Macroeconomics of Mental Health*

Boaz Abramson
Columbia Business School

Job Boerma
University of Wisconsin-Madison

Aleh Tsyvinski
Yale University

January 2025

Abstract

We develop a quantitative macroeconomic theory of mental health. The theory is grounded in classic and modern psychiatric literature, is disciplined with micro data, and is formalized in a life-cycle heterogeneous agent framework. In our model, individuals experiencing mental illness have negative expectations and lose time due to rumination. As a result, they work less, consume less, invest less in risky assets, and forego treatment which in turn reinforces their mental illness. We use the model to evaluate the effects of prominent mental health policies. We show that expanding the availability of treatment services and improving treatment of mental illness in late adolescence substantially improve mental health and welfare.

*We thank our discussants Karen Kopecky at the OIGI Conference and Dirk Krueger at the NBER EFG Meeting in Fall 2024 for helpful comments. We thank Anmol Bhandari, Adam Blandin, Corina Boar, Jaroslav Borovička, Diego Daruich, Mariacristina de Nardi, Eric French, Jeremy Greenwood, Roozbeh Hosseini, Quentin Huys, Rishabh Kirpalani, Rory McGee, Ellen McGrattan, Luigi Pistaferri, José-Victor Rios-Rull, Tom Sargent, Ananth Seshadri, Martin Schneider, Kjetil Storesletten, and Pierre Yared for useful discussions.

1 Introduction

Mental illness is widespread and costly. In the U.S., more than 20 percent of adults live with mental illness and approximately 5.5 percent experience serious mental illness (SAMHSA, 2022). Depression and anxiety, the most common mental illnesses, account for 8 percent of years lived with disability globally (GBD, 2018). Policymakers are increasingly implementing policy initiatives to improve mental health, for example by expanding access to treatment or by lowering out-of-pocket costs of mental health services.

We develop a quantitative macroeconomic theory of mental health. The theory is based on classic and modern psychiatric literature and is formalized in a dynamic life-cycle heterogeneous agent economy. We discipline the theory with micro-level data and show how mental illness alters consumption, savings, portfolio choice, and labor supply. We use this framework to evaluate prominent policy proposals.

Our economic framework of mental health builds on both the psychiatric literature and on quantitative macroeconomic models of health. Psychiatric theories identify three common features of mental illness: negative thinking, rumination, and reinforcement through behavior.[1] First, we model negative thinking as individuals having negative expectations. Second, we model rumination, a repetitive and uncontrollable preoccupation with negative thoughts, as loss of available time. The third feature is that mental illness reinforces itself through behavior. For example, individuals experiencing mental illness can choose to seek treatment, but negative thinking about its efficacy and rumination may deter them, perpetuating mental illness. We model treatment decisions, which generates self-reinforcing behavior of mental illness. Consistent with the psychiatric and structural macroeconomic health literature, we further extend the model with the stochastic evolution of mental health, the impact of mental health on labor productivity, the influence of labor market experiences on mental health, and stigma costs associated with treatment.

We formalize our economic theory of mental illness in a lifecycle model with heterogenous agents. Individuals choose consumption, labor supply, and investments in risk-free and risky assets. Mental health is a stochastic state variable that governs negative thinking, rumination, treatment efficacy, and labor productivity. We model negative thinking building on the cognitive model of depression (Beck, 1967, 1976, 2002, 2008) and the clinical and neuroscience literature supporting it (Clark et al., 2000; Mathews and MacLeod, 2005; Disner, Beevers, Haigh, and Beck, 2011; Beck and Bredemeier, 2016).[2] Individuals

[1] We model mental illness focusing on depression and anxiety, the most prevalent mental illnesses around the world (GBD, 2018). Our model also captures salient aspects of a variety of other mental conditions, such as impulse control disorders and PTSD, as they share mechanisms and symptoms with and are comorbid to depression and anxiety (Kessler, Chiu, Demler, and Walters, 2005).

[2] Mental health is distinct from physical health in two main dimensions (see also Dirk Krueger's NBER EFG "Discussion on Macroeconomics of Mental Health" (web.sas.upenn.edu/dkrueger/). First, the cognitive distortion is the distinguishing feature of mental health relative to physical health. Second, mental illness is more prevalent

experiencing mental illness have more negative expectations over the realizations of uncertain outcomes. Mental illness decreases the subjective probability assigned to favorable outcomes while increasing the subjective probabilities assigned to less favorable outcomes. Individuals experiencing mental illness expect lower productivity, lower returns on risky investments, and have a negative view of the efficacy of mental health treatment. The second aspect of mental illness is rumination (Nolen-Hoeksema, 1991; Just and Alloy, 1997; Nolen-Hoeksema, 2000; Nolen-Hoeksema, Wisco, and Lyubomirsky, 2008; Singer and Dobson, 2007) which we model as losing a portion of available time. The third aspect of mental illness is reinforcement through behavior. We model this through the treatment decision. Treatment increases the probability of transitioning into better mental health but is costly. Individuals experiencing mental illness may choose to not seek treatment as they expect treatment to be ineffective due to negative thinking and since they have less time due to rumination. Mental illness is thus reinforced through the treatment decisions. The fourth element of mental health that we incorporate is productivity losses associated with mental illness.

We next quantify the model. We model three mental health states: healthy, mild illness, and serious illness. First, we parameterize the extent of negative thinking for each mental health state. The idea is to use differences in subjective probabilities by mental health state in the data to inform negative thinking in the model. We operationalize this by estimating differences in subjective loss probabilities by mental health status using micro-level data from RAND's American Life Panel (ALP). The subjective loss probabilities are elicited using the classic Ellsberg urn problem. We show that subjective loss probabilities increase with the severity of mental illness. Individuals experiencing mild (serious) mental illness have a 3.1 (6.7) percentage point higher subjective loss probability. Using these estimates, we calibrate the extent of negative thinking to be 3.1 (6.7) percent for individuals experiencing mild (serious) mental illness.

Second, we estimate the mental health transition probability matrix, which depends on the treatment decision and idiosyncratic productivity. For quantification, we use biannual transition probabilities between mental health states from the Panel Study for Income Dynamics (PSID), population shares and treatment propensities across mental health states obtained from the National Institute of Mental Health (NIMH) as well as estimates on the impact of treatment from the medical literature (Ekers, Richards, and Gilbody, 2008; Barth et al., 2016). Consistent with these treatment effects, the first takeaway of our estimation is that treatment is effective. For example, the probability to transition from serious mental illness to the healthy state is 12.4 percent without treatment, while 33.2 percent with treatment. The

early in life, relative to most physical illnesses, which are more common later in life.

second takeaway is that negative labor market shocks increase the likelihood to experience mental illness in the future. For example, the likelihood to transition from the healthy state to the serious (mild) illness state is 1.1 (5.5) percent in normal productivity states but equals 1.9 (7.3) percent in low productivity states.

We calibrate remaining parameters so that model moments align with the data. Using the panel structure of the PSID data, we calibrate the impact of rumination on available time to match observed changes in working hours associated with changes in mental health. Individuals experiencing mild (serious) mental illness work 4.7 (12.7) percent fewer hours in the data. With 6.7 (11.1) hours per week of lost time to rumination for mild (serious) mental illness, the model aligns with the data. We calibrate the utility cost of treatment to match the share of seriously ill who receive treatment. We target the estimate of the NIMH that 65.4 percent of those who are seriously ill receive treatment. We assume treatment is not available to a fraction of individuals when they are mildly ill. This captures the fact that availability is one of the most commonly cited barriers to treatment.[3] We calibrate this fraction so that the share of mildly ill receiving treatment is equal to 41.4 percent as in the data. This implies that one-third of the population does not have access to treatment when mildly ill.[4]

We validate the model by evaluating how it compares to non-targeted moments that describe the relation between mental health and economic outcomes. First, we evaluate the model predictions for average consumption, hours worked, income, wealth and risky investments by mental health status. The model captures almost perfectly consumption, hours worked, income levels, and risky participation rates by mental health group. The model generates two thirds of the decrease in wealth by mental health. Second, the model captures well the distributions of consumption, income, and portfolio allocations by mental health. For example, the income distribution among healthy individuals is skewed to the right, while the income distribution among individuals experiencing serious illness is skewed to the left due to working fewer hours. Finally, we validate the model with regression evidence from the PSID. The conditional correlations between consumption and mental health, and between risky investments and mental health, align with the data. For example, all else equal, individuals with mild mental illness consume 2.2 (1.9) percent less and with serious illness consume 6.5 (3.3) percent less than healthy individuals in the model (data). Importantly, we show that without negative thinking there is no motive for individuals

[3]See the White House Fact Sheets (www.whitehouse.gov/s1, www.whitehouse.gov/s2, www.whitehouse.gov/s3) and workforce data from the United States Department of Health and Human Services (www.hrsa.gov) and from the American Psychological Association (www.apa.org).

[4]This coincides with estimates of the number of individuals whose treatment needs are not met according to the United States Department of Health and Human Services that we discuss in footnote 37.

experiencing mental illness to consume less or invest less in risky assets, and these regression coefficients would instead be close to zero.

Having quantified our theory of mental illness, we discuss its implications. In order to have a benchmark to evaluate the effects of policies, we first estimate the aggregate welfare costs of mental illness. We find an aggregate cost of mental illness equal to 1.2 percent of consumption annually. The average cost of mental illness for individuals with serious mental illness is 13.3 percent and is 6.9 percent for those with mild illness.

We then evaluate the effects of three prominent mental health policies: expanding the availability of treatment, lowering out-of-pocket treatment costs, and improving mental health in late adolescence and young adulthood. First, we consider expanding the availability of treatment. We evaluate a policy that makes treatment available to all individuals. Expanding availability of mental health treatment services reduces mental illness by 1.6 percentage points. This reduction in mental illness is driven by a strong increase in the treatment share among individuals experiencing mild illness, to 69.9 percent from 41.4 percent. The welfare benefits of providing full access to treatment services is equivalent to 0.31 percent of aggregate consumption. The welfare gains are largest for individuals who are mildly ill and do not have access to treatment in the benchmark economy. Importantly, healthy individuals also experience gains due to improved access in case they experience mental illness in the future.

Second, we consider the implications of a policy under which individuals do not pay out-of-pocket for their treatment. We find that the welfare benefit of reducing out-of-pocket costs is equivalent to 0.16 percent of aggregate consumption. By comparing this result with the significant welfare benefits of increasing treatment availability, we conclude that lack of availability rather than affordability is the most salient barrier for mental health treatment. Third, we consider a policy that improves mental health treatment in late adolescence and young adulthood. Specifically, we change the initial distribution of mental health assuming all individuals between age 16 and 25 receive treatment when they experience mental illness. Treatment of young adults improves the mental health of 25 year olds, which translates into an aggregate consumption equivalent gain of 0.95 percent annually among young adults.

We show that the quantitative results are robust to various model specifications and parameter choices. The results are robust to including negative thinking for all individuals (that is, not only those who experience mental illness), to incorporating utility penalties of mental illness, and to allowing for ex-ante unobserved heterogeneity in types. Moreover, the results are largely invariant to changes in the level of borrowing constraints and the labor productivity elasticity. We further show that the welfare costs of mental illness and the benefits of policies such as expanded treatment availability are primarily driven by negative thinking and rumination.

Finally, we quantify the value of improving the efficacy of mental health treatment, for example due to advances in therapy or anti-depressant medication. We re-estimate the mental health transition matrix when treatment is 25 percent more effective. The aggregate consumption equivalent gain of this improvement in treatment is 0.7 percent.

Literature. The main contribution of our paper is to develop a quantitative macroeconomic model of mental health. There is a rich literature, starting with Grossman (1972), that studies macroeconomic models of health (Hubbard, Skinner, and Zeldes, 1995; French, 2005; Hall and Jones, 2007; Low, Meghir, and Pistaferri, 2010; De Nardi, French, and Jones, 2010; French and Jones, 2011; Kopecky and Koreshkova, 2014; Low and Pistaferri, 2015; De Nardi, French, and Jones, 2016; Braun, Kopecky, and Koreshkova, 2017, 2019; Cole, Kim, and Krueger, 2019; Ameriks, Briggs, Caplin, Shapiro, and Tonetti, 2020; Fang and Krueger, 2022; Greenwood, Guner, and Kopecky, 2022; Hosseini, Kopecky, and Zhao, 2024). We build on this literature and explicitly incorporate the key cognitive distortions associated with mental illness identified by the psychiatric literature. Agents experiencing mental illness hold pessimistic expectations on the future, and this negative view contributes to the observed symptoms of mental illness and to its persistence. By modeling mental illness as a cognitive distortion, we merge a rich psychiatric literature that perceives mental illness as predominantly characterized by negative cognitive biases with the macroeconomics and health literature. We discuss the psychiatric literature that provides the foundation for our model in Section 2.

Our economic theory of mental health is related to the literature on multiple priors and ambiguity aversion (Gilboa and Schmeidler, 1989; Epstein and Schneider, 2003; Ilut and Schneider, 2014; Ilut, Valchev, and Vincent, 2020; Ilut and Valchev, 2023; Bhandari, Borovička, and Ho, 2024). In our model, individuals experiencing more severe mental illness behave as if they are more ambiguity averse. That is, they consider a larger set of multiple priors regarding the probability distribution of future states and evaluate their choices according to the worst prior in this set. Modeling more negative expectations as a key feature of mental illness is motivated by classic and modern psychiatric theories emphasizing that individuals who experience mental illness deem negative outcomes to be more likely relative to healthy individuals (see Section 2). Using survey data, we show that mental illness is positively associated with ambiguity aversion. We then use these empirical moments to identify the dependence of ambiguity aversion on mental health in our quantitative model. We also allow for ambiguity aversion among healthy individuals. In line with the data, individuals think more negatively, or are more ambiguity averse, when they experience mental illness.

Our quantitative framework builds more broadly on a large literature on consumption, savings, and labor supply over the life-cycle (Rios-Rull, 1996; Gourinchas and Parker, 2002; Cocco, Gomes, and Maenhout, 2005; Gomes and Michaelides, 2005; Heathcote, Storesletten, and Violante, 2010; Low, Meghir, and Pistaferri, 2010; Huggett, Ventura, and Yaron, 2011; Heathcote, Storesletten, and Violante, 2014; Blundell, Pistaferri, and Saporta-Eksten, 2016; Fagereng, Gottlieb, and Guiso, 2017; Boar, 2021). We contribute to this literature by introducing mental health into an otherwise standard life-cycle model of consumption, savings, portfolio choice, and labor supply, and by quantifying the relationship between mental health and these economic outcomes.

Finally, by evaluating the aggregate welfare costs of mental illness, we contribute to the epidemiological literature that quantifies the aggregate costs of mental disorders (Greenberg et al., 2003; Kessler et al., 2009; Greenberg et al., 2015). Our results suggest that these estimates, which are frequently cited by policymakers to provide justification for increasing funding for mental health services, are downward biased. The epidemiological literature focuses primarily on the static income penalty associated with mental illness and the monetary costs associated with treating mental illness. By developing an economic model of mental health, we are able to quantify not only these costs, but also how mental health affects consumption, job choice, savings and portfolio choice, how this dynamically translates to improved lifetime trajectories, and how individuals value better mental health. Our estimates imply that the welfare costs of mental illness is 24 percent larger than the estimates from the epidemiological literature.

2 Psychiatric Literature

This section provides an overview of elements in the psychiatric literature that provide the foundation for our economic model of mental illness.

Negative Thinking. The first feature of mental illness that is highlighted by the psychiatric literature is negative thinking. The predominant psychiatric theory of depression is Beck's cognitive model of depression. Beck's theory posits that depression is primarily a cognitive disorder characterized by negative thinking (Beck, 1967, 1976, 2002, 2008). Depressed patients hold a negative view of the self, the future, and the past – commonly referred to as the negative cognitive triad. These negative thoughts are responsible for many of the observed symptoms of depression such as inaction, sadness, hopelessness, and loss of initiative. Negative thinking is not only a hallmark of depression but is also considered a key cognitive bias in other mental disorders such as anxiety disorders, PTSD, and psychosis (Beck, Emery, and Greenberg, 1985; Eysenck, 2014; Ehring and Watkins, 2008; Beck and Clark, 1991).

Clinical research in psychology provides extensive empirical support for negative thinking among individuals experiencing mental illness (see Clark et al. (2000), Mathews and MacLeod (2005), and Beck (2008) for reviews). Depressed and anxious patients negatively interpret ambiguous stimuli (Butler and Mathews, 1983; Muris and van der Heiden, 2006), suffer from repetitive negative thinking (Watkins, 2008), selectively attend to negative aspects of experiences (Derryberry and Reed, 2002; Gotlib, Krasnoperova, Yue, and Joormann, 2004), and overgeneralize and self-attribute negative realizations (Phillips, Hine, and Thorsteinsson, 2010). A recent literature in behavioral genetics and cognitive neuroscience has provided further support for the cognitive model. Due to advances in genetics and neuroimaging, this literature has identified a number of neurobiological correlates of depression that associate with negative thinking (see Disner, Beevers, Haigh, and Beck (2011) and Beck and Bredemeier (2016) for reviews). We build on the psychiatric theory and the clinical and neuroscience literature that supports it and model negative thinking as a feature of mental illness.

Rumination. A second feature of mental illness that is highlighted by the psychiatric literature is rumination. Formalized by the response styles theory (Nolen-Hoeksema, 1991; Nolen-Hoeksema, Wisco, and Lyubomirsky, 2008), rumination is defined as an uncontrollable and repetitive preoccupation with one's negative thoughts. The theory posits that individuals experiencing mental illness spend excessive amounts of time ruminating on negative thoughts. Rumination in turn disrupts behavior and decision making and is recognized as a main driver of the symptoms of depression. More recent psychiatric theories of cognitive control also support the link between rumination and depression. Rumination is regarded as a maladaptive emotion regulation strategy that is due to deficits in cognitive control (Gotlib and Joormann, 2010; Le Moult and Gotlib, 2019). Depressed individuals experience difficulties in controlling the content of their working memory – a cognitive system with a limited capacity that is important for reasoning and behavior (Williams, Watts, MacLeod, and Mathews, 1988; Mathews and MacLeod, 2005). Instead of disengaging from negative information, depressed individuals spend their time ruminating on it. Similar impairments in cognitive control that manifest through rumination are observed in other mental disorders, such as anxiety, schizophrenia, and personality disorders (Burt, Zembar, and Niederehe, 1995; Nolen-Hoeksema, 2000; Watkins, 2008).

A large body of work provides empirical support for the key role of rumination in mental disorders. Research in clinical psychology has connected rumination with the onset and duration of depression (see, for example, Just and Alloy (1997), Nolen-Hoeksema (2000), Singer and Dobson (2007)). Individuals who ruminate more about their negative mood experience longer and more severe depression spells. Rumination has also been shown to predict the severity of anxiety symptoms and the duration of anxiety spells (Ehring and Watkins, 2008). Recent advances in cognitive neuroscience provide further empirical evidence for the connection between rumination and depression. For example, rumination is strongly associated with neurobiological correlates of depression (Disner, Beevers, Haigh, and Beck, 2011). We build on the psychiatric theory and the clinical and neuroscience literature that supports it and model rumination as a feature of mental illness that leads to time loss.

Reinforcement Through Behavior. A third feature identified by both classic and modern psychiatric theories is that mental illness reinforces itself through behavior. In Beck's cognitive model of depression, individuals experiencing mental illness exhibit reduced motivation to engage in goal-directed or problem-solving activities due to negative expectations over the outcome of such activities. In theories of rumination, excessive elaboration on one's negative thoughts similarly discourages individuals from taking action that might benefit their mental health (Nolen-Hoeksema, 1991). For example, individuals who experience mental illness might not seek treatment because they think negatively about the efficacy of mental health services and because ruminating preoccupies their time. This inaction in turn reinforces mental illness.

Reinforcement through behavior is also at the center of computational psychiatry. This interdisciplinary field combines computational and mathematical tools with neuroimaging and clinical data to study mental illness (see Adams, Huys, and Roiser (2015), Huys, Maia, and Frank (2016) and Bishop and Gagne (2018) for reviews). In computational psychiatry, mental illness is characterized as a range of distortions in the evaluation of costs and benefits of actions that persist through self-reinforcement. Individuals experiencing mental illness hold negative expectations of future outcomes − they underestimate the likelihood of positive outcomes and overestimate the likelihood of negative outcomes. This leads to inaction which in turn implies that negative thinking is reinforced. In line with the classic and modern psychiatric theories, mental illness in our model reinforces itself through behavior. Individuals experiencing mental illness can choose to seek treatment but negative thinking and rumination may deter them from doing so.

Additional Features. While our theory focuses on the role of three key features of mental illness − negative thinking, rumination, and reinforcement through behavior − mental illness is multi-faceted and complex. We further capture and discuss several prominent additional psychiatric features of mental illness in our model.

First, the psychiatric literature emphasizes that mental illness is stochastic and triggered by adverse events such as dissolution of relationships, death of a loved one, or economic hardship (Caspi et al., 2003; Kendler et al., 2005; Beck, 2008). This literature is complemented by an economic literature that documents that negative economic shocks undermine mental health (see Ridley, Rao, Schilbach, and Patel (2020) for a review).[5] In our framework, mental health is stochastic, and negative economic shocks lead to worse mental health.

Second, mental illness might impact flow utility. Self-reported anhedonia, the inability to derive pleasure, is a diagnostic criteria of mental disorders in the American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). At the same time, a number of studies finds that individuals experiencing mental illness are not less sensitive to pleasure or more sensitive to pain (Amsterdam et al., 1987; Schaefer et al., 2010). Computational psychiatry (Dichter et al., 2010; Huys, Daw, and Dayan, 2015; Bishop and Gagne, 2018) reconciles these findings with the DSM-IV by highlighting that anhedonia is not due to deficits in flow utility, but rather due to low expected utility over future outcomes. Consistent with this literature, in our main model specification we do not incorporate a direct utility penalty of mental illness. We then consider an extension where flow utility directly depends on mental health in line with a number of structural economic models where utility directly depends on physical health (French, 2005; De Nardi, French, and Jones, 2010).

Third, mental illness might not only be a consequence of negative labor market experience but can also impact labor market outcomes (Currie and Madrian, 1999; Ridley, Rao, Schilbach, and Patel, 2020). Our model incorporates this feature in two ways. First, consistent with Beck's cognitive model and with theories of rumination, individuals with mental illness have negative expectations over their future labor productivity and lose time due to rumination, which leads them to choose less demanding jobs and to work less hours. Second, consistent with the structural macro health literature (French, 2005; French and Jones, 2011; De Nardi, Pashchenko, and Porapakkarm, 2024), our model incorporates a direct labor productivity penalty due to mental illness.

Finally, psychiatric theory documents that there is heterogeneity across individuals in their vulnerability to mental illness (Beck, 2002; Mathews and MacLeod, 2005). Vulnerability depends on genetic factors as well as early childhood experience. This view is supported by epidemiological evidence showing that mental illness tends to relapse (Hardeveld et al., 2010; Richards, 2011; Kessler et al., 2012). Our theory therefore incorporates both observed and unobserved sources of heterogeneity in individuals'

[5]Negative economic shocks also lead to worse physical health outcomes (Currie and Madrian, 1999; Sullivan and Von Wachter, 2009; Davis and Wachter, 2011).

vulnerability to mental illness.[6]

Treatment. Cognitive behavioral therapy (CBT), the current standard in psychotherapy, is grounded in Beck's cognitive model and in theories of rumination. CBT aims to change negative thinking patterns by helping patients understand their thinking and behavior, and by providing tools to change distorted beliefs (Beck, 1976; Dobson and Dozois, 2019). CBT guides them to disengage from negative information and regain cognitive control. Consistent with CBT, treatment in our model, if successful, reduces negative thinking and rumination.

A vast medical literature estimates the effects of psychotherapy treatment, as well as of anti-depressant medication, on mental health using randomized trials. The treatment effect sizes are typically standardized to facilitate comparison across different studies. Specifically, they are reported in terms of the standardized mean difference (SMD). Both therapy and anti-depressants are generally found to be effective treatment options, with therapy being more effective than anti-depressants. A meta-analysis by Ekers, Richards, and Gilbody (2008) reports an average SMD of -0.70 for behavioral psychotherapy. For antidepressants, the meta-analysis of Turner et al. (2008) shows an average SMD of -0.37. These treatment effects inform the efficacy of treatment in our quantitative model.

Despite the efficacy of mental health treatment, a relatively low share of individuals experiencing mental illness seek treatment. The NIMH estimates that, in 2021, only 65.4 (41.4) percent of individuals with serious (mild) mental illness receive treatment.[7] The medical literature identifies several possible explanations for the low take-up. First, individuals experiencing mental illness may have negative expectations over the efficacy of treatment. Second, lack of availability of mental health services is one of the most commonly cited barriers to treatment (see Section 5.2.1). Third, even when mental health treatment is available, it might be unaffordable (see Section 5.2.2). Fourth, stigma is an important factor contributing to low treatment rates of mental illness despite the efficacy of treatment (see, for example, Corrigan (2004) and Clement et al. (2015)). Our model incorporates these barriers to mental health treatment

[6]Our estimates of the welfare costs of mental illness and the benefits of mental health policies are underestimates to the extent that we omit several factors related and influenced by mental health – substance abuse (Greenwood, Guner, and Kopecky, 2022), homelessness (Abramson, 2024; Abramson and van Nieuwerburgh, 2024; Imrohoroglu and Zhao, 2024; Corbae, Glover, and Nattinger, 2024), and suicide (Greenberg et al., 2003, 2015). Opioid misuse is 10.3 percent among individuals experiencing severe mental illness and 2.2 percent among individuals who are healthy (SAMHSA, 2022). According to the 2022 Annual Homelessness Assessment Report to Congress, among the 0.5 percent of individuals who experience homelessness 21 percent experiences severe mental illness. The suicide rate in the U.S. is 0.014 percent.

[7]The 2021 National Survey on Drug Use and Health documents that 22.8 percent of U.S. adults experience any mental illness, for which 47.2 percent receives treatment. Furthermore, 5.5 percent of adults experience a serious mental illness, for which 65.4 percent receives treatment. As a consequence, 41.4 percent of adults experiencing a mild illness receives treatment as $\frac{5.5}{22.8} \times 65.4 + (1 - \frac{5.5}{22.8}) \times 0.41 = 47.2$.

and we use the model to evaluate the efficacy of interventions designed to alleviate these barriers.

3 Model

We formalize our economic theory of mental health in a lifecycle model with heterogeneous agents. We consider an infinite horizon economy populated by overlapping generations, each of mass one. Individuals live for T periods. Time is discrete. Age is denoted by $t = 1, 2, \ldots, T$.[8]

Preferences. Individuals derive flow utility $u(c, \ell)$ from consumption c and leisure ℓ. Individuals have preferences which are separable in time and discount the future with a constant discount factor β. Total time each period is normalized to one.

Mental Health. Mental health status is denoted by $m \in \mathcal{M}$, where $\mathcal{M}$ is a finite set. We consider a specification with three mental health states: a healthy state m_0, a mild illness state m_1, and a serious illness state m_2. Individuals draw an initial mental health state from a distribution π_m. Mental health evolves according to a first-order Markov chain with conditional transition probabilities $\Gamma_m(\tau_t, \nu_t)$ that depend on the treatment choice τ_t and the idiosyncratic labor productivity ν_t.[9] Negative labor market shocks can thus affect mental health. Mental health governs negative thinking, rumination, the efficacy of treatment, labor productivity, and a flow utility cost.

Negative Thinking. The distinctive feature of mental health relative to physical health is negative thinking. Building on the cognitive model of mental illness discussed in Section 2, we model negative thinking as negative expectations over random outcomes.

