The Macroeconomics of Intergenerational Mental Health Dynamics*

Boaz Abramson
Columbia GSB

Job Boerma
University of Wisconsin-Madison

Diego Daruich
University of Southern California

Aleh Tsyvinski
Yale University

May 2026

Abstract

We develop a quantitative macroeconomic theory of child mental health. The theory is grounded in child psychiatry, formalized in a life-cycle heterogeneous agent model of child development, and disciplined using micro data on mental health of children and parents. Intergenerational transmission of mental illness arises due to both biological factors and parental behavior. Parents experiencing mental illness have negative expectations and lose time due to rumination. As a result, they invest less in their child's mental health. We use the model to evaluate policies designed to improve child mental health. We show that subsidizing mental health treatment for children generates sizable welfare gains.

JEL-Codes: D31, I12, I18, J13, J24
Keywords: Mental Health, Child Development, Intergenerational Dynamics

*We thank Francesco Agostinelli and Anmol Bhandari for useful discussions and Nicholas Garvey for excellent research assistance.

1 Introduction

The U.S. Surgeon General has described child mental health as "the defining public health crisis of our time" (Murthy, 2021). In the U.S., more than one in five adolescents is diagnosed with a mental health condition at a given point in time (Sappenfield et al., 2023) and roughly half of all adolescents experience a mental health problem at some point in their life (Merikangas et al., 2010b).[1] The costs of mental illness in childhood and adolescence are substantial. Children experiencing mental illness are more likely to experience mental illness in adulthood (Kim-Cohen et al., 2003; Kessler, Chiu, Demler, and Walters, 2005), are less likely to complete high school, to enroll in college and to graduate from college (Fergusson and Woodward, 2002; Currie and Stabile, 2006; Breslau, Lane, Sampson, and Kessler, 2008), and experience worse labor market outcomes later in life (Weissman et al., 1999; Goodman, Joyce, and Smith, 2011; Clayborne, Varin, and Colman, 2019).

This paper develops a quantitative macroeconomic model of child mental health and uses it to study the determinants of the child mental health crisis and policy responses to it. Our model integrates the psychiatric literature on child mental health with quantitative macroeconomic models. The psychiatric literature identifies four key determinants of child mental health. First, both biological factors and environmental factors shape child mental health. Second, among environmental factors, parent behavior is a main determinant of child mental health. Third, parental behavior depends on parental mental health. Fourth, mental health treatment is effective but taken up by a relatively small share of children experiencing mental illness. We use our model to evaluate public policies that are often proposed to address child mental illness. We consider policies that expand mental health treatment services for children and that subsidize treatment for parents.

We formalize our economic theory of child mental health in an overlapping generations economy with heterogeneous households and child development. Child mental health is a stochastic state variable that depends on biological and environmental factors. Children are born with an initial mental health state that probabilistically depends on their parents' mental health. Parents can invest both time and money to improve child mental health, and also choose whether to seek mental health treatment for their children. Treatment increases the probability of transitioning into better mental health but is costly. Children are also endowed with cognitive skills that develop as a function of parental investments. Child mental health and cognitive skills are important for educational attainment and income later in life.

[1] Currie (2025) shows that rates of child mental illness have been relatively stable in the past decades to argue that the child mental health crisis is not a new phenomenon. The most common disorders are anxiety disorders, behavior disorders and mood disorders (including depressive disorders).

In our framework, adults also experience mental illness. Following Abramson, Boerma, and Tsyvinski (2024), adults experiencing mental illness hold a negative view of the future (Beck, 1967b, 1976, 2002, 2008) and spend time ruminating on their negative thoughts (Nolen-Hoeksema, 1991; Just and Alloy, 1997; Nolen-Hoeksema, 2000; Nolen-Hoeksema, Wisco, and Lyubomirsky, 2008; Singer and Dobson, 2007). Negative thinking implies parents experiencing mental illness have negative expectations over the returns to risky parental investments and child mental health treatment. Rumination implies parents experiencing mental illness have less time to spend with their children. Due to negative thinking and rumination, parents experiencing mental illness may invest less in their child's mental health. Overall, in line with the child psychiatry literature, mental illness in our model persists across generations through both exogenous biological factors as well as environmental factors, parental behavior behavior plays a key role in shaping child mental health, and parent behavior in turn depends on parents' own mental health.

We quantify the model using micro data on mental health of children and parents. We estimate the child development technology for mental health and cognitive skills using the National Longitudinal Survey of Youth (NLSY), which measures parents and children mental health and cognitive skills, as well as parental investments. We assume child mental health and cognitive skills develop as a function of current child mental health and cognitive skills, parents mental health and cognitive skills, parental investments and mental health treatment. We estimate the parameters of the development technology building on the framework of Cunha, Heckman, and Schennach (2010). The distribution from which initial child mental health is drawn is quantified as part of the estimation of the child development technology using the empirical relationship between parents' mental health and newborns' mental health.

The estimation of the child mental health development function delivers the following new insights. First, self-productivity of child mental health is large and increasing with age. Second, parent mental health plays an important role in terms of the mental health and cognitive skill development of children. While the effect of parent mental health on cognitive skills sharply falls with age, the effect on child mental health remains pronounced. Third, the inputs into child mental health development are complementary at each development stage and the complementarity increases with age. Fourth, biological factors matter. Newborns whose parents experience mental illness are more likely to experience mental illness.

We quantify the impact of mental health treatment on child mental health using medical estimates on the impact of child psychotherapy. In line with the medical literature (Weisz et al., 2017), we set the treatment efficacy for children experiencing mental illness such that the model-implied treatment effect, in terms of standard mean differences (SMD), equals 0.46. We quantify the extent of negative thinking associated with mental illness so that the model matches observed differences in consumption by mental health status in the PSID data. In the model, more pronounced negative thinking induces more precautionary savings and thus lower consumption. We calibrate the extent of rumination associated with mental illness such that the model matches working hours by mental health status in the data. Controlling for individual level fixed effects, we estimate that individuals experiencing mild (serious) mental illness work 2.7 (15.9) percent fewer hours. With 1.1 (6.3) hours per week of rumination for those experiencing mild (serious) mental illness, the model aligns with the data. The productivity costs of mental illness are estimated from wage regressions that control for individual fixed effects. We estimate a decrease of 2.2 percent in hourly wages associated with mild mental illness, and 5.5 percent decrease associated with serious mental illness. We assume that mental illness imposes a flow utility cost on individuals and that individuals who seek treatment incur an additional stigma utility cost. These utility costs are calibrated so that the model matches treatment rates over the life cycle.

We validate the model by evaluating model predictions against non-targeted moments that are important for child mental health. First, our model accurately predicts the empirical relationship between parental mental health and parental investments in child mental health. Using PSID data, we show that parental mental illness is associated with a 15 percent drop in the number of hours spent with children. The model is qualitatively and quantitatively consistent with this important elasticity. Second, the model closely matches the intergenerational transmission of mental health measured in the data. In both the model and the data, children of parents who experience mental illness are 25 percentage points more likely to experience mental illness. Third, in both the model and in the data individuals experiencing mental illness are less likely to enroll in and to graduate college. In both the model and in the data, individuals who experience mental illness are more pessimistic about their graduation prospects conditional on enrollment.

We use the model to assess the costs associated with parental mental illness. For individuals experiencing serious (mild) mental illness, eliminating mental illness yields a 40 (17) percent annual consumption equivalent welfare gain. In line with the psychiatric literature, the welfare effects are largely driven by the lack of negative thinking for healthy individuals. In terms of aggregate output, eliminating parental mental illness increases output by 8.3 percent. The output effect is largely driven by improved parenting quality. When parents are healthy their children are less likely to experience mental illness when they grow up, in part because healthy parents invest more in their children. Subsequently, children of healthy parents are more likely to enroll in college and are more productive.

We evaluate the equilibrium effects of policies that aim to address child mental health. A main policy that is often proposed is to expand the take-up of mental health services among children, for example by providing school-based mental health services or by providing access to treatment through community health clinics. Motivated by these policies, we study the effects of fully subsidizing mental health services for children. We find that this increases the treatment rate among children experiencing mental illness from 11 percent in the baseline economy to 36 percent under the policy. This translates to improved mental health for children and adolescents, and to an average welfare gains of 0.78 percent of annual consumption.

