329 | 0.118 | 0.323 | 0.128 | 0.335 |
| occupation ==Others | 0.237 | 0.426 | 0.223 | 0.417 | 0.251 | 0.434 |
| IUIPC | 6.037 | 0.713 | 6.003 | 0.708 | 6.070 | 0.718 |
| ranking in-person services | 1.917 | 0.789 | 1.873 | 0.784 | 1.960 | 0.792 |
| ranking online services | 1.598 | 0.603 | 1.610 | 0.608 | 1.587 | 0.600 |
| ranking AI chatbot | 2.485 | 0.783 | 2.517 | 0.765 | 2.453 | 0.801 |
| # of obs. | 650 | | 323 | | 327 | |

*Note:* Low IPR stands for low interpersonal privacy concerns, the treatment group, and high IPR stands for high interpersonal privacy concerns, the control group.

Table F2 display descriptive statistics for several pre-treatment respondent characteristics as well as dependent variables. We don't observe significant difference in these characteristics between treatment and control groups. We do not include these characteristics in the regressions; they are reported only to demonstrate pre-treatment balance. This is because in the main analysis of the ranking experiment, we include respondent fixed effects, and individual characteristics are invariant across services. For consistency across all specifications of the ranking experiment, we therefore exclude these characteristics

even in the wide-format regressions where respondent fixed effects are not included.

Pre-treatment respondent characteristics include % perceived stigma, which measure participants' initial stigma beliefs. These are based on their estimates of the percentage of other participants who agree with the statement *Others would view me negatively if they believed that I experienced a mental health issue*. Further, we include measures of participants' mental health status (two dummy variables indicating generalized anxiety or depression, based on two pairs of patient survey items that are used in clinical practice to identify mental health challenges among patients, namely GAD-2 or PHQ-2), a dummy variable indicating whether the participant reports ever having made use of mental health service (Ever used services), and a measure of AI utilization (the participant's frequency of AI tool usage in the last month, rated 1–5, ranging from 1 = never, to 5 = several times a day). Finally, demographic controls include age groups (ranging from 1 = 18–24 years, to 6 = 65+ years), annual household income bracket (ranging from 1 = less than $25,000, to 6 = $150,000+), gender, race, educational achievement, marital status, and occupation.

We use the Internet Users' Information Privacy Concerns (IUIPC) scale to measure individuals' informational concerns (Malhotra et al., 2004). The scale is constructed by averaging responses to 10 items measured on a 1 (strongly disagree)–7 (strongly agree) scale and captures informational privacy concerns across three dimensions—collection, control, and awareness—which closely correspond to the types of informational privacy concerns faced by users of AI-based services (see Appendix Appendix E for the full texts of the scale items).

The dependent variables are rankings of three services—in-person, online, and AI chatbot—ranging from 1 to 3. In the regressions, we use a reversed ranking (defined as 3 − ranking) so that larger values represent higher utility.

# Appendix G Heterogeneity Analysis across the IUIPC Dimensions

Table G3: The Role of Information Privacy Concerns in Preferences for AI Services (Ranking Experiment: Awareness)

| Sample: | Full Sample | IPR (High) | IPR (Low) | Full Sample |
|---|---|---|---|---|
| DV: r.ranking of AI | (1) | (2) | (3) | (4) |
| iuipc_aware | -0.129*** | -0.093 | -0.158*** | -0.156*** |
| | (0.035) | (0.058) | (0.044) | (0.045) |
| convenience | 0.002** | 0.002 | 0.002* | 0.002** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| effectiveness | 0.013*** | 0.013*** | 0.013*** | 0.013*** |
| | (0.001) | (0.002) | (0.002) | (0.001) |
| privacy | 0.002* | 0.002 | 0.002 | 0.002* |
| | (0.001) | (0.002) | (0.001) | (0.001) |
| IPR | | | | 0.057 |
| | | | | (0.049) |
| iuipc × IPR | | | | 0.062 |
| | | | | (0.071) |
| Constant | 1.515*** | 1.542*** | 1.485*** | 1.485*** |
| | (0.024) | (0.036) | (0.033) | (0.034) |
| # Observations / Respondents | 650 | 327 | 323 | 650 |
| $R^2$ | 0.388 | 0.371 | 0.408 | 0.390 |

*Note:* The dependent variables for all columns are r.ranking, measuring preference for AI chatbots using a reversed ranking, where 1 indicates the lowest preference and 3 indicates the highest preference. iuipc_aware measures the awareness dimension of informational privacy concern using the Internet Users' Information Privacy Concerns (IUIPC) scale. The independent variable IPR refers to the experimental arm with high interpersonal privacy concerns, and is an indicator equal to 1 for interpersonal privacy concerns and 0 otherwise. convenience, effectiveness and privacy measure respondents' perceptions about the extent to which AI is convenient, of high quality, of high data privacy and free from stigma, measured on a 1–100 scale (higher values indicate higher convenience, effectiveness, data privacy and lower perceived stigma). convenience, effectiveness, privacy and iuipc_aware are demeaned to facilitate interpretation of the coefficients. ***, ** and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table G4: The Role of Information Privacy Concerns in Preferences for AI Services (Ranking Experiment: Control)