Consider the following example to illustrate how we model negative thinking. Let $w(\chi)$ denote the value associated with a random outcome χ in a finite set of realizations Ω_χ. Let $q(\chi)$ be the objective probability of the outcome. Negative thinking is represented by individuals forming their expectation over the random outcome according to:

$$\min_p \; \mathbb{E}_p w(\chi) = \min_{p(\chi)} \sum_{\chi \in \Omega_\chi} p(\chi) w(\chi). \tag{1}$$

That is, individuals form expectations based on the subjective probability distribution that minimizes their expected value. If this minimization problem was unconstrained, individuals would put a probability

[8] We consider a stationary economy, hence, time is left implicit and variables are indexed only by age t.

[9] Jolivet and Postel-Vinay (2024) propose a search model of labor market trajectories and mental health dynamics and study the bi-directional relationship between mental health and labor market outcomes and the effects of job loss, mental health shocks and job stress shocks.

one on the worst state, $\underline{\chi}$. Minimization, however, is subject to the following total variation constraint that limits the choice of subjective probabilities to those that are close to the objective probabilities:[10]

$$\frac{1}{2} \sum_{\chi \in \Omega_\chi} \left| p(\chi) - q(\chi) \right| \leqslant \kappa(m). \tag{2}$$

The extent to which subjective probabilities can differ from objective probabilities is thus governed by κ, which represents the degree of negative thinking. For example, if $\kappa = 0$, subjective probabilities are equal to the objective probabilities, and there is no negative thinking. The solution to the minimization problem is that individuals put as much mass as possible on the lowest state by lowering the probability of the best outcomes. The subjective probability of the worst state is:

$$p^*(\underline{\chi}) = q(\underline{\chi}) + \kappa(m). \tag{3}$$

The larger κ, the more negatively individuals think about the future. Notably, the degree of negative thinking is a function of mental health. In the calibration, individuals experiencing more severe mental illness think more negatively about the future, that is, κ is increasing with the severity of mental illness.

Importantly, a powerful feature of this approach is that it does not require the underlying dimension of uncertainty to be unidimensional as it can be applied to any joint distributions over outcomes. We exploit this feature in the decision problem where individuals face uncertainty about returns on risky assets and uncertainty about the evolution of mental health.[11]

Rumination. Available time varies with mental health. As discussed in Section 2, a prominent feature of mental illness is rumination. We model rumination as a reduction of time available for work, leisure, and treatment. Specifically, individuals with mental health m lose $n_r(m)$ hours due to rumination. Available time for work, leisure, and treatment is therefore $1 - n_r(m)$.

Treatment Choice. Individuals decide whether to get treatment. We denote by $\tau_t = 0$ if the individual does not undertake treatment, and by $\tau_t = 1$ if the individual undertakes treatment. Treatment increases the probability of transitioning into better mental health states. An individual going into treatment incurs a time cost n_τ, a financial cost φ_τ, and a utility cost ξ_τ. As a result, time available for leisure and

[10] The total variation distance between probability measures P and Q is $\delta(P,Q) = \max |P(A) - Q(A)|$, that is, the largest possible difference between the probabilities that the two probability measures assign to some event A. For our discrete domain, this is equivalent to half of the taxicab distance between the probability mass functions. One could employ alternative distances between probability distributions such as the relative entropy. The conceptual framework – modeling negative thinking of individuals experiencing mental illness building on tools from the ambiguity aversion literature and using subjective probability measures to discipline the extent of negative thinking by mental health status – applies generally, independent of the choice of statistical distance.

[11] We discuss the solution to the negative thinking minimization problem in more detail in Appendix A.

work is $\bar{n}(m_t, \tau_t) = 1 - n_r(m_t) - n_\tau \tau_t$. We introduce the utility cost ξ_τ to model stigma. The psychiatric literature identifies stigma as an important factor contributing to low treatment rates of mental illness despite the efficacy of treatment (see, for example, Corrigan (2004) and Clement et al. (2015)).

A fraction ω_τ of all individuals has access to treatment when experiencing mild illness. This limited availability captures the fact that access to mental health services is an important barrier to treatment. Let $\omega = 1$ denote that an individual has access to treatment when experiencing mild illness, and $\omega = 0$ otherwise. Access to treatment ω is a permanent type. All individuals have access to treatment when experiencing serious illness.[12]

Productivity. Individuals can work for the first T_w periods of life and are retired for the remaining periods. During retirement, individuals receive a constant pension income y_t^p. During working life, individuals face idiosyncratic productivity risk. As in French (2005) and Bick, Blandin, and Rogerson (2022), labor productivity is given by:

$$\log z_t = \log \zeta_t + \Lambda(m_t) + \theta(n_t) \log n_t + \Phi(n_t) + \log \nu_t. \tag{4}$$

The first component, $\log \zeta_t$, is a deterministic life-cycle component. The second term, $\Lambda(m_t)$, captures how labor productivity is affected by mental health m_t. The component $\theta(n_t)$ captures the elasticity of labor productivity with respect to hours worked, which varies with hours worked. We follow Bick, Blandin, and Rogerson (2022) and specify $\theta(n_t)$ as a step function so the relationship between labor productivity and hours is piecewise log-linear. The function $\Phi(n_t)$ preserves continuity of labor productivity with respect to hours worked despite discontinuities in the step function $\theta(n_t)$. The idiosyncratic persistent component $\log \nu_t$ follows a discretized AR(1) process with persistence ρ_ν and variance of innovations σ_ν^2. Denote by Ω_ν the finite set of realizations that ν_t takes and by Γ_ν the corresponding transition matrix.

Labor Supply. Each period, individuals choose a job j before their labor productivity is realized. After choosing a job, productivity is realized, and individuals choose the number of hours to work. A job j is described by an up-to-task production technology which is parameterized by a job-specific up-to-task requirement y_j. Consider an individual who chooses a job j. If the individual's effective labor input, which is the product of productivity z and working hours n, exceeds the job requirement y_j, then the worker is up to the task and income is equal to y_j. If the individual's effective labor input is less than the job requirement y_j, then the worker is not up to the task and income is zero. The individual's income y

[12] Holding constant overall access to treatment, the welfare costs of mental illness and the benefit of mental health policies increase if we restrict access to individuals experiencing serious rather than mild mental illness. As a result, we obtain a lower estimate for the cost of mental illness and the benefit of mental health policies.

is therefore a function of effective labor input and the job requirement:

$$y(zn, j) = \begin{cases} y_j & \text{if } zn \geqslant y_j. \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Given the up-to-task production technology, a worker in job j either works zero hours, or chooses hours to exactly meet the requirement y_j. In the latter case, the worker's income equals $y_j = zn$. That is, when individuals work a positive amount of hours, the hourly wage y_j/n is equal to labor productivity z. Going forward, we refer to z as the hourly wage and as labor productivity interchangeably.[13]

We use the up-to-task production technology (5) and assume that jobs are chosen before productivity is realized to introduce the psychiatric notion of inaction into the labor market. Negative thinking means that individuals who experience mental illness hold more negative expectations over future productivity. Expecting their productivity may not be high enough to fulfill a job with high task requirements, individuals who think negatively select into jobs with lower up-to-task requirements. Individuals experiencing mental illness may thus choose lower paying jobs as they may underestimate their capabilities.[14] Inaction is an important symptom highlighted in psychiatric theories of mental illness (Beck, 1967, 2008; Huys, Maia, and Frank, 2016). According to these theories, negative thinking induces low valuation of future rewards which in turn deters individuals from taking an action. In our labor market setting, individuals experiencing mental illness do not pick demanding jobs since they think they may not be able to fulfill the requirements. This inaction in turn means they have less resources to spend on mental health treatment, thereby reinforcing their mental illness.

Assets. Individuals can save in risk-free and risky assets. The risk-free asset is a one-period bond that earns a gross return R_f. Denote by $r_f = \log R_f$ the log return on the risk-free asset. The log return on the risky asset is given by:

$$r_t = r_f + r_p + \upsilon_t, \tag{6}$$

[13]The specification of the up-to-task labor technology (5) follows two strands of literature. Similar to the search and matching literature (Shi, 2002; Albrecht and Vroman, 2002; Jarosch and Pilossoph, 2019; Braxton and Taska, 2023), our technology (5) specifies that worker inputs need to meet the standards to generate income and output. Different from these papers, the worker input, zn, is endogenous in our framework due to the worker's labor supply decision. Similar to Goldin (2014), an hours choice determines whether the requirements for a job are met.

[14]In order to illustrate this mechanism, consider an individual who chooses a job with requirement y_j prior to the realization of productivity z to maximize $\mathbb{E}\big[\log c - \frac{1}{2}n^2\big]$ subject to a budget constraint $c = y$ and up-to-task production technology (5). The optimal job choice is characterized by $y_j = (\mathbb{E}\big[\frac{1}{z^2}\big])^{-\frac{1}{2}}$. Under negative thinking, the subjective probability of low productivity realizations is higher and the subjective probabilities for high realizations are lower, thus increasing $\mathbb{E}\big[\frac{1}{z^2}\big]$ and lowering the job requirement, income, and hours $n = y_j/z$.

where r_p is the expected risk premium over the risk-free asset, and υ_t is an innovation drawn from a discretized normal distribution $\mathcal{N}(0, \sigma_\upsilon^2)$. We denote the finite set of realizations that υ_t can take by Ω_υ and denote the risky asset's gross return by $R_t = \exp(r_t)$.

Individuals choose savings s_t and how to allocate savings between risk-free and risky assets. To invest in risky assets, individuals incur a per-period participation cost φ_k. Denote by $k_t \in [0, 1]$ the share of savings invested in risky assets. Given a savings choice s_t, a portfolio choice k_t and a realized return on risky assets R_t, an individual's wealth at the beginning of period $t+1$ is given by:

$$a_{t+1} = s_t R_t^s(k_t), \tag{7}$$

where $R_t^s(k_t) = k_t R_t + (1 - k_t) R_f$. Individuals can borrow up to an amount $\underline{s}$, that is $s_t \geqslant \underline{s}$.[15]

Timing Within a Period. The state of an individual at the beginning a period is age t, wealth a_t, lagged idiosyncratic labor productivity component ν_{t-1}, and mental health state m_t. Within a period individuals choose job j_t before idiosyncratic productivity ν_t realizes. After idiosyncratic productivity realizes, individuals choose consumption c_t, labor supply n_t, and allocate savings s_t towards risky and risk-free assets as well as decide whether to go into treatment. At the end of the period, returns on risky assets R_t realize, determining next period wealth, and next period's mental health m_{t+1} realizes.

Job Choice. We next formalize the individual's job choice. The individual chooses job j_t before idiosyncratic productivity realizes. Let the value of working in job j_t with wealth a_t, idiosyncratic productivity ν_t, and mental health m_t for an individual with access to treatment ω be given by $w_t(j_t, a_t, \nu_t, m_t, \omega)$. The indirect utility associated with the optimal job choice is denoted $v_t(a_t, \nu_{t-1}, m_t, \omega)$ and is given by:

$$v_t(a_t, \nu_{t-1}, m_t, \omega) = \max_{j_t} \min_{p_t} \mathbb{E}_{p_t} w_t(j_t, a_t, \nu_t, m_t, \omega) = \max_{j_t} \min_{p_t} \sum_{\nu_t \in \Omega_\nu} p_t(\nu_t) w_t(j_t, a_t, \nu_t, m_t, \omega), \tag{8}$$

where the subjective probabilities $p_t(\nu_t)$ are chosen subject to the total variation constraint:

$$\frac{1}{2} \sum_{\nu_t \in \Omega_\nu} \left| p_t(\nu_t) - q_t(\nu_t) \right| \leqslant \kappa(m_t), \tag{9}$$

where $q_t(\nu_t)$ is the objective conditional probability of idiosyncratic productivity realization ν_t given ν_{t-1}, and $\kappa(m_t)$ governs the degree of negative thinking.

An individual selects a job j_t together with the probability distribution p_t that minimizes the expected payoffs in that job among the probability distributions that are within a distance $\kappa(m_t)$ from the objective

[15] Sergeyev, Lian, and Gorodnichenko (2024) develop a model of financial stress where individuals lose time when they are close to borrowing constraint.

probability distributions. This minimization problem over probability distributions represents negative thinking.

Decision Problem. After choosing job j_t and after the realization of productivity ν_t, individuals decide how much to consume c_t, work n_t, save s_t, make a portfolio choice k_t for their savings, and make a treatment choice τ_t. The budget constraint is:

$$c_t + \varphi_\tau \tau_t + \varphi_k \mathbf{1}_{k_t} + s_t \leqslant a_t + y_t(z_t n_t, j_t). \tag{10}$$

The individual pays a cost φ_τ for treatment. If the individual allocates positive savings to risky assets at date t, there is a fixed participation fee φ_k – the indicator variable $\mathbf{1}_{k_t}$ takes the value one if $k_t > 0$, and takes the value zero otherwise.

The problem of an individual with job j_t, wealth a_t, productivity ν_t, mental health m_t, and access to treatment ω is to choose consumption c_t, hours worked n_t, treatment τ_t, savings s_t, and portfolio share $k_t \in [0, 1]$ and is given by:

$$w_t(j_t, a_t, \nu_t, m_t, \omega) = \max \Big\{ u\big(c_t, \bar{n}(m_t, \tau_t) - n_t\big) - \xi_m(m_t) - \xi_\tau \tau_t + \beta \min_{p_t} \mathbb{E}_{p_t} v_{t+1}(a_{t+1}, \nu_t, m_{t+1}, \omega) \Big\}, \tag{11}$$

where $\xi_m(m_t)$ is a direct utility cost of mental illness and ξ_τ is the stigma cost associated with treatment. Optimization is subject to the asset accumulation equation (7), the budget constraint (10), the borrowing condition $s_t \geqslant \underline{s}$, and to negative thinking. Subjective probabilities $p_t(a_{t+1}, m_{t+1})$ minimize the expected continuation value $\mathbb{E}_{p_t} v_{t+1}(a_{t+1}, \nu_t, m_{t+1}, \omega) = \sum p_t(a_{t+1}, m_{t+1}) v_{t+1}(a_{t+1}, \nu_t, m_{t+1}, \omega)$ subject to the total variation constraint:

$$\frac{1}{2} \sum_{\Omega_v \times \Omega_m} \big| p_t(a_{t+1}, m_{t+1}) - q_t(a_{t+1}, m_{t+1}) \big| \leqslant \kappa(m_t), \tag{12}$$

where $q_t(a_{t+1}, m_{t+1})$ is the objective probability of state (a_{t+1}, m_{t+1}) induced by the distribution of risky returns and the mental health transition matrix. The continuation value of choosing a job at the beginning of $t+1$ with wealth a_{t+1}, productivity ν_t, and mental health m_{t+1} is given by $v_{t+1}(a_{t+1}, \nu_t, m_{t+1}, \omega)$ and described by (8). The mental health status determines the degree of negative thinking $\kappa(m_t)$ in (12). Negative thinking in this consumption and saving problem is with respect to joint uncertainty over the returns on the risky investment and the future mental health status.[16] That is, individuals experiencing

[16] We model negative thinking by mental health status as a structural parameter that is invariant to the source of risk, similar to risk aversion. The model can be generalized to allow for differences in negative thinking by source of risk. The conceptual idea to use subjective probabilities to discipline negative thinking extends to different sources of risk. Alternatively, we could use additional moments to determine negative thinking by source of risk. For

mental illness think negatively about both returns on risky investments and their mental health evolution. Importantly, individuals experiencing mental illness think negatively about the benefits of treatment, and thus may not seek treatment.

As discussed in Section 2, an important feature of mental illness is reinforcement through behavior. In our model, mental illness reinforces itself through several channels. First, negative thinking over the efficacy of treatment reduces the propensity to get mental health treatment. Second, due to rumination individuals with mental illness have less time to seek treatment. Third, negative thinking over their future productivity, together with rumination, leads individuals experiencing mental illness to choose less demanding jobs and to work less, thereby providing them with less financial resources to seek treatment. Finally, negative thinking about the performance of risky investments discourages individuals experiencing mental illness from making high-return investments, which further reduces their ability to pay for treatment.

4 Model Quantification and Validation

This section quantifies the model.

4.1 Data on Mental Health and Economic Outcomes

A main data source we use to quantify the model is the Panel Study for Income Dynamics (PSID). A key feature of the PSID is that it records the mental health of respondents, which allows to quantify the relationship between mental health and economic outcomes such as consumption, savings and portfolio choice, and labor supply.

The PSID reports the mental health of respondents using the Kessler Psychological Distress Scale. The Kessler Psychological Distress Scale (K6 scale) is widely used by the epidemiological and psychiatric literature to assess the prevalence and severity of mental illness, and is the primary mental health measure used in U.S. government administered health surveys as well as the WHO World Mental Health Surveys.[17] The K6 scale has been extensively validated against clinical mental health diagnoses and has been shown to

example, we could use the conditional correlation between mental health status and the risky investment share to determine the extent of negative thinking with respect to risky returns (see Table 7 and Table 8 below). Since these conditional correlations align well between the model and the data under the benchmark model, the assumption that negative thinking is invariant to the source of risk is not significantly restrictive.

[17] The K6 scale is calculated using respondents' answers to six questions (Kessler et al., 2002, 2003). In particular, respondents are asked: "In the past 30 days, about how often did you feel (1) sadness, (2) nervous, (3) restless or fidgety, (4) hopeless, (5) that everything was an effort, and (6) worthless". To each question, individuals respond (0) none of the time, (1) a little of the time, (2) some of the time, (3) most of the time, or (4) all of the time. The K6 scale is computed as the sum of respondents' answers to the six questions.

consistently predict clinical diagnoses of mood and anxiety disorders (Kessler et al., 2002, 2003; Furukawa, Kessler, Slade, and Andrews, 2003; Cairney et al., 2007). We classify individuals into the three mental health states (healthy, mild, severe) based on the K6 scale following Kessler et al. (2008).[18]

We use PSID waves between 2000 and 2020 since earlier waves lack information on respondents' mental health. Our measure of income is individual labor income over the past year. Hours worked are total hours worked including overtime. Hourly wage rates are computed as individual income divided by hours worked. Our benchmark measure of consumption is annual nondurable expenditures which include expenditures on food, utilities, child care, clothing, home insurance, telecommunications, home maintenance, and variable transportation costs.[19] All dollar values are reported in 2015 values.

We categorize equity holdings, business assets and liabilities, and real estate assets and liabilities as risky. We classify checking accounts, vehicles, certificates of deposit, government bonds and debt balances (except for business loans and real estate debt) as safe. Individual retirement accounts and other assets are labeled mixed investments. Total wealth is the sum of risky, safe, and mixed assets net of liabilities. The risky investment share measures the share of risky assets and liabilities in a portfolio. We provide more detail, and discuss the construction of our PSID sample, in Appendix B.

4.2 Exogenous Parameters

We next describe the parameters that are exogenously calibrated based on direct empirical evidence or existing literature. Table 1 summarizes the exogenous model parameters.

Demographics. A model periods corresponds to two years. Individuals start adult life at age 25 and can choose to work for up to $T_w = 20$ periods, which corresponds to age 65. Individuals die deterministically at period $T = 30$, which corresponds to age 85, the average life expectancy conditional on reaching the normal retirement age.

Productivity. One unit of time corresponds to 100 hours per week. We calibrate the dependence of wages on working hours, $\theta(n_t)$, using data from the CPS-ORG documented in Bick, Blandin, and Rogerson (2022). We consider three regions for the wage elasticity: θ_S for short hours (less than 40 hours per

[18]Individuals with a K6 score between 13 and 24 are classified as experiencing serious mental illness, individuals with a K6 score between 8 and 12 are classified as experiencing mild mental illness, and individuals with K6 scores between 0 and 7 are classified as healthy.

[19]Our benchmark measure of consumption is closest to measures used by Aguiar and Hurst (2013) and Boerma and Karabarbounis (2021). Since detailed consumption expenditures are available in the PSID starting from 2004, we restrict the analysis with respect to consumption to this period. De Quidt and Haushofer (2016) is a stylized model of static actions of how depression affects food, non-food, and sleep consumption through negative beliefs and leads to overeating and under-sleeping.

Table 1: Exogenous General Parameters

Parameter	Target	Value
	Demographics	
Retirement age T_w	Normal retirement age	65
Terminal age T	Life expectancy	85
	Labor Markets	
Wage elasticity for short hours θ_S	Bick, Blandin, and Rogerson (2022)	0.40
Wage elasticity for medium hours θ_M	Bick, Blandin, and Rogerson (2022)	0.58
Wage elasticity for long hours θ_L	Bick, Blandin, and Rogerson (2022)	−0.76
Frisch elasticity of labor supply η	Chetty, Guren, Manoli, and Weber (2012)	0.281
Persistence of productivity ρ_ν	Persistence of residual wages	0.949
Variance of productivity σ_ν^2	Variance of innovation in residual wages	0.112
Retirement income y^p in dollars	Average retirement income	14,100
	Asset Markets	
Risk-free rate r_f	Return on safe assets	0.0186
Standard deviation of risky returns σ_v	Standard deviation on risky assets	0.0830
Risk premium r_p	Risk premium for risky assets	0.0258
Borrowing constraint $\underline{s}$		0

Table 1 presents the values of model parameters that are set exogenously. The first columns shows the parameters. The second column describes the empirical moment that directly informs the parameter value. The third column shows the parameter value. The parameter values for the productivity process, income and returns are annualized.

week, or $n_t \leqslant 0.4$), θ_M for medium hours (between 40 and 50 hours per week, or $0.4 < n_t \leqslant 0.5$), and θ_H for long hours (exceeding 50 hours per week, or $n_t > 0.5$). Using the data underlying Figure 3 of Bick, Blandin, and Rogerson (2022), we estimate the wage elasticities $\theta_S = 0.40$, $\theta_M = 0.58$, and $\theta_L = -0.76$.[20]

We quantify the remaining productivity parameters by analyzing residual wages in the PSID. In order

[20] To illustrate the identification, evaluate earnings growth in Bick, Blandin, and Rogerson (2022) between 20 and 40, between 40 and 50, and between 50 and 80 hours. This gives elasticities $\frac{0.81}{\log(35/20)} - 1 = 0.45$, $\frac{0.19}{\log(45/40)} - 1 = 0.59$ and $\frac{0.10}{\log(80/50)} - 1 = -0.79$. We set the step function $\Phi(n_t)$ in equation (4) by choosing $\Phi_M = -\theta_M \log(0.4)$ such that there is no wage penalty when individuals work medium hours (full-time), and select Φ_S and Φ_H to ensure continuity of the wage penalty $\theta(n_t)\log n_t + \Phi(n_t)$.

to obtain residual wages, we regress log hourly wages on mental health and log hours worked, where the elasticity of wages to working hours as well as the intercept vary by the short, medium, and long hours regions, consistent with wage equation (4).

We quantify productivity effects of mental illness by using the panel structure of the PSID to assess the change in residual wages in response to a change in mental health. We find that individuals experiencing mild (serious) mental illness have 1.3 (3.2) percent lower productivity, or $\Lambda(m_1) = -0.013$ and $\Lambda(m_2) = -0.032$. The estimates of the mental health effects on productivity are small and not statistically significant. This aligns with the evidence of the psychiatric literature that depression is characterized by impaired cognitive control (manifested as rumination) rather than by cognitive deficits (Gotlib and Joormann, 2010) − individuals with depression perform on par with healthy individuals once their attention is controlled and they cannot ruminate (Hertel, 2004). We extract the deterministic life-cycle profile ζ_t by fitting a third-order polynomial through the age effects on the remaining variation, and estimate the annual persistence ρ_ν and variance of productivity shocks σ_ν^2 to align the model-implied and empirical auto-covariation between residual wages. We find $\rho_\nu = 0.949$ and $\sigma_\nu^2 = 0.112$. Retirement income y^p is equal to 14,100 dollars, which is the average retirement benefit.

Preferences. Individuals have flow utility over consumption c and leisure ℓ given by:

$$u(c, \ell) = \log c + \psi \frac{\ell^{1-\frac{1}{\eta}} - 1}{1 - \frac{1}{\eta}} \,, \tag{13}$$

where $\eta \geqslant 0$ governs the curvature with respect to leisure hours, and $\psi \geqslant 0$ governs the value of leisure. We choose the parameter η so that the Frisch elasticity of labor supply for an average healthy worker, who works $\bar{n} = 0.403$ hours, equals 0.55 following Chetty, Guren, Manoli, and Weber (2012). To align with the Frisch elasticity of labor supply for these workers in the model, we require $\eta = \frac{\bar{n}}{1-\bar{n}} \frac{1}{\frac{1}{0.55} + \theta_M} = 0.281$.[21]

Assets. We set the logarithmic return on the risk-free asset to $r_f = 0.0186$, corresponding to the annual real returns on safe assets reported by Jordà, Knoll, Kuvshinov, Schularick, and Taylor (2019) between 2001 and 2020. The risk premium is set to $r_p = 0.0258$ per year, which is the observed return differential between risky assets and government bonds. We set the standard deviation of log risky returns σ_v to 0.0830.[22] The borrowing constraint is set so that individuals cannot borrow, $\underline{s} = 0$.

[21]Using the first-order conditions for labor supply, we express the Frisch elasticity for workers working $\bar{n} = 0.403$ hours as $1/\big(\frac{\bar{n}}{1-\bar{n}}\frac{1}{\eta} - \theta(n)\big)$. Given that an average healthy individual works $\bar{n}$ hours, we obtain a Frisch elasticity for healthy individuals working average hours equal to 0.55 when $\eta = \frac{\bar{n}}{1-\bar{n}}/\big(\frac{1}{0.55} + \theta_M\big) = \frac{0.403}{1-0.403}/\big(\frac{1}{0.55} + 0.58\big) = 0.281$.

[22]The returns on risky assets are distributed with a lognormal distribution. The mean returns on the risky assets in logarithms is set equal to $r_p + r_f - \sigma_v^2/2$.

Mental Health. For each mental health status, we need to calibrate the extent of negative thinking $\kappa(m)$. The main idea is to use differences in subjective probabilities across mental health states in the data to inform negative thinking in the model.[23] We operationalize this idea by focusing on the differences in subjective loss probabilities, which is the subjective probability of losing in a lottery with two outcomes. Let p_i denote an individual i's subjective loss probability for a given lottery and similarly let $p_{i'}$ denote the subjective loss probability of individual i' for the same lottery. Within our model, the subjective loss probability is governed by equation (3) for the case of a random variable with two outcomes. Hence, when individuals i and i' face the same lottery, the difference in the subjective loss probabilities reflects the difference in the extent of negative thinking:

$$p_i - p_{i'} = \kappa_i - \kappa_{i'}. \tag{14}$$

Intuitively, individuals who have higher subjective loss probabilities think more negatively about uncertain outcomes. Equation (14) shows that differences in subjective loss probabilities inform differences in the extent of negative thinking.

In the data, we measure differences in subjective loss probabilities across mental health states using the RAND American Life Panel (ALP), a nationally representative survey of U.S. adults. We combine two different ALP modules. The first module, designed by Dimmock, Kouwenberg, Mitchell, and Peijnenburg (2016, 2021) and implemented between March and April 2012, elicits respondents' subjective loss probabilities by presenting respondents with a series of classic Ellsberg urn problems (Ellsberg, 1961). This approach is routinely used by the ambiguity aversion literature to elicit ambiguity aversion. We merge it with a second ALP module that asks the same respondents about their mental health, and was implemented between May and August 2012.[24]

The RAND ALP elicits an individual's subjective loss probability as the point of indifference between a gamble on an unknown urn and a gamble on a known urn with an objective loss probability $\hat{q}$. In order to illustrate, let w_1 denote the value when losing the gamble, and let $w_2 > w_1$ denote the value when winning.[25] Consider an individual i. The expected value from a gamble on the outcome of the known urn is given by $(1 - \hat{q})w_2 + qw_1$. The expected value from a gamble on the outcome of the unknown urn is $(1 - p_i)w_2 + p_i w_1$, where p_i is the individual's subjective loss probability. By presenting an individual

[23] This idea applies across different statistical distances between probability distributions that one could use to model negative thinking. The exact mapping between the subjective probabilities and the structural parameters that govern negative thinking would differ.

[24] We provide details in Appendix C.