Policymakers recognize the importance of parents' mental health for children's mental health. For example, according to the U.S. Surgeon General, parental mental health is a child healthcare issue (Murthy, 2021). Policies targeting parents' mental health primarily focus on improving maternal mental health during and immediately post pregnancy. Motivated by these policies, we study the effects of subsidizing mental health treatment for adults before and immediately after a child is born. We find that this increases the treatment rate among the target audience by 3.2 percentage points, and as a result increases parental investments in children. This translates to an average welfare gains of 0.13 percent of annual consumption. Finally, we consider a policy that subsidizes treatment for both children and parents. We find that it yields an average welfare gains of 1.04 percent of annual consumption

Literature. The main contribution of our paper is to develop a quantitative macroeconomic model of child mental health. There is a rich literature studying macroeconomic models of health (Grossman, 1972; 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). Recent work studies the macroeconomics of mental health (Abramson, Boerma, and Tsyvinski, 2024). While previous work has focused on mental health during adulthood, our focus is on child mental health. In contrast to physical illness, mental illness tends to onset early in life and is widespread among children and adolescents (Kessler et al., 2005; Caspi et al., 2020). We develop a framework that explicitly incorporates the key factors that shape child mental health according to the psychiatric literature and use it to study policy proposals to intended to mitigate the child mental health crisis. We discuss the child psychiatric literature that provides the foundation for our model in Section 2.

Our work relates to the macroeconomic literature on parenting. Starting with the seminal work of Becker and Tomes (1979, 1986), this literature studies how parental behavior shapes child outcomes. A common finding is that parent investments, in terms of money and time, account for a large share of the intergenerational persistence of earnings and wealth (Restuccia and Urrutia, 2004; Lee and Seshadri, 2019; Yum, 2023). Models of parent investments have also been used to study the effects of government policy interventions on parental behavior, child outcomes and welfare. For example, Krueger and Ludwig (2013) and Abbott, Gallipoli, Meghir, and Violante (2019) study the effects of college financial aid programs, Daruich (2024) analyzes the equilibrium effects of government investment in early childhood programs, while Caucutt and Lochner (2020) consider the effect of relaxing borrowing constraints.[2] Doepke and Zilibotti (2017), Doepke, Sorrenti, and Zilibotti (2019), and Agostinelli, Doepke, Sorrenti, and Zilibotti (2026) emphasize how different parenting preferences translate to different parental behavior, children's welfare and economic outcomes.[3] The contribution of our paper is to develop a model of parent investments that incorporates mental health of parents and children and to use the model to study how parent mental illness shapes parenting behavior and how government interventions targeted at both children and parents affect child mental health.

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, 2025). In our model, parents experiencing more serious 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 Abramson, Boerma, and Tsyvinski (2024) for a discussion). In line with the psychiatric literature (see Section 2), in our model negative thinking is one source of intergenerational transmission of mental health. Parents experiencing mental illness hold a negative view on the efficacy of parental investment and mental health treatment, which leads them to invest less in the mental health of their child.

[2] Gu and Zhang (2026) study educational competition in China with parental investments in tutoring for college entrance exams. Distinctive from the child development literature, investments in tutoring increase cognitive skills, while reducing mental health, which in their framework affect consumption utility and income, akin to models of physical health.

[3] Adamopoulou, Cavalcanti, Greenwood, and Santos (2026) study the relationship between video gaming in early adulthood and family formation as well as the probability of being depressed later in adulthood.

2 Literature on Child Psychiatry

The psychiatric literature emphasizes the role of both biological and environmental factors in shaping the mental health of children and adolescents.

Genes. Psychiatric genetics, which studies the genetic etiology of psychiatric disorders, has demonstrated that psychiatric disorders are heritable (see Sullivan, Daly, and O'Donovan (2012) and Andreassen, Hindley, Frei, and Smeland (2023) for recent reviews): a portion of the variation in the risk of developing mental illness is attributable to genetic factors.[4] Estimates of heritability tend to be highest for schizophrenia and lowest for mood disorders (Sullivan and Geschwind, 2019).

Maternal Prenatal Mental Health. Maternal mental health during pregnancy also plays an important role in shaping child mental health. A large body of literature documents that prenatal depression and anxiety are negatively associated with child mental health (see Madigan et al. (2018) and Rogers et al. (2020) for recent reviews). This relationship is not explained by hereditary factors (Rice et al., 2010; Rajyaguru, Kwong, Braithwaite, and Pearson, 2021) and remains strong after controlling for postnatal maternal mental health (O'Connor, Heron, Golding, Beveridge, and Glover, 2002). The literature indicates neuro-biological mechanisms through which prenatal maternal mental illness negatively impacts child mental health.[5] Motivated by this body of work, children in our model are born with an initial mental health that probabilistically depends on parental mental health.

Parenting. Parenting quality has been recognized as a critical determinant of child mental health since the very early days of psychoanalysis (James, 1890; Breuer and Freud, 1895). A large body of work in clinical psychology shows that parental involvement, sensitivity to distress, warmth, responsiveness, and engagement are positively associated with child mental health (see Fearon, Bakermans-Kranenburg, van IJzendoorn, Lapsley, and Roisman (2010), Groh, Roisman, van IJzendoorn, Bakermans-Kranenburg, and Fearon (2012), Pinquart (2017), and Spruit et al. (2020) for recent reviews). Attachment theory, the predominant psychiatric theory for parent-child interactions (Bowlby, 1969; Ainsworth, Blehar, Waters, and Wall, 1978), posits that early childhood experiences with caregivers shape a child's internal working

[4]Through classical twin and adoption studies, and advances in the sequencing of the human genome and genome-wide association studies, it has been established that common genetic variants influence the risk of mental illness (Polderman et al., 2015; Visscher, Yengo, Cox, and Wray, 2021; Grotzinger et al., 2025).

[5]Prenatal depression and anxiety have been shown to increase fetal cortisol concentration (Seckl and Meaney, 2004), which leads to changes in fetal brain function (Talge, Neal, and Glover, 2007), a reduction in the flow of oxygen and nutrients to the fetus (Teixeira, Fisk, and Glover, 1999), and, ultimately, alterations in fetal brain development. Prenatal depression and anxiety can also lead to epigenetic dysregulation in serotonin transmission that delays offspring cognitive development (Oberlander et al., 2008).

model – a set of cognitive structures that govern how children perceive their self-worth and relationships with others. When a caregiver is repeatedly unavailable to a child's needs, the child is at risk of developing a negative view of the self and negative expectations about relationships with others. These negative cognitive biases enhance the risk of mental disorders (Beck, 1967a). Building on the psychological and psychiatric literature, we model parental mental health and parental investment as a key driver of child mental health.

Parent behavior depends on parental mental health. Parents experiencing mental illness spend less quality time with their children, are less sensitive and responsive, less involved and engaged, and display less warmth (Weissman and Paykel, 1974; Field, Healy, Goldstein, and Guthertz, 1990; Lovejoy, Graczyk, O'Hare, and Neuman, 2000; Wilson and Durbin, 2010). As discussed in Abramson, Boerma, and Tsyvinski (2024), two key features of mental illness are negative thinking and rumination.[6] These features can rationalize why parents experiencing mental illness exhibit more negative parenting behavior. Negative thinking biases parents' appraisal of their ability to parent effectively, which can lead to withdrawal, inaction, and helplessness (Abramson, Seligman, and Teasdale, 1978; Cummings and Davies, 1994; Dix and Meunier, 2009). Rumination makes it difficult for parents to engage with their children and respond to their needs (Lovejoy, Graczyk, O'Hare, and Neuman, 2000; Stein et al., 2012; DeJong, Fox, and Stein, 2016). We build on the psychiatric literature and model negative thinking and rumination as key features of mental illness. In turn, negative thinking and rumination induce lower parental investments in children. Overall, and consistent with the psychiatric literature, parents in our model shape child mental health through both biological factors and parenting behavior.