| Sample: | Full Sample | IPR (High) | IPR (Low) | Full Sample |
|---|---|---|---|---|
| DV: r.ranking of AI | (1) | (2) | (3) | (4) |
| iuipc_control | -0.020 | -0.011 | -0.032 | -0.031 |
| | (0.025) | (0.037) | (0.034) | (0.035) |
| convenience | 0.002** | 0.002 | 0.002 | 0.002** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| effectiveness | 0.013*** | 0.013*** | 0.013*** | 0.013*** |
| | (0.001) | (0.002) | (0.002) | (0.001) |
| privacy | 0.003** | 0.002 | 0.003* | 0.003** |
| | (0.001) | (0.002) | (0.001) | (0.001) |
| IPR | | | | 0.047 |
| | | | | (0.049) |
| iuipc × IPR | | | | 0.021 |
| | | | | (0.050) |
| Constant | 1.515*** | 1.538*** | 1.492*** | 1.492*** |
| | (0.024) | (0.036) | (0.034) | (0.035) |
| # Observations / Respondents | 650 | 327 | 323 | 650 |
| $R^2$ | 0.376 | 0.367 | 0.386 | 0.377 |

*Note:* The dependent variables for all columns are r.ranking, measuring preference for AI chatbots using a reversed ranking, where 1 indicates the lowest preference and 3 indicates the highest preference. The independent variable IPR refers to the experimental arm with high interpersonal privacy concerns, and is an indicator equal to 1 for interpersonal privacy concerns and 0 otherwise. iuipc_control measures the control dimension of informational privacy concern using the Internet Users' Information Privacy Concerns (IUIPC) scale. convenience, effectiveness and privacy measure respondents' perceptions about the extent to which AI is convenient, of high quality, of high data privacy and free from stigma, measured on a 1–100 scale (higher values indicate higher convenience, effectiveness, data privacy and lower perceived stigma). convenience, effectiveness, privacy and iuipc_control are demeaned to facilitate interpretation of the coefficients. ***, ** and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Table G5: The Role of Information Privacy Concerns in Preferences for AI Services (Ranking Experiment: Collection)

| Sample: | Full Sample | IPR (High) | IPR (Low) | Full Sample |
|---|---|---|---|---|
| DV: r.ranking of AI | (1) | (2) | (3) | (4) |
| iuipc_collect | -0.053** | -0.016 | -0.101*** | -0.098*** |
| | (0.022) | (0.031) | (0.033) | (0.033) |
| convenience | 0.002** | 0.002 | 0.002* | 0.002** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| effectiveness | 0.013*** | 0.013*** | 0.013*** | 0.013*** |
| | (0.001) | (0.002) | (0.002) | (0.001) |
| privacy | 0.002** | 0.002 | 0.002 | 0.002* |
| | (0.001) | (0.002) | (0.001) | (0.001) |
| IPR | | | | 0.051 |
| | | | | (0.049) |
| iuipc × IPR | | | | 0.081* |
| | | | | (0.044) |
| Constant | 1.515*** | 1.539*** | 1.488*** | 1.488*** |
| | (0.024) | (0.036) | (0.033) | (0.034) |
| # Observations / Respondents | 650 | 327 | 323 | 650 |
| $R^2$ | 0.380 | 0.367 | 0.403 | 0.385 |

*Note:* The dependent variables for all columns are r.ranking, measuring preference for AI chatbots using a reversed ranking, where 1 indicates the lowest preference and 3 indicates the highest preference. The independent variable IPR refers to the experimental arm with high interpersonal privacy concerns, and is an indicator equal to 1 for interpersonal privacy concerns and 0 otherwise. iuipc_collect measures the collection dimension of informational privacy concern using the Internet Users' Information Privacy Concerns (IUIPC) scale. convenience, effectiveness and privacy measure respondents' perceptions about the extent to which AI is convenient, of high quality, of high data privacy and free from stigma, measured on a 1–100 scale (higher values indicate higher convenience, effectiveness, data privacy and lower perceived stigma). convenience, effectiveness, privacy and iuipc_collect are demeaned to facilitate interpretation of the coefficients. ***, ** and * represent statistical significance at the 1%, 5%, and 10% levels, respectively.

# Appendix H Rank-Ordered and Ordered Logit Results for Online Ranking Experiment

We use a rank-ordered logit model with the long-form ranking data to analyze how interpersonal privacy concerns affect the relative ranking of AI and telehealth services compared to in-person services. The dependent variable is the full ranking over the available service options. Conceptually, the model treats each respondent's ranking as a sequence of conditional choices. At the first stage, the model estimates the probability that a given alternative (e.g., AI, telehealth, in-person) is ranked first among all available options. Conditional on that first choice being "removed" from the choice set, the second stage estimates the probability of being ranked second among the remaining options, and so on. Thus, a complete ranking is decomposed into a series of staged multinomial logit choices. The likelihood is constructed as the product of these conditional choice probabilities across ranking positions, which allows us to estimate how interpersonal privacy concerns shift the relative utility of each service option. (Punj and Staelin, 1978; Beggs et al., 1981)

This model is known by many names. In economics, it is referred to as the rank-ordered logit model (Beggs et al., 1981); in marketing research, as the exploded logit model (Punj and Staelin, 1978); in applied conjoint analysis, as the choice-based conjoint model (Hair et al., 2006); and in probability theory and statistics, as the Plackett–Luce model (Marden, 1996). Despite different terminology, these approaches are formally equivalent and rely on the same sequential random utility framework.