[25] While the utility from winning a gamble may differ by individual i, we suppress the notation since preference heterogeneity does not affect the measurement.

Table 2: Negative Thinking and Mental Illness Severity

Mild κ_1	3.4	3.4	3.4	3.5	3.1
	(1.3)	(1.3)	(1.3)	(1.3)	(1.3)
Serious κ_2	6.4	7.2	7.2	7.2	6.7
	(2.2)	(2.2)	(2.2)	(2.2)	(2.1)
Controls	None	+ Income, Age	+ Education	+ Race, Gender	All
R^2	0.02	0.03	0.03	0.04	0.10

Table 2 displays the regression coefficients κ_1 (first row) and κ_2 (third row) estimated from equation (15) as well as their corresponding standard errors (in rows 2 and 4). The control variables include income, age, education, race, gender, employment, and risk aversion. Table 2 shows how negative thinking varies with mental health status. From the first to the final column, we add additional control variables. All estimates are statistically significant as implied by the standard errors, which are reported in parentheses below the regression coefficients. The average subjective loss probability is 0.474. The number of observations is equal to 2,973.

with a series of Ellsberg urn problems that differ by the objective loss probability $\hat{q}$, the ALP elicits the indifference probability such that the individual is indifferent between a gamble on the known urn with objective loss probability $\hat{q} = q$, and a gamble on the unknown urn. In other words, the Ellsberg module elicits q such that $(1-q)w_2 + qw_1 = (1-p_i)w_2 + p_i w_1$. The elicited indifference probability q is exactly the individual's subjective loss probability for the gamble on the unknown urn: $p_i = q$. An individual thinks more negatively if the subjective loss probability for the unknown urn is higher. Faced with the same objective uncertainty, an individual who thinks more negatively has lower expectations of winning.

In order to evaluate how the subjective loss probability varies with the severity of mental illness, we estimate the following regression:

$$p_i = \kappa_1 D_{1i} + \kappa_2 D_{2i} + \kappa_x X_i + \varepsilon_i, \tag{15}$$

where p_i is the subjective loss probability of individual i, and D_{1i} (D_{2i}) is a dummy variable taking the value one when individual i is classified as experiencing mild (serious) illness. Control variables X_i include age, sex, education, race, risk aversion, household income, employment status, and a constant.[26] The coefficients κ capture how the subjective loss probability varies with mental health.

Tables 2, C.2, and C.3 show how subjective loss probabilities vary by mental health. Each column corresponds to a regression that differs in the controls that are included. From the first to the final

[26] In Appendix C, we estimate separate regressions with risk aversion as the dependent variable to show that risk aversion does not vary systematically with mental health. In line with psychiatric theory (Beck, 1967, 2008), these estimates indicate that differences in negative thinking rather than risk aversion is a key feature of mental illness.

column, we add control variables. For example, the first column in Table 2 shows that without controls, we find that individuals experiencing mild mental illness have a subjective loss probability that is 3.4 percentage point higher relative to healthy individuals (first row), while individuals experiencing serious mental illness have a subjective loss probability that is a 6.4 percentage point higher (third row). The final column shows that this finding is robust to the inclusion of all control variables. Individuals with mild (serious) mental illness have a subjective loss probability that is 3.1 (6.7) percentage point higher relative to healthy individuals.[27] In sum, individuals experiencing mental illness have higher subjective loss probabilities, and the extent of negative thinking increases with the severity of mental illness. Normalizing $\kappa(m_0) = 0$, these empirical estimates together with (14) thus determine the total variation budget for individuals who experiences mild illness $\kappa(m_1) = 0.031$, and for individuals who experience serious illness $\kappa(m_2) = 0.067$.[28] In Section 5.2.4, we conduct sensitivity analysis for an economy with $\kappa(m_0) > 0$, i.e. where healthy individuals are ambiguity averse.

Treatment. We now describe how to quantify the mental health transition probability matrix $\Gamma_m(\tau, \nu)$.[29] We assume treatment does not benefit healthy individuals, that is, the transition probabilities for healthy individuals are independent of treatment. This assumption is motivated by the finding that healthy individuals rarely receive treatment (Cronin, Forsstrom, and Papageorge, 2024).[30] In addition, we assume that transitions for healthy individuals depend only on whether or not idiosyncratic productivity is below a threshold $\underline{\nu}$, while transitions from mild and serious mental illness do not depend on idiosyncratic productivity. We set the threshold $\underline{\nu}$ to be the bottom quartile of the invariant idiosyncratic productivity distribution to capture in a parsimonious way that bad labor market shocks might deteriorate future mental health. Given these assumptions, the mental health transition matrices with and without treatment for normal and low productivity shocks require the quantification of 12 transition probabilities.

The moments that we use to quantify the transition probabilities are population shares across mental health status from the NIMH (two moments), estimates of the efficacy of treatment from the medical literature (two moments), and unconditional biannual transition probabilities between the three mental

[27]By comparing the partial R^2 of the explanatory variables, we evaluate the relative significance of mental health in accounting for the variation in the subjective loss probability. We find that mental health has the largest partial R^2 among all explanatory variables.

[28]We model mental illness featuring negative thinking and adopt the tools from the ambiguity aversion literature to model negative thinking. From the literature on ambiguity aversion, we know ambiguity aversion applies across all individuals.

[29]We provide the details in Appendix D.

[30]Cronin, Forsstrom, and Papageorge (2024) develop a structural model of dynamic treatment choices to study the reluctance to use talk therapy observed in the data.

Table 3: Mental Health Transition Matrix

No Treatment	Healthy	Mild	Serious	*Treatment*	Healthy	Mild	Serious
Healthy ($\nu < \underline{\nu}$)	0.908	0.073	0.019	Healthy ($\nu < \underline{\nu}$)	0.908	0.073	0.019
Healthy ($\nu \geqslant \underline{\nu}$)	0.934	0.055	0.011	Healthy ($\nu \geqslant \underline{\nu}$)	0.934	0.055	0.011
Mild	0.395	0.467	0.138	Mild	0.766	0.161	0.073
Serious	0.124	0.241	0.635	Serious	0.332	0.360	0.308

Table 3 presents the mental health transition matrix for individuals who receive treatment and who do not receive treatment. Rows correspond to the current mental health m, and columns correspond to mental health status two years ahead m'.

health states obtained from the PSID, where we differentiate between healthy individuals with normal and low productivity states (eight moments). Estimates of the efficacy of treatment are typically reported by the medical literature in terms of the standardized mean difference (SMD). The more negative is the SMD, the larger is the drop in a mental illness measure in terms of its pooled standard deviation among the treated group relative to the control group, or in other words the more effective is treatment. For our calibration, we use a meta-analysis by Ekers, Richards, and Gilbody (2008) who report an average SMD of -0.70 for behavioral psychotherapy.

Table 3 reports the estimated mental health transition matrix as a function of idiosyncratic labor market shocks and treatment. The first takeaway is that, consistent with the medical literature, treatment is effective. For example, the probability to transition from serious mental illness to the healthy state is 12.4 percent without treatment, while 33.2 percent with treatment. The second takeaway is that bad labor market shocks increase the likelihood to experience mental illness in the future consistent with the unconditional transitional probabilities from the PSID. For example, the likelihood to transition from the healthy state into serious (mild) illness is 1.1 (5.5) percent in normal productivity states, while it is 1.9 (7.3) percent in low productivity states.

We set the monetary cost of treatment based on Cronin, Forsstrom, and Papageorge (2024), who report an out-of-pocket expenditure on psychotherapy of 24 dollars per visit.[31] We consider an average of one visit per week per year to arrive at an annual treatment cost of φ_τ of 1,250 dollars. We calibrate the time cost to two hours per week $n_\tau = 0.02$. Monetary and time costs do not vary by mental health.

In the baseline calibration we do not incorporate a flow utility penalty associated with mental illness,

[31]The total expenditure, including out-of-pocket payments and insurer payments, is reported to be 126 dollars. Individuals from the 1996 to 2011 cohorts of the Medical Expenditure Panel Survey (MEPS) thus pay $\frac{24}{126} = 0.19$ of the treatment costs out-of-pocket and 0.81 is covered by insurance.

Table 4: Exogenous Mental Health Parameters

Parameter	Target	Value
Productivity loss, mild $\Lambda(m_1)$	Δ Residual wage regressions, mild	-0.013
Productivity loss, serious $\Lambda(m_2)$	Δ Residual wage regressions, serious	-0.032
Negative thinking, mild $\kappa(m_1)$	Ambiguity index regressions, mild	0.031
Negative thinking, serious $\kappa(m_2)$	Ambiguity index regressions, serious	0.067
Monetary cost of treatment φ_τ	Cronin, Forsstrom, and Papageorge (2024)	1,250
Time cost of treatment n_τ	One-hour session and commute per week	0.020

Table 4 presents the values of mental health parameters that we set without solving the model. The first column shows the parameters. The second column describes the empirical moment that informs the parameter value. The third column shows the parameter value.

i.e. we calibrate $\xi_m = 0$. This aligns with the psychiatric literature that posits that depressed mood is due to expected utility being low due to negative thinking on the probabilities of future outcomes, which is captured by the depressed continuation value in our model, and not due to deficits in primary (flow) utility (Amsterdam et al., 1987; Berlin et al., 1998; Dichter et al., 2010; Huys, Daw, and Dayan, 2015). In the sensitivity analysis in Section 5.2.4, we show that flow utility penalties increase the cost of mental illness and the benefits of mental health policies. That is, the benchmark model provides a conservative estimate for the cost of mental illness and the gains from policies that improve mental health.

4.3 Endogenous Parameters

We calibrate remaining parameters so that the model matches data moments related to labor supply, savings and portfolio choice, and to mental health treatment. Table 5 summarizes the endogenous parameters and data moments. Parameters are presented together with the data moment that determines them most quantitatively. We conduct a parameter sensitivity analysis in Appendix E to illustrate which moments structurally inform which parameters.

We set the discount factor β to 0.967 to match average wealth of working age individuals in the PSID data, which is 288 thousand dollars. The annual participation fee required for investing in the risky asset, φ_k, is estimated to be $3,500$ dollars. It is identified from the average share of savings invested in risky assets in the data, which is 0.557.

We calibrate the disutility from work to $\psi = 0.29$ to match average hours worked in the cross section, which is 40 hours per week. We exploit the panel structure of the PSID to calibrate the time loss due to

Table 5: Endogenous Parameters

Parameter	Value	Moment (mean of)	Data	Model
Discount factor β	0.967	Wealth in dollars	288,000	288,000
Risky investments costs φ_k	3,500	Risky investment share	0.557	0.556
Disutility from work ψ	0.290	Hours worked	0.399	0.400
Rumination, mild $n_r(m_1)$	0.067	Δ Hours worked, mild	-0.047	-0.047
Rumination, serious $n_r(m_2)$	0.111	Δ Hours worked, serious	-0.127	-0.127
Utility cost of treatment ξ_τ	0.032	Treatment share, serious	0.656	0.657
Treatment availability ω_τ	0.682	Treatment share, mild	0.414	0.414

Table 5 presents the parameter values set to match model-generated moments to their data analog. The first three columns present the parameters and their values. The fourth column describes a moment that informs the parameter value. The fifth and sixth column present the model-generated moment and the data-equivalent.

rumination to match the changes in hours in response to changes in mental health. We estimate what happens to (log) hours for a given individual as they transition between mental health states using a regression with individual fixed effects.[32] Individuals experiencing mild mental health problems work 4.7 percent fewer hours, while individuals experiencing serious mental health problems work 12.7 percent fewer hours. With rumination of 6.7 hours per week for mild mental illness, and 11.1 hours per week for serious mental illness, the regression coefficients of the model match the regression coefficients in the data.

We calibrate the utility cost of treatment, ξ_τ, so that the model matches the share of individuals with severe illness who receive treatment in the data. According to the National Institute of Mental Health (NIMH), 65.4 percent of those who are seriously ill receive treatment during the year (see Section 2). We obtain an estimate of $\xi_\tau = 0.032$. Similarly, we calibrate the share of individuals who have access to treatment when experiencing mild illness, ω_τ, so that the model matches the share of individuals with mild illness who receive treatment in the data. With $\omega_\tau = 0.682$, the share of individuals with mild illness who get treated is equal to 0.414, which matches the share reported by the NIMH (see footnote 7).

[32]That is, we regress (log) hours worked on indicator variables for mild and severe mental illness with individual fixed effects and control variables for education, age, sex, race, household composition, wealth, and physical illness.

Table 6: Validation: Averages

	Data			Model		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Consumption	51	48	45	51	47	43
Hours	0.403	0.380	0.357	0.404	0.383	0.351
Income	65	57	48	64	56	46
Wealth	312	232	208	292	262	236
Risky investment share	0.581	0.512	0.466	0.572	0.461	0.384
Risky participation rate	0.662	0.576	0.529	0.601	0.516	0.449

Table 6 shows average consumption, hours, income, wealth, and risky investment by mental health status. Consumption, income, and wealth holdings are in thousands of dollars. The risky investment share is the average share of total assets invested in risky assets. The risky participation rate measures the share of the population that holds more than half of their portfolio in risky assets.

4.4 Model Validation

Having quantified the model, we evaluate its fit to non-targeted moments. We first show that the model matches average consumption, hours, income, wealth, risky investment share, and risky participation rate by mental health status.[33] The first three columns of Table 6 show the non-targeted averages in the PSID data, and the final three columns show the model generated averages. The model matches almost perfectly average consumption, average hours, average income as well as the average risky investment share within each of the mental health groups. Both the data and the model show a decrease in wealth, with the model capturing this decrease at a somewhat lower rate. The model captures the risky participation rate, which we define as the share of individuals who invest more than half of their savings in risky assets. We choose this threshold since in the model it is only worth paying the fixed participation cost φ_k if the risky savings $k_t s_t$ are sufficiently large upon participation.

We next assess the ability of the model to fit the observed distributions of these variables by mental health status. Figure 1 displays the distribution of consumption by mental health status in the model and in the data. The histogram displays the within-group percentage of individuals that consumes at a given level, displayed on the horizontal axis in thousand dollars. The figure shows that the model captures the empirical consumption patterns. For example, healthy individuals are overrepresented at

[33] We scale nondurables expenditures in the PSID by a constant factor such that aggregate personal expenditures in our model align with aggregate consumption expenditures in the national accounts.

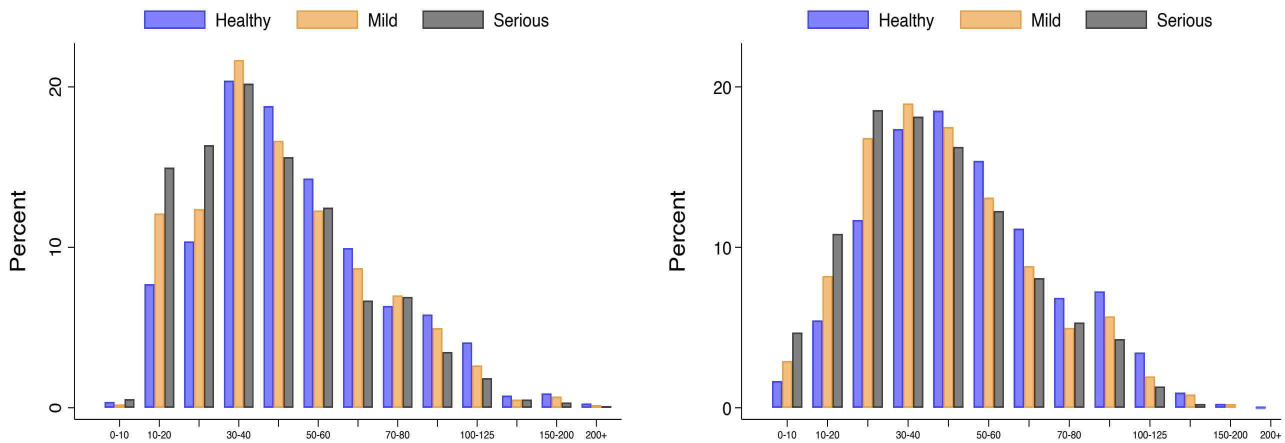

Figure 1: Consumption by Mental Health in the Model and the Data

Figure 1 shows the distribution of consumption by mental health status in the model (left panel) and in the data (right panel). The height of the bars capture the fraction of individuals consuming a particular amount within each mental health status – healthy (blue), mild illness (orange), and serious illness (black).

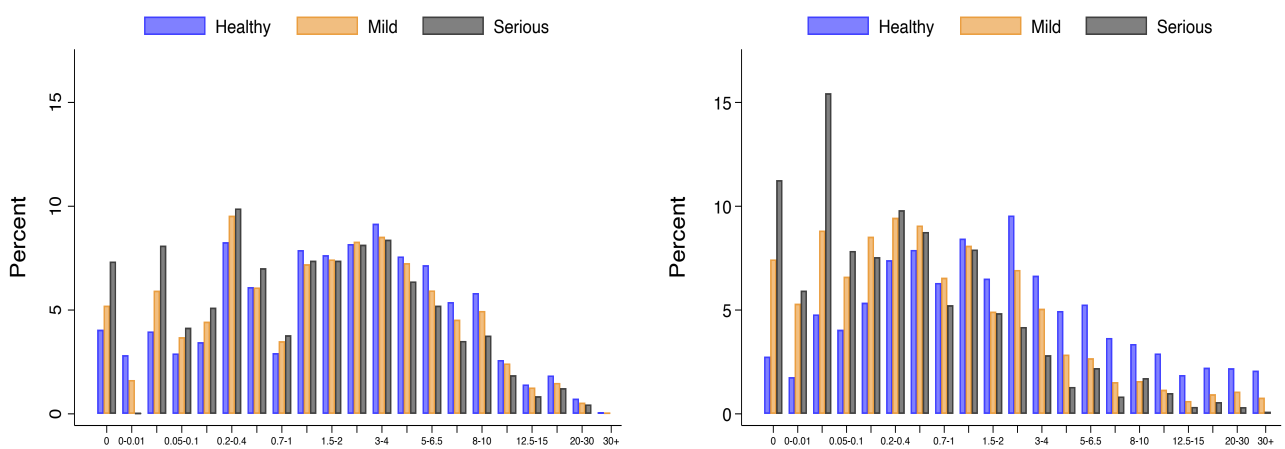

Figure 2: Savings in the Model and the Data

Figure 2 shows the distribution of savings by mental health status in the model (left panel) and in the data (right panel). The height of the bars captures the fraction of individuals holding a particular amount of savings within each mental health status – healthy (blue), mild illness (orange), and serious illness (black).

high consumption levels, and individuals with serious mental illness tend to be concentrated at low levels of consumption.

Figure 2 shows the distribution of savings by mental health status in the model and the data. The histogram displays the within-group percentage of individuals that holds a given level of savings, displayed on the horizontal axis in hundred thousand dollars. Both in the model and in the data, the savings

Table 7: Validation: Consumption and Portfolio Choice Regressions

Variable	Log Consumption		Risky Investment Share	
	Data	Model	Data	Model
Mild $\hat{\gamma}_1$	-2.2	-1.9	-3.6	-4.8
	(0.8)	(0.5)	(0.6)	(0.5)
Serious $\hat{\gamma}_2$	-6.5	-3.3	-5.6	-5.5
	(1.2)	(0.8)	(0.9)	(0.9)
Observations	35,153	35,153	36,334	36,334
R^2	0.53	0.80	0.31	0.61

Table 7 reports the regression coefficients estimated from (16) in the model and the data. For each dependent variable the first column shows the estimated coefficients from the data, while the second columns reports the model estimates.

distribution of healthy individuals is more skewed to the right and the savings distribution of individuals experiencing serious illness is more skewed to the left. The fraction of individuals with mild mental illness lies in between the fraction of individuals with serious mental illness and healthy individuals at nearly all wealth levels in the data as in the model.

Figure F.1 shows the distribution of income by mental health status. It shows that the model captures the patterns of the empirical income distribution. As in the data, healthy individuals are overrepresented in the top categories, while individuals experiencing serious mental health problems earn less.

Table F.1 summarizes the distribution of the risky investment share by mental health status. Both in the model and in the data a significant fraction of households do not hold risky investments. The risky participation rate of healthy individuals is 0.60 in the model, which aligns with 0.66 in the data. Conditional on investing in risky assets, individuals save a significant fraction of their wealth in risky assets. In the data, the 75th percentile among healthy individuals invests 0.90 of their wealth in risky assets compared to 0.98 in the model. In the model, this is due to the fixed cost of participation, which is only worth paying if a sufficiently large share of savings is invested in the risky asset. In both the model and the data, the fraction of individuals that does not participate in risky investments is higher for individuals with worse mental health and is somewhat stronger in the data than in the model. Healthy individuals invest larger shares of their savings in risky investments both in the model and the data.

We also validate the model by analyzing the extent to which consumption and portfolio choice vary with mental health conditional on other characteristics. Specifically, we estimate the following regressions in the model and in the PSID data. Let Y_{it} be the dependent variable for individual i in year t, which is either log consumption or the risky investment share. Let D_{1it} be an indicator variable taking the value one when individual i experiences mild illness in year t. Let D_{2it} be an indicator taking the value one if individual i experiences serious mental illness in year t. The regressions also include a vector of additional controls X_{it} such as the individual's age, sex, education, race, household composition, income, and wealth, and time fixed effects γ_t. We estimate the following regression:

$$Y_{it} = \gamma_t + \gamma_1 D_{1it} + \gamma_2 D_{2it} + \gamma_x X_{it} + \varepsilon_{it}. \tag{16}$$

The coefficients γ_1 and γ_2 respectively measure how the dependent variable varies with mild and serious mental illness.

In the model, we estimate equation (16) on simulated data from the economy's stationary distribution. We then compare the regression coefficients $\hat{\gamma} = (\hat{\gamma}_1, \hat{\gamma}_2)$ to their empirical counterparts, which we discuss in Appendix B.

Table 7 reports the estimated regression coefficients in the model and in the data. For each dependent variable, the first column shows the estimated coefficients in the data, while the second columns reports model estimates. The model matches the conditional correlations between consumption, portfolio choice, and mental health observed in the data. In the data, individuals experiencing mild illness consume on average 2.2 percent less than healthy individuals, and individuals experiencing serious illness consume 6.5 percent less. In the model, individuals with mild illness consume 1.9 percent less and with serious illness 3.3 percent less. In the data, individuals experiencing mild (serious) mental illness invest 3.6 (5.6) less of their savings in risky assets relative to healthy individuals. In the model, individuals experiencing mental illness also invest less in risky assets: individuals experiencing mild illness invest 4.8 percent less, while individuals experiencing severe illness invest 5.5 percent less, relative to healthy individuals.[34]

4.5 Evaluating the Mechanisms of Mental Illness

We now quantitatively evaluate the mechanisms through which mental illness affects economic outcomes. We examine how negative thinking, rumination, stigma costs of treatment, and productivity losses affect consumption, labor supply, income, wealth, and portfolio choice.

We first evaluate the impact of negative thinking. Table 8 compares the benchmark economy to an economy where mental health is not associated with negative thinking, i.e. $\kappa(m) = 0$. Without negative

[34] The results are robust to controlling for individual physical health with regression coefficients $\hat{\gamma}_c = (-2.2, -6.6)$ and $\hat{\gamma}_k = (-3.5, -5.3)$. The raw correlation between physical health and mental health in our sample is 0.06.

Table 8: Effects of Negative Thinking

	Benchmark			No negative thinking $\kappa = 0$		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Hours worked	0.404	0.383	0.351	0.403	0.386	0.357
Δ Hours worked	0.0	−4.7	−12.7	0.0	−4.0	−11.3
Income (in thousands)	64	56	46	64	56	47
Wealth (in thousands)	292	262	236	281	250	228
Risky investment share	0.572	0.461	0.384	0.560	0.495	0.444
Risky participation rate	0.601	0.516	0.449	0.589	0.520	0.467
Consumption coefficient $\hat{\gamma}_c$	—	−1.9	−3.3	—	−0.0	0.0
Investment coefficient $\hat{\gamma}_k$	—	−4.8	−5.5	—	−0.0	1.2
Treatment shares	0.000	0.414	0.657	0.000	0.268	0.438

Table 8 reports moments from the baseline economy with negative thinking and an economy without negative thinking. The coefficients $\hat{\gamma}_c$ and $\hat{\gamma}_k$ denote the estimates for coefficients γ_1 and γ_2 in (16) when log consumption and the risky investment share are the dependent variable, respectively.

thinking, individuals experiencing mental illness work slightly more, as shown in the first three rows. Due to rumination, they have less hours available to work and work and earn less than healthy individuals. The absence of negative thinking reduces the precautionary savings motive, which lowers wealth across all mental health groups.

Re-estimating equation (16) in the economy without negative thinking shows that negative thinking is key for the conditional correlations between consumption, portfolio choice, and mental health. Absent negative thinking, there is no motive for individuals with mental illness to consume less or invest less in risky assets after conditioning on age, income, and wealth. The bottom rows of Table 8 show that the model regression coefficients on consumption and portfolio choice are close to zero.[35] Since the cost of mental illness is lower without negative thinking, individuals seek less treatment relative to the benchmark economy. In Section 5, we analyze the welfare effects of different mechanisms of mental health.

We next evaluate the effects of rumination. Table 9 compares the benchmark economy to an economy

[35]This shows we could alternatively use the regression coefficients from the data to discipline negative thinking in the model. The subjective loss probabilities by mental health status in Table 2 could then be used for model validation. Since the coefficients align well between model and data in the benchmark model (see Table 7), this would yield similar results.

Table 9: Effects of Rumination

	Benchmark			No rumination $n_r = 0$		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Hours worked	0.404	0.383	0.351	0.404	0.396	0.391
Δ Hours worked	0.0	-4.7	-12.7	0.0	-1.4	-2.3
Income (in thousands)	64	56	46	64	59	55
Wealth (in thousands)	292	262	236	294	267	242
Risky investment share	0.572	0.461	0.384	0.572	0.471	0.401
Risky participation rate	0.601	0.516	0.449	0.601	0.527	0.469
Consumption coefficient $\hat{\gamma}_c$	—	-1.9	-3.3	—	-2.1	-3.5
Investment coefficient $\hat{\gamma}_k$	—	-4.8	-5.5	—	-5.4	-9.2
Treatment shares	0.000	0.414	0.657	0.000	0.367	0.573

Table 9 reports moments from the baseline economy with rumination and an economy without rumination. The coefficients $\hat{\gamma}_c$ and $\hat{\gamma}_k$ denote the estimates for coefficients γ_1 and γ_2 in (16) when log consumption and the risky investment share are the dependent variable, respectively.

where mental illness is not associated with rumination, or $n_r(m) = 0$. Rumination decreases the total number of hours available to an individual and reduces work hours. Without rumination, individuals with mental illness do not lose available time, and work similar hours as healthy individuals as shown in the second and third row. Individuals with mental illness think negatively about their future productivity, and choose less demanding jobs even without rumination. As a result, their income and wealth are lower relative to healthy individuals. Individuals with mental illness are wealthier than in the benchmark economy as they have more time and work more. This increases risky investments on both intensive and extensive margins. Estimating equation (16) in the economy without rumination shows that the regression coefficients for consumption and investment remain constant. Without rumination, individuals who experience mental illness seek less treatment, even though they have more time because the costs of experiencing mental illness are lower.

In addition to evaluating the consequences of negative thinking and rumination on economic outcomes, we evaluate the effects of stigma costs of treatment and productivity losses. Table F.2 shows that stigma costs of treatment, ξ_τ, are an important barrier to treatment, lowering the treatment rate of individuals experiencing serious (mild) mental illness by 13.2 (5.8) percentage points. Stigma costs have insignificant impact on other outcomes. Table F.3 shows that direct productivity losses, $\Lambda(m)$, have negligible impact on economic outcomes.

5 Quantitative Results

In this section, we evaluate the consequences of a number of prominent mental health policy proposals as well as the aggregate welfare costs of mental illness.