Other Environmental Factors. Other environmental factors influence child mental health. Psychiatric literature emphasizes that mental illness can be triggered by adverse events such as the death of a loved one (Dowdney, 2000; Jensen and Zhang, 2026), peer victimization and bullying (Moore et al., 2017; Arseneault, 2018), sexual abuse (Widom, DuMont, and Czaja, 2007; Danese et al., 2009), and neighborhood violence (Fowler, Tompsett, Braciszewski, Jacques-Tiura, and Bastes, 2009). Consistent with this, mental health in our framework evolves stochastically and is subject to shocks that are independent of parental behavior.

Treatment. The main way to improving child mental health is treatment. A vast medical literature estimates the effects of psychotherapy and antidepressants on child mental health using randomized

[6] Individuals experiencing mental illness tend to hold a negative view of the self, of others, and of the future (Beck, 1967a, 1976), and they tend to spend excessive amounts of time ruminating on negative thoughts (Nolen-Hoeksema, 1991; Nolen-Hoeksema, Wisco, and Lyubomirsky, 2008).

trials. Treatment effect sizes are typically reported in terms of the standardized mean difference (SMD) to facilitate comparison across studies. Psychotherapy is generally found to be effective: a meta-analysis by Weisz et al. (2017) reports an average SMD of 0.46 for behavioral psychotherapy.[7] Our model features mental health treatment and uses these treatment effects to inform its efficacy.

Despite the efficacy of treatment, a relatively low share of children and adolescents who experience mental illness receive treatment. The NIMH estimates that, in 2021, only 41 percent of 12-17 year-olds experiencing major depression received treatment.[8] The medical literature identifies several possible explanations for the low take-up. First, the lack of availability of mental health services is one of the most commonly cited barriers to treatment. Second, even when mental health treatment is available, it might be unaffordable or require significant parental time investments (Pavuluri, Luk, and McGee, 1996). Third, stigma is an important factor contributing to low treatment rates among children and adolescents (see, for example, Gulliver, Griffiths, and Christensen (2010) and Clement et al. (2015)). Our model incorporates such barriers to treatment, and we use the model to evaluate the efficacy of interventions designed to alleviate them.

The psychiatric literature on child mental health shares notable commonalities with the economics literature on child development (Becker and Tomes, 1979; Todd and Wolpin, 2003; Cunha and Heckman, 2008; Cunha, Heckman, and Schennach, 2010). Both literatures emphasize the fundamental role played by parents in shaping child outcomes. Consistent with the psychiatric literature, the child development literature highlights that parents shape child skills through both biological factors and parental investments (Cunha and Heckman, 2007), that quality time is a key component of parental investment (Del Boca, Flinn, and Wiswall, 2014; Caucutt and Lochner, 2020), that parental investments are particularly effective in early childhood (Cunha, Heckman, and Schennach, 2010), and that skills developed in childhood are important for subsequent outcomes in adulthood (Abbott, Gallipoli, Meghir, and Violante, 2019; Daruich, 2024). In line with the psychiatric literature, recent work on child development argues that subjective expectations about the efficacy of parental investments are an important source of variation in parental investments (Attanasio, Cunha, and Jervis, 2019; Kinsler and Pavan, 2021; Attanasio, Boneva, and Rauh, 2022). Motivated by these commonalities, we model child mental health dynamics by building on the economic child development framework.

[7] Antidepressants are generally found to be less effective. A meta-analysis by Cipriani et al. (2016) finds that out of 14 antidepressant treatments, only fluoxetine was more effective than a placebo, with a treatment effect that was marginally statistically significant.

[8] See www.nimh.nih.gov. Merikangas et al. (2010a) and Merikangas et al. (2011) report similar treatment rates for other mental disorders.

3 Model

We formalize our theory of child mental health in an overlapping generations model with heterogeneous agents. We consider an infinite horizon economy where each generation is of mass one. Individuals live for 20 periods $t = 1, 2, 3, \ldots, 20$ and die deterministically thereafter.[9] Each period represents a period of four calendar years. There is no population growth.

The model consists of two main blocks. The first is a child development block where biological factors and parental behavior shape child mental health. The second is a heterogeneous agent life-cycle model in which mental health impacts human capital accumulation, labor market outcomes, and well-being.

Preferences. Adults derive utility $u(c, \ell)$ from consumption c and leisure ℓ. They have preferences that are separable in time and discount the future with a discount factor β. Total time within each period is normalized to one. Adults also derive utility from their children's consumption and leisure, as discussed below in more detail.

Adult Mental Health. Adults' mental health state is denoted by m. We consider three mental health states: healthy ($m = m_0$), mild illness ($m = m_1$), and serious illness ($m = m_2$). During adulthood, mental health evolves according to a first-order Markov chain with conditional transition probabilities $\Gamma_m(\tau_t, z_t)$ that depend on the adult's treatment choice τ_t and on idiosyncratic labor productivity z_t. Negative labor market shocks can thus affect mental health. Adult mental health governs negative thinking, rumination, labor productivity, hours worked, treatment efficacy, flow utility, and the development of child mental health and cognitive skills.

Negative Thinking. A key feature of mental illness according to the psychiatric literature is negative thinking. Building on the cognitive model of depression (Beck, 1967a, 1976, 2002), we model negative thinking as negative expectations over random outcomes. Consider the following example to illustrate how we formalize 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 means individuals form expectations 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 $p(\chi)$ that minimizes

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

the expected value. Minimization is subject to a relative entropy constraint:

$$\mathcal{R}(p\|q) = \sum_{\chi \in \Omega_\chi} p(\chi) \log \frac{p(\chi)}{q(\chi)} \leqslant \kappa(m). \tag{2}$$

The relative entropy constraint limits the choice of subjective probabilities to those that are close enough to the objective probabilities. The extent to which subjective probabilities can differ from objective probabilities is governed by κ, which represents the extent 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 place more weight on bad outcomes and less weight on good outcomes. Specifically, the subjective probability of state χ is given by:

$$p(\chi) = \frac{q(\chi)\exp(-\lambda(m)w(\chi))}{\sum q(\chi)\exp(-\lambda(m)w(\chi))}, \tag{3}$$

where λ is the inverse of the Lagrange multiplier with respect to the relative entropy constraint. Adverse outcomes are states with low values $w(\chi)$ and λ governs the distance between the subjective and objective probability distributions. An increase in λ corresponds to more negative thinking. A value of $\lambda = 0$ implies $p(\chi) = q(\chi)$, i.e., that the subjective probabilities coincide with the objective probabilities. The relationship between λ and the extent of negative thinking κ is monotonic, implying that the larger κ is, 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 serious mental illness think more negatively about the future. That is, κ increases with the severity of mental illness.[10]

Rumination. A second prominent feature of mental illness is rumination (Nolen-Hoeksema, 1991, 2000). We model rumination as a reduction in 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 thus $1 - n_r(m)$. In the calibrated model, individuals experiencing more serious mental illness lose more time due to rumination. That is, $n_r(m)$ increases with the severity of mental illness.

Income. Mental health affects wages and hours worked. Working-age households earn an hourly wage rate $w_t(m, z_t, e, \theta)$. The hourly wage is the product of the wage per hourly efficiency unit w_e, which varies by education level e, and the individual's hourly efficiency units. Hourly efficiency units are composed of a deterministic age component ζ_t^e (which depends on education), mental health m, cognitive ability θ,

[10] A useful feature of this approach is that it does not require the dimension of uncertainty to be unidimensional. We exploit this feature in the household decision problems in which households face uncertainty with respect to the joint evolution of stochastic variables.

and an idiosyncratic productivity component z_t. Specifically, log hourly wage is given by:

$$\log w_t(m, z_t, e, \theta) = \log w_e + \log \zeta_t^e + \lambda_m(m) + \lambda_\theta^e \log \theta + \log z_t\,, \tag{4}$$

where $\lambda_m(m)$ captures the effect of mental health on wages and λ_θ^e captures the effect of cognitive skills on wages which can vary by educational attainment. Idiosyncratic productivity z_t follows a first-order autoregressive process in logs:

$$\log z_t = \rho^e \log z_{t-1} + \upsilon_t^e\,, \tag{5}$$

where $\upsilon_t^e \sim N(0, \sigma^e)$. The process for idiosyncratic productivity z_t is persistent and evolves stochastically, following an education-dependent process.