To study the effect of informational privacy concerns on rankings of AI (and, as a falsification test, on rankings of in-person services), we use an ordered logit model with the wide-form data. In this specification, the dependent variable is the reversed ranking of a given service (e.g., 3 = most preferred, 1 = least preferred). We reverse the original ranking so that higher values correspond to higher preference, which simplifies interpretation. The ordered logit assumes an underlying latent utility that determines the observed ordinal outcome through a set of threshold cutoffs. The model estimates how informational privacy concerns shift this latent utility, thereby changing the probability that a service falls into a higher versus lower ranking category.

The coefficients estimated in Table H6, H7 and H8 are are exponentiated and reported as odds ratios for ease of interpretation. Standard errors and p-values are taken from the original logit model on the log-odds scale, because standard errors and hypothesis tests are defined for the estimated coefficients ($\beta$), not for their nonlinear transformation $\exp(\beta)$. Odds ratios themselves do not have separately defined standard errors or p-values; statistical inference is therefore based on the underlying log-odds estimates.

To illustrate the interpretation, in column (4) of Table H6, the odds ratio for $\mathsf{AI} \times \mathsf{IPR}$ is 1.476, indicating that high (vs. low) interpersonal privacy concerns increase the odds of AI being chosen relative to in-person services by approximately 47.6%.

Table H6: The Role of Interpersonal Privacy Concerns in Preferences for Individual Services (Rank-Ordered Logit Results for Online Ranking Experiment)

| Sample: | Full Sample | IPR (High) | IPR (Low) 2 | Full Sample |
|---|---|---|---|---|
| DV: ranking; Coef: exp($\beta$) | (1) | (2) | (3) | (4) |
| AI | 0.138*** | 0.156*** | 0.121*** | 0.113*** |
| | (0.141) | (0.196) | (0.203) | (0.184) |
| Online | 2.009*** | 2.281*** | 1.794*** | 1.800*** |
| | (0.090) | (0.130) | (0.125) | (0.126) |
| convenience | 1.030*** | 1.034*** | 1.026*** | 1.030*** |
| | (0.003) | (0.004) | (0.004) | (0.003) |
| effectiveness | 1.058*** | 1.056*** | 1.061*** | 1.058*** |
| | (0.004) | (0.006) | (0.006) | (0.004) |
| privacy | 1.016*** | 1.020*** | 1.012*** | 1.016*** |
| | (0.003) | (0.004) | (0.004) | (0.003) |
| AI × IPR | | | | 1.476* |
| | | | | (0.219) |
| Online × IPR | | | | 1.245 |
| | | | | (0.177) |
| # Respondents | 650 | 327 | 323 | 650 |
| # Observations | 1,950 | 981 | 969 | 1,950 |

*Note:* Coefficients are exponentiated and reported as odds ratios for ease of interpretation. Standard errors and p-values are taken from the original logit model on the log-odds scale, because standard errors and hypothesis tests are defined for the estimated coefficients ($\beta$), not for their nonlinear transformation exp($\beta$). Odds ratios themselves do not have separately defined standard errors or p-values; statistical inference is therefore based on the underlying log-odds estimates. *p<0.1; **p<0.05; ***p<0.01. The reference alternative for AI and Online is the in-person service. The dependent variables for all columns are ranking, where 1 indicates the highest preference and 3 indicates the lowest preference (see Punj and Staelin (1978); Beggs et al. (1981) for more details). The independent variable IPR refers to the experimental arm with high interpersonal privacy concerns, and is an indicator equal to 1 for interpersonal privacy concerns and 0 otherwise, and IPR is omitted when it does not vary across alternatives. convenience, effectiveness and privacy measure respondents' perceptions about the extent to which each service is convenient, of high quality, of high data privacy and free from stigma, measured on a 1–100 scale (higher values indicate higher convenience, effectiveness, data privacy and lower perceived stigma). convenience, effectiveness, privacy are demeaned.

Table H7: The Role of Information Privacy Concerns in Preferences for AI Services (Ordered Logit Results for Online Ranking Experiment)

| Sample:<br>DV: r.ranking of AI; Coef: exp($\beta$) | Full Sample<br>(1) | IPR (High)<br>(2) | IPR (Low) 2<br>(3) | Full Sample<br>(4) |
|---|---|---|---|---|
| iuipc | $0.633^{***}$ | 0.790 | $0.469^{***}$ | $0.506^{***}$ |
| | (0.131) | (0.175) | (0.202) | (0.188) |
| convenience | $1.024^{***}$ | $1.018^{**}$ | $1.034^{***}$ | $1.024^{***}$ |
| | (0.006) | (0.007) | (0.010) | (0.006) |
| effectiveness | $1.045^{***}$ | $1.041^{***}$ | $1.051^{***}$ | $1.044^{***}$ |
| | (0.005) | (0.006) | (0.007) | (0.005) |
| privacy | $1.007^{*}$ | 1.008 | 1.005 | $1.007^{*}$ |
| | (0.004) | (0.005) | (0.006) | (0.004) |
| IPR | | | | 1.208 |
| | | | | (0.199) |
| iuipc × IPR | | | | $1.524^{*}$ |
| | | | | (0.256) |
| # Observations / Respondents | 650 | 327 | 323 | 650 |