5.1 The Aggregate Welfare Cost of Mental Illness

In order to have a benchmark to evaluate the effects of policies, we first estimate the aggregate welfare costs of mental illness as the consumption equivalent welfare gain Δ_i^m of being mentally healthy for each individual i. The consumption equivalent welfare gain is such that the individual is indifferent between being in the healthy state and a per period consumption increase Δ_i^m. This is the cost of mental illness for individual i.

With logarithmic preferences for consumption, the individual consumption equivalent welfare gain of being healthy is given by:

$$\log \Delta_i^m = \beta_t \big(v_t(a_{it}, \nu_{it-1}, m_0, \omega_i) - v_t(a_{it}, \nu_{it-1}, m_{it}, \omega_i) \big), \tag{17}$$

where $\beta_t = 1/(1 + \beta + \cdots + \beta^{T-t})$. The aggregate welfare cost of mental illness Δ^m is the average of individual consumption equivalent gains.

We find an aggregate consumption equivalent cost of mental illness Δ^m of 1.2 percent of consumption. The aggregate welfare cost of mental illness masks substantial heterogeneity in the cross-section. Figure 3 shows the distribution of the consumption equivalent welfare costs by mental health status. The height of the bar is the fraction of individuals with a particular welfare cost within each mental health status: healthy (blue), mild illness (orange), serious illness (black). Since, by definition, $\Delta_i^m = 0$ for individuals who are healthy, no blue bars appear in Figure 3. The welfare effects are driven by individuals who are not healthy, which is 13.4 percent of the population: 3.7 percent experience serious mental illness, and 9.7 percent experience mild illness. The average welfare cost of mental illness for individuals experiencing serious mental illness is equivalent to 13.3 percent of consumption, while the average consumption equivalent cost of mental illness for individuals experiencing mild illness is 6.9 percent. Taken together, this yields an aggregate welfare cost of $0.037 \times 13.3 + 0.097 \times 6.9 = 1.2$ percent.[36]

[36] In Figure F.2, we display heterogeneity in the welfare costs of mental illness by age and wealth groups. We show that the welfare costs are larger for younger individuals than for older individuals. Younger individuals experience an average welfare cost of 1.6 percent while older individuals experience a welfare cost of 0.8 percent.

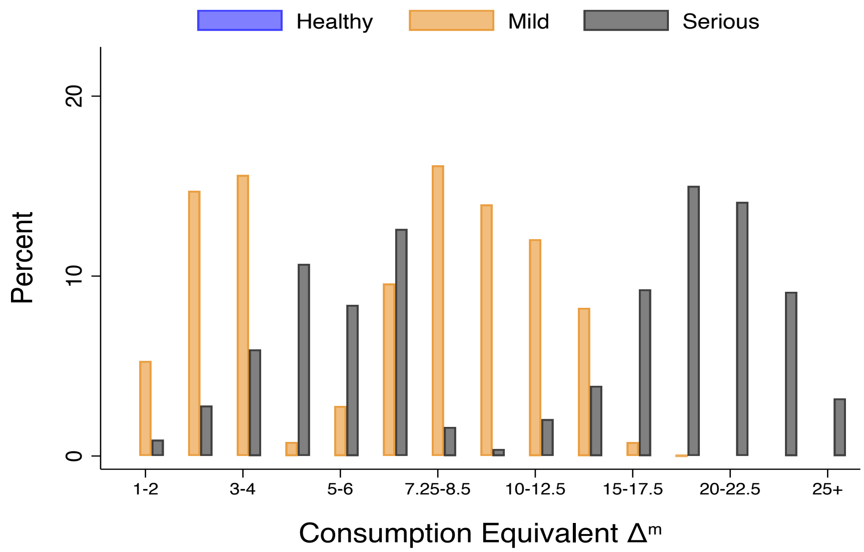

Figure 3: Welfare Cost of Mental Illness

Figure 3 shows the distribution of the consumption equivalent welfare costs of mental illness Δ_i^m by mental health status. The height of the bars captures the fraction of individuals with a particular welfare cost for each health status: healthy (blue), mild illness (orange), serious illness (black). Since healthy individuals do not experience a gain from becoming healthy, the blue bars record a value of zero.

5.2 Mental Health Policies

We now evaluate the effects of three widely discussed public policies that aim to improve mental health: expanding availability of mental health services, lowering out-of-pocket costs of treatment, and improving mental health of adolescents and young adults. In Appendix G, we provide more detail on specific policies.

5.2.1 Expanding Availability of Mental Health Services

We consider the consequences of expanding availability of treatment. Lack of availability of mental health services is one of the most commonly cited barriers to treatment.[37] According to the U.S. Department of Health and Human Services, in 2023, approximately 165 million Americans live in Health Professional Shortage Areas (HPSA), which are geographic areas that experience a shortage of mental health professionals.[38] In these areas, the number of mental health professionals is only 27.2 percent of the required capacity to meet the population's treatment needs on average. Given that the U.S. population size is 341 million, this implies that the share of Americans who do not have access to mental health services is

[37] See the White House Fact Sheets (www.whitehouse.gov/s1, www.whitehouse.gov/s2, www.whitehouse.gov/s3).

[38] For statistics on the number of Americans living in HPSA, see www.kff.org. The fraction of treatment needs met is calculated by the HPSA as the number of psychiatrists available to serve a population group divided by the number of psychiatrists that is needed to completely eliminate the shortage of mental health professionals to this population group, where the required number of psychiatrists is one for every 30,000 individuals. For more detail see www.kff.org.

Table 10: The Effects of Expanding Availability to Mental Health Services

	Benchmark $\omega_\tau = 0.682$			Increased availability $\omega_\tau = 1$		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Mental health shares	0.866	0.097	0.037	0.882	0.085	0.033
Treatment shares	0.000	0.414	0.657	0.000	0.699	0.632
Hours worked	0.404	0.383	0.351	0.404	0.377	0.350
Income (in thousands)	64	56	46	64	54	46
Wealth (in thousands)	292	262	236	290	255	228
Risky investment share	0.572	0.461	0.385	0.569	0.438	0.378
Risky participation rate	0.601	0.516	0.449	0.598	0.497	0.441

Table 10 reports the effects of expanding availability of mental health services. The first three columns report the averages by mental health status in the benchmark economy where a fraction $\omega_\tau = 0.682$ of individuals has access to mental health services when mildly ill. The final three columns report the moments of a counterfactual economy where all individuals have access to mental health services when mildly ill, $\omega_\tau = 1$.

$(1 - 0.272) \times \frac{165}{341} = 0.35$. This estimate aligns well with the limited availability of $1 - \omega_\tau = 0.32$ that we estimate internally in our model.

Various policies are considered to increase the availability of treatment in response to its shortage. One prominent policy is to increase the supply of mental health care professionals. A second policy is to expand access to treatment through community health clinics. A third set of policies aims to expand access to treatment through virtual mental health care.

We evaluate a policy that makes treatment available to all. This corresponds to an economy where all individuals can choose to get treated when they experience mild mental illness. That is, we consider an increase of ω_τ from 0.682 to 1. Table 10 presents the results. Expanding availability of mental health services reduces the share of individuals who experience mental illness by 1.6 percentage points relative to the benchmark economy. The share of individuals with serious illness decreases by 0.4 percentage points, from 3.7 to 3.3 percent, while the share of individuals with mild illness decreases by 1.2 percentage points, from 9.7 to 8.5 percent. This reduction in mental illness is driven by a significant increase in the treatment share among individuals experiencing mild illness, from 41.4 to 69.9 percent.

The increase in the treatment share among individuals experiencing mild illness is driven by both compositional and direct effects. When treatment is available, the distribution over mental health states is $(0.882, 0.085, 0.033)$ with corresponding treatment shares 0.699 and 0.632 for mild and serious illness. When treatment is not available, the stationary distribution over mental health is $(0.833, 0.122, 0.046)$ with 0.694 of the individuals experiencing serious illness seeking treatment. When about a third of the population gains access to mental treatment services, the treatment share among individuals with mild illness increases for two reasons. First, the group that originally had access to treatment conditional on experiencing mild illness increases from $\frac{0.682\times0.085}{0.682\times0.085+0.318\times0.122}$ to 0.682, which increases the treatment share from 0.414 to $0.682 \times 0.699 = 0.477$. Second, the direct effect increases the treatment share among individuals with mild illness from 0.526 by 0.318×0.699 to 0.699.

When treatment is available to all, average hours worked and average income slightly decrease among individuals experiencing mild illness. These decreases are driven by increased treatment. The treatment share increases by 28.5 percent. Since treatment is associated with a time cost of $n_\tau = 0.02$, this reduces working hours by about 0.57 hours per week. Individuals save less as the precautionary motive for savings is lower when mental illness spells are shorter due to increased availability of treatment. While, all else equal, better mental health increases individuals' risky investment share, the fact that individuals are now less wealthy leads them to invest on average less in risky assets.

In order to evaluate the welfare benefits of expanding availability, we calculate the consumption equivalent welfare gain for the cross-section of individuals. The consumption equivalent measure of providing full availability to mental health services Δ_i^ω is given by:

$$\log \Delta_i^\omega = \beta_t\big(v_t(a_{it}, \nu_{it-1}, m_{it}, \omega_i = 1) - v_t(a_{it}, \nu_{it-1}, m_{it}, \omega_i)\big), \tag{18}$$

where $\omega_i \in \{0, 1\}$ indicates whether individual i has access to treatment when experiencing mild illness. In the counterfactual economy, all individuals have access to treatment as indicated by $\omega_i = 1$. The consumption equivalent welfare gain is such that individual i is indifferent between a per period consumption increase Δ_i^ω and between having full access to treatment.

The average welfare benefit of providing full availability of treatment services, Δ^ω, is 0.31 percent of aggregate consumption, or 41 billion dollars annually.[39] This aggregate welfare benefit masks heterogeneity in the cross-section. For individuals who have access to treatment in the baseline, this policy yields no welfare gain, or $\Delta_i^\omega = 0$. Since 68.2 percent of the population have access to mental health services in the baseline, the welfare gains are driven by the 31.8 percent of the population which gains access due to the policy.

Figure 4 shows the welfare gains from full accessibility to treatment by mental health status. The

[39] Aggregate consumption expenditures in NIPA in 2011 amount to 13,222 billion dollars.

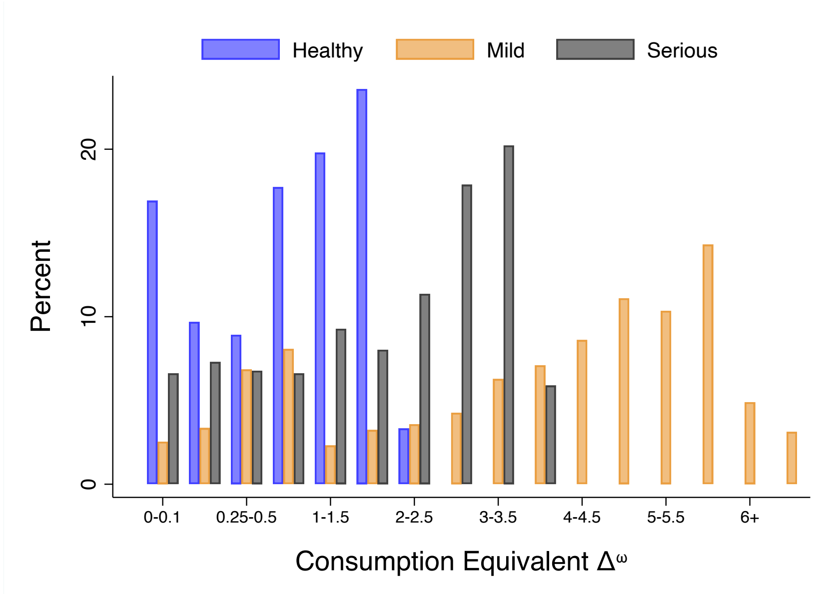

Figure 4: Welfare Gains from Increased Availability of Mental Health Services

Figure 4 displays the distribution of the consumption equivalent welfare gains of increased availability of treatment services by mental health status. The height of the bars captures the percentage of individuals with a particular welfare gain within each mental health status: healthy (blue), mild illness (orange), and serious illness (black).

welfare gains of full availability increase with the expected use of treatment services among those who do not have access. The welfare gains are largest for individuals who are mildly ill, yet do not have access to treatment services in the baseline (as indicated by the orange bars). The welfare gains for these individuals are concentrated between consumption equivalents of 3 and 7 percent. Welfare gains are also large for individuals who experience serious mental illness. Even though these individuals have access to treatment given their serious illness, some expect to lose access if their mental illness becomes mild. When access to mental health services is provided to all, they can continue receiving treatment when mildly ill. Importantly, healthy individuals who do not have access to treatment if they become mildly ill also benefit from increased access. These individuals are now less likely to experience serious mental illness, and the length of the illness is shorter since they can get treatment if they become mildly ill.

Figure F.3 shows the distribution of welfare gains by age and wealth levels in the baseline economy. Increased availability most strongly benefits individuals for whom the cost of mental illness is highest. These are individuals who are younger and from the middle class. The average welfare gain of increased availability to younger individuals is 0.6 percent of consumption, relative to 0.1 percent for older individuals. In addition to calculating the consumption equivalent gains from expanding access to mental health services to all, we estimate the welfare gains of partial expansions. The welfare gain from providing access to mental health treatment services accrue linearly. Each additional 10 percentage point increase in treatment availability translates into an average consumption equivalent welfare gain of $0.31 \times \frac{0.10}{0.318} = 0.1$ percent. Finally, in Appendix H, we estimate that the costs of expanding treatment availability are at most 3.8 billion dollars per year. This suggests that the resulting benefits of 41 billion dollars per year significantly exceed the costs.[40]

5.2.2 Reducing Treatment Costs

The second policy we evaluate is reducing the private out-of-pocket costs of mental health treatment. In the United States, the out-of-pocket cost of mental health services was reduced through the expansion of Medicaid, and through mental health parity laws which require health insurers to cover mental health care in parity with physical health care, enacted by the Affordable Care Act in 2010. We consider a further reduction of the out-of-pocket costs of mental health services, specifically, a policy under which individuals do not pay out of pocket for their treatment, or $\varphi_\tau = 0$.

We find that the average welfare gain of eliminating the out-of-pocket cost of treatment Δ^φ is 0.16 percent of consumption.[41] We summarize the implications in Table F.4. Treatment shares increase for individuals experiencing mild and serious mental illness, which translates to a reduction in mental illness in the population. Our model features four motives for low uptake in treatment discussed in the psychiatric literature (see Section 2): negative thinking, lack of availability, affordability, and stigma. By comparing these results with the sizable welfare benefits of increasing treatment availability, we conclude that lack of availability rather than affordability is the more salient policy barrier for mental health treatment.[42] The smaller response to lower treatment costs is similar to Cronin, Forsstrom, and Papageorge (2024), and is driven by a relative low monetary cost of treatment in the benchmark economy.

[40] A potential concern is that the costs of expanding treatment availability do not extrapolate linearly because it may become more difficult to draw in more able providers into mental health occupations. Since both the benefits and costs of expanding treatment availability accrue linearly in our framework, we capture correctly the cost-benefit tradeoff for small increases in availability. In addition, since the overall benefits, which we consider a lower bound, outweigh the costs by a factor of 8, expanding treatment availability still passes the cost-benefit tradeoff when costs increase nonlinearly.

[41] We construct the aggregate welfare benefit by aggregating the individual consumption equivalent welfare gain Δ_i^φ, which satisfies $\log \Delta_i^\varphi = \beta_t(v_t(a_{it}, \nu_{it-1}, m_{it}, \omega_i; \varphi_\tau = 0) - v_t(a_{it}, \nu_{it-1}, m_{it}, \omega_i; \varphi_\tau))$, where φ_τ indicates the explicit dependence of welfare on the out-of-pocket cost of treatment.

[42] The increase in non-private treatment cost due to eliminating private treatment cost is 5.5 billion dollars. The fraction of adults that get treated every period under the counterfactual economy is $0.095 \times 45.8 + 0.035 \times 75.3 = 6.99$ percent relative to 6.45 percent in the benchmark economy. Treatment increases by 0.55 percent of 230 million US adults, for whom the full private cost of 2.0 billion dollars ($= 1,250 \times 1.6$ million adults) for treatment is incurred. For the remaining 6.45 percent, the non-private costs increase by 3.5 billion dollars ($= 0.19 \times 1,250 \times 14.8$ million adults).

5.2.3 Improving Mental Health of Young Adults

A third policy that is proposed is to improve the mental health of adolescents and young adults. Examples of such policy measures are increasing the number of mental health professionals in schools and providing better mental health education.

In order to evaluate the implications of increased treatment of young adults, we consider a policy that changes the initial distribution over mental health states. Mental health treatment in late adolescence and young adulthood under the proposed policies takes place before age 25, when individuals enter our model, and hence alters the initial distribution of mental health states. Specifically, we consider a counterfactual economy with an initial distribution that would emerge at age 25 if: (1) the distribution of mental illness at age 16 is identical to our baseline distribution of mental illness at age 25, and (2) all individuals who experience mental illness between ages 16 and 25 receive treatment. To assess the welfare implications of treatment for individuals in adolescence and young adulthood, we consider the consumption equivalent welfare gains for 25 year olds in the model. The consumption equivalent $\Delta_i^{\tau_0}$ is such that a 25-year old individual i is indifferent between a per period consumption increase of $\Delta_i^{\tau_0}$ and living in the counterfactual economy. It satisfies:

$$\log \Delta_i^{\tau_0} = \beta_1 \left(v_1(0, \nu_{i0}, \tilde{m}_{i1}, \omega_i) - v_1(0, \nu_{i0}, m_{i1}, \omega_i) \right), \tag{19}$$

where $\tilde{m}_{i1}$ is the mental health of individual i under the counterfactual policy and m_{i1} corresponds to the mental health of individual i in the baseline.

We find that the average consumption equivalent gain of treatment in young adulthood Δ^{τ_0} is equal to 0.96 percent, or 127 billion dollars. In order to understand the gain of 0.96 percent, we note that the consumption equivalent gain of being healthy is 8.1 percent for 25 year olds with mild mental illness and 18.0 percent for individuals with serious illness. Treatment of young adults improves the mental health distribution of 25 year olds. The share of healthy individuals increases by 8.5 percentage points to 90.1 percent, with a corresponding reduction in individuals with mild illness from 13.5 percent to 7.5 percent (a decrease of 6.0 percentage points) and a reduction in individuals with serious illness from 5.1 percent to 2.5 percent (a decrease of 2.6 percentage points). As a result, the consumption equivalent welfare gain is $0.060 \times 8.1 + 0.026 \times 18.0 = 0.96$ percent. We conclude that improving mental health of adolescents and young adults leads to significant welfare benefits.[43]

Table 11: Sensitivity Analyses

	Parameter	Δ^m	Δ^ω	Δ^φ	Δ^{τ_0}
1	Baseline	1.17	0.31	0.16	0.95
	Model Specification				
2	Ambiguity $\kappa(m_0) = 0.025$	1.01	0.24	0.25	0.86
3	Utility penalty	1.45	0.33	0.16	1.06
4	Heterogeneous types	0.95	0.23	0.15	0.62
5	Borrowing up to 20,000 dollars	1.18	0.29	0.16	0.95
6	Borrowing up to 50,000 dollars	1.21	0.29	0.16	0.96
7	Productivity elasticity $\theta = 0.40$	0.96	0.22	0.16	0.86
	Model Mechanisms				
8	Negative thinking	0.52	0.05	0.09	0.26
9	Rumination	0.76	0.27	0.12	0.75
10	Productivity loss	1.14	0.29	0.15	0.94
11	Stigma	1.11	0.35	0.19	0.92

Table 11 shows sensitivity analysis for the welfare costs of mental illness and the welfare benefits of mental health policies. Each row corresponds to a particular sensitivity analysis. For each specification, we show the welfare cost of mental illness Δ^m in the first column. The second to fourth column show the welfare benefits of increased availability of mental health treatment services Δ^ω, of reducing treatment costs Δ^φ, and of increasing the mental health of young adults Δ^{τ_0}. In each row for the sections on model specification, we recalibrate our economy to target the same moments as in Table 5.

5.2.4 Sensitivity Analysis

In this section we present sensitivity analyses for our quantitative results to alternative model specifications.

Each row in Table 11 corresponds to an alternative model specification. For each specification, we report the aggregate welfare cost of mental illness Δ^m in the first column. The second to fourth column show the welfare benefits of increased availability of mental health treatment services Δ^ω, of reducing treatment costs Δ^φ, and of increasing the mental health of young adults Δ^{τ_0}. In each row for the sections on model specification, we recalibrate our economy to target the same moments as in Table 5. We present the model fit for each recalibrated model in Appendix I.

The first row repeats our results for the baseline model and serves as a benchmark. In rows 2 to 4, we vary the model specification. The second row incorporates ambiguity aversion for healthy individuals by setting $\kappa(m_0) = 0.025$, which is the average subjective loss probability, or ambiguity aversion, in the ALP data (see Table 2). Correspondingly, we set $\kappa(m_1) = 0.056$ and $\kappa(m_2) = 0.092$ to preserve the extent of negative thinking observed in the data. Individuals experiencing mental illness still exhibit more negative thinking, or, in other words, are more ambiguity averse. The welfare cost of mental illness increases when healthy individuals are also ambiguity averse. Negative thinking becomes a more important barrier to treatment, which reduces the welfare benefit of expanding treatment availability.

The third row considers a model with a direct utility penalty of mental illness. We set the utility penalty to 5 percent of consumption per year, i.e. $\xi_m(m_1) = \xi_m(m_2) = 0.05$. Since the benefits of treatment are higher, we increase the utility cost of treatment to obtain the empirical treatment shares by mental health. The welfare cost of mental health consequently increases from 1.17 to 1.45 percent of aggregate consumption, indicating that the benchmark welfare costs are a lower bound.

The fourth row augments the baseline model to incorporate unobserved individual heterogeneity. Doing so, we capture the idea that individuals are heterogeneous in their vulnerability to mental illness (Beck, 2002; Mathews and MacLeod, 2005). We consider a model specification with innate types, where the initial distribution and the transition probabilities for mental health as well as the productivity process depend on the individual's type. We identify the classes of individual types by k-means clustering following Bonhomme, Lamadon, and Manresa (2019, 2022) and Jolivet and Postel-Vinay (2024).[44] Allowing for unobserved heterogeneity reduces welfare costs of mental illness and the welfare benefits of mental health policies.

[43] In Appendix H, we estimate an annual cost of improving mental health of young adults of 11.9 billion dollars.
[44] We provide details in Appendix J.

Rows 5 and 6 show that the results are not sensitive to the borrowing constraints. With a borrowing limit of 20 and 50 thousand dollars, the welfare numbers remain unchanged. Row 7 considers a model specification with a constant elasticity of labor productivity with respect to hours worked. Specifically, we consider an elasticity of productivity with respect to hours of $\theta = \theta_S = \theta_M = \theta_L = 0.4$ following French (2005). This eliminates the productivity penalty of working long hours, and mutes the welfare effects of mental illness and mental health policies.

Finally, rows 8 to 11 evaluate the role of different model mechanisms for the aggregate cost of mental illness and for policy evaluation. In order to do so, we eliminate one mental health mechanism in each row and resolve the model. In row 8 we set $\kappa(m) = 0$, and in row 9 we set $n_r(m) = 0$. The results show that negative thinking and rumination are the main features of mental health in terms of the welfare costs and the benefit of expanding treatment availability. Absent negative thinking, the welfare cost of mental illness decreases to 0.52 percent of consumption, while the benefit of expanding treatment availability drops to 0.05 percent. Absent rumination, the aggregate welfare cost of mental illness decreases to 0.76 percent of consumption, while the welfare gain of expanding treatment availability decreases to 0.17 percent of consumption. Finally, in row 10 we set the productivity penalty associated with mental illness to zero, $\Lambda(m) = 0$, and in row 11 we set the stigma cost of treatment to zero, or $\xi_\tau = 0$. The results show that the direct productivity loss due to mental illness and the stigma costs of treatment do not have large welfare consequences. The welfare costs of mental illness and the policy benefits remain largely unchanged as we eliminate these channels.

5.2.5 Comparison to Epidemiological Literature

We now compare our aggregate welfare costs of mental illness to the estimates of the burden of mental illness from the epidemiological literature. Our estimate of the welfare costs of 1.2 percent of aggregate consumption is equivalent to 154 billion dollars annually, which is 24 percent larger than established estimates from the epidemiological literature. The epidemiological literature focuses on three types of mental health costs: costs due to impaired functioning in the workplace, direct medical expenditures, and suicide-related costs.[45] Greenberg et al. (2015) estimate a workplace costs of 105 billion, direct expenditures on medical and pharmaceutical services of 102 billion, and suicide-related costs of 10 billion.[46]

[45] Workplace costs are typically estimated by assessing the cost of missed work days and the cost of hours where the individual is at work but not working. The estimates abstract from other work-related cost such as unemployment costs. The cost of selection into lower-earning jobs is also not accounted for since the cost of missed hours of work is computed assuming that the wage that the individual would have earned during these hours is the average wage in the economy. Suicide-related costs are estimated as lifetime earnings lost due to mental health related suicides.

[46] We use the estimates of the economic cost of mental illness for 2010 in Greenberg et al. (2015), which is the middle of our period of analysis. Together with Greenberg et al. (2021), this paper estimates that the societal cost

Our aggregate consumption equivalent welfare cost is the analog of the workplace costs together with the privately incurred healthcare costs, which Greenberg et al. (2015) thus estimate as $105 + 0.19 \times 102 = 124$ billion.[47] Our estimate of 154 billion dollars per year thus exceeds the epidemiological estimate of 124 billion dollars by 24 percent.[48]

We remark that our estimate of the societal cost of mental illness takes into account the stochastic life-cycle evolution of mental illness and optimal static and dynamic responses when experiencing mental illness. First, our welfare measure takes into consideration that being healthy today lowers the likelihood of experiencing mental illness later in life. Second, our welfare measure takes into account that mental health affects the labor supply decisions both in terms of job choice and in terms of hours worked. Third, our estimate incorporates the effect of mental illness on dynamic savings decisions and portfolio choices. Improving mental health today improves future well-being through improved mental health in the future, increased savings and increases returns on savings by changing the portfolio allocation towards higher expected-return investments. Finally, our welfare cost of mental illness takes into account the psychological costs of the cognitive distortion of mental illness − negative thinking.

Relative to the epidemiological literature, our approach has the advantage of being able to account for optimal static and dynamic responses to mental illness regarding for consumption, labor, and portfolio decisions. Our structural approach is also important for quantitative policy evaluation. In order to illustrate, in Section 5.2.1, we estimate the welfare benefit of full treatment availability to be 41 billion dollars annually. However, according to the epidemiological calculation the cost of mental health increases with expanded treatment availability. Table 10 shows that the overall treatment share strongly increases and that hours worked also decrease. By evaluating policy using the epidemiological costs of mental health one would erroneously conclude that increased treatment availability increases the costs of mental illness. Instead, we find that increased treatment availability generates sizable welfare gains. In other words, while the societal costs of mental illness may be a useful measure for the societal costs when policy does not change, it cannot be used for policy evaluation.

of mental illness has increased from 179 billion to 299 billion dollars between 2005 and 2018.

[47]We assume that individuals pay 19 percent of the total mental healthcare costs out-of-pocket while 81 percent is covered by insurance (Cronin, Forsstrom, and Papageorge, 2024).

[48]Given a U.S. adult population of 230 million, this corresponds to a welfare cost of mental illness of 670 dollars per person. De Nardi, Pashchenko, and Porapakkarm (2024) calculate a lifetime cost of bad health of 1,500 dollars.

Table 12: Transition Matrix with Improved Mental Health Treatment

Treatment	Healthy	Mild	Serious	*Improved Tech*	Healthy	Mild	Serious
Healthy	0.934	0.055	0.011	Healthy	0.934	0.055	0.011
Mild	0.766	0.161	0.073	Mild	0.855	0.073	0.072
Serious	0.332	0.360	0.308	Serious	0.370	0.401	0.229

Table 12 presents the mental health transition matrix for individuals receiving treatment in the benchmark economy, which is displayed on the left, and for individuals receiving treatment in the economy with a treatment technology that is 25 percent more effective, which is displayed on the right.