Labor income for working-age individuals is the product of the hourly wage rate w_t and working hours $n_w(m, s)$. Working hours vary by mental health and depend on whether the individual is currently enrolled in college as a student. College students $(s = 1)$ spend n_s hours studying and thus work less than those not enrolled $(s = 0)$: $n_w(m, 1) = n_w(m, 0) - n_s$. Household labor income before taxes is denoted by $y_t(m, z_t, e, \theta) = w_t(m_t, z_t, e, \theta) n_w(m_t, s_t)$. In retirement, households receive pension income $y_t^p(\theta, e)$ which depends on their cognitive skills and educational attainment.

Mental Health Treatment. Adults decide whether to seek mental health treatment. Let $\tau_t = 0$ denote an adult who does not undertake treatment and let $\tau_t = 1$ denote an adult undertakes treatment. Treatment increases the probability of transitioning into better mental health states. Treatment is associated with a time cost n_τ, a financial cost φ_τ, and a utility cost $\xi_{\tau,t}$ that can depend on age. The utility cost of treatment to captures the stigma associated with mental health treatment, which is an important driver of treatment take-up according to the psychiatric literature (see, for example, Corrigan (2004) and Clement et al. (2015)).

Adults also decide whether to seek mental health treatment for their children. Let $\tau_t^k = 0$ denote that a child does not get treated and $\tau_t^k = 1$ denote that a child does get treated. As with adults, treatment probabilistically improves child mental health. When a child gets treated, the parent incurs a time cost, n_τ^k, a financial cost, φ_τ^k, and a utility cost, ξ_τ^k.

Leisure. Leisure is a function of mental health, of time devoted to mental health treatment, and of time devoted to college. It is given by

$$\ell(m, s, \tau, \tau_k) = 1 - n_r(m) - n_w(m, s) - n_\tau \tau - n_\tau^k \tau_k - n_s s. \tag{6}$$

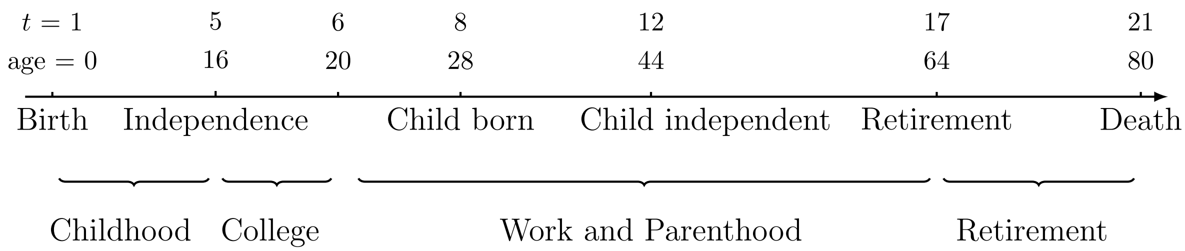


Figure 1: Life-cycle

Figure 1 illustrates key events and the four main stages of life for an agent in the model.

Assets. We consider a small open economy with a risk-free interest rate R. Households that save receive a gross interest rate R, while households that borrow pay a gross interest rate $R_b \geqslant R$. The spread captures the cost of unsecured borrowing. Households are subject to a borrowing constraint which varies by age. That is,

$$a_{t+1} \geqslant \underline{a}_{t+1}, \tag{7}$$

where $a_{t+1} > 0$ denotes savings and $a_{t+1} < 0$ denotes borrowing. The borrowing limit $\underline{a}_{t+1}$ is equal to the natural borrowing limit (see Section 4.2). Student loans are discussed in detail below.

3.1 Life-Cycle Problem

The life-cycle is divided into four stages: childhood, college, work and parenthood, and retirement. Figure 1 presents the timing of these stages. Each time period corresponds to four calendar years. That is, period $t = 1$ refers to ages 0−3, period $t = 2$ refers to ages 4−7, and so on. The childhood stage spans period $t = 1$ through $t = 4$. In childhood, individuals live with their parents and do not make any economic decisions. Individuals become independent and begin making their own choices at the college stage, which spans period $t = 5$. The first decision individuals make at the college stage is whether or not to enroll in college. The work and parenthood stage spans periods $t = 6$ to $t = 16$. Individuals become parents at age $t = 8$. Individuals retire at the beginning of period $t = 17$ (when they are 64) and die deterministically at the beginning of period 21 (when they are 80).

3.2 College Stage

Individuals begin making decisions in period $t = 5$. During the college stage, they choose their consumption, savings, mental health treatment decision, as well as whether to go to college for one period. Enrolling in college (a decision denoted by $s_t = 1$) entails a monetary cost τ_s and lowers working hours by

n_s. The benefit of enrolling is that college graduates earn higher wages. Let $e = 1$ denote an individual with a college degree and $e = 0$ an individual without a college degree. Individuals who choose not to enroll ($s_t = 0$) enter the labor market without a college degree. Individuals' state variables in period $t = 5$ are their assets a_t, mental health m_t, cognitive skills θ, an indicator for whether their parents are college graduates e_p, and a vector of school taste shocks χ. Individuals can finance their studies out of pocket or with subsidized student loans. Students loans carry a gross interest rate R_e such that $R \leqslant R_e \leqslant R_b$.

Enrolling in college is risky. Conditional on enrollment, individuals graduate with an objective probability that depends on their cognitive skills and mental health $p_e(\theta, m)$. With a complementary probability, $1 - p_e(\theta, m)$, they drop out of college and begin the next period as non-college graduates. Note that individuals experiencing mental illness think negatively about their probability of graduating which can affect their enrollment decisions.

The value of an individual choosing to enroll in college is:

$$v_5^{s=1}(a_5, m_5, \theta) = \max_{c_5, a_6, \tau_5} \Big\{ u\big(c_5, \ell_t(m_5, 1, \tau_5, 0)\big) - \xi\tau_5 + \beta \min_p \mathbb{E}_p v_6\big(\tilde{a}_s(a_6), m_6, z_6, e, \theta\big) \Big\}$$

subject to the entropy constraint (2), the borrowing constraint (7), and the budget constraint:

$$c_5 + a_6 + \tau_s + \varphi_\tau \tau_5 = Ra_5 + y_5 - T(y_5, a_5, c_5).$$

Subjective expectations are formed over future mental health, m_6, the graduation outcome, e, and the initial productivity in the labor market post-college, z_6. Individuals who enroll in college earn a non-graduate hourly wage while in school. Their idiosyncratic productivity during college is given by $\log z_5 = 0$. Their income is therefore given by $y_5 = w_5(m_5, 1, 0, \theta) n_w(m_5, 1)$. Initial idiosyncratic productivity post-college is drawn according to $\log z_6 \sim N(0, \sigma_6^e)$, where the distribution depends on the graduation outcome e_6. College students can access subsidized loans at a rate R_e, and these loans are subject to the borrowing constraint (7). To simplify computation, we follow Abbott, Gallipoli, Meghir, and Violante (2019) and assume that college student debt is refinanced into a single bond with fixed payments over 5 periods (20 years) following graduation that carries an interest rate R_b, where $\tilde{a}_s(a_6)$ is the function performing this transformation (see Appendix E.1 for details). Income, consumption, and dividend taxes are captured by the function T, which depends on labor income, assets, and consumption.

Individuals who decide not to enroll in college enter the labor market as non-graduates. Upon entering the labor market, individuals draw an initial idiosyncratic productivity shock z_5 from a log-normal distribution with mean zero and standard deviation σ_6^0. Individuals then choose consumption, savings, and whether to seek mental health treatment. The value of choosing not to enroll in college given a realization of idiosyncratic productivity is given by:

$$v_5^{s=0}(a_5, m_5, z_5, \theta) = \max_{c_5, a_6, \tau_5} \Big\{ u(c_5, \ell_t(m_5, 0, \tau_5, 0)) - \xi_\tau \tau_5 + \beta \min_p \mathbb{E}_p v_6\big(a_6, m_6, z_6, 0, \theta\big) \Big\}. \tag{8}$$

The maximization is subject to the entropy constraint (2) where subjective expectations are formed over future mental health and idiosyncratic productivity, to the borrowing constraint (7), and to the budget constraint:

$$c_5 + a_6 + \varphi_\tau \tau_5 = y_5 + Ra_5 - T(y_5, a_5, c_5),$$

where $y_5 = y_5(m_5, z_5, 0, \theta)$.