*Note:* Coefficients are exponentiated and reported as odds ratios for ease of interpretation. Standard errors and p-values are taken from the original logit model on the log-odds scale, because standard errors and hypothesis tests are defined for the estimated coefficients ($\beta$), not for their nonlinear transformation exp($\beta$). Odds ratios themselves do not have separately defined standard errors or p-values; statistical inference is therefore based on the underlying log-odds estimates. $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01. The dependent variables for all columns are r.ranking, measuring preference for AI chatbots using a reversed ranking, where 1 indicates the lowest preference and 3 indicates the highest preference. The independent variable IPR refers to the experimental arm with high interpersonal privacy concerns, and is an indicator equal to 1 for interpersonal privacy concerns and 0 otherwise, and IPR is omitted when it does not vary across alternatives. iuipc measures informational privacy concern using the Internet Users' Information Privacy Concerns (IUIPC) scale, which comprises three dimensions—collection, control, and awareness—and is constructed by averaging responses to 10 items measured on a 1 (strongly disagree)–7 (strongly agree) scale. convenience, effectiveness and privacy measure respondents' perceptions about the extent to which AI is convenient, of high quality, of high data privacy and free from stigma, measured on a 1–100 scale (higher values indicate higher convenience, effectiveness, data privacy and lower perceived stigma). convenience, effectiveness, privacy and iuipc are demeaned to facilitate interpretation of the coefficients.

Table H8: Falsification Test: The Role of Information Privacy Concerns in Preferences for In-person Services (Ordered Logit Results for Online Ranking Experiment)

| Sample: | Full Sample | IPR (High) | IPR (Low) 2 | Full Sample |
|---|---|---|---|---|
| DV: r.ranking of In-person; Coef: exp($\beta$) | (1) | (2) | (3) | (4) |
| iuipc | 1.548*** | 1.638*** | 1.458** | 1.505*** |
| | (0.105) | (0.149) | (0.149) | (0.147) |
| convenience | 1.026*** | 1.028*** | 1.024*** | 1.026*** |
| | (0.003) | (0.005) | (0.004) | (0.003) |
| effectiveness | 1.006 | 0.998 | 1.017** | 1.007 |
| | (0.005) | (0.008) | (0.008) | (0.005) |
| privacy | 0.998 | 1.001 | 0.993 | 0.998 |
| | (0.005) | (0.006) | (0.007) | (0.005) |
| IPR | | | | 0.771* |
| | | | | (0.150) |
| iuipc × IPR | | | | 1.069 |
| | | | | (0.204) |
| # Observations / Respondents | 650 | 327 | 323 | 650 |

*Note:* Coefficients are exponentiated and reported as odds ratios for ease of interpretation. Standard errors and p-values are taken from the original logit model on the log-odds scale, because standard errors and hypothesis tests are defined for the estimated coefficients ($\beta$), not for their nonlinear transformation exp($\beta$). Odds ratios themselves do not have separately defined standard errors or p-values; statistical inference is therefore based on the underlying log-odds estimates. *p<0.1; **p<0.05; ***p<0.01. The dependent variables for all columns are r.ranking, measuring preference for in-person services using a reversed ranking, where 1 indicates the lowest preference and 3 indicates the highest preference. The independent variable IPR refers to the experimental arm with high interpersonal privacy concerns, and is an indicator equal to 1 for interpersonal privacy concerns and 0 otherwise, and IPR is omitted when it does not vary across alternatives. iuipc measures informational privacy concern using the Internet Users' Information Privacy Concerns (IUIPC) scale, which comprises three dimensions—collection, control, and awareness—and is constructed by averaging responses to 10 items measured on a 1 (strongly disagree)–7 (strongly agree) scale. convenience, effectiveness and privacy measure respondents' perceptions about the extent to which AI is convenient, of high quality, of high data privacy and free from stigma, measured on a 1–100 scale (higher values indicate higher convenience, effectiveness, data privacy and lower perceived stigma). convenience, effectiveness, privacy and iuipc are demeaned to facilitate interpretation of the coefficients.

# Appendix I Manipulation of Interpersonal Privacy Concerns

This experiment includes a small but meaningful modification to the manipulation used in the ranking experiment. In this experiment, the high-stigma condition presented fabricated information stating that 85% of others held stigmatizing views, whereas the low-stigma condition stated that only 15% held stigmatizing views. In the ranking experiment, the high-stigma condition provides no information about others' stigmatizing views, while the low-stigma condition presents fabricated information that 35% of others hold stigmatizing views.

To implement this manipulation, participants are first asked to respond to two statements about mental health stigma and then estimate the percentage of others who agree with those same statements:

> *Statement 1*: I would view others negatively if I believed that they experienced mental health issues;
>
> *Statement 2*: Others would view me negatively if they believed that I experienced a mental health issue.

Next, participants are told that several hundred responses have already been collected, and they are asked to estimate the percentage of others who agreed with these statements. Participants are then randomly assigned to one of two experimental conditions, where we present them with fabricated agreement percentages to manipulate their perceptions:

> *Reduced stigma group*: Participants see their own responses alongside a fabricated low percentage of agreement among others–5% for Statement 1 and 15% for Statement 2. If a participant overestimates the prevalence of stigma, they are informed of the discrepancy. They are also shown results from a national poll that aligns with these fabricated values. For example, if a participant estimates that 50% of others agree with Statement 2, they are shown that they have overestimated by 35%.
>
> *Heightened stigma group*: Participants see their responses alongside a fabricated high percentage of agreement among others–75% for Statement 1 and 85% for Statement 2. If they

underestimate the prevalence of stigma, they are informed of the difference between their estimate and the fabricated figures. They are also shown national poll results that support these manipulated statistics. For instance, if a participant estimates that only 50% of others agree with Statement 2, they are informed that they underestimated by 35%.