5.3 Improving Mental Health Treatment

We next quantify the welfare consequences of improving the efficacy of mental health treatment. More efficient treatment corresponds to technological or medical advances in therapy and anti-depressant medication. We consider a counterfactual economy where treatment is 25 percent more effective.

In order to evaluate the impact of improved treatment efficacy, we re-estimate the transition matrix between mental health states, conditional on treatment, to match an SMD of -0.875, relative to an SMD of -0.7 in the baseline economy discussed in Section 4. We assume that improved treatment implies that the likelihood that mental health improves following treatment is a factor δ_+ higher than in the baseline, and that the likelihood that mental health deteriorates following treatment is a factor δ_- lower. We estimate the parameters δ_+ and δ_- such that the model implied SMD given mild illness and the model implied SMD given serious illness both equal -0.875. We obtain $\delta_+ = 1.12$ and $\delta_- = 0.98$. The right panel of Table 12 shows the mental health transition matrix under the improved treatment technology. Relative to the baseline in the left panel, the entries in the strictly lower triangular part of the transition matrix increase by a factor $\delta_+ = 1.12$ under the improved treatment technology. For example, the transition probability from serious to mild illness is 0.360 in the baseline economy and $0.360 \times 1.12 = 0.401$ under the improved technology.

In order to quantify the implications of improved mental health treatment, we evaluate the welfare gain for 25 year olds for the economy with improved treatment. The average consumption equivalent gain of a 25 percent increase in treatment efficacy is 0.7 percent, or 91 billion dollars annually. The results are approximately linear in the extent of the improvement of mental healthcare. For example, a 10 percent increase in treatment efficacy translates into a consumption equivalent gain of 0.3 percent, whereas a 10 percent decrease in treatment efficacy translates into a consumption equivalent loss of 0.3 percent. We conclude that improving the efficacy of treatment leads to significant welfare gains.

6 Conclusion

This paper develops a quantitative macroeconomic theory of mental health. Based on classic and modern psychiatric theories, we model mental illness as a state of negative thinking and rumination which is reinforced through behavior. In the model, agents who experience mental illness have negative expectations of future productivity, risky returns, and the efficacy of mental health treatment, and lose time due to rumination. As a result, they work less, consume less, invest less in risky assets, and forego treatment. Foregoing treatment, in turn, reinforces their mental illness.

We discipline our model using micro data on mental health. We quantify the extent of negative thinking among individuals with mental illness from subjective loss probabilities, which are elicited using survey data. We estimate parameters that govern rumination, the efficacy and availability of treatment, and its costs so that the model matches the prevalence of mental illness, transition dynamics of mental health, observed treatment shares, and labor choices among individuals with mental illness. We validate our model by showing that it also matches non-targeted moments that describe the relation between mental health, consumption, income, wealth, and portfolio choice.

We use our model to evaluate the welfare costs of mental illness and the effects of mental health policies. We find the aggregate welfare cost of mental illness to be 1.2 percent of aggregate consumption per year. Our policy analysis shows that expanding the availability of mental health services substantially improves mental health and welfare. Reducing the out-of-pocket cost of mental health services has a significantly smaller welfare impact. Finally, we find that policies that promote treatment of mental illness among adolescents and young adults can substantially improve welfare.

References

Abramson, B. (2024): “The Equilibrium Effects of Eviction Policies,” Discussion paper, Columbia Business School Working Paper.

Abramson, B. and van Nieuwerburgh, S. (2024): “Rent Guarantee Insurance,” Discussion paper, NBER Working Paper No. 32582.

Adams, R.A., Huys, Q.J. and Roiser, J.P. (2015): “Computational Psychiatry: Towards a Mathematically Informed Understanding of Mental Illness,” *Journal of Neurology, Neurosurgery & Psychiatry*, 87, 53–63.

Aguiar, M. and Hurst, E. (2013): “Deconstructing Life Cycle Expenditure,” *Journal of Political Economy*, 121(3), 437–492.

Albrecht, J. and Vroman, S. (2002): “A Matching Model with Endogenous Skill Requirements,” *International Economic Review*, 43(1), 283–305.

Ameriks, J. et al. (2020): “Long-Term-Care Utility and Late-in-Life Saving,” *Journal of Political Economy*, 128(6), 2375–2451.

Amsterdam, J. et al. (1987): “Taste and Smell Perception in Depression,” *Biological Psychiatry*, 22(12), 1481–1485.

Barth, J. et al. (2016): “Comparative Efficacy of Seven Psychotherapeutic Interventions for Patients with Depression: A Network Meta-analysis,” *Focus*, 14(2), 229–243.

Beck, A.T. (1967): *The Diagnosis and Management of Depression*. University of Pennsylvania Press.

——— (1976): *Cognitive Therapy and the Emotional Disorders*. New York: International Universities Press.

——— (2002): “Cognitive Models of Depression,” *Clinical Advances in Cognitive Psychotherapy: Theory and Application*, 14(1), 29–61.

——— (2008): “The Evolution of the Cognitive Model of Depression and Its Neurobiological Correlates,” *American Journal of Psychiatry*, 165(8), 969–977.

Beck, A.T. and Bredemeier, K. (2016): "A Unified Model of Depression: Integrating Clinical, Cognitive, Biological, and Evolutionary Perspectives," *Clinical Psychological Science*, 4(4), 596–619.

Beck, A.T. and Clark, D.A. (1991): "Anxiety and Depression: An Information Processing Perspective," in *Anxiety and Self-Focused Attention*, pp. 41–54.

Beck, A.T., Emery, G. and Greenberg, R.L. (1985): *Anxiety Disorders and Phobias: A Cognitive Perspective*. New York: Basic Books.

Berlin, I. et al. (1998): "Measures of Anhedonia and Hedonic Responses to Sucrose in Depressive and Schizophrenic Patients in Comparison with Healthy Subjects," *European Psychiatry*, 13(6), 303–309.

Bhandari, A., Borovička, J. and Ho, P. (2024): "Survey Data and Subjective Beliefs in Business Cycle Models," *Review of Economic Studies*, p. rdae054.

Bick, A., Blandin, A. and Rogerson, R. (2022): "Hours and Wages," *Quarterly Journal of Economics*, 137(3), 1901–1962.

Bishop, S.J. and Gagne, C. (2018): "Anxiety, Depression, and Decision Making: A Computational Perspective," *Annual Review of Neuroscience*, 41, 371–388.

Blundell, R., Pistaferri, L. and Saporta-Eksten, I. (2016): "Consumption Inequality and Family Labor Supply," *American Economic Review*, 106(2), 387–435.

Boar, C. (2021): "Dynastic Precautionary Savings," *Review of Economic Studies*, 88(6), 2735–2765.

Boerma, J. and Karabarbounis, L. (2021): "Inferring Inequality with Home Production," *Econometrica*, 89(5), 2517–2556.

Bonhomme, S., Lamadon, T. and Manresa, E. (2019): "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 87(3), 699–739.

——— (2022): "Discretizing Unobserved Heterogeneity," *Econometrica*, 90(2), 625–643.

Braun, R.A., Kopecky, K.A. and Koreshkova, T. (2017): "Old, Sick, Alone, and Poor: A Welfare Analysis of Old-Age Social Insurance Programmes," *Review of Economic Studies*, 84(2), 580–612.

——— (2019): "Old, Frail, and Uninsured: Accounting for Features of the US Long-Term Care Insurance Market," *Econometrica*, 87(3), 981–1019.

Braxton, J.C. and Taska, B. (2023): "Technological Change and the Consequences of Job Loss," *American Economic Review*, 113(2), 279–316.

Burt, D.B., Zembar, M.J. and Niederehe, G. (1995): "Depression and Memory Impairment: A Meta-Analysis of the Association, Its Pattern, and Specificity," *Psychological Bulletin*, 117(2), 285.

Butler, G. and Mathews, A. (1983): "Cognitive Processes in Anxiety," *Advances in Behaviour Research and Therapy*, 5(1), 51–62.

Cairney, J. et al. (2007): "Evaluation of 2 Measures of Psychological Distress as Screeners for Depression in the General Population," *Canadian Journal of Psychiatry*, 52(2), 111–120.

Caspi, A. et al. (2003): "Influence of Life Stress on Depression: Moderation by a Polymorphism in the 5-HTT Gene," *Science*, 301(5631), 386–389.

Chetty, R. et al. (2012): "Does Indivisible Labor Explain the Difference Between Micro and Macro Elasticities? A Meta-Analysis of Extensive Margin Elasticities," *NBER Macroeconomics Annual*, 27, 1–56.

Clark, D.A. et al. (2000): "Scientific Foundations of Cognitive Theory and Therapy of Depression," *Journal of Cognitive Psychotherapy*, 14(1), 100–106.

Clement, S. et al. (2015): "What is the Impact of Mental Health-Related Stigma on Help-Seeking? A Systematic Review of Quantitative and Qualitative Studies," *Psychological Medicine*, 45(1), 11–27.

Cocco, J.F., Gomes, F.J. and Maenhout, P.J. (2005): "Consumption and Portfolio Choice over the Life Cycle," *Review of Financial Studies*, 18(2), 491–533.

Cole, H.L., Kim, S. and Krueger, D. (2019): "Analysing the Effects of Insuring Health Risks: On the Trade-off Between Short-Run Insurance Benefits Versus Long-Run Incentive Costs," *Review of Economic Studies*, 86(3), 1123–1169.

Corbae, D., Glover, A. and Nattinger, M. (2024): "Equilibrium Evictions," Discussion paper, NBER Working Paper No. 32898.

Corrigan, P. (2004): "How Stigma Interferes with Mental Health Care," *American Psychologist*, 59(7), 614.

Cronin, C.J., Forsstrom, M.P. and Papageorge, N.W. (2024): "What Good Are Treatment Effects without Treatment? Mental Health and the Reluctance to Use Talk Therapy," *Review of Economic Studies*, p. rdae061.

Currie, J. and Madrian, B.C. (1999): "Health, Health Insurance and the Labor Market," in *Handbook of Labor Economics*, vol. 3, pp. 3309–3416.

Davis, S.J. and Wachter, T.V. (2011): "Recessions and the Costs of Job Loss," *Brookings Papers on Economic Activity*, (2), 1–72.

De Nardi, M., French, E. and Jones, J.B. (2010): "Why Do The Elderly Save? The Role of Medical Expenses," *Journal of Political Economy*, 118(1), 39–75.

De Nardi, M., French, E. and Jones, J.B. (2016): "Medicaid Insurance in Old Age," *American Economic Review*, 106(11), 3480–3520.

De Nardi, M., Pashchenko, S. and Porapakkarm, P. (2024): "The Lifetime Costs of Bad Health," *Review of Economic Studies*, p. rdae080.

De Quidt, J. and Haushofer, J. (2016): "Depression for Economists," Discussion paper, NBER Working Paper No. 22973.

Derryberry, D. and Reed, M.A. (2002): "Anxiety-Related Attentional Biases and Their Regulation by Attentional Control," *Journal of Abnormal Psychology*, 111(2), 225.

Dichter, G. et al. (2010): "Unipolar Depression Does Not Moderate Responses to the Sweet Taste Test," *Depression and Anxiety*, 27(9), 859–863.

Dimmock, S.G. et al. (2016): "Ambiguity Aversion and Household Portfolio Choice Puzzles: Empirical Evidence," *Journal of Financial Economics*, 119(3), 559–577.

——— (2021): "Household Portfolio Underdiversification and Probability Weighting: Evidence from the Field," *Review of Financial Studies*, 34(9), 4524–4563.

Disner, S.G. et al. (2011): "Neural Mechanisms of the Cognitive Model of Depression," *Nature Reviews Neuroscience*, 12(8), 467–477.

Dobson, K.S. and Dozois, D.J. (2019): *Handbook of Cognitive-Behavioral Therapies*. Guilford Publications.

Ehring, T. and Watkins, E.R. (2008): "Repetitive Negative Thinking as a Transdiagnostic Process," *International Journal of Cognitive Therapy*, 1(3), 192–205.

Ekers, D., Richards, D. and Gilbody, S. (2008): "A Meta-Analysis of Randomized Trials of Behavioural Treatment of Depression," *Psychological Medicine*, 38(5), 611–623.

Ellsberg, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75(4), 643–669.

Epstein, L.G. and Schneider, M. (2003): "Recursive Multiple-Priors," *Journal of Economic Theory*, 113(1), 1–31.

Eysenck, M. (2014): *Anxiety and Cognition: A Unified Theory*. Psychology Press.

Fagereng, A., Gottlieb, C. and Guiso, L. (2017): "Asset Market Participation and Portfolio Choice over the Life-Cycle," *Journal of Finance*, 72(2), 705–750.

Fang, H. and Krueger, D. (2022): "The Affordable Care Act After a Decade: Its Impact on the Labor Market and the Macro Economy," *Annual Review of Economics*, 14, 453–494.

French, E. (2005): "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour," *Review of Economic Studies*, 72(2), 395–427.

French, E. and Jones, J.B. (2011): "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," *Econometrica*, 79(3), 693–732.

Furukawa, T.A. et al. (2003): "The Performance of the K6 and K10 Screening Scales for Psychological Distress in the Australian National Survey of Mental Health and Well-Being," *Psychological Medicine*, 33(2), 357–362.

GBD (2018): "Global, Regional, and National Incidence, Prevalence, and Years Lived with Disability for 354 Diseases and Injuries for 195 Countries and Territories, 1990−2017: A Systematic Analysis for the Global Burden of Disease Study 2017," *The Lancet*, 392(10159), 1789–1858.

Gilboa, I. and Schmeidler, D. (1989): "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, 18(2), 141–153.

Goldin, C. (2014): "A Grand Gender Convergence: Its Last Chapter," *American Economic Review*, 104(4), 1091–1119.

Gomes, F. and Michaelides, A. (2005): “Optimal Life-cycle Asset Allocation: Understanding the Empirical Evidence,” *Journal of Finance*, 60(2), 869–904.

Gotlib, I.H. and Joormann, J. (2010): “Cognition and Depression: Current Status and Future Directions,” *Annual Review of Clinical Psychology*, 6, 285–312.

Gotlib, I.H. et al. (2004): “Attentional Biases for Negative Interpersonal Stimuli in Clinical Depression,” *Journal of Abnormal Psychology*, 113(1), 127.

Gourinchas, P.O. and Parker, J.A. (2002): “Consumption over the Life Cycle,” *Econometrica*, 70(1), 47–89.

Greenberg, P.E. et al. (2003): “The Economic Burden of Depression in the United States: How Did it Change Between 1990 and 2000?,” *Journal of Clinical Psychiatry*, 64(12), 1465–1475.

——— (2015): “The Economic Burden of Adults with Major Depressive Disorder in the United States (2005 and 2010),” *Journal of Clinical Psychiatry*, 76(2), 155–162.

——— (2021): “The Economic Burden of Adults with Major Depressive Disorder in the United States (2010 and 2018),” *Pharmacoeconomics*, 39(6), 653–665.

Greenwood, J., Guner, N. and Kopecky, K.A. (2022): “The Downward Spiral,” Discussion paper, NBER Working Paper No. 29764.

Grossman, M. (1972): “On the Concept of Health Capital and the Demand for Health,” *Journal of Political Economy*, 80(2), 223–255.

Hall, R.E. and Jones, C.I. (2007): “The Value of Life and the Rise in Health Spending,” *Quarterly Journal of Economics*, 122(1), 39–72.

Hardeveld, F. et al. (2010): “Prevalence and Predictors of Recurrence of Major Depressive Disorder in the Adult Population,” *Acta Psychiatrica Scandinavica*, 122(3), 184–191.

Heathcote, J., Storesletten, K. and Violante, G.L. (2010): “The Macroeconomic Implications of Rising Wage Inequality in the United States,” *Journal of Political Economy*, 118(4), 681–722.

——— (2014): “Consumption and Labor Supply with Partial Insurance: An Analytical Framework,” *American Economic Review*, 104(7), 2075–2126.

Hertel, P. (2004): "Memory for Emotional and Nonemotional Events in Depression," in *Memory and Emotion*, pp. 186–216.

Hosseini, R., Kopecky, K.A. and Zhao, K. (2024): "How Important is Health Inequality for Lifetime Earnings Inequality?," Discussion paper, University of Georgia Working Paper.

Hubbard, R.G., Skinner, J. and Zeldes, S.P. (1995): "Precautionary Saving and Social Insurance," *Journal of Political Economy*, 103(2), 360–399.

Huggett, M., Ventura, G. and Yaron, A. (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101(7), 2923–2954.

Huys, Q.J., Daw, N.D. and Dayan, P. (2015): "Depression: A Decision-Theoretic Analysis," *Annual Review of Neuroscience*, 38, 1–23.

Huys, Q.J., Maia, T.V. and Frank, M.J. (2016): "Computational Psychiatry as a Bridge from Neuroscience to Clinical Applications," *Nature Neuroscience*, 19(3), 404–413.

Ilut, C., Valchev, R. and Vincent, N. (2020): "Paralyzed by Fear: Rigid and Discrete Pricing Under Demand Uncertainty," *Econometrica*, 88(5), 1899–1938.

Ilut, C.L. and Schneider, M. (2014): "Ambiguous Business Cycles," *American Economic Review*, 104(8), 2368–2399.

Ilut, C.L. and Valchev, R. (2023): "Economic Agents as Imperfect Problem Solvers," *Quarterly Journal of Economics*, 138(1), 313–362.

Imrohoroglu, A. and Zhao, K. (2024): "Homelessness," .

Jarosch, G. and Pilossoph, L. (2019): "Statistical Discrimination and Duration Dependence in the Job Finding Rate," *Review of Economic Studies*, 86(4), 1631–1665.

Jolivet, G. and Postel-Vinay, F. (2024): "A Structural Analysis of Health and Labor Market Trajectories," *Review of Economic Studies*, p. rdae071.

Jordà, Ò. et al. (2019): "The Rate of Return on Everything, 1870–2015," *Quarterly Journal of Economics*, 134(3), 1225–1298.

Just, N. and Alloy, L.B. (1997): "The Response Styles Theory of Depression: Tests and an Extension of the Theory," *Journal of Abnormal Psychology*, 106(2), 221–229.

Kendler, K.S. et al. (2005): "The Interaction of Stressful Life Events and a Serotonin Transporter Polymorphism in the Prediction of Episodes of Major Depression: A Replication," *Archives of General Psychiatry*, 62(5), 529–535.

Kessler, R.C. et al. (2005): "Prevalence, Severity, and Comorbidity of 12-month DSM-IV Disorders in the National Comorbidity Survey Replication," *Archives of General Psychiatry*, 62(6), 617–627.

Kessler, R.C. et al. (2002): "Short Screening Scales to Monitor Population Prevalences and Trends in Non-Specific Psychological Distress," *Psychological Medicine*, 32(6), 959–976.

——— (2003): "Screening for Serious Mental Illness in the General Population," *Archives of General Psychiatry*, 60(2), 184–189.

——— (2008): "Trends in Mental Illness and Suicidality After Hurricane Katrina," *Molecular Psychiatry*, 13(4), 374–384.

——— (2009): "The Global Burden of Mental Disorders: An Update From The WHO World Mental Health (WMH) Surveys," *Epidemiology and Psychiatric Sciences*, 18(1), 23–33.

Kessler, R.C. et al. (2012): "Prevalence, Persistence, and Sociodemographic Correlates of DSM-IV Disorders in the National Comorbidity Survey Replication Adolescent Supplement," *Archives of General Psychiatry*, 69(4), 372–380.

Kopecky, K.A. and Koreshkova, T. (2014): "The Impact of Medical and Nursing Home Expenses on Savings," *American Economic Journal: Macroeconomics*, 6(3), 29–72.

Krueger, D. and Perri, F. (2006): "Does Income Inequality Lead to Consumption Inequality? Evidence and Theory," *Review of Economic Studies*, 73(1), 163–193.

Le Moult, J. and Gotlib, I.H. (2019): "Depression: A Cognitive Perspective," *Clinical Psychology Review*, 69, 51–66.

Low, H., Meghir, C. and Pistaferri, L. (2010): "Wage Risk and Employment Risk over the Life Cycle," *American Economic Review*, 100(4), 1432–1467.

LOW, H. AND PISTAFERRI, L. (2015): "Disability Insurance and the Dynamics of the Incentive Insurance Trade-off," *American Economic Review*, 105(10), 2986–3029.

MATHEWS, A. AND MACLEOD, C. (2005): "Cognitive Vulnerability to Emotional Disorders," *Annual Review of Clinical Psychology*, 1, 167–195.

MURIS, P. AND VAN DER HEIDEN, S. (2006): "Anxiety, Depression, and Judgments About The Probability of Future Negative and Positive Events in Children," *Journal of Anxiety Disorders*, 20(2), 252–261.

NOLEN-HOEKSEMA, S. (1991): "Responses to Depression and Their Effects on the Duration of Depressive Episodes," *Journal of Abnormal Psychology*, 100(4), 569–582.

——— (2000): "The Role of Rumination in Depressive Disorders and Mixed Anxiety/Depressive Symptoms," *Journal of Abnormal Psychology*, 109(3), 504–511.

NOLEN-HOEKSEMA, S., WISCO, B.E. AND LYUBOMIRSKY, S. (2008): "Rethinking Rumination," *Perspectives on Psychological Science*, 3(5), 400–424.

OJIO, Y. ET AL. (2021): "Innovative Approach to Adolescent Mental Health in Japan: School-Based Education About Mental Health Literacy," *Early Intervention in Psychiatry*, 15(1), 174–182.

PHILLIPS, W.J., HINE, D.W. AND THORSTEINSSON, E.B. (2010): "Implicit Cognition and Depression: A Meta-Analysis," *Clinical Psychology Review*, 30(6), 691–709.

RICHARDS, D. (2011): "Prevalence and Clinical Course of Depression: A Review," *Clinical Psychology review*, 31(7), 1117–1125.

RIDLEY, M. ET AL. (2020): "Poverty, Depression, and Anxiety: Causal Evidence and Mechanisms," *Science*, 370(6522).

RIOS-RULL, J.V. (1996): "Life-Cycle Economies and Aggregate Fluctuations," *Review of Economic Studies*, 63(3), 465–489.

SAMHSA (2022): "Key Substance Use and Mental Health Indicators in the United States: Results from the 2021 National Survey on Drug Use and Health," Discussion paper, HHS Publication No. PEP22−07−01−005, NSDUH Series H−57.

Schaefer, K.L. et al. (2010): "Perception of Facial Emotion in Adults with Bipolar or Unipolar Depression and Controls," *Journal of Psychiatric Research*, 44(16), 1229–1235.

Sergeyev, D., Lian, C. and Gorodnichenko, Y. (2024): "The Economics of Financial Stress," *Review of Economic Studies*, p. rdae110.

Shi, S. (2002): "A Directed Search Model of Inequality with Heterogeneous Skills and Skill-Biased Technology," *Review of Economic Studies*, 69(2), 467–491.

Singer, A.R. and Dobson, K.S. (2007): "An Experimental Investigation of the Cognitive Vulnerability to Depression," *Behaviour Research and Therapy*, 45(3), 563–575.

Sullivan, D. and Von Wachter, T. (2009): "Job Displacement and Mortality: An Analysis Using Administrative Data," *Quarterly Journal of Economics*, 124(3), 1265–1306.

Tanaka, T., Camerer, C.F. and Nguyen, Q. (2010): "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam," *American Economic Review*, 100(1), 557–571.

Turner, E.H. et al. (2008): "Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy," *New England Journal of Medicine*, 358(3), 252–260.

Watkins, E.R. (2008): "Constructive and Unconstructive Repetitive Thought," *Psychological Bulletin*, 134(2), 163.

Williams, J.M.G. et al. (1988): *Cognitive Psychology and Emotional Disorders*. John Wiley & Sons.

Macroeconomics of Mental Health

Online Appendix

Boaz Abramson, Job Boerma, and Aleh Tsyvinski

January 2025

A Solving the Negative Thinking Problem

We illustrate the solution to the negative thinking minimization problem by considering a simple example. Specifically, consider an example with values w_1 and $w_2 > w_1$, where the objective probability of the low outcome is q. Negative thinking is modeled as selecting a subjective probability p which solves $\min_p (1-p)w_2 + pw_1$ subject to $|p-q| \leqslant \kappa$. The solution to this problem satisfies:

$$p^* = q + \kappa. \tag{A.1}$$

The parameter κ, the total variation budget, represents the degree to which the subjective probability of the worst state exceeds the corresponding objective probability, which is the extent of negative thinking.

The general negative thinking problem is solved identically. For any set of N ordered values associated with events e, $w_1 < w_2 < \cdots < w_N$, consider choosing a probability distribution $(p_1, p_2, \ldots, p_N)$ satisfying the total variation constraint (9) that minimizes the expected value $\sum p_e w_e$. The solution to this program, negative thinking, consists of two parts. First, negative thinking maximally increases the subjective worst-case probability, that is, $p_1 = q_1 + \kappa$. The observation that differences in negative thinking are identified by differences in subjective probabilities of the worst outcomes thus also applies to the general case with N possible random outcomes. Second, negative thinking sequentially decreases the probabilities associated with the best outcomes. Specifically, negative thinking first decreases the probability of the best outcome by δ_N such that $p_N - \delta_N \geqslant 0$, then decreases the subjective probability of the second best outcome by δ_{N-1} such that $p_{N-1} - \delta_{N-1} \geqslant 0$, and so on until the total variation constraint binds, $\delta_N + \delta_{N-1} + \cdots = \kappa$.

B Empirical Evidence

In this appendix, we present empirical results on the relationship between mental health and consumption and portfolio choice.

Data. We quantify the relationship between mental health and consumption and portfolio choice using the Panel Study of Income Dynamics (PSID). We incorporate data from all waves from 2000 to 2020. Earlier waves lack information on respondents' mental health. Our analysis focuses on heads of households between 25 and 65 years of age. All dollar values are reported in 2015 values. Our measure of income is individual labor income over the past calendar year. Hours worked are measured as total hours worked including overtime. Hourly wage rates are computed as individual income divided by hours worked. Our benchmark measure of consumption is annual nondurable expenditures which include expenditures on food, utilities, child care, clothing, home insurance, telecommunications, home maintenance, and variable transportation costs.[49] For all analyses, we use the sample weights provided by the surveys.[50]

The PSID reports the mental health status of respondents using the Kessler Psychological Distress Scale. The Kessler Psychological Distress Scale (K6 scale) is widely used by the epidemiological literature to measure the mental health of survey respondents.[51] We classify individuals into three groups based on the K6 scale following Kessler et al. (2008). Individuals with a K6 score between 13 and 24 are classified as experiencing serious mental illness, individuals with a K6 score between 8 and 12 are classified as experiencing mild mental illness, and individuals with K6 scores between 0 and 7 are classified as healthy. The K6 scale is included in all PSID waves conducted between 2000 and 2020 expect for 2004.

We next describe the wealth variables. We categorize equity holdings, business assets and liabilities, and real estate assets and liabilities as risky, which we denote as the set $\mathtt{R}$. We classify checking accounts, vehicles, certificates of deposit, government bonds and debt balances (except for business loans and real estate debt) as safe, which we denote by the set $\mathtt{S}$. Individual retirement accounts and other assets are

[49] Our benchmark measure of consumption is closest to the consumption measures used by Aguiar and Hurst (2013) and Boerma and Karabarbounis (2021). Since detailed consumption expenditures are available in the PSID starting from 2004, we restrict the analysis with respect to consumption to this period.

[50] We drop observations where the head of the household is a student; where reported consumption expenditure is in the top and bottom 1 percent of the consumption distribution; and where reported wealth is in the top 0.1 percent and bottom 1 percent of the wealth distribution. We drop observations with an hourly wage below 3 dollars or above 300 dollars in 2015 dollars, and observations where respondents reported working less than 20 hours per week or more than 92 hours per week.