At the beginning of period 5, individuals make their college enrollment decision. Following Heckman, Lochner, and Todd (2006) and Abbott, Gallipoli, Meghir, and Violante (2019), school taste affects the utility of going to college in an additively separable fashion. The value function of an individual entering period $t = 5$ is given by:

$$v_5(a_5, m_5, \theta, e_p, \chi) = \max_s \Big\{ \min_p \mathbb{E}_p v_5^{s=0}(a_5, m_5, z_5, \theta) + \chi_0, v_5^{s=1}(a_5, m_5, \theta) + \Upsilon(\theta, e_p) + \chi_1 \Big\}, \tag{9}$$

where $\Upsilon(\theta, e_p)$ is a direct flow utility from college that depends on cognitive skills θ and parental education e_p, and χ_0 and χ_1 are additive school taste shocks that are realized between period $t = 4$ and $t = 5$.[11] The maximization problem is subject to the entropy constraint (2), where subjective expectations are formed on the realization of idiosyncratic productivity conditional on non-enrollment z_5.

3.3 Work and Parenthood

Adults are childless in periods $t = 6$ and $t = 7$. They choose their consumption, savings, and whether to seek mental health treatment. The optimization problem in these periods is given by:

$$v_t(a_t, m_t, z_t, e, \theta) = \max_{c_t, a_{t+1}, \tau_t} \Big\{ u(c_t, \ell_t(m_t, e, \tau_t, 0)) - \xi_\tau \tau_t + \beta \min_p \mathbb{E}_p v_{t+1}\big(\Omega_{t+1}\big) \Big\} \tag{10}$$

subject to the entropy constraint (2), the borrowing constraint (7), and the budget constraint:

$$c_t + a_{t+1} + \varphi_\tau \tau_t = y_t + Ra_t - T(y_t, a_t, c_t),$$

where $y_t = y_t(m_t, z_t, e_t, \theta)$. Subjective expectations are formed over the vector of state variables in period $t + 1$, denoted by Ω_{t+1}. In period $t = 7$, the state variables are as in period $t = 6$, that is $\Omega_7 = \{a_7, m_7, z_7, e, \theta\}$.

[11] School taste shocks are introduced to account for observed variation in educational patterns.

Adults deterministically have one child at the beginning of period $t = 8$. Children are born with cognitive skills θ_{kt} and mental health m_{kt}, which are drawn from a joint distribution that depends on parental cognitive ability θ_t and parental mental health m_t. The dependence of a child's initial mental health on their parent's mental health state is motivated by the psychiatric literature on the role of genes and maternal prenatal mental health in shaping children's mental health (Section 2). Children can either be healthy ($m_{kt} = m_{k0}$) or experience mental illness ($m_{kt} = m_{k1}$). The child's cognitive skills and mental health are part of the parent's vector of state variables. That is $\Omega_8 = \{a_8, m_8, z_8, e, \theta, m_{k8}, \theta_{k8}\}$.

During periods $t = 8$ to $t = 11$, which correspond to ages 0 to 15 for the child, parents can invest time n_{kt} and financial resources x_t into developing their child's skills and mental health. They also choose the amount of consumption goods c_{kt} to purchase for their child. Children have the same utility function as adults but they do not work. Parents care about the utility of their child with an altruism factor δ.

Child Development. The development of a child's mental health and cognitive skills is modeled using a CES specification. A child's future mental health depends on their current mental health and cognitive skills, on their parent's mental health and cognitive skills, on parental time and financial investments, and on whether the child receives mental health treatment. The law of motion for child mental health is given by:

$$m_{kt+1} = \left(\alpha_{1mt}\theta_{kt}^{\rho_{mt}} + \alpha_{2mt}m_{kt}^{\rho_{mt}} + \alpha_{3mt}\theta^{\rho_{mt}} + \alpha_{4mt}m_t^{\rho_{mt}} + \alpha_{5mt}\iota_t^{\rho_{mt}}\right)^{\frac{1}{\rho_{mt}}} e^{\upsilon_{mt}+\mu(m_{kt})\tau_{kt}}, \tag{11}$$

where $\upsilon_{mt} \sim N(0, \sigma^2_{t\upsilon_m})$ is an i.i.d. mental health shock drawn from an age-dependent distribution. The efficacy of treatment for children is given by $\mu(m_{kt})$ and depends on the child's mental health. The investment aggregate ι_t is given by:

$$\iota_t = A_\iota\left(\alpha_\iota x_t^\gamma + (1-\alpha_\iota){n_{kt}}^\gamma\right)^{\frac{1}{\gamma}}. \tag{12}$$

The law of motion for child mental health is motivated by the psychiatric literature (Section 2). In line with this literature, child mental health is persistent, parental investments and parental mental health shape child mental health, and treatment improves mental health.[12]

The evolution of a child's cognitive skills follows an analogous specification to mental health, but there is no direct impact of mental health treatment. The law of motion for child cognitive skills is given by:

$$\theta_{kt+1} = \left(\alpha_{1\theta t}\theta_{kt}^{\rho_{\theta t}} + \alpha_{2\theta t}m_{kt}^{\rho_{\theta t}} + \alpha_{3\theta t}\theta^{\rho_{\theta t}} + \alpha_{4\theta t}m_t^{\rho_{\theta t}} + \alpha_{5\theta t}\iota_t^{\rho_{\theta t}}\right)^{\frac{1}{\rho_{\theta t}}} e^{\upsilon_{\theta t}}, \tag{13}$$

[12]Since child mental health is a binary variable, we define a cutoff $\underline{m}_k$ such that $m_{kt+1} = m_{k0}$ ($m_{kt+1} = m_{k1}$) if the right hand side of Equation 11 is higher (lower or equal) than $\underline{m}_k$.

where $v_{\theta t} \sim N(0, \sigma^2_{tv_\theta})$ is an independently drawn shock.

The problem of adults during parenthood ($t = 8, 9, 10, 11$) is to choose consumption c_t, consumption for the child c_{kt}, savings a_{t+1}, parental financial investments x_t, parental time investments n_{kt}, as well as whether to seek mental health treatment for themselves τ_t and for their child τ_{kt}. Their Bellman equation is given by:[13]

$$
\begin{aligned}
v_t(a_t, m_t, z_t, e, \theta, m_{kt}, \theta_{kt}) = \max \Big\{ & u\big(c_t, \ell_t(m_t, 0, \tau_t, \tau_{kt}) - \nu n_{kt}\big) - \xi_\tau \tau_t + \delta u(c_{kt}, 1 - n_\tau \tau_{kt}) - \xi^k_\tau \tau_{kt} \\
& - \xi_k m_{kt} + \beta \min_p \mathbb{E}_p v_{t+1}(a_{t+1}, m_{t+1}, z_{t+1}, e, \theta, m_{kt+1}, \theta_{kt+1}) \Big\}
\end{aligned}
$$

subject to the entropy constraint (2), the borrowing constraint (7), the childhood skill development production function, and the budget constraint:

$$
c_t + c_{kt} + a_{t+1} + x_t + \varphi_\tau \tau_t + \varphi_{k\tau} \tau_{kt} = y_t + R a_t - T(y_t, a_t, c_t).
$$

The parameter $\nu \in [0, 1]$ captures the extent to which time spent on childcare counts as leisure. Parents experience a direct utility cost ξ_k when their children experience mental illness. This parameter allows the model to match the age pattern of child mental health treatment in the data.