# Appendix J Data and Descriptive Results

Descriptive statistics of several pre-treatment covariates, as well as dependent variables by experiment arms are displayed.

Pre-treatment covariates include negative view of others and negative view of self, which measure participants' initial stigma beliefs. These are based on their agreement with two statements, rated on a 1–5 scale (1 = strongly disagree, 2=somewhat disagree, 3 = neither agree nor disagree, 4=somewhat agree, 5 = strongly agree). Additional covariates include the percentage of participants who report holding negative views of individuals who have mental health difficulties (i.e., their own beliefs), as well as participants' reported estimates of the percentage of peers that they believe will hold negative views of individuals who have mental health difficulties (i.e., perceptions of others' beliefs). Further, we include measures of participants' mental health status (two dummy variables indicating generalized anxiety or depression, based on two pairs of patient survey items that are used in clinical practice to identify mental health challenges among patients, namely GAD-2 or PHQ-2), a dummy variable indicating whether the participant reports ever having made use of mental health service (Ever used services), and a measure of AI utilization (the participant's frequency of AI tool usage in the last month, rated 1–5, ranging from 1 = never, to 5 = several times a day). Finally, demographic controls include age groups (ranging from 1 = 18–24 years, to 6 = 65+ years), annual household income bracket (ranging from 1 = less than $25,000, to 6 = $150,000+), gender, race, educational achievement, marital status, and occupation.

Table J9: Descriptive Statistics (Online Bundled Experiment)

| | All | | L+I | | L+O | | L+AN | | L+AL | | L+AH | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variables: | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| *Covariates* | | | | | | | | | | | | |
| self stigma 1 | 1.427 | 0.823 | 1.451 | 0.804 | 1.404 | 0.757 | 1.447 | 0.883 | 1.447 | 0.801 | 1.471 | 0.903 |
| self stigma 2 | 3.081 | 1.185 | 3.176 | 1.189 | 3.144 | 1.210 | 3.068 | 1.247 | 3.194 | 1.121 | 3.067 | 1.151 |
| % perceived stigma 1 | 28.605 | 21.447 | 30.029 | 25.521 | 28.510 | 20.389 | 27.204 | 21.115 | 25.427 | 19.886 | 29.596 | 20.012 |
| % perceived stigma 2 | 57.377 | 22.382 | 58.451 | 23.170 | 57.981 | 23.703 | 55.320 | 25.489 | 57.961 | 22.037 | 58.135 | 21.227 |
| GAD | 0.483 | 0.500 | 0.500 | 0.502 | 0.519 | 0.502 | 0.534 | 0.501 | 0.544 | 0.501 | 0.452 | 0.500 |
| PHQ | 0.408 | 0.492 | 0.402 | 0.493 | 0.404 | 0.493 | 0.437 | 0.498 | 0.447 | 0.500 | 0.423 | 0.496 |
| Ever used services | 0.761 | 0.427 | 0.686 | 0.466 | 0.788 | 0.410 | 0.767 | 0.425 | 0.786 | 0.412 | 0.827 | 0.380 |