[51] The K6 scale is calculated based on respondents' answers to six questions (Kessler et al., 2002, 2003). In particular, respondents are asked the following: "In the past 30 days, about how often did you feel (1) sadness, (2) nervous, (3) restless or fidgety, (4) hopeless, (5) that everything was an effort, and (6) worthless". For each question, the individual responds (0) none of the time, (1) a little of the time, (2) some of the time, (3) most of the time, or (4) all of the time. The K6 scale is computed as the sum of respondents' answers to the six questions.

labeled mixed investments which we denote by the set $\mathtt{M}$. Total wealth is the sum of risky, safe, and mixed investments net of liabilities. The total set of assets and liabilities $\mathtt{A}$ is the union of sets $\mathtt{R}$, $\mathtt{S}$, and $\mathtt{M}$.

The risky investment share measures the share of risky assets and liabilities in a portfolio. It is the sum of absolute values of risky assets and liabilities and a half of mixed investments relative to the sum of absolute values over all assets and liabilities:

$$\text{Risky Investment Share}_i = \left(\sum_{h \in \mathtt{R}} |a_{hi}| + \frac{1}{2} \sum_{h \in \mathtt{M}} |a_{hi}| \right) \Big/ \sum_{h \in \mathtt{A}} |a_{hi}|, \tag{B.1}$$

where a_{hi} denotes asset and liabilities in category h for an individual i. When the risky investment share is strictly positive, the individual is exposed to risk in financial markets.

Estimation. In order to assess the extent to which consumption and portfolio choices vary with mental health, we estimate the following regressions. Let Y_{it} be the dependent variable of interest for individual i in year t. The variables of interest are log consumption and the risky investment share. Let D_{1it} be an indicator variable taking the value one when individual i experiences mild illness in year t. Let D_{2it} be an indicator variable taking the value one if individual i experiences serious mental illness in year t. The regressions further include a vector of individual controls X_{it}, such as the individual's age, sex, education, race, and household composition, income, and wealth.[52] We estimate the following regression:

$$Y_{it} = \gamma_t + \gamma_1 D_{1it} + \gamma_2 D_{2it} + \gamma_x X_{it} + \varepsilon_{it}. \tag{16}$$

All regressions include time fixed effects γ_t. The coefficients γ_1 and γ_2 measure how the dependent variable varies with mild and serious mental illness.

Consumption. Table B.1 reports our findings on the relation between mental health and consumption. The first column presents the results from estimating equation (16) when the dependent variable is log consumption. This column shows the results for nondurable consumption, which include expenditures on food, utilities, child care, clothing, home insurance, telecommunications, home maintenance, and variable transportation costs. Individuals who experience mild mental illness are estimated to consume 2.2 percent less relative to healthy individuals. Individuals experiencing serious mental illness consume 6.5 percent less relative to healthy individuals, or 1,500 dollars.

[52] We control for education by including dummies for whether the individual is a high-school dropout, a high-school graduate or a college graduate. We control for race by including dummy variables for white, Black, and others. We control for household composition by including dummy variables for the number of adults as well as the number of children in the household, each up to a maximum of five. We control for household wealth and for logarithmic household income in all regressions.

Table B.1: Consumption and Mental Health

Variable (in logs)	Non-durables	+ Education	+ Recreation	+ Durables
Mild γ_1	−2.2	−2.6	−3.5	−4.2
	(0.8)	(0.7)	(0.8)	(0.8)
Serious γ_2	−6.5	−7.4	−8.7	−9.5
	(1.2)	(1.2)	(1.3)	(1.3)
Observations	35,153	35,153	35,153	35,153
R^2	0.53	0.54	0.56	0.56
Mean (in levels)	23,000	24,100	26,800	30,700

Table B.1 displays the regression results using individual data from the PSID. The set of control variables include dummies for education, age, sex of the household head, time, race, household composition as well as household wealth and income.

The second to fourth column in Table B.1 show the robustness of the regression results. Specifically, we estimate equation (16) using broader measures of consumption. In the second column, we add education expenditures into the consumption measure. The results indicate that individuals with mild mental illness consume 2.6 percent less, while individuals with serious illness consume 7.4 percent less. The third column further adds vacation and recreation expenditures into the measure of consumption, similar to Krueger and Perri (2006). With this measure, individuals with mild mental illness consume 3.5 percent less while individuals with serious illness consume 8.7 percent less. In the final column, we add expenditures on durables by including payments on car loans, car down payments, car leases, and furniture. Individuals with a mild mental illness consume 4.2 percent less and individuals with a serious illness consume 9.5 percent less.

Portfolio Choice. Table B.2 reports how portfolio choices vary by mental health. We find that, relative to healthy individuals, individuals experiencing mild mental illness invest 3.6 percentage points less of their portfolio in risky assets, while individuals experiencing serious illness invest 5.6 percentage points less in risky assets. The second column presents the results of estimating equation (16) with the dependent variable being the extensive margin of risky investments. As discussed in Section 4, an individual is said to participate in risky investments if the share of their portfolio invested in risky instruments exceeds 0.5. This regression shows that individuals with serious and mental illness are less likely to invest in risky

Table B.2: Portfolio Allocation and Mental Health

Variable	Intensive Margin	Extensive Margin
Mild	−3.6	−4.6
	(0.6)	(0.7)
Serious	−5.6	−6.5
	(0.9)	(1.1)
Observations	36,334	36,334
R^2	0.31	0.31
Mean	0.56	0.64

Table B.2 reports regression coefficients estimated from equation (16). The set of control variables include dummies for education, age, sex of the household head, time, race, household composition, household income and wealth.

assets. Individuals with mild illness are 4.6 percent less likely to invest in risky investments, whereas individuals with serious illness are 6.5 percent less likely to invest in risky investments.

C Quantifying Negative Thinking

To quantify the relationship between negative thinking and mental illness, we use the RAND American Life Panel (ALP), a nationally representative survey of U.S. adults. Specifically, we merge two different ALP modules. The first module, implemented between March and April 2012, was designed by Dimmock, Kouwenberg, Mitchell, and Peijnenburg (2016, 2021). This module, which we call the Ellsberg module, elicits respondents' subjective loss probability. It does so by presenting them with a sequence of classic Ellsberg urn problems (Ellsberg, 1961).

The Ellsberg module elicits an individual's indifference point between a gamble on an urn with unknown winning probabilities $\mathcal{U}$, and a gamble on an urn with known winning probabilities $\mathcal{K}$. Each urn contains balls that are purple or yellow. For the urn with known winning probabilities, the individual knows the exact proportion q of yellow balls. The urn with unknown winning probabilities contains purple and yellow balls in unknown proportion. Individuals are asked to choose between the urn $\mathcal{K}$ and the urn $\mathcal{U}$. One ball is then drawn from the selected urn, and the individual wins a prize of 15 dollars if a purple ball is drawn.

The Ellsberg module presents individuals with a series of Ellsberg urn problems that differ by the known winning probability q. First, the respondent faces an Ellsberg problem with $q = \frac{1}{2}$. If the individual reports to prefer urn $\mathcal{K}$, then this urn is subsequently made less attractive by increasing the proportion of yellow balls q. If the respondent again prefers the known urn, it is made less attractive again. If the unknown urn is chosen, the known urn is made more attractive by lowering q. The process continues until a point of indifference is attained.[53] This point of indifference is exactly the individual's subjective loss probability.

The Ellsberg module does not contain information on mental health. We merge it with the second ALP module that asks respondents about their mental health. This module, which we refer to as the well-being module, consists of two ALP surveys conducted between May and July 2012 and between May and August 2012, in close proximity to the Ellsberg module. We merge the Ellsberg module with the well-being module, exploiting the structure of the ALP which allows identifying respondents across ALP surveys. By combining these modules, we quantify differences in subjective loss probabilities across mental health states.

The well-being module contains three questions about respondents' mental health. First, respondents are asked whether they experienced depression. Our first measure of mental illness is a dummy variable that takes the value of one if the answer to this question is yes, and zero otherwise. We refer to this indicator as Depression Indicator I. Second, respondents are asked whether they felt depressed, and can reply not at all, a little, somewhat, quite a bit, or very. Our second measure of mental illness is a dummy variable that takes the value of one if the respondents answered quite a bit or very depressed, and zero otherwise. We refer to this indicator as Depression Indicator II. Both depression related questions are asked only to subsamples of the well-being surveys. Third, respondents are asked to describe how anxious they feel on a scale from 0 to 10, where 0 corresponds to not anxious at all and 10 corresponds to completely anxious. Our third measure of mental illness is a dummy variable that takes the value of one if the answer to this question exceeds 5. We refer to this indicator as the Anxiety Indicator. The question on anxiety is fielded to all survey respondents.

To assess the relationship between mental illness and subjective loss probabilities, we first estimate

[53] The updating scheme follows a bisection algorithm. In the first round, the known urn has a proportion $q = 0.5$ of yellow balls. If the individual prefers the known urn $\mathcal{K}$ in the first round, the subjective probability p is above 0.5 and the proportion of yellow balls increases to $q = 0.75 = \frac{1}{2} \times (0.5 + 1)$. If the individual prefers the unknown urn $\mathcal{U}$ in the next round, the subjective probability p is below 0.75 and the proportion of yellow balls in the known urn decreases to $q = 0.625 = \frac{1}{2} \times (0.5 + 0.75)$. The difference between the upper and lower bound in the subjective probability is cut in two in each round. The maximum number of rounds without reaching the point of indifference is four. In this case, the average of the remaining upper and lower bound is the subjective probability.

Table C.1: Negative Thinking and Mental Illness Indicators

	Depression Indicator I	Depression Indicator II	Anxiety Indicator
κ	4.8	3.9	4.1
	(1.5)	(2.0)	(1.1)
Observations	1,636	1,651	2,974
R^2	0.09	0.06	0.06
Mean	47.3	47.3	47.4

Table C.1 displays regression coefficients on indicator variables of mental illness with respect to negative thinking in equation (C.1). The dependent variable across all specifications is the subjective loss probability. Columns correspond to different regression specifications that vary by the independent dummy variable D. Standard errors are reported in parenthesis in the second row. The control variables include education, age, sex, race, income, employment, and risk aversion.

the following regression. Let p_i be the subjective loss probability of individual i, elicited from the Ellsberg module. Let D_i be one of our three mental illness indicators. We consider the following regression:

$$p_i = \kappa D_i + \kappa_x X_i + \varepsilon_i, \tag{C.1}$$

where X_i are controls, such as age, sex, education, race, risk aversion, household income, employment status, and a constant. The regression coefficient κ captures how the subjective loss probability varies with mental health.

Table C.1 shows that individuals experiencing mental illness think more negatively. Across different measures of mental illness, represented by the different columns, we find that mental illness is associated with a higher subjective loss probability. Quantitatively, the subjective loss probability is 4 to 5 percentage points higher for individuals experiencing mental illness.

In order to evaluate how the subjective loss probability varies with the severity of mental illness we construct a new categorical variable indicating whether an individual is healthy, experiences mild mental illness, or experiences serious mental illness. We do so using the anxiety question that is fielded to all survey respondents.[54] We classify an individual as experiencing serious mental illness if the reported anxiousness exceeds an upper threshold a_s in both the well-being surveys. We choose the threshold a_s such that the proportion of individuals classified as experiencing serious mental illness aligns with the proportion of adults experiencing serious mental illness in the population, as reported by the NIMH.

[54]The anxiety question is a part of a block of questions that is fielded to all ALP respondents. The depression questions are fielded only to subsets of the respondents.

Table C.2: Negative Thinking and Mental Illness Severity (I)

Mild κ_1	3.4	3.1	3.4	3.4	3.6	3.5
	(1.3)	(1.3)	(1.3)	(1.3)	(1.3)	(1.3)
Serious κ_2	6.4	6.6	7.2	7.2	7.2	7.2
	(2.2)	(2.2)	(2.2)	(2.2)	(2.2)	(2.2)
Controls	None	+ Age	+ Income	+ Education	+ Race	+ Gender
R^2	0.02	0.02	0.03	0.03	0.04	0.04

Table C.2 displays the regression coefficients κ_1 (first row) and κ_2 (third row) estimated from equation (15) as well as their corresponding standard errors (in rows 2 and 4). The control variables include income, age, education, race, gender, employment, and risk aversion. From the first to the final column, we incorporate additional control variables. All numbers are statistically significant as implied by the standard errors, which are reported in parentheses below the regression coefficients. The mean loss probability is 0.474. The number of observations is equal to 2,973.

We classify an individual as experiencing mild mental illness if the reported anxiousness exceeds a lower threshold a_m in both the well-being surveys, and the individual is not classified as experiencing serious illness. We select the threshold a_m so that the proportion of individuals classified as experiencing mild mental illness is closest to their proportion in the population reported by the NIMH.[55] In order to evaluate how negative thinking varies with the severity of mental illness, we estimate (15).

Table C.2 and Table C.3 shows how negative thinking varies with mental health. Each column corresponds to a regression that differs in the controls that are included. From the first to the fifth column, we add control variables. For example, the first column of Table C.2 shows that without controls, we find that individuals experiencing mild mental illness have a subjective loss probability that is 3.4 percentage point higher relative to healthy individuals (first row), while individuals experiencing serious mental illness have a subjective loss probability that is a 6.4 percentage point higher (third row). The final column of Table C.3 shows that this finding is robust to the inclusion of all control variables. Individuals with mild (serious) mental illness have a subjective loss probability that is 3.1 (6.7) percentage point higher relative to healthy individuals.

[55] The Substance Abuse and Mental Health Services Administration's 2012 National Survey on Drug Use and Health reports that 13.9 percent of adults in the US experience mild illness, and 4.1 percent of adults experience serious mental illness. We classify individuals with an anxiety score greater than or equal to $a_s = 7$ as experiencing serious mental illness, and individuals with an anxiety score of 5 or 6 as experiencing mild mental illness, $a_m = 5$. With these cutoffs, 10.0 percent of ALP respondents experience mild mental illness and 3.1 percent of adults experience serious mental illness.

Table C.3: Negative Thinking and Mental Illness Severity (II)

Mild κ_1	3.5	3.4	3.5	3.5	3.1
	(1.3)	(1.3)	(1.3)	(1.3)	(1.3)
Serious κ_2	7.2	7.2	7.2	7.2	6.7
	(2.2)	(2.2)	(2.2)	(2.2)	(2.1)
Controls	+ Marital Status	+ Household Composition	+ Retirement Benefits	+ Financial Literacy	All
R^2	0.04	0.04	0.04	0.05	0.10

Table C.3 displays the regression coefficients κ_1 (first row) and κ_2 (third row) estimated from equation (15) as well as their corresponding standard errors (in rows 2 and 4). The control variables include income, age, education, race, gender, employment, and risk aversion. From the first to the final column, we incorporate additional control variables. All numbers are statistically significant as implied by the standard errors, which are reported in parentheses below the regression coefficients. The mean loss probability is 0.474. The number of observations is equal to 2,973.

Risk Aversion. We next establish that risk aversion does not vary systematically with mental health. In order to see how risk aversion varies with the severity of mental illness, we estimate the following regression:

$$\text{Risk Aversion}_i = \kappa_1 D_{1i} + \kappa_2 D_{2i} + \kappa_x X_i + \varepsilon_i, \tag{C.2}$$

where Risk Aversion$_i$ is the measure of risk aversion in the Ellsberg module for individual i, D_{1i} is a dummy variable taking the value one when individual i experiences mild illness, and D_{2i} is a dummy variable taking the value one when individual i experiences serious illness.[56]

Table C.4 and Table C.5 establish how risk aversion varies with mental health, where each column corresponds to a regression that differs in the controls that are included. The main result of the tables is that the differences in risk aversion between healthy individuals and those experiencing mental illness are not statistically significant. Similarly, the difference in risk aversion between individuals experiencing mild and serious mental is not statistically significant. The finding is robust across all columns. In sum, risk aversion does not vary systematically with mental illness.

[56]The measure of risk aversion in the Ellsberg module builds on Tanaka, Camerer, and Nguyen (2010). To measure risk aversion, the indifference point between a certain payoff, and a gamble with a known probability of losing q, is elicited. When the individual prefers the certain outcome, the probability of losing is decreased in the next round. When the gamble is preferred, the probability of losing is increased in the next round. The updating scheme follows a bisection algorithm, and stops when the respondent is indifferent.

Table C.4: Risk Aversion and Mental Illness Severity (I)

Mild κ_1	0.7	2.7	2.3	2.3	2.8	2.8
	(2.7)	(2.7)	(2.7)	(2.7)	(2.7)	(2.7)
Serious κ_2	−0.3	−0.9	−1.4	−1.6	−1.2	−1.2
	(4.6)	(4.6)	(4.6)	(4.6)	(4.6)	(4.6)
Controls	None	+ Age	+ Income	+ Education	+ Race	+ Gender
R^2	0.01	0.02	0.03	0.03	0.04	0.04

Table C.4 displays the regression coefficients κ_1 (first row) and κ_2 (third row) estimated from equation (C.2) and their corresponding standard errors (in rows 2 and 4). The control variables include income, age, education, race, gender, employment, and the subjective loss probability. From the first to the final column, we incorporate additional control variables. All numbers are statistically insignificant as implied by the standard errors, which are reported in parentheses below the regression coefficients.

Table C.5: Risk Aversion and Mental Illness Severity (II)

Mild κ_1	2.8	2.8	2.6	2.6	1.4
	(2.7)	(2.7)	(2.7)	(2.7)	(2.7)
Serious κ_2	−1.2	−1.2	−1.2	−1.2	−3.1
	(4.6)	(4.6)	(4.6)	(4.6)	(4.5)
Controls	+ Marital Status	+ Household Composition	+ Retirement Benefits	+ Financial Literacy	All
R^2	0.03	0.03	0.04	0.04	0.08

Table C.5 displays the regression coefficients κ_1 (first row) and κ_2 (third row) estimated from equation (C.2) and their corresponding standard errors (in rows 2 and 4). The control variables include income, age, education, race, gender, employment, and the subjective loss probability. From the first to the final column, we incorporate additional control variables. All numbers are statistically insignificant as implied by the standard errors, which are reported in parentheses below the regression coefficients.

D Mental Health Transition Matrix

We estimate the transition rates between mental health states as a function of the individual's treatment decision and idiosyncratic productivity state. We denote the transition probability from state m to m', conditional on the treatment decision τ and idiosyncratic productivity state ν, by $\Gamma_m(m' \mid m, \tau, \nu)$. In this appendix, we drop the subscript m on Γ_m to simplify notation.

We make several assumptions. First, we assume that treatment does not yield any benefits for healthy individuals. That is, $\Gamma(m' \mid m_0, 1, \nu) = \Gamma(m' \mid m_0, 0, \nu)$ for every m' and ν. This assumption is motivated by the fact that, in the data, healthy individuals rarely receive treatment (see, e.g., Cronin, Forsstrom, and Papageorge (2024)). Second, we assume that transitions from mild and serious mental illness do not depend on idiosyncratic productivity, that is $\Gamma(m' \mid m, \tau, \nu) = \Gamma(m' \mid m, \tau, \nu')$ for every $m = \{m_1, m_2\}$, τ and (ν, ν'). Third, we assume that transitions from the healthy state depend only on whether or not idiosyncratic productivity is below or above a threshold $\underline{\nu}$, which we set in the calibration to be the bottom quartile of the invariant productivity distribution based on the estimated productivity parameters ρ_ν and σ_ν^2 reported in Table 1. The last two assumptions allow us to capture, in a parsimonious way, the idea that negative income shocks deteriorate future mental health.

Data. We begin by describing the data moments used for the estimation. First, we compute biannual transition probabilities between mental health states from the PSID sample discussed in Section 4. Specifically, for every $m \in \{m_1, m_2\}$ and $m' \in \{m_0, m_1, m_2\}$, we compute the share of individuals who transition from state m to state m' two years later. Denote these empirical transition rates by $\Gamma^d(m' \mid m)$, where d labels data. These empirical transition probabilities are not conditional on treatment, since treatment is not observed in the PSID, and are unconditional on idiosyncratic productivity. We compute transitions from the healthy state separately for individuals who have normal idiosyncratic productivity (i.e., $\nu_i \geqslant \underline{\nu}$) and for individuals who have low idiosyncratic productivity (i.e., $\nu_i < \underline{\nu}$).[57] These transitions are denoted by $\Gamma^d(m' \mid m_0, \nu \geqslant \underline{\nu})$, and $\Gamma^d(m' \mid m_0, \nu < \underline{\nu})$. The empirical transition probabilities from the healthy state are independent of treatment.

Second, we compute the population shares by mental health state using the 2021 PSID wave. In this wave, 5.1 percent of individuals are classified as experiencing serious illness, and 13.5 percent are classified as mildly ill. The remaining 81.4 percent are classified as healthy. These empirical shares are denoted $\pi_d(m)$ for $m \in \{m_0, m_1, m_2\}$.

Third, we obtain treatment shares by mental health status from the 2021 National Survey on Drug

[57] An individual's productivity state ν is their residual wage from the PSID wage regression described in Section 4.

Use and Health of the Substance Abuse (see footnote 7). The report shows that 41.4 percent of all adults with mild mental illness receive treatment, while 65.4 percent of individuals experiencing serious mental illness receive treatment. We denote the empirical share of individuals by treatment status τ given mental health status m as $\pi_d^\tau(\tau \mid m)$. Since we assume that healthy adults do not receive treatment, we set $\pi_d^\tau(1 \mid m_0) = 0$. Fourth, we obtain the share of healthy individuals who have an idiosyncratic productivity above $\underline{\nu}$ in our PSID sample, which we denote by $\pi_d^{\underline{\nu}}$.

Finally, we use estimates on the efficacy of mental health treatment from the medical literature. A large body of work in psychology and psychiatry estimates the effects of treatment on mental health using randomized trials. The effect sizes are typically standardized to facilitate comparison across different studies. Specifically, they are reported in terms of the standardized mean difference (SMD), defined as the mean effect divided by the combined standard deviation of the outcome, that is, $SMD = \frac{\mu_T - \mu_C}{\sqrt{\frac{1}{2}(\sigma_T^2 + \sigma_C^2)}}$, where μ_T is the average outcome in the treatment group, μ_C is the average outcome in the control group, σ_T^2 is the variance of the outcome in the treatment group, and σ_C^2 is the variance of the outcome in the control group. As discussed in the main text, we use an intermediate value of -0.70.

Estimation. To estimate the transition probabilities $\Gamma(m' \mid m, \tau, \nu)$, we solve a system of 18 unknowns and 18 equations. The 18 unknowns are 6 transition probabilities from the healthy state that depend on idiosyncratic productivity but do not depend on the treatment decision ($\Gamma(m' \mid m_0, \tau, \nu \geqslant \underline{\nu})$ and $\Gamma(m' \mid m_0, \tau, \nu < \underline{\nu})$ for every $m' \in \{m_0, m_1, m_2\}$, where the dependence on τ is redundant), and 12 transition probabilities from the mild and serious illness states that depend on the treatment decision but do not depend on idiosyncratic productivity ($\Gamma(m' \mid m, \tau = 0, \nu)$ and $\Gamma(m' \mid m, \tau = 1, \nu)$ for every $m \in \{m_1, m_2\}$ and $m' \in \{m_0, m_1, m_2\}$, where the dependence on ν is redundant).

We next describe the 18 equations we use in our estimation. First, transition probabilities from every state sum to one. Since transitions from the healthy state depend on whether idiosyncratic productivity is below or above $\underline{\nu}$, and since transitions from mild and serious illness depend on the treatment decision, this gives six transition equations.

Second, for each mental health state $m = \{m_1, m_2\}$ and $m' = \{m_1, m_2\}$, we equate the empirical transition rate between m and m' (which is unconditional on treatment) to the model-derived (unconditional) transition rate as implied by the empirical treatment shares. This gives additional four equations that ensure consistency between the unconditional transition probabilities in the data and the model:

$$\Gamma^d(m' \mid m) = \pi_d^\tau(1 \mid m)\Gamma(m' \mid m, 1, \nu) + \pi_d^\tau(0 \mid m)\Gamma(m' \mid m, 0, \nu). \tag{D.1}$$

Third, for each mental health state $m' = \{m_1, m_2\}$, we equate the empirical transition rate between

m_0 and m', conditional on whether idiosyncratic productivity is above or below $\underline{\nu}$, to the equivalent model transition rate. This gives four additional equations:

$$\Gamma^d(m' \mid m = m_0, \nu \geqslant \underline{\nu}) = \Gamma(m' \mid m = m_0, \tau, \nu \geqslant \underline{\nu}),$$
$$\Gamma^d(m' \mid m = m_0, \nu < \underline{\nu}) = \Gamma(m' \mid m = m_0, \tau, \nu < \underline{\nu}). \tag{D.2}$$

Fourth, we assume that the observed shares of individuals across mental health states correspond to steady state shares. That is, that the model-derived distribution across mental states, as implied by the observed distribution across idiosyncratic states and treatment decisions, is equal to the observed shares of individuals across mental health states. This gives rise to two additional equations for $m = \{m_1, m_2\}$:

$$\begin{aligned}\pi_d(m) = {} & \pi_d(m_0)\pi_d^{\nu}\Gamma(m' \mid m_0, \tau, \nu \geqslant \underline{\nu}) + \pi_d(m_0)(1 - \pi_d^{\nu})\Gamma(m' \mid m_0, \tau, \nu < \underline{\nu}) \\ & + \pi_d(m_1)\pi_d^{\tau}(1 \mid m_1)\Gamma(m \mid m_1, 1, \nu) + \pi_d(m_1)\pi_d^{\tau}(0 \mid m_1)\Gamma(m \mid m_1, 0, \nu) \\ & + \pi_d(m_2)\pi_d^{\tau}(1 \mid m_2)\Gamma(m \mid m_2, 1, \nu) + \pi_d(m_2)\pi_d^{\tau}(0 \mid m_2)\Gamma(m \mid m_2, 0, \nu).\end{aligned} \tag{D.3}$$

Finally, we compute the model-implied SMD and equate it to its empirical counterpart from the medical literature. To be consistent with the outcomes measured in the medical literature, we measure the model-implied SMD in terms of a depression severity rating. In particular, we use the K6 scale discussed in Appendix C. Individuals with mental health state m are assigned $\overline{\mathrm{K6}}(m)$, the median K6 scale for individuals in state m in the PSID. For each state $m = \{m_1, m_2\}$, we align the model-implied SMD to its empirical counterpart, which we denote by SMD_d, giving the remaining two equations:

$$\mathrm{SMD}_d(m) = \frac{\mathbb{E}\left[\overline{\mathrm{K6}}(m') \mid m, \tau = 1\right] - \mathbb{E}\left[\overline{\mathrm{K6}}(m') \mid m, \tau = 0\right]}{\sqrt{\frac{1}{2}\left(V\left[\overline{\mathrm{K6}}(m') \mid m, \tau = 1\right] + V\left[\overline{\mathrm{K6}}(m') \mid m, \tau = 0\right]\right)}}, \tag{D.4}$$

with the conditional mean and the conditional variance respectively given by:

$$\mathbb{E}[\overline{K6}(m') \mid m, \tau] = \mathbb{E}[\overline{K6}(m') \mid m, \tau, \nu] = \sum_{m'} \Gamma(m' \mid m, \tau, \nu)\overline{\mathrm{K6}}(m') \tag{D.5}$$

$$V\left[\overline{K6}(m') \mid m, \tau\right] = V\left[\overline{K6}(m') \mid m, \tau, \nu\right] = \sum_{m'} \Gamma(m' \mid m, \tau, \nu)\overline{\mathrm{K6}}(m')^2 - \left(\mathbb{E}\left[\overline{\mathrm{K6}}(m') \mid m, \tau\right]\right)^2 \tag{D.6}$$

The resulting estimated mental health transition matrix $\Gamma_m(\tau, \nu)$ is reported in the main text in Table 3.