Inter Vivos Transfer. The final channel through which a parent can impact a child's outcomes is financial transfers. Shortly before period 12 (i.e., before the child becomes independent), parents can transfer $\hat{a} \geqslant 0$ to their child.[14] Parents transfer at most all their wealth and can borrow up to the borrowing constraint, that is, $\hat{a} \leqslant a_{12} - \underline{a}_{12}$. After receiving the transfer, the child becomes independent and decides whether to go to college. The transfer is made in a sub-period where the parent knows the state $(a_{12}, m_{12}, z_{12}, e, \theta, m_{k12}, \theta_{k12})$ but before the realization of their child's college preference shocks χ.[15] The inter vivos transfer problem is:

$$
v^{\text{iv}}(a_{12}, m_{12}, z_{12}, e, \theta, m_{k12}, \theta_{k12}) = \max_{\hat{a} \geqslant 0} \Big\{ v_{12}(a_{12} - \hat{a}, m_{12}, z_{12}, e, \theta) + \delta \min_p \mathbb{E}_p v_5(\hat{a}, m_{k12}, \theta_{k12}, e_p, \chi) \Big\}, \tag{14}
$$

subject to the entropy constraint (2) where subjective expectations are formed over the child's school taste shocks χ. It is useful to note that since school taste shocks follow an EV1 distribution, parents'

[13] The Bellman equation at $t = 11$ is slightly different in that the continuation value is $v^{\text{iv}}(\cdot)$ (see Equation 14).

[14] The non-negativity condition on transfer $\hat{a}$ rules out that parents borrow against their child's future income.

[15] After the child's mental health state realizes between period 11 and and period 12 (but before the inter vivos sub-period), it is mapped to adult mental health via a transformation that is discussed in Section 4.2. The outcome of this transformation is m_{k12}.

subjective expectations over the realization of these shock does not impact their transfer decision (see details in Appendix E.10). Parents' negative thinking can still impact the transfer decision due to subjective expectations over the evolution of their own future income. Specifically, negative thinking over their future income inflates the expected marginal utility of parents' future consumption at the time of the transfer decision. This raises the value of retaining wealth and lowers inter vivos transfers.

After making the transfer, the individual reverts to the standard consumption and savings problem without a child (10) in the periods leading up to retirement.

3.4 Retirement Stage

Individuals are retired between period $t = 17$ and period $t = 21$. They supply no labor in retirement and earn pension income $y_t^p(e, \theta)$. Their Bellman equation is given by:[16]

$$v_t(a_t, m_t, e, \theta) = \max_{c_t, a_{t+1}, \tau_t} \Big\{ u(c_t, \ell_t(m_t, 0, \tau_t, 0)) - \xi \tau_t + \beta \min_p \mathbb{E}_p v_{t+1}\big(a_{t+1}, m_{t+1}, e, \theta\big) \Big\}$$

subject to the entropy constraint (2), the borrowing constraint(7), and the budget constraint:

$$c_t + a_{t+1} + \varphi_\tau \tau_t = y^p(e, \theta) + R a_t - T(y^p(e, \theta), a_t, c_t).$$

3.5 Aggregate Production

A representative firm produces consumption goods using a Cobb-Douglas technology with a capital share α. Aggregate output is given by $Y = AK^\alpha N^{1-\alpha}$, where A is total factor productivity and K denotes capital. Aggregate labor N is a CES aggregate over the labor inputs of the different education groups, following Katz and Murphy (1992). Specifically, $N = (\varpi N_0^\Omega + (1-\varpi) N_1^\Omega)^{\frac{1}{\Omega}}$, where N_0 is the aggregate labor efficiency units of non-college graduates, and N_1 is the aggregate labor efficiency units of college graduates. Firms rent capital and hire labor in competitive factor markets. Capital depreciates at a rate δ_K. In the small open economy, the capital-labor ratio is chosen so that the marginal product of capital is equal to its marginal cost, i.e. to the interest rate R plus the depreciation rate δ_K. Education-specific wages per hourly efficiency unit, w_e, are determined in general equilibrium to clear the education-specific labor market. The government budget constraint and equilibrium definition are described in Appendix B. When we evaluate counterfactual policies in Section 5.2, government expenditures are held constant at their baseline steady-state level and income taxes adjust to balance the government budget.

[16] Since individuals die at the beginning of period $t = 21$, $v_{21}(\cdot) = 0$.

4 Model Quantification

We quantify the model using U.S. data. All monetary values are expressed in 2015 dollars. In line with the model, economic variables are measured at the per-adult basis within the household. We begin by describing the data sources we use for the model quantification. We then discuss the externally estimated parameters with a particular focus on the child development technology. Finally, we discuss the parameters that we estimate internally via simulated method of moments.

4.1 Data

We use four main data sources to quantify the model: (1) the Panel Study of Income Dynamics (PSID); (2) the Child Development Supplement (CDS) of the PSID, which records information on children of PSID respondents from birth to age 18; (3) the Transition into Adulthood Supplement (TAS) of the PSID, which records information on children of PSID respondents after the age of 18, and (4) the 1979 cohort of the National Longitudinal Survey of Youth (NLSY). Appendix C describes the data in detail.

A key feature of the PSID is that it records the mental health of respondents, which allows quantifying the relationship between adult mental health and economic outcomes such as consumption, labor supply and wealth. A key feature of the CDS and TAS is that they record measures of mental health for children of PSID respondents throughout childhood and early adulthood. By linking parents in the PSID to their children in the CDS and TAS, we can measure the relationship between parent and child mental health. The CDS also records time investments of parents in children and whether a child receives mental health treatment, which allows us to quantify how these investments vary with parent and child mental health. The TAS records young adults' expectations regarding college graduation as well as their realized college enrollment and graduation outcomes, allowing us to quantify the relationship between these educational attainment measures and adolescence mental health. A key feature of the NLSY is that it measures mothers' and children's cognitive skills and mental health across time as well as metrics of parental investments. This allows us to estimate the parameters of the child development technology in the model. Our measures for parental investments, maternal cognitive skills, and child cognitive skills in the NLSY are constructed as in Cunha, Heckman, and Schennach (2010).

We measure adult mental health, in both the PSID and TAS, using the Kessler Psychological Distress Scale (K6 Scale). The K6 scale is widely used by the epidemiological and psychiatric literature to assess the prevalence and severity of mental illness, and is a primary mental health measure used, for example, in U.S. government administered health surveys and in the World Health Organization surveys.[17] The K6 scale has been extensively validated against clinical mental health diagnoses and has been shown to consistently predict clinical diagnoses of mental disorders (Kessler et al., 2002, 2003; Furukawa, Kessler, Slade, and Andrews, 2003; Cairney et al., 2007). We classify adults into three mental health states - healthy, mild illness, and serious illness - based on the K6 scale following Kessler et al. (2008).[18] The NLSY does not record the K6 scale. In the NLSY, we measure maternal mental health using the Center for Epidemiologic Studies Depression Scale (CES-D).[19]

We measure child mental health in the CDS and in the NLSY using the Behavioral Problem Index (BPI). The BPI measures the frequency and types of behavior problems for children aged four and above (Peterson and Zill, 1986). It is based on the Achenbach Child Behavior Checklist (Achenbach and Edelbrock, 1981) but is shorter and easier to administer in an interview setting.[20] The BPI is widely used by the epidemiological literature to assess the prevalence and severity of children's behavioral problems (Korenman, Miller, and Sjaastad, 1995; Brand and Brinich, 1999; Pettit, Laird, Dodge, Bates, and Criss, 2001; McLeod and Kaiser, 2004; Aughinbaugh, Pierret, and Rothstein, 2005; McCormick et al., 2006), and is a primary behavioral health measure used in government administered health surveys (e.g., in the NLSY, the US National Health Interview Survey (NHIS), and the UK National Child Development Study (NCDS)). The Child Behavior Checklist and the Behavioral Problem Index have been extensively validated against clinical mental health diagnoses (Achenbach, 1983; Gortmaker, Walker, Weitzman, and Sobol, 1990). We classify a child as experiencing mental illness if their BPI score is below a threshold. This threshold is set such that the share of 16-17 year-olds classified as experiencing mental illness according to the BPI in the CDS aligns with the share of 18-19 year-olds classified as experiencing mental illness

[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.

[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] The CES-D scale is a widely used self-report questionnaire designed to measure depressive symptoms in the general population (Radloff, 1977). It asks respondents how often over the past week they experienced different symptoms associated with depression. The CES-D is commonly used to screen depression in epidemiological studies.