| AI Usage | 1.740 | 1.252 | 1.794 | 1.300 | 1.625 | 1.294 | 1.816 | 1.127 | 1.583 | 1.241 | 1.760 | 1.326 |
| age groups | 1.909 | 1.170 | 2.000 | 1.135 | 1.990 | 1.318 | 1.796 | 1.070 | 1.864 | 1.205 | 1.981 | 1.088 |
| income groups | 2.394 | 1.632 | 2.441 | 1.596 | 2.587 | 1.599 | 2.136 | 1.528 | 2.350 | 1.770 | 2.442 | 1.630 |
| gender ==Male | 0.487 | 0.500 | 0.529 | 0.502 | 0.500 | 0.502 | 0.495 | 0.502 | 0.476 | 0.502 | 0.413 | 0.495 |
| gender ==Female | 0.479 | 0.500 | 0.441 | 0.499 | 0.481 | 0.502 | 0.456 | 0.501 | 0.515 | 0.502 | 0.538 | 0.501 |
| gender ==Others | 0.034 | 0.182 | 0.029 | 0.170 | 0.019 | 0.138 | 0.049 | 0.216 | 0.010 | 0.099 | 0.048 | 0.215 |
| races ==White | 0.778 | 0.416 | 0.775 | 0.420 | 0.856 | 0.353 | 0.767 | 0.425 | 0.777 | 0.418 | 0.750 | 0.435 |
| races ==Black | 0.104 | 0.306 | 0.078 | 0.270 | 0.077 | 0.268 | 0.117 | 0.322 | 0.097 | 0.298 | 0.106 | 0.309 |
| races ==Asian | 0.037 | 0.189 | 0.039 | 0.195 | 0.038 | 0.193 | 0.039 | 0.194 | 0.029 | 0.169 | 0.038 | 0.193 |
| races ==Others | 0.081 | 0.273 | 0.108 | 0.312 | 0.029 | 0.168 | 0.078 | 0.269 | 0.097 | 0.298 | 0.106 | 0.309 |
| educ == College | 0.141 | 0.349 | 0.127 | 0.335 | 0.173 | 0.380 | 0.117 | 0.322 | 0.165 | 0.373 | 0.125 | 0.332 |
| educ == Bachelor or above | 0.336 | 0.472 | 0.373 | 0.486 | 0.288 | 0.455 | 0.350 | 0.479 | 0.330 | 0.473 | 0.337 | 0.475 |
| educ ==Others | 0.523 | 0.500 | 0.500 | 0.502 | 0.538 | 0.501 | 0.534 | 0.501 | 0.505 | 0.502 | 0.538 | 0.501 |
| marital ==Never Married | 0.461 | 0.499 | 0.363 | 0.483 | 0.490 | 0.502 | 0.466 | 0.501 | 0.456 | 0.501 | 0.433 | 0.498 |
| marital ==Married | 0.387 | 0.487 | 0.480 | 0.502 | 0.394 | 0.491 | 0.398 | 0.492 | 0.340 | 0.476 | 0.404 | 0.493 |
| marital ==Others | 0.151 | 0.358 | 0.157 | 0.365 | 0.115 | 0.321 | 0.136 | 0.344 | 0.204 | 0.405 | 0.163 | 0.372 |
| occupation ==Working | 0.749 | 0.434 | 0.725 | 0.448 | 0.750 | 0.435 | 0.709 | 0.457 | 0.718 | 0.452 | 0.750 | 0.435 |
| occupation ==Unemployed | 0.073 | 0.261 | 0.069 | 0.254 | 0.067 | 0.252 | 0.078 | 0.269 | 0.136 | 0.344 | 0.087 | 0.283 |
| occupation ==Others | 0.178 | 0.382 | 0.206 | 0.406 | 0.183 | 0.388 | 0.214 | 0.412 | 0.146 | 0.354 | 0.163 | 0.372 |
| *Dependent Variables* | | | | | | | | | | | | |
| ITU Any Services | 0.757 | 0.429 | 0.647 | 0.480 | 0.750 | 0.435 | 0.786 | 0.412 | 0.718 | 0.452 | 0.856 | 0.353 |
| ITU In-person Services | 0.589 | 0.492 | 0.647 | 0.480 | 0.529 | 0.502 | 0.544 | 0.501 | 0.641 | 0.482 | 0.635 | 0.484 |
| ITU Online Services | 0.690 | 0.463 | - | - | 0.606 | 0.491 | 0.718 | 0.452 | 0.660 | 0.476 | 0.808 | 0.396 |
| ITU AI Chatbot | 0.322 | 0.468 | - | - | - | - | 0.417 | 0.496 | 0.311 | 0.465 | 0.346 | 0.478 |
| # of obs. | 438 | | 48 | | 47 | | 46 | | 52 | | 42 | |