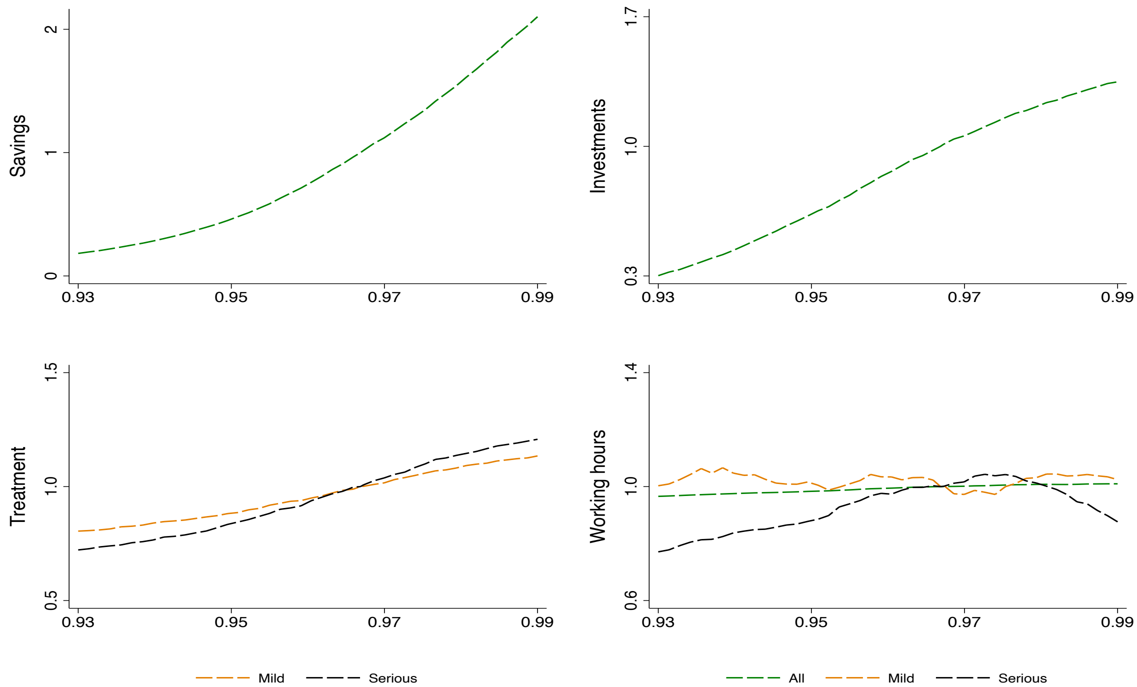

Figure E.1: Sensitivity of Moments to Discount Factor β

Figure E.1 illustrates the sensitivity of model moments with respect to changes in the discount factor β between 0.93 and 0.99. The baseline parameter value for the discount factor is equal to 0.967.

E Sensitivity to Model Parameters

In Section 4.3, we estimate endogenous model parameters so that the model matches data moments related to labor supply, savings and portfolio choice, and to mental health treatment. In this appendix, we analyze how sensitive the model moments are to changes in the endogenous parameters. This illustrates which moments structurally identify which parameter. Since we calibrate seven parameters to seven data moments, we show how each of the data moments vary with a change of each parameter.

All sensitivity figures adopt an identical structure. The top left panel shows how average wealth in the model varies with changes in the parameter value; the top right panel shows the sensitivity of the average risky investment share; the bottom left panel shows the sensitivity of treatment share by mental health state; while the bottom right panel shows the sensitivity of labor supply by mental health. Moments are normalized to one for the baseline parameter value. On the horizontal axis, we display the different values for the parameter of interest, where the interval we consider is centered around the baseline parameter value.

Discount Factor. We first analyze the sensitivity of model moments to variation in the discount factor.

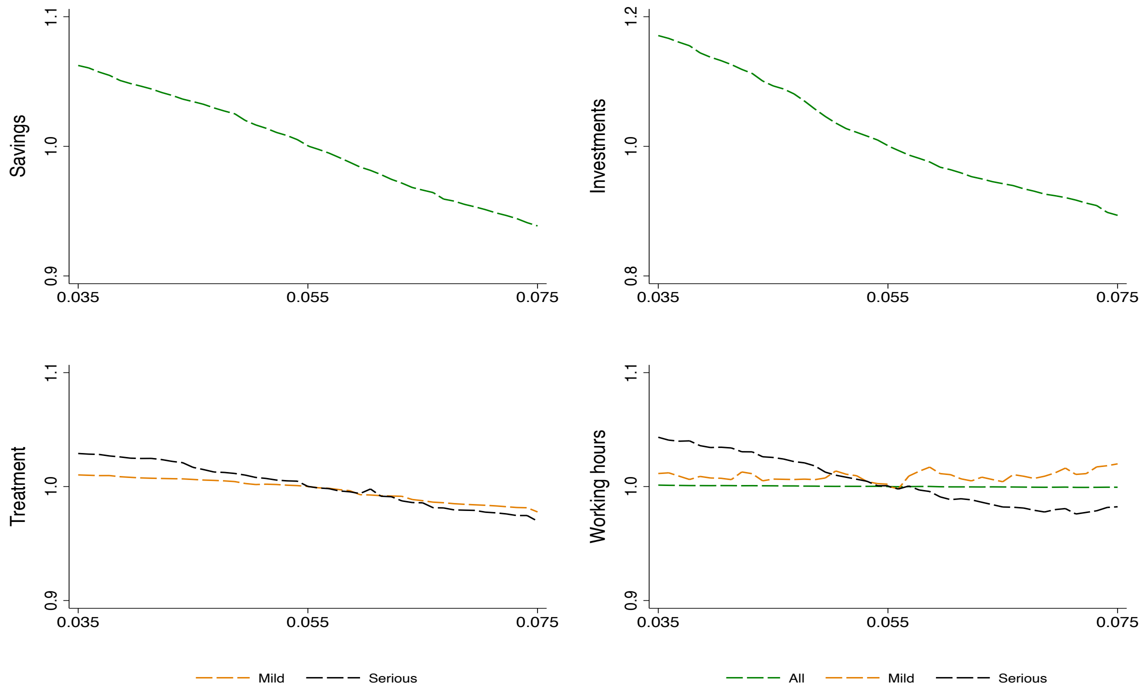

Figure E.2: Sensitivity of Moments to Participation Cost φ_k

Figure E.2 displays the sensitivity of model moments to changes in the participation costs for risky investments φ_k between about 2,200 dollars (corresponding to 0.035) and 4,750 dollars (corresponding to 0.075).

The result is shown in Figure E.1. The discount factor has a pronounced impact on savings, investments, and the treatment share. As individuals become increasingly patient, they save more. Holding fixed the costs of investing in risky assets, the share of savings invested in risky assets increases as shown in the top right panel. As individuals become increasingly patient, the cost of negative thinking about the future rise. The benefit from receiving treatment thus increases, which leads to an uptake in treatment shown in the bottom left panel. The bottom right panel shows that the response in labor supply is small relative to the other moments.

Figure E.2 shows the sensitivity of the model to changes in the participation costs between about 2,200 and 3,750 dollars, corresponding to the values 0.035 and 0.075 on the horizontal axis. The figure shows that the participation costs for risky assets governs the extent to which individuals invest in risky assets, while having negligible impact on the other moments. Reducing the participation costs to 2,200 dollars per period increases the share of savings in risky assets by 10 percentage points.

Figure E.3 shows the sensitivity of the model to changes in the disutility of work, which is governed by the parameter φ. We vary the disutility cost of work between 0.20 and 0.35 as displayed on the horizontal axis. Increasing the disutility from working decreases hours worked across all mental health groups, with labor supply of individuals with serious illness being most strongly affected. Since a decrease in the utility

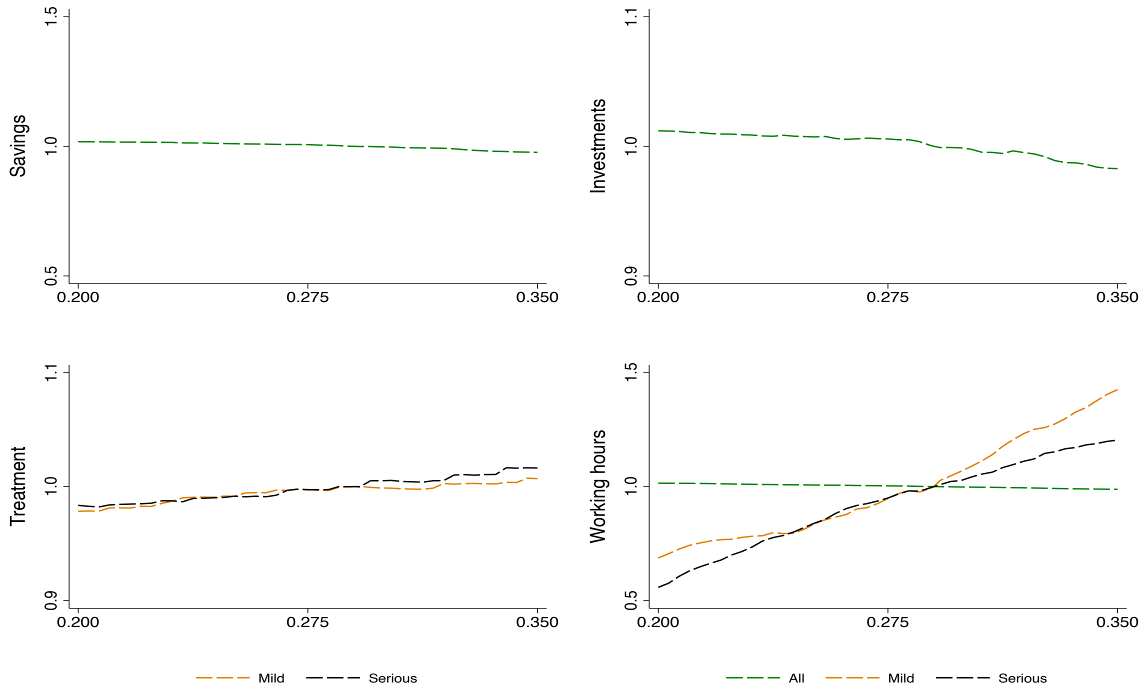

Figure E.3: Sensitivity of Moments to the Disutility of Work φ

Figure E.3 shows the sensitivity of the model to changes in the disutility cost of work, which is governed by the parameter φ. We vary the disutility cost between 0.20 and 0.35 as displayed on the horizontal axis.

cost from work increases labor supply, the marginal cost of undergoing treatment increases. As a result, the decrease in the utility cost from working decreases the propensity of seeking treatment for individuals with mental illness as is illustrated in the bottom left panel.

Figure E.4 reports the sensitivity of the model to changes in the utility cost of treatment ξ_τ. We vary the disutility cost ξ_τ between 0.01 and 0.05 as displayed on the horizontal axis, around the model parameter value of 0.032. The utility costs of treatment has no implications for aggregate savings, portfolio choice, and labor supply, while impacting the rate at which individuals undergo treatment. Lowering the utility costs of treatment increases the propensity of undergoing treatment for individuals with mental illness as is illustrated in the bottom left panel. The sensitivity of the propensity of undergoing treatment increases with the severity of illness.

We next analyze the sensitivity of the model moments to changes in the rumination parameters. First, we analyze the sensitivity to rumination among individuals with mild mental illness. Figure E.5 shows that rumination when mildly ill, $n_r(m_1)$, predominantly affects the labor supply of individuals with mild mental illness, illustrated by the increasing orange dashed line in the bottom right panel in Figure E.5.

Second, we evaluate the sensitivity of model moments to rumination among individuals with serious mental illness. Figure E.6 shows that rumination for those experiencing serious illness, $n_r(m_2)$, mostly

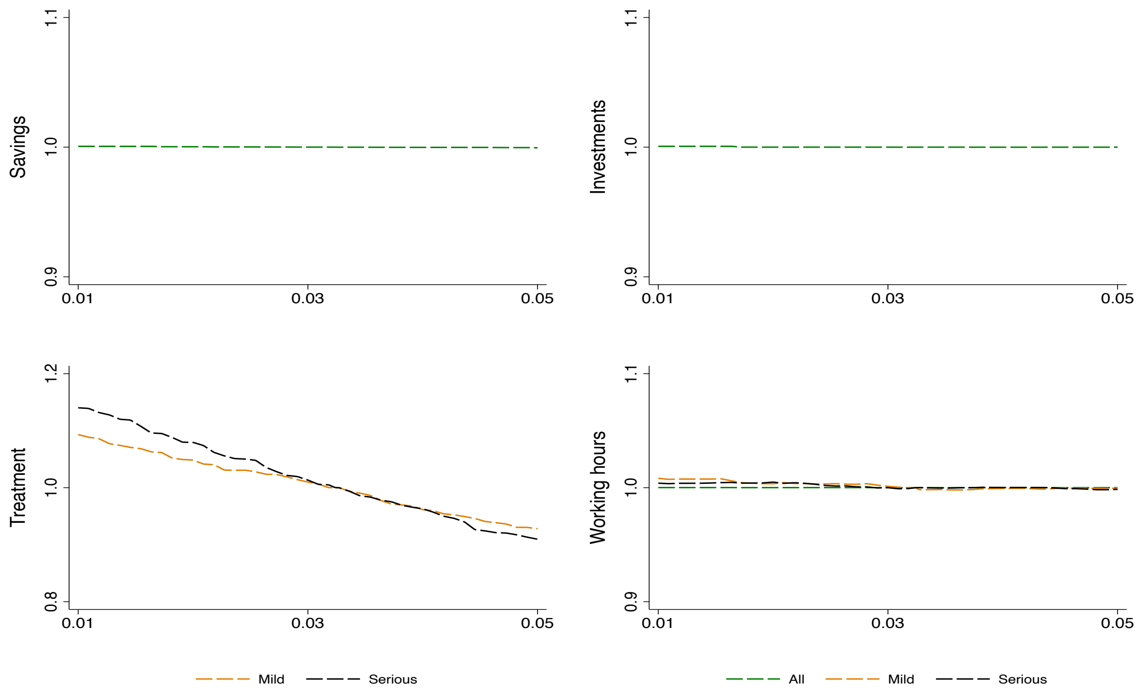

Figure E.4: Sensitivity of Moments to the Utility Cost of Treatment ξ_τ

Figure E.4 reports the sensitivity of the model to changes in the utility cost of treatment ξ_τ. We vary the disutility cost ξ_τ between 0.01 and 0.05 as displayed on the horizontal axis, around the model parameter value of 0.032. The utility costs of treatment has no implications for aggregate savings, portfolio choice, and labor supply, while impacting the rate at which individuals undergo treatment.

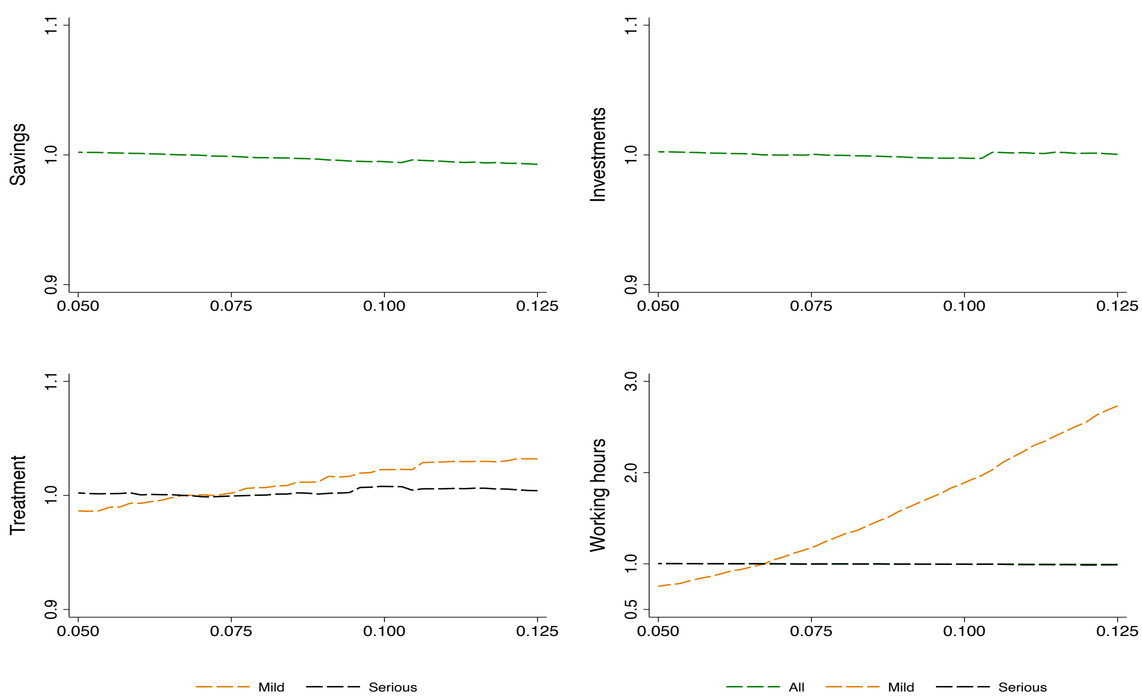

Figure E.5: Sensitivity of Moments to Rumination of Mild $n_r(m_1)$

Figure E.5 analyzes the sensitivity of the model moments to changes in rumination when individuals experience mild mental illness.

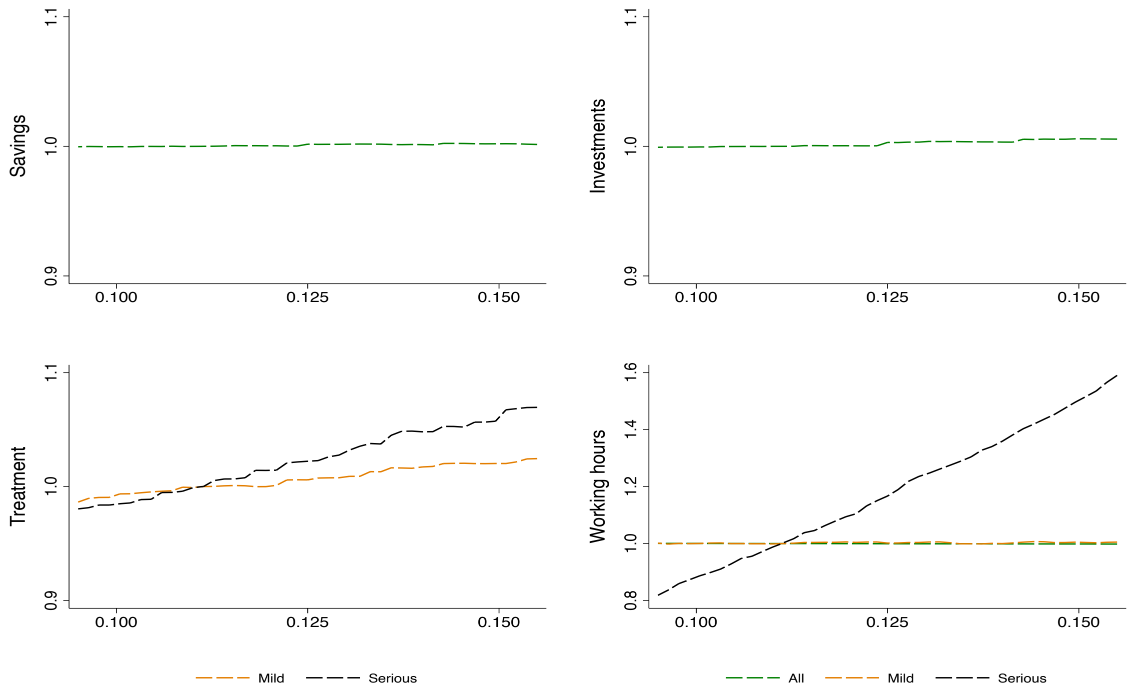

Figure E.6: Sensitivity of Moments to Rumination of Serious $n_r(m_2)$

Figure E.6 analyzes the sensitivity of the model moments to changes in rumination when individuals experience serious mental illness.

affects labor supply of individuals with serious mental illness, illustrated by the increasing black dashed line in the bottom right panel in Figure E.6. Since serious mental illness becomes more costly as rumination increases, the propensity to get treatment also increases despite the reduction in available time, as shown in the bottom left panel.

Finally, we evaluate the sensitivity of model moments with respect to the availability of treatment. Figure E.7 shows that expanding the availability for treatment strongly affects the treatment share among individuals experiencing mild mental illness. As treatment availability expands, individuals gain access to mental health treatment services when experiencing mild mental illness, increasing the treatment rate in this mental health state. This is seen in the bottom left panel of Figure E.7. Since increased treatment consumes time, increased treatment also reduces working hours of individuals experiencing mild mental illness as seen in the bottom right panel.

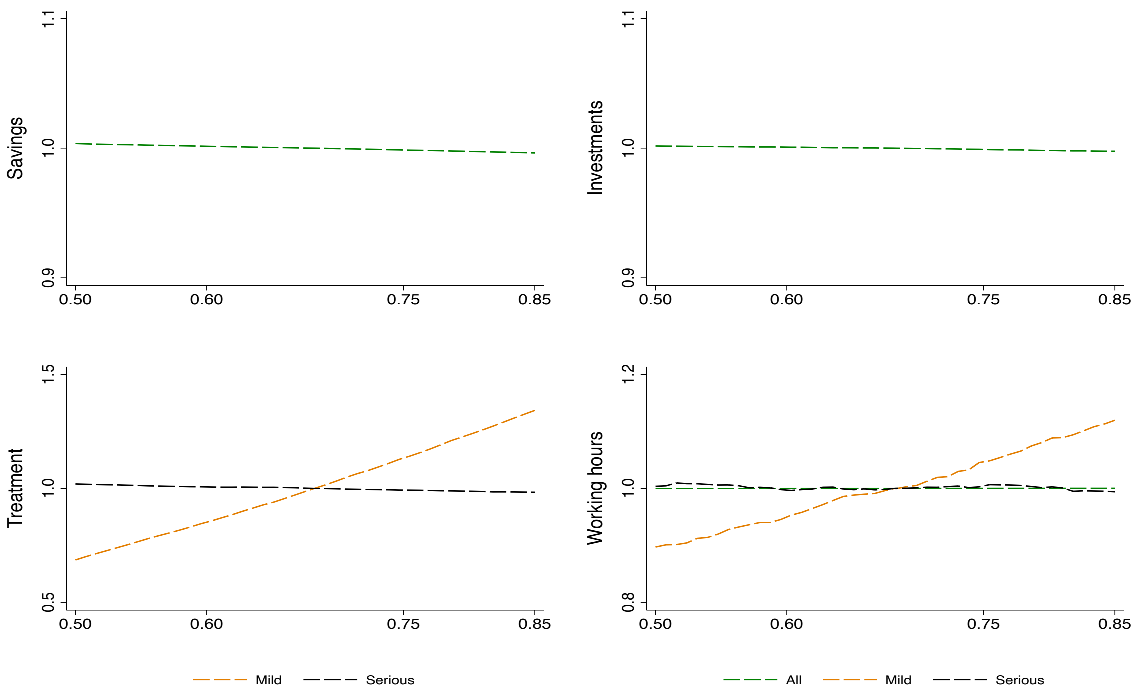

Figure E.7: Sensitivity of Moments to the Availability of Treatment ω_τ

Figure E.7 shows the sensitivity of the model moments to the availability of treatment ω_τ. We vary the availability of treatment when individuals experience mild illness from 0.50 to 0.85, around the calibrated parameter value $\omega_\tau = 0.682$.

F Additional Figures and Tables

This appendix provides additional figures and tables.

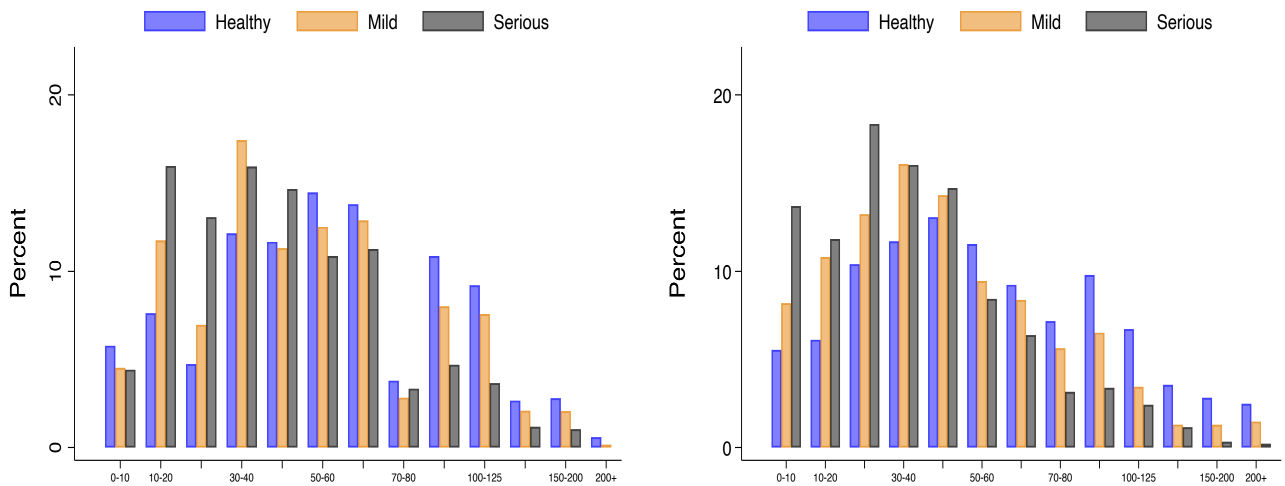

Figure F.1: Income by Mental Health in the Model and the Data

Figure F.1 shows the distribution of labor income by mental health status in the model (left panel) and in the data (right panel). The height of the bars capture the fraction of individuals earning a particular income within each mental health status.

Table F.1: Validation: Investment Share by Mental Health Status

	Data			Model		
Percentile	Healthy	Mild	Serious	Healthy	Mild	Serious
5	0	0	0	0	0	0
10	0	0	0	0	0	0
25	0	0	0	0	0	0
50	0.76	0.23	0	0.90	0.70	0
75	0.90	0.87	0.81	0.98	0.92	0.84
90	0.95	0.95	0.95	1.00	0.98	0.94
95	0.97	0.97	0.98	1.00	1.00	0.96
99	1.00	1.00	1.00	1.00	1.00	1.00

Table F.1 summarizes the distribution of the risky investment share by mental health status in the model and in the data.

Table F.2: Eliminating Stigma of Treatment ξ_τ

	Benchmark			Treatment stigma $\xi_\tau = 0$		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Mental health shares	0.866	0.097	0.037	0.871	0.095	0.034
Treatment shares	0.000	0.414	0.657	0.000	0.472	0.789
Hours worked	0.404	0.383	0.351	0.404	0.383	0.351
Income (in thousands)	64	56	46	64	56	46
Wealth (in thousands)	292	262	236	293	262	237
Risky investment share	0.572	0.461	0.385	0.573	0.461	0.386
Risky participation rate	0.601	0.516	0.449	0.602	0.516	0.450

Table F.2 displays the effects of eliminating stigma costs of mental health treatment. The first three columns display the averages by mental health group in the benchmark economy where the utility cost of treatment is equal to $\xi_\tau = 0.032$. The final three columns report the moments of the counterfactual economy where the utility costs are equal to zero.

Table F.3: Eliminating Productivity Loss

	Benchmark			Productivity loss $\Lambda(m) = 0$		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Mental health shares	0.866	0.097	0.037	0.866	0.097	0.037
Treatment shares	0.000	0.414	0.657	0.000	0.414	0.656
Hours worked	0.404	0.383	0.351	0.404	0.384	0.352
Income (in thousands)	64	56	46	64	57	48
Wealth (in thousands)	292	262	236	293	263	238
Risky investment share	0.572	0.461	0.385	0.572	0.463	0.391
Risky participation rate	0.601	0.516	0.449	0.601	0.517	0.455

Table F.3 displays the effects of eliminating the productivity losses associated with mental illness. The first three columns display the averages by mental health group in the benchmark economy where the productivity losses equal $\Lambda(m_1) = -0.013$ and $\Lambda(m_2) = -0.032$. The final three columns report the moments of the counterfactual economy where the productivity loss is zero.