[20] The BPI includes 28 questions and is administered to the child's primary caregiver. Each question asks whether it is "often true", "sometimes true", or "not true" that the child exhibited a specific behavior in the previous three months, for example whether the child is unhappy, sad, or depressed and whether the child is too fearful or anxious. Based on these questions, an overall BPI score is computed. To obtain the BPI score, responses to each question are recoded as an indicator variable that is equal to one if the response is "often true" or "sometimes true" and is equal to zero otherwise. The overall BPI score is then the sum of these indicators.

according to the K6 scale in the TAS.

4.2 Externally Estimated Parameters

Taxes and Transfers. Taxes $T(y, a, c)$ are the sum of taxes on labor earnings, taxes on interest income and taxes on consumption. Following Feldstein (1969), Benabou (2002) and Heathcote, Storesletten, and Violante (2014), labor earnings are taxed according to a nonlinear function where after-tax labor earnings $\tilde{y}$ are a power function of pre-tax labor earnings y, that is $\tilde{y} = (1-\tau_0)y^{1-\tau_1}$. The progressivity parameter $\tau_1 = 0.185$ is taken from Heathcote, Storesletten, and Violante (2014). The scale parameter τ_0 is estimated internally to match the size of government spending relative to GDP in the data. Consumption is taxed at a linear rate $\tau_c = 0.05$ and interest income, $(R-1)a$, is taxed at a linear rate $\tau_a = 0.36$ when positive following Trabandt and Uhlig (2011). Pension income $y^p(e, \theta)$ is calibrated using the Old Age Insurance of the US Social Security System (Appendix E.3). Finally, all households receive a lump-sum transfer ω which is estimated internally to match the amount of redistribution in the data.

Time Allocation. Individuals are endowed with one unit of time every period, which corresponds to 100 hours per week. We calibrate working hours using PSID data following Abramson, Boerma, and Tsyvinski (2024). Healthy individuals who are not enrolled in college work on average 33.75 hours per week. We therefore set $n_w(m_0, 0) = 0.3375$. To determine the working hours of non-student individuals experiencing mental illness, we estimate how working hours change for a given individual as they transition between mental health states. In particular, we regress (log) working hours on individual fixed effects controlling for wealth, household composition as well as age and year fixed effects. Individuals experiencing mild mental illness work on average 2.6 percent fewer hours, while individuals experiencing serious mental illness work on average 15.8 percent fewer hours. We therefore set $n_w(m_1, 0) = 0.3287$ and $n_w(m_2, 0) = 0.2883$. In the model, individuals enrolled in college spend n_s hours on school work. We set $n_s = 0.3375 - 0.104$ to match the average working hours of college students in the data.[21] In terms of rumination, we assume that healthy individuals do not ruminate, that is $n_r(m_0) = 0$. Following Abramson, Boerma, and Tsyvinski (2024), individuals who experience mild and serious illness ruminate 1.3 and 3 hours per week, that is $n_r(m_1) = 0.013$ and $n_r(m_2) = 0.03$. We calibrate the time cost of mental health treatment for both adults and children to two hours per week, that is $n_\tau = n_\tau^k = 0.02$.

[21] According to the National Center for Education Statistics, 42 percent of full-time college students work while in college. We infer that full-time students work on average about 10.4 hours per week in 2015 (National Center for Education Statistics, 2015).

Wages. The process for hourly wages $w_t(m_t, z_t, e, \theta)$ is specified in equation (4). The hourly wage is the product of wage per hourly efficient unit, w_e, and hourly efficiency units. The wage per hourly efficiency unit is determined in equilibrium. To estimate the components of hourly efficiency units, we use PSID and NLSY data.

Using PSID data, we estimate a regression of (log) wages on a third-order age polynomial that can vary by whether the individual is a college graduate and on indicator variables for mild and serious mental illness. We control for individual fixed effects, year fixed effects, and measures of physical illness. The individual-level fixed effects absorb time invariant characteristics at the individual level (including non-cognitive skills and education levels). The estimated third-order age polynomial maps to the deterministic life-cycle component in the model, $\log \zeta_t^e$, while the estimated coefficients on the indicator variables for mild and serious mental illness capture how hourly wages vary with mental health. We estimate a decrease of 2.2 (5.5) percent in hourly wages associated with mild (serious) mental illness. We therefore set $\lambda_m(m_0) = 0$, $\lambda_m(m_1) = -0.022$ and $\lambda_m(m_2) = -0.055$.[22] We estimate the remaining two components of the wage process using NLSY79 data. First, we estimate the dependence of wages on standardized cognitive skills controlling for potential selection bias (Heckman, 1979). We find that a one-standard deviation increase in cognitive skills increases hourly wages by $\lambda_\theta^1 = 0.336$ log points for college educated workers, and by $\lambda_\theta^0 = 0.204$ log points for non-college educated workers. We finally estimate the residual wage process by education group. For college educated workers, we obtain persistence $\rho^1 = 1$, innovation standard deviation variance $\sigma^1 = 0.194$, and an initial variance of $\sigma_6^1 = 0.366$. For non-college educated workers, we obtain persistence $\rho^0 = 0.991$, innovation standard deviation variance $\sigma^0 = 0.158$, and an initial variance of $\sigma_6^0 = 0.288$.

Assets. We set $R = 1.0524^4$ to match the annual real returns on safe assets between 2001 and 2020 (Jordà, Knoll, Kuvshinov, Schularick, and Taylor, 2019). We set $R_b = (R + 0.12)^4$ to reflect the average annual credit card borrowing spread reported by Gross and Souleles (2002). Based on Daruich and Kozlowski (2020), we set $R_e = (R + 0.009)^4$ to reflect the weighted average annual interest rate on different types of Stafford loans, which are the dominant source of student loans.

Individuals can borrow up to the natural borrowing limit $\underline{a}_t$. This means they are not allowed to

[22]The effects of mental health on hourly wages align with the estimated effects of hourly wages with respect to hours worked of 0.4 in French (2005) and Bick, Blandin, and Rogerson (2022). The 2.6 and 15.8 percent reduction in working hours imply hourly wage decrease by $0.4 \times 2.6 = 1.1$ percent and $0.4 \times 15.8 = 6.3$ percent. The implied direct mental health effects on productivity are therefore small. This is in line with the psychiatric literature that finds that depression is characterized by impaired cognitive control (manifested as rumination) rather than by cognitive deficits (Hertel, 2004; Gotlib and Joormann, 2010).

borrow in the final period, i.e. $\underline{a}_{21} = 0$. In period $t = 19$, individuals are allowed to borrow up to the lump-sum government transfer, i.e. $\underline{a}_{20} = -\omega/R_b$, and generally for $t \leqslant 18$:

$$\underline{a}_{t+1} = -\frac{\omega}{R_b} \sum_{j=t}^{19} \left(\frac{1}{R_b}\right)^{19-j}. \tag{15}$$

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

$$u(c, \ell) = \log c + \psi \frac{\ell^{1-\frac{1}{\eta}} - 1}{1 - \frac{1}{\eta}}, \tag{16}$$

where $\eta \geqslant 0$ governs the curvature with respect to leisure hours, and $\psi \geqslant 0$ governs the value of leisure. Children have the same utility function over consumption and parents value their child's consumption with altruism factor δ.[23] We choose the parameter η so that the Frisch elasticity of labor supply for healthy non-student individuals (who work $n_w(m_0, 0) = 0.3375$ hours) equals 0.55, following Chetty, Guren, Manoli, and Weber (2012). To align with the Frisch elasticity of labor supply for these individuals in the model, we require $\eta = \frac{n_w(m_0,0)}{1-n_w(m_0,0)} \frac{0.55}{1-\tau_1(1+0.55)} = 0.521$, where τ_1 governs the progressivity of the tax schedule.

College. The direct flow utility from going to college is given by:

$$\Upsilon(\theta, e_p) = \chi_s \big[\alpha_0 + \alpha_1 \mathbf{1}_{\{e_p=1\}} + \alpha_\theta \log \theta\big]. \tag{17}$$

The college taste shock χ is distributed according to a Type I extreme value distribution with scale parameter χ_s. The direct utility from college is scaled by χ_s, the scale parameter of the taste shock distribution. This normalization ensures that the parameters α_0, α_1, and α_θ are estimated in units comparable to that of the taste shock. The college preference parameters, α_0, α_1, α_θ, and χ_s are estimated internally (Section 4.3).