*Note:* 'L' - Reduced stigma, 'H' - Heightened stigma, 'I' - In-person bundle, 'O' - Online bundle, 'AN' - AI with no privacy information bundle, 'AL' - AI with low privacy bundle, 'AH' - AI with high privacy bundle, ITU - intent to use. We restrict the focal sample to participants with moderate or severe mental health issues, as measured by the PHQ-4.

Continued

| | H+I | | H+O | | H+AN | | H+AL | | H+AH | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variables: | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| *Covariates* | | | | | | | | | | |
| self stigma 1 | 1.333 | 0.674 | 1.333 | 0.869 | 1.500 | 0.875 | 1.543 | 1.067 | 1.556 | 0.939 |
| self stigma 2 | 3.178 | 0.936 | 3.462 | 1.022 | 3.521 | 1.091 | 3.829 | 0.923 | 3.444 | 1.206 |
| % perceived stigma 1 | 29.356 | 22.509 | 28.974 | 19.020 | 27.750 | 18.626 | 29.571 | 16.863 | 28.722 | 24.428 |
| % perceived stigma 2 | 56.844 | 22.623 | 63.949 | 19.453 | 63.042 | 21.035 | 63.743 | 18.271 | 58.889 | 22.488 |
| GAD | 4.156 | 1.364 | 4.154 | 1.443 | 4.354 | 1.451 | 4.514 | 1.314 | 4.611 | 1.103 |
| PHQ | 3.756 | 1.448 | 3.923 | 1.421 | 4.271 | 1.349 | 4.543 | 1.221 | 3.861 | 1.417 |
| Ever used services | 0.844 | 0.367 | 0.821 | 0.389 | 0.750 | 0.438 | 0.800 | 0.406 | 0.750 | 0.439 |
| AI utilization | 1.822 | 1.353 | 1.897 | 1.392 | 1.917 | 1.182 | 1.571 | 1.243 | 1.722 | 1.233 |
| age groups | 1.556 | 1.119 | 1.974 | 1.246 | 1.708 | 1.051 | 1.771 | 1.031 | 1.528 | 0.878 |
| income brackets | 1.956 | 1.551 | 1.949 | 1.450 | 2.188 | 1.468 | 2.029 | 1.505 | 2.361 | 1.676 |
| gender ==Male | 0.444 | 0.503 | 0.487 | 0.506 | 0.479 | 0.505 | 0.514 | 0.507 | 0.472 | 0.506 |
| gender ==Female | 0.489 | 0.506 | 0.513 | 0.506 | 0.500 | 0.505 | 0.400 | 0.497 | 0.528 | 0.506 |
| gender ==Others | 0.067 | 0.252 | 0.000 | 0.000 | 0.021 | 0.144 | 0.086 | 0.284 | 0.000 | 0.000 |
| races ==White | 0.800 | 0.405 | 0.667 | 0.478 | 0.771 | 0.425 | 0.857 | 0.355 | 0.861 | 0.351 |
| races ==Black | 0.111 | 0.318 | 0.128 | 0.339 | 0.083 | 0.279 | 0.114 | 0.323 | 0.000 | 0.000 |
| races ==Asian | 0.044 | 0.208 | 0.128 | 0.339 | 0.042 | 0.202 | 0.000 | 0.000 | 0.000 | 0.000 |
| races ==Others | 0.044 | 0.208 | 0.077 | 0.270 | 0.104 | 0.309 | 0.029 | 0.169 | 0.139 | 0.351 |
| educ ==Associate degree | 0.400 | 0.495 | 0.308 | 0.468 | 0.438 | 0.501 | 0.400 | 0.497 | 0.306 | 0.467 |
| educ ==Bachelor or above | 0.422 | 0.499 | 0.615 | 0.493 | 0.354 | 0.483 | 0.514 | 0.507 | 0.528 | 0.506 |
| educ ==Others | 0.178 | 0.387 | 0.077 | 0.270 | 0.208 | 0.410 | 0.086 | 0.284 | 0.167 | 0.378 |
| marital ==Never Married | 0.578 | 0.499 | 0.410 | 0.498 | 0.417 | 0.498 | 0.543 | 0.505 | 0.611 | 0.494 |
| marital ==Married | 0.333 | 0.477 | 0.410 | 0.498 | 0.438 | 0.501 | 0.314 | 0.471 | 0.306 | 0.467 |
| marital ==Others | 0.089 | 0.288 | 0.179 | 0.389 | 0.146 | 0.357 | 0.143 | 0.355 | 0.083 | 0.280 |
| occupation ==Working | 0.578 | 0.499 | 0.718 | 0.456 | 0.813 | 0.394 | 0.743 | 0.443 | 0.806 | 0.401 |
| occupation ==Unemployed | 0.089 | 0.288 | 0.103 | 0.307 | 0.021 | 0.144 | 0.057 | 0.236 | 0.111 | 0.319 |
| occupation ==Others | 0.333 | 0.477 | 0.179 | 0.389 | 0.167 | 0.377 | 0.200 | 0.406 | 0.083 | 0.280 |
| *Dependent Variables* | | | | | | | | | | |
| ITU Any Services | 0.644 | 0.484 | 0.769 | 0.427 | 0.792 | 0.410 | 0.886 | 0.323 | 0.778 | 0.422 |
| ITU In-person Services | 0.644 | 0.484 | 0.487 | 0.506 | 0.625 | 0.489 | 0.600 | 0.497 | 0.667 | 0.478 |
| ITU Online Services | - | - | 0.769 | 0.427 | 0.604 | 0.494 | 0.714 | 0.458 | 0.694 | 0.467 |
| ITU AI Services | - | - | - | - | 0.271 | 0.449 | 0.371 | 0.490 | 0.194 | 0.401 |
| # of obs. | 45 | | 39 | | 48 | | 35 | | 36 | |

*Note:* 'L' - Reduced stigma, 'H' - Heightened stigma, 'I' - In-person bundle, 'O' - Online bundle, 'AN' - AI with no privacy information bundle, 'AL' - AI with low privacy bundle, 'AH' - AI with high privacy bundle, ITU - intent to use. We restrict the focal sample to participants with moderate or severe mental health issues, as measured by the PHQ-4.

# Appendix K Substitution and Complementarity across Service Modalities

Table K10: Substitution and Complementarity across Service Modalities

| Ref. Group: In-person | DV: ITU of In-person (1) | DV: ITU of (+ Online) (2) |
|---|---|---|
| + Online | -0.206*** | |
| | (0.072) | |
| + Online + AI NoP | -0.037 | -0.049 |
| | (0.070) | (0.072) |
| Constant | 0.792** | 0.186 |
| | (0.311) | (0.398) |
| Control Variables | ✓ | ✓ |
| # Observations | 273 | 180 |
| $R^2$ | 0.188 | 0.164 |

*Note:* Standard errors are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Stigma beliefs, perceptions of others, mental status, previous mental health services utilization, previous AI utilization, and other demographic characteristics (age, gender, education, household income, race, marital status, occupation) are incorporated as control variables. NoP stands for no AI privacy information.

## K1 Sample Description of Field Experiment

Table K11 presents the summary statistics for the field experiment. The final sample consists of 333 student cohorts, with student cohort leaders completing the message retrieval survey, resulting in a participation rate of 45% [17]. The average student cohort-level CTR (calculated as number of students who click divided by cohort size) is 1.01%, which is comparable with the average click-through rate on Facebook ads [18]. Covariates incorporated in the regressions include students' year of enrollment (YoE), ranging from 21 (year 2021) to 24 (year 2024), and the percentage of female students in each cohort (Female %).