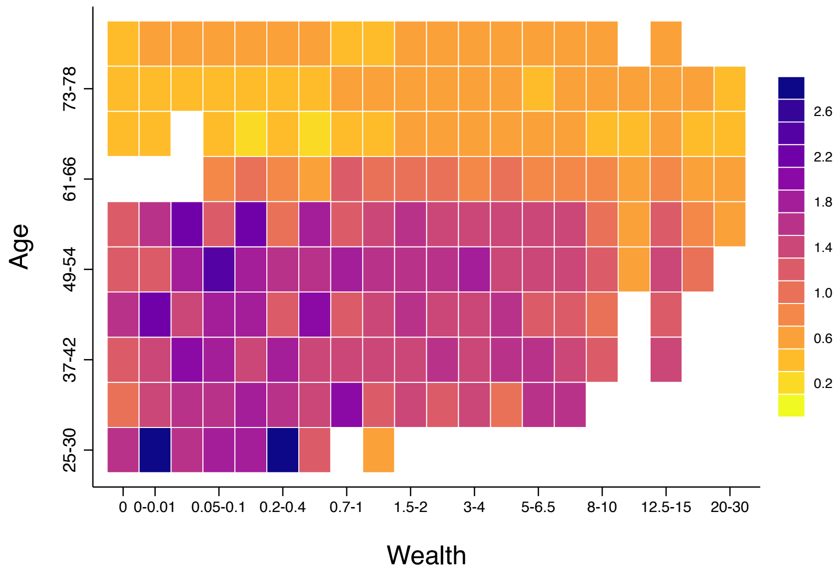

Figure F.2: Welfare Cost of Mental Illness by Age and Wealth

Figure F.2 displays the average Δ_i^m by age bracket (vertical axis) and by wealth bracket (horizontal axis). The different colors indicate different levels of Δ_i^m: darker shades indicate a higher average Δ_i^m, and lighter shades indicate a lower average Δ_i^m. Figure F.2 shows that the welfare costs are higher for younger individuals than for older individuals. Younger individuals (below age 55) experience an average welfare cost of 1.6 percent, while older individuals (above age 55) experience a welfare cost of 0.8 percent.

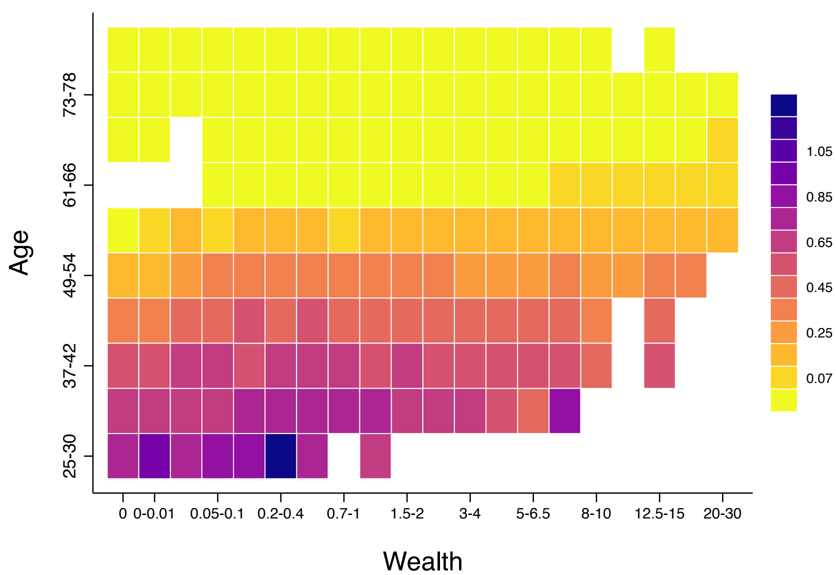

Figure F.3: Welfare Gains from Increased Availability of Treatment Services by Age and Wealth

Figure F.3 displays the average consumption equivalent welfare gain from increased availability of treatment by age (vertical axis) and by wealth (horizontal axis). Different colors indicate different levels of welfare gains: dark shades indicate larger welfare gains of increased availability, and light shades indicate lower welfare benefits.

Table F.4: Eliminating Out-of-Pocket Treatment Costs

	Benchmark			Treatment costs $\varphi_\tau = 0$		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Mental health shares	0.866	0.097	0.037	0.870	0.095	0.035
Treatment shares	0.000	0.414	0.657	0.000	0.458	0.753
Hours worked	0.404	0.383	0.351	0.404	0.383	0.351
Income (in thousands)	64	56	46	64	56	46
Wealth (in thousands)	292	262	236	292	261	237
Risky investment share	0.572	0.461	0.385	0.572	0.464	0.387
Risky participation rate	0.601	0.516	0.449	0.601	0.519	0.453

Table F.4 displays the effects of eliminating out-of-pocket costs for mental health services. The first three columns display the averages by mental health group in the benchmark economy where the cost of treatment is equal to 1,250 dollars. The final three columns report the moments of the counterfactual economy where the out-of-pocket costs are equal to zero.

G Mental Health Policies Examples

We provide examples of mental health policies. Mental health policies can be roughly classified into one of three main categories: policies targeted towards expanding availability of mental health services, policies reducing out-of-pocket treatment costs, and policies targeted at improving mental health of young adults.

Expanding Availability of Mental Health Services.

- Lack of availability of mental health services

  A shortage of mental health services is a challenge faced not only in the United States. Countries across the world are considering policies to close the accessibility gap. For example, in the United Kingdom, the National Audit Office describes the shortage of mental health staff as the main constraint to improving mental treatment services and to reducing treatment gaps (see www.nao.org.uk). In Canada, access to services is a major constraint according to the Centre for Addiction and Mental Health (see www.camh.ca).

- Increasing the supply of mental health professionals

  One popular solution to the shortage of mental health services is to increase the number of mental health professionals. In the U.S., in 2023, a total of 700 million dollars was invested into programs that provide training, access to scholarships and loan repayment to mental health clinicians. Further investments are made in addressing burnout and strengthening resiliency among health care workers and in programs that aim to train community health workers (see www.whitehouse.gov/s1). In the United Kingdom, the National Health Service Long Term Workforce Plan similarly sets out to increase training places for mental health nursing, as well as to increase the number of clinical psychologists and adolescent psychotherapists (see www.england.nhs.uk).

- Expanding access to treatment through community health clinics

  A second solution that is proposed for the shortage of mental health services is to expand the capacity of community health centers. The World Health Organization (see www.who.int) recommends decentralizing mental health services to the community settings. In the United States, Certified Community Behavioral Health Clinics (CCBHCs) are designed to ensure access to comprehensive behavioral health care. These health clinics are funded by the state and federal government and are required to serve anyone who requests care for mental health or substance use, regardless of ability to pay, residence, or age. In Belgium, a 2022 preventative care reform also aims to improve access to mental health services at the community level (see www.brusselstimes.com).

- Expanding access to treatment through virtual mental health care

  Finally, a more recent proposed solution to the shortage problem is to expand the capacity of virtual mental health services. The U.S. government stated that it will ensure coverage of virtual mental health care across health plans (see www.whitehouse.gov/s1). In Scotland, the National Health Service provides free access to therapeutics apps to help individuals experiencing anxiety (see www.nhslothian.scot). German doctors can prescribe mental health apps to individuals through the Digital Healthcare Act of 2019 with costs reimbursed through public health insurance (see www.bfarm.de).

Reducing Out-Of-Pocket Treatment Costs. Across the world, governments use policy to reduce out-of-pocket expenses for mental health care. In the United States, the Biden administration proposed to expand mental health parity laws (see www.whitehouse.gov/s3). In France, the government launched an initiative that covers therapy costs (see www.weforum.org). In Germany, a patient can request reimbursement for outpatient psychotherapy "if the treatment cannot be carried out in a timely manner or at an acceptable distance for the patient" (see www.pksh.de).

Improving Mental Health of Young Adults. Recent years have seen rising concerns over the mental health of young adults. In the U.S., the Biden administration proposed investing one billion dollars to double the number of school-based mental health professionals such as counselors, social workers, and school psychologists (see www.whitehouse.gov/s1). The UK government announced it would allocate funds to community hubs to deliver mental support for children and young adults (see www.gov.uk). In Japan, education about mental illness has been included in the high school curriculum (Ojio et al., 2021).

H Costs of Mental Health Policies

In this appendix, we calculate the costs associated with expanding availability of mental health treatment services and with improving mental health of young adults discussed in Section 5.2.

H.1 Expanding Availability of Mental Health Services

We estimate an upper bound of 3.8 billion dollars per year for the cost of expanding treatment availability. This suggests that the assessed policy benefit of 41 billion dollars per year significantly outweighs its costs.

The cost of expanding the availability of mental health services from $\omega_\tau = 0.682$ to 1 is composed of two components. First, we consider the cost of training professionals that are required to expand treatment availability. According to the Department of Health and Human Services (see www.kff.org), this amounts to training about 6,250 professionals. Second, since additional treatment services are demanded relative to the baseline, we have to account for the additional cost of treatment that is not paid out-of-pocket.

We consider the median four-year cost of private medical schools of about 360 thousand dollars as an upper bound for the training costs of mental health professionals.[58] An upper bound for the cost of training 6,250 mental health professionals is 2.25 billion dollars overall. Assuming these costs are discounted over 30 years, this implies a negligible annual cost of 75 million dollars.

For the benchmark economy, Table 5 shows that $0.097 \times 41.4 + 0.037 \times 65.7 = 6.45$ percent of the adult population get treatment. In the counterfactual, $0.085 \times 69.9 + 0.033 \times 63.2 = 8.03$ percent get treated. Given an adult population of 230 million adults, this implies that 3.7 million more people get treated every year. Since 81 percent of the treatment cost is not paid out of pocket amounts, the additional costs outside the model is $1250 \times 0.81 = 1012.5$ per person, or about 3.75 billion dollars per year.

H.2 Improving Mental Health of Young Adults

The cost of treating adolescents and young adults consists of the cost of treatment between ages 16 and 25. Starting with the invariant distribution, the fraction of individuals getting treatment at age 16 equals 18.6 percent, at age 18 is 12.4 percent, at age 20 is 10.8 percent, at age 22 is 10.2 percent, and at age 24 is 10.0 percent. Since the treatment annual cost is 1,250 dollars, this corresponds to an average cost of $2 \times 1,250 \times (0.186 + 0.124 + 0.108 + 0.102 + 0.100) = 1,550$ per person. Given an adult population of 230 million adults, assuming costs are discounted over 30 years, this implies an annual cost of 11.9 billion.

[58] This cost includes tuition, textbooks and supplies, and living expenses (see www.princetonreview.com).

I Alternative Model Specifications

In Section 5.2.4, we evaluate the sensitivity of our policy results to alternative model specifications. In particular, rows 2 to 7 in Table 11 correspond to alternative model specifications where the alternative model is recalibrated to target the moments in Table 5. In this appendix, we report the endogenously estimated parameters for each alternative model. We also compare the performance of the estimated model in terms of non-targeted moments to the performance of the baseline model. Table I.1 and Table I.2 correspond to the model where healthy individuals are also ambiguity averse, Table I.3 and Table I.4 correspond to the model with a utility penalty of mental illness, Table I.5 and Table I.6 (Table I.7 and Table I.8) correspond to the model where the borrowing constraint is 20,000 dollars (50,000 dollars), and Table I.9 and Table I.10 correspond to the model with a constant elasticity of labor productivity with respect to hours worked.

Table I.1: Endogenous Parameters for Model with Ambiguity $\kappa(m_0) = 0.025$

Parameter	Value	Moment (mean of)	Data	Model
Discount factor β	0.962	Wealth in dollars	288,000	290,000
Risky investments costs φ_k	2,700	Risky investment share	0.557	0.553
Disutility from work ψ	0.290	Hours worked	0.399	0.400
Rumination, mild $n_r(m_1)$	0.061	Δ Hours worked, mild	-0.047	-0.047
Rumination, serious $n_r(m_2)$	0.105	Δ Hours worked, serious	-0.127	-0.127
Utility cost of treatment ξ_τ	0.020	Treatment share, serious	0.656	0.657
Treatment availability ω_τ	0.690	Treatment share, mild	0.414	0.415

Table I.1 presents the parameter values set to match model-generated moments to their data analog. The first three columns present the parameters and their values. The fourth column describes a moment that informs the parameter value. The fifth and sixth column present the model-generated moment and the data-equivalent.

Table I.2: Validation of Averages for Model with Ambiguity $\kappa(m_0) = 0.025$

	Benchmark Model			Model		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Consumption	51	47	43	53	49	45
Hours	0.404	0.383	0.351	0.404	0.384	0.352
Income	64	56	46	65	57	47
Wealth	292	262	236	295	264	239
Risky investment share	0.572	0.461	0.384	0.569	0.462	0.390
Risky participation rate	0.601	0.516	0.449	0.633	0.533	0.470

Table I.2 shows average consumption, hours, income, wealth, and risky investment by mental health status. Consumption, income, and wealth holdings are in thousands of dollars. The risky investment share is the average share of total assets invested in risky assets. The risky participation rate measures the share of the population that holds more than half of their portfolio in risky assets.

Table I.3: Endogenous Parameters for Model with Utility Penalty $\xi_m(m_1) = \xi_m(m_2) = 0.05$

Parameter	Value	Moment (mean of)	Data	Model
Discount factor β	0.967	Wealth in dollars	288,000	290,000
Risky investments costs φ_k	3,500	Risky investment share	0.557	0.556
Disutility from work ψ	0.290	Hours worked	0.399	0.400
Rumination, mild $n_r(m_1)$	0.067	Δ Hours worked, mild	-0.047	-0.047
Rumination, serious $n_r(m_2)$	0.111	Δ Hours worked, serious	-0.127	-0.127
Utility cost of treatment ξ_τ	0.068	Treatment share, serious	0.656	0.657
Treatment availability ω_τ	0.673	Treatment share, mild	0.414	0.415

Table I.3 presents the parameter values set to match model-generated moments to their data analog. The first three columns present the parameters and their values. The fourth column describes a moment that informs the parameter value. The fifth and sixth column present the model-generated moment and the data-equivalent.

Table I.4: Validation of Averages for Model with Utility Penalty $\xi_m(m_1) = \xi_m(m_2) = 0.05$

	Benchmark Model			Model		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Consumption	51	47	43	52	47	43
Hours	0.404	0.383	0.351	0.404	0.383	0.352
Income	64	56	46	65	56	47
Wealth	292	262	236	295	264	237
Risky investment share	0.572	0.461	0.384	0.572	0.462	0.386
Risky participation rate	0.601	0.516	0.449	0.601	0.517	0.450

Table I.4 shows average consumption, hours, income, wealth, and risky investment by mental health status. Consumption, income, and wealth holdings are in thousands of dollars. The risky investment share is the average share of total assets invested in risky assets. The risky participation rate measures the share of the population that holds more than half of their portfolio in risky assets.

Table I.5: Endogenous Parameters for Model with Borrowing up to 20,000 dollars

Parameter	Value	Moment (mean of)	Data	Model
Discount factor β	0.932	Wealth in dollars	288,000	290,000
Risky investments costs φ_k	4,200	Risky investment share	0.557	0.555
Disutility from work ψ	0.375	Hours worked	0.399	0.399
Rumination, mild $n_r(m_1)$	0.067	Δ Hours worked, mild	-0.047	-0.048
Rumination, serious $n_r(m_2)$	0.111	Δ Hours worked, serious	-0.127	-0.127
Utility cost of treatment ξ_τ	0.036	Treatment share, serious	0.656	0.654
Treatment availability ω_τ	0.690	Treatment share, mild	0.414	0.414

Table I.5 presents the parameter values set to match model-generated moments to their data analog. The first three columns present the parameters and their values. The fourth column describes a moment that informs the parameter value. The fifth and sixth column present the model-generated moment and the data-equivalent.

Table I.6: Validation of Averages for Model with Borrowing up to 20,000 dollars

	Benchmark Model			Model		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Consumption	51	47	43	52	47	44
Hours	0.404	0.383	0.351	0.403	0.382	0.351
Income	64	56	46	65	56	46
Wealth	292	262	236	295	265	241
Risky investment share	0.572	0.461	0.384	0.578	0.434	0.339
Risky participation rate	0.601	0.516	0.449	0.601	0.487	0.400

Table I.6 shows average consumption, hours, income, wealth, and risky investment by mental health status. Consumption, income, and wealth holdings are in thousands of dollars. The risky investment share is the average share of total assets invested in risky assets. The risky participation rate measures the share of the population that holds more than half of their portfolio in risky assets.

Table I.7: Endogenous Parameters for Model with Borrowing up to 50,000 dollars

Parameter	Value	Moment (mean of)	Data	Model
Discount factor β	0.927	Wealth in dollars	288,000	290,000
Risky investments costs φ_k	4,828	Risky investment share	0.557	0.555
Disutility from work ψ	0.375	Hours worked	0.399	0.399
Rumination, mild $n_r(m_1)$	0.068	Δ Hours worked, mild	-0.047	-0.047
Rumination, serious $n_r(m_2)$	0.111	Δ Hours worked, serious	-0.127	-0.127
Utility cost of treatment ξ_τ	0.036	Treatment share, serious	0.656	0.656
Treatment availability ω_τ	0.690	Treatment share, mild	0.414	0.415

Table I.7 presents the parameter values set to match model-generated moments to their data analog. The first three columns present the parameters and their values. The fourth column describes a moment that informs the parameter value. The fifth and sixth column present the model-generated moment and the data-equivalent.

Table I.8: Validation of Averages for Model with Borrowing up to 50,000 dollars

	Data			Model		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Consumption	51	47	43	52	48	44
Hours	0.404	0.383	0.351	0.400	0.380	0.348
Income	64	56	46	64	56	46
Wealth	292	262	236	294	267	245
Risky investment share	0.572	0.461	0.384	0.574	0.444	0.334
Risky participation rate	0.601	0.516	0.449	0.600	0.484	0.387

Table I.8 shows average consumption, hours, income, wealth, and risky investment by mental health status. Consumption, income, and wealth holdings are in thousands of dollars. The risky investment share is the average share of total assets invested in risky assets. The risky participation rate measures the share of the population that holds more than half of their portfolio in risky assets.

Table I.9: Endogenous Parameters for Model with Productivity Elasticity $\theta = 0.4$ (French, 2005)

Parameter	Value	Moment (mean of)	Data	Model
Discount factor β	0.966	Wealth in dollars	288,000	290,000
Risky investments costs φ_k	2,900	Risky investment share	0.557	0.556
Disutility from work ψ	0.375	Hours worked	0.399	0.404
Rumination, mild $n_r(m_1)$	0.021	Δ Hours worked, mild	-0.047	-0.048
Rumination, serious $n_r(m_2)$	0.060	Δ Hours worked, serious	-0.127	-0.127
Utility cost of treatment ξ_τ	0.025	Treatment share, serious	0.656	0.655
Treatment availability ω_τ	0.697	Treatment share, mild	0.414	0.413

Table I.9 presents the parameter values set to match model-generated moments to their data analog. The first three columns present the parameters and their values. The fourth column describes a moment that informs the parameter value. The fifth and sixth column present the model-generated moment and the data-equivalent.

Table I.10: Validation of Averages for Model with Productivity Elasticity $\theta = 0.4$ (French, 2005)

	Benchmark Model			Model		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Consumption	51	47	43	53	49	45
Hours	0.404	0.383	0.351	0.408	0.386	0.355
Income	64	56	46	65	57	48
Wealth	292	262	236	294	265	240
Risky investment share	0.572	0.461	0.384	0.574	0.454	0.386
Risky participation rate	0.601	0.516	0.449	0.605	0.507	0.449

Table I.10 shows average consumption, hours, income, wealth, and risky investment by mental health status. Consumption, income, and wealth holdings are in thousands of dollars. The risky investment share is the average share of total assets invested in risky assets. The risky participation rate measures the share of the population that holds more than half of their portfolio in risky assets.

J Model with Unobserved Heterogeneity

In this appendix, we augment our baseline model with unobserved individual heterogeneity. This captures the idea that unobserved confounders may be driving both mental health and labor market outcomes. In this extension of our model, each household is born with an innate type from a discrete set $x \in \{1, ..., K\}$. Individual type impacts (1) the distribution from which the initial mental health state is drawn, (2) the mental health transition matrix, (3) the deterministic life-cycle component of productivity, and (4) the persistence and variance of innovations of the idiosyncratic productivity component. Other parameters are common across types.

J.1 Estimation

We estimate the model with unobserved heterogeneity in four steps. First, exogenous model parameters (see Section 4.2), except for the ones that depend on individual type x, are set to their baseline values. Second, we estimate the classes of individual heterogeneity by k-means clustering following Bonhomme, Lamadon, and Manresa (2019, 2022) and Jolivet and Postel-Vinay (2024). Third, we estimate the exogenous parameters that depend on individual heterogeneity x. Fourth, we re-estimate the endogenous parameters as discussed in Section 4.3.

Estimation of Unobserved Heterogeneity. We estimate the classes of individual heterogeneity by k-means clustering following the approach of Bonhomme, Lamadon, and Manresa (2019, 2022). The idea is to partition individuals into classes by minimizing the dissimilarity between individuals that are assigned to the same class. In particular, denote by m_i a vector of M outcome variables for individual i. A partition assigns a class $x_i \in \{1, ..., K\}$ to each individual i. The partition minimizes $\sum_i \|m_i - \overline{m}(x_i)\|$, where $\overline{m}(x_i)$ is the average of the vector m across all individuals that are assigned to class x_i, and $\|\cdot\|$ denotes the Euclidean distance.

Bonhomme, Lamadon, and Manresa (2019, 2022) highlight that the outcome variables in the vector m should be informative about unobserved individual heterogeneity. Since the main sources of heterogeneity in our application are labor productivity and mental health, we include variables that summarize the individual's wage and mental health as in Jolivet and Postel-Vinay (2024). Specifically, the vector m includes the average wage and the share of time the individual is healthy, experiences mild mental illness, and experiences severe mental illness. All moments are computed using PSID data. The average wage is based on residualized wages from a regression of log hourly wages on log hours worked, education, gender, mental health, family composition, race, and time fixed effects. Following Jolivet and Postel-

Table J.1: Heterogeneity Classes

Class	Share	Average Wage	Healthy	Mild	Severe
1	0.126	-0.31	28.7	53.2	18.1
2	0.452	-0.40	94.6	4.2	1.2
3	0.422	0.34	94.4	4.4	1.2

Table J.1 presents the partition of individuals into classes. The second column reports the fraction of individuals in each class. The third column reports the average wage within each class. Columns 4 to 6 report the percent of time spent in each of the three mental health states for each class.

Vinay (2024), we modify the classification procedure to account for the fact that our PSID panel data follows different cohorts for a limited number of years.

Table J.1 describes the resulting partition into three classes. It reports the share of individuals in each class and the moment averages for each class. Our partition results in groups that differ along the productivity dimension and the health dimension. Class 3 is a group with high productivity and in good mental health. Class 1 is a group with low productivity and in low mental health. Class 2 lies in between Class 1 and 3, and consists of individuals with low productivity but in good mental health.

Estimation of Exogenous Parameters. Taking the classification as given, we proceed to estimate the class-dependent parameters: (1) the distribution from which the initial mental health is drawn, (2) the mental health transition matrix, (3) the deterministic life-cycle component of productivity, and (4) the persistence and variance of innovations of the idiosyncratic productivity component.

The initial distribution of mental health states π_m is estimated from the observed distribution in the PSID data at age 25. For each class, we calculate the share of 25 year-olds that are in each of the three mental health states. For the mental health transitions, we assume that individuals in Class 1 and Class 3, which are characterized by low productivity, face the transition probabilities $\Gamma_m(m' \mid m, \tau, \nu < \underline{\nu})$ estimated in the baseline calibration (see Appendix D). In contrast, individuals in Class 2, which are characterized by high productivity, face the transition probabilities that correspond to the high productivity states, that is $\Gamma_m(m' \mid m, \tau, \nu \geqslant \underline{\nu})$.

We estimate the deterministic life-cycle component of productivity ζ_t, as well as the persistence ρ_ν and variance of innovations σ_ν^2 of the idiosyncratic productivity component, by analyzing residual wages in the PSID by class. Consistent with the baseline calibration and with the wage equation (4), we regress log hourly wages on log hours worked, where the elasticity of wages to hours as well as the intercept may vary

Table J.2: Endogenous Parameters for Model with Unobserved Heterogeneity

Parameter	Value	Moment (mean of)	Data	Model
Discount factor β	0.971	Wealth in dollars	288,000	289,000
Risky investments costs φ_k	4,100	Risky investment share	0.557	0.558
Disutility from work ψ	0.305	Hours worked	0.399	0.409
Rumination, mild $n_r(m_1)$	0.060	Δ Hours worked, mild	-0.047	-0.047
Rumination, serious $n_r(m_2)$	0.111	Δ Hours worked, serious	-0.127	-0.127
Utility cost of treatment ξ_τ	0.039	Treatment share, serious	0.656	0.655
Treatment availability ω_τ	0.689	Treatment share, mild	0.414	0.415

Table J.2 presents the parameter values set to match model-generated moments to their data analog. The first three columns present the parameters and their values. The fourth column describes a moment that informs the parameter value. The fifth and sixth column present the model-generated moment and the data-equivalent.

by the short, medium, and long hours regions. We extract a deterministic life-cycle profile ζ_t by fitting a third-order polynomial through the age effects of the remaining variation, and estimate the persistence ρ_ν and the variance of productivity shocks σ_ν^2 to align the model-implied and empirical auto-covariation of residual wages. We find persistence parameters $\rho_\nu(x=1)=0.940$, $\rho_\nu(x=2)=0.917$, $\rho_\nu(x=3)=0.900$, and variance of the innovations $\sigma_\nu^2(x=1)=0.089$, $\sigma_\nu^2(x=2)=0.096$, $\sigma_\nu^2(x=3)=0.084$.

Estimation of Endogenous Parameters. We estimate the endogenous model parameters to match the data moments specified in Table 5. The results are reported in Table J.2. Parameters are largely unchanged relative to the baseline estimation.

J.2 Quantitative Results

We evaluate the robustness of our quantitative analysis (Section 5) to incorporating unobserved household heterogeneity.

Welfare Cost of Mental Illness. We find that the aggregate consumption equivalent cost of mental illness Δ^m is 1.0 percent of consumption. The average welfare cost of mental illness for individuals experiencing serious mental illness is equivalent to 10.5 percent of consumption, while the average consumption equivalent cost of mental illness for individuals experiencing mild mental illness is 5.4 percent. These

Table J.3: Validation: Averages for Model with Unobserved Heterogeneity

	Benchmark Model			Model		
	Healthy	Mild	Serious	Healthy	Mild	Serious
Consumption	51	47	43	52	47	44
Hours	0.404	0.383	0.351	0.412	0.393	0.365
Income	64	56	46	65	56	48
Wealth	292	262	236	295	261	243
Risky investment share	0.572	0.461	0.384	0.574	0.464	0.400
Risky participation rate	0.601	0.516	0.449	0.602	0.522	0.468

Table J.3 shows the average consumption, hours, income, wealth, and risky investment shares by mental health status. Consumption, income, and wealth holdings are in thousands of dollars. The risky investment share is the average share of total assets invested in risky assets. The risky participation rate measures the share of the population that holds more than half of their portfolio in risky assets.

estimates are in line with the results obtained from the baseline model without unobserved household heterogeneity.

Mental Health Policies. As in our baseline model, we evaluate a policy that makes treatment available to all. That is, we consider an increase of ω_τ from 0.689 to 1. In line with our baseline results, we find that the average welfare benefit of providing full availability of treatment services Δ^ω is 0.24 percent of aggregate consumption. As in the baseline, we also consider a policy that improves the mental health of adolescents and young adults. We find that the average consumption equivalent gain of treatment in young adulthood Δ^{τ_0} is equal to 0.62 percent.