"'latex "'

Table K11: Descriptive Statistics (Field Experiment)

| | | Cohort Size | Female % | YoE | CTR | N (Cohorts) |
|---|---|---|---|---|---|---|
| All | Mean | 36.048 | 0.445 | 22.153 | 0.010 | 333 |
| | SD | 4.541 | 0.253 | 0.809 | 0.025 | |
| L+I | Mean | 36.820 | 0.452 | 22.060 | 0.010 | 50 |
| | SD | 4.059 | 0.261 | 0.767 | 0.017 | |
| L+O | Mean | 35.175 | 0.438 | 22.150 | 0.006 | 40 |
| | SD | 4.695 | 0.239 | 0.802 | 0.021 | |
| L+AL | Mean | 36.244 | 0.452 | 22.024 | 0.003 | 41 |
| | SD | 4.570 | 0.262 | 0.758 | 0.009 | |
| L+AH | Mean | 36.351 | 0.407 | 22.297 | 0.009 | 37 |
| | SD | 5.443 | 0.254 | 0.878 | 0.015 | |
| H+I | Mean | 35.895 | 0.441 | 22.000 | 0.012 | 38 |
| | SD | 4.235 | 0.236 | 0.870 | 0.019 | |
| H+O | Mean | 35.381 | 0.440 | 22.167 | 0.012 | 42 |
| | SD | 5.446 | 0.268 | 0.794 | 0.036 | |
| H+AL | Mean | 35.837 | 0.484 | 22.512 | 0.016 | 43 |
| | SD | 3.909 | 0.240 | 0.736 | 0.037 | |
| H+AH | Mean | 36.524 | 0.435 | 22.024 | 0.011 | 42 |
| | SD | 4.008 | 0.273 | 0.811 | 0.030 | |

*Note:* 'L' - reduced stigma, 'H' - heightened stigma, 'I' - In-person consultation, 'O' - Online consultation, 'AL' - AI with low privacy, 'AH' - AI with high privacy. YoE represents year of enrollment, ranging from 21 to 24. CTR represents click-through rate.

[17] There is no significant difference in the student cohort-level characteristics, such as student cohort gender ratio and average years of enrollment between the participants and non-participants.

[18] Mark Irvine, "Facebook Ad Benchmarks for YOUR Industry," WordStream, August 27, 2019, https://www.wordstream.com/blog/ws/2017/02/28/facebook-advertising-benchmarks.

# Appendix L Generalizability for Bundled Experiment: The Effects of AI Informational Privacy Policies

Table L12: The Role of Information Privacy Concerns in Preferences for Bundled Services (Field Experiment)

| Ref. Group: AI Privacy<br>DV: CTR Any Services | Full Sample<br>(1) | IPR (High)<br>(2) | IPR (Low)<br>(3) | Full Sample<br>(4) |
|---|---|---|---|---|
| AI Personalization | -0.000 | 0.004 | -0.006** | -0.005* |
| | (0.005) | (0.009) | (0.002) | (0.003) |
| IPR | | | | 0.001 |
| | | | | (0.005) |
| AI Personalization × IPR | | | | 0.009 |
| | | | | (0.008) |
| Constant | 0.010*** | 0.011** | 0.010*** | 0.009*** |
| | (0.003) | (0.005) | (0.001) | (0.002) |
| Control Variables | ✓ | ✓ | ✓ | ✓ |
| # Student Cohorts | 162 | 84 | 76 | 162 |
| $R^2$ | 0.093 | 0.134 | 0.159 | 0.115 |

*Note:* *AI Personalization* includes three services: In-person and telehealth services, plus a conversational AI chatbot emphasizing informational privacy assurances. *AI Privacy* includes three services: In-person and telehealth services, plus a conversational AI chatbot emphasizing personalization at the expenses of information privacy (see appendix for full texts of privacy policy manipulation). This regression is conducted at the student cohort level. Dependent variable CTR is the student cohort-level click-through rate, calculated as number of students who click divided by cohort size. Standard errors in parentheses, are clustered at the school (college) level. ***, ** and * represent statistical significance at the 1%, 5% and 10% levels, respectively. Control variables include school (college) fixed effects. 1, 1, 2 and 1 singleton observations are dropped for column (1), (2), (3) and (4) respectively due to adding fixed effects. IPR refers to the experimental arm with high interpersonal privacy concerns.

# Appendix M Excerpts from Reddit Posts and Google Play Reviews on Mental Health AI Chatbots

**AI is An Effective Therapist**

- ChatGPT (v4) was a really good therapist. I could share my traumatic memories and talk about my anxiety and it would reply spot on like a well trained therapist. I felt very often so relieved after a short "session" with it. https://www.reddit.com/r/ChatGPT/comments/14hfr5y/i_felt_so_blessed_i_can_use_chatgpt_as_my/

**AI's Non-Judgmental Nature in Combating Stigma**

- I feel you, it's like: you can be vulnerable, you can ask everything, no one judges you AND you learn. A lot of perspectives and other stuff. https://www.reddit.com/r/ChatGPT/comments/1f85dl7/comment/llcn7ba/

**Data Privacy Concerns**

- The privacy policy and terms and conditions state that they can personally identify you and can share all your data. Just don't tell it anything you wouldn't want out in the public domain. https://www.reddit.com/r/ChatGPT/comments/11enzge/comment/jagq9m8/

**AI as a Complement not a Substitute**

- ...this will shift to a compulsatory need to consult an AI or LLM before beginning a therapy. https://www.reddit.com/r/ChatGPT/comments/12cr4cw/comment/jf3hcqi/

- ...LLMs should not replace a trained therapist (yet). https://www.reddit.com/r/ChatGPT/comments/174k4eg/comment/k4afkp0/

**Overlooking Privacy to Mitigate Stigma**

- I am grateful there are no humans in this loop. I am fine with my data being aggregated with others for study. It is great talking with professionals, and that is offered. The AI alone, however, is an incredible tool for self reflection if used honestly. https://play.google.com/store/apps/details?id=bot.touchkin&hl=en_US