We estimate the graduation probabilities, $p_e(\theta, m)$, using TAS data. Focusing on individuals above the age of 28 who ever enrolled in college, we estimate a linear probability model for the likelihood of graduation as a function of cognitive skills and mental health. Cognitive skills are measured using the Applied Problems raw score during childhood and standardized to have a mean-zero and unit variance. Mental health corresponds to an individual's mental health state when she first appears in the TAS (typically at age 18). Based on the estimated linear probability model, we set $p_e(\theta, m) = 0.5 + 0.236\theta$.[24]

[23] In Appendix E.2 we show that this does not introduce an additional choice variable in the numerical analysis. Consumption for parent and child can be collapsed into a single choice variable as in Lee and Seshadri (2019).

[24] Since the coefficients on the mental health indicators are economically and statistically insignificant, we calibrate graduation probabilities in the model to be independent of mental health.

Table 1: Mental Health Transition Matrix

No Treatment	Healthy	Mild	Serious	*Treatment*	Healthy	Mild	Serious
Healthy ($z < \underline{z}$)	0.894	0.078	0.028	Healthy ($z < \underline{z}$)	0.894	0.078	0.028
Healthy ($z \geqslant \underline{z}$)	0.901	0.077	0.022	Healthy ($z \geqslant \underline{z}$)	0.901	0.077	0.022
Mild	0.611	0.255	0.134	Mild	0.893	0.076	0.031
Serious	0.331	0.316	0.353	Serious	0.635	0.220	0.145

Table 1 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 four years ahead m'.

Adult Mental Health Treatment. We quantify the mental health transition matrix for adults $\Gamma_m(\tau_t, z_t)$ following Abramson, Boerma, and Tsyvinski (2024). Table 1 reports the calibrated mental health transition matrix. The first takeaway is that treatment is effective. For example, the probability to transition from serious mental illness to the healthy state is 33.1 percent without treatment, while 63.5 percent with treatment. This finding is driven by the efficacy of psychotherapy as estimated by the psychiatric literature. For our calibration, we use an average SMD of -0.7 from the meta-analysis by Ekers, Richards, and Gilbody (2008). The second takeaway is that bad labor market shocks increase the likelihood to experience mental illness in the future consistent with the unconditional transition probabilities from the PSID. For example, the likelihood to transition from the healthy state into serious (mild) illness is 2.2 (7.7) percent in normal productivity states ($z \geqslant \underline{z}$), while it is 2.8 (7.8) percent in low productivity states ($z < \underline{z}$).[25] Finally, the estimation assumes mental health treatment does not impact mental health for those who are healthy. This is motivated by the observation that in the data healthy individuals rarely seek treatment (Cronin, Forsstrom, and Papageorge, 2025).

The financial cost of treatment is set based on Cronin, Forsstrom, and Papageorge (2025) who, using data from the Medical Expenditure Panel Survey (MEPS), report an out-of-pocket expenditure on psychotherapy of 24 dollars per visit. The total expenditure, including insurer payments, is 126 dollars. Individuals thus pay $\frac{24}{126} = 0.19$ of the treatment costs out-of-pocket and 0.81 is covered by insurance. We consider an average of one visit per week per year to arrive at an annual treatment cost of $\varphi_\tau = \varphi_\tau^k = \$1,250$. When we later evaluate policies that subsidize mental health treatment in Section 5.2, some of these costs will be subsidized and funded through income taxes.

Production. The labor share α is set to 0.692 to match the ratio of employees total compensation to

[25] The cutoff $\underline{z}$ is set to the bottom quartile of the invariant idiosyncratic productivity distribution.

national income net of proprietors income between 2000 and 2020 in the BEA's national income and product accounts. We set $\Omega = 1/3$ so that the elasticity of substitution between college and non-college labor is equal to 1.5 in line with the literature (Katz and Murphy, 1992; Heckman, Lochner, and Taber, 1998; Goldin and Katz, 2008). The rate of depreciation is set $\delta_K = 0.2385$ to match the average annual depreciation rate of fixed assets in the BEA's national income and product accounts, which is 6.5 percent.

4.2.1 Child Development

We estimate the parameters of the child development technology ((11) for mental health and (13) for cognitive skills) using the framework of Cunha, Heckman, and Schennach (2010) and NLSY data. Consistent with Cunha, Heckman, and Schennach (2010) and with the biannual nature of the NLSY data, in the empirical analysis a period corresponds to two years.[26] We consider two stages of development: from age 0 to age 6 and from age 7 to age 14. We allow parameters of the development function to vary across stages and assume they are fixed across periods within the same stage. We first discuss the empirical measures of mental health we use for our estimation. We then present our empirical estimates for the biannual version of the development technology. Finally, we discuss the mapping of the biannual parameters to the model's 4-year frequency.

Mental Health Measures. Our measurements for child mental health are the BPI sub-scales for anxiety/depression, hyperactivity, and conduct disorder following (Currie and Stabile, 2007). We adopt these measurements for child mental health from age 3 on. Before age 3, our measurements for child mental health are given by difficulty and friendliness for age 0 and by compliance, insecurity, sociability, difficulty and friendliness for ages 1-2.[27] Our measurements for maternal mental health are items of the Center for Epidemiologic Studies Depression Scale (CES-D), namely the extent to which the caregiver experiences poor appetite, has trouble keeping her mind on tasks, feels depressed, experiences restless sleep, feels sad, and can not get "going".[28]

Results. Table 2 presents our estimates for the biannual development technology parameter estimates and the corresponding standard errors. A few results are worth noting. First, self-productivity of child mental health is large and increasing with age. This is illustrated by the weight on child mental health

[26] The first period corresponds to age 0, the second period corresponds to ages 1-2, the third period corresponds to ages 3-4, and so on. The eight period corresponds to ages 13-14.

[27] Our measurements before age 3 overlap with the measurement variables for non-cognitive skills in Cunha, Heckman, and Schennach (2010).

[28] Following Cunha, Heckman, and Schennach (2010), we use five controls for each measurement equation: child sex, a dummy for whether the child is born after 1987, a dummy for whether the mother was less than 20 years old at the time of the first birth, the child's age at the assessment, and a constant.

Table 2: Child Development Function

	Mental Health ($s = m$)		Cognitive Skill ($s = \theta$)	
	1st Stage	2nd Stage	1st Stage	2nd Stage
Cognitive skill (α_{1st})	0.000	0.000	0.491	0.857
	(0.028)	(0.091)	(0.027)	(0.012)
Mental health (α_{2st})	0.531	0.797	0.000	0.000
	(0.035)	(0.012)	(0.026)	(0.005)
Parent cognitive skill (α_{3st})	0.020	0.000	0.044	0.062
	(0.013)	(0.007)	(0.013)	(0.008)
Parent mental health (α_{4st})	0.348	0.166	0.297	0.040
	(0.032)	(0.015)	(0.027)	(0.011)
Investments (α_{5st})	0.102	0.037	0.168	0.041
	(0.023)	(0.008)	(0.015)	(0.007)
Complementarity (ρ_{st})	-0.357	-0.742	0.358	-1.365
	(0.189)	(0.202)	(0.119)	(0.141)
Shocks variance ($\sigma^2_{t\upsilon_s}$)	0.225	0.098	0.169	0.086
	(0.012)	(0.003)	(0.007)	(0.003)

Table 2 presents estimates for the biannual version of the child development technology for mental health and for cognitive skills.

(α_{2mt}) in the both stages of childhood. Similarly, the self-productivity of cognitive skills is large and increasing with age as shown by the weight on cognitive skill ($\alpha_{1\theta t}$) in the third and fourth columns.

Second, the estimation suggests that child cognitive skills do not affect child mental health. This is illustrated by the weight on cognitive skill (α_{1mt}) in both childhood stages. Similarly, child mental health does also not affect cognitive skill development, as illustrated by $\alpha_{2\theta t}$. In sum, there appears to be no cross-productivity between child mental health and cognitive skill.

Third, as illustrated by the estimated α_{4st}, parental mental health plays an important role in terms of both the mental health and cognitive skill development of the child. While the effect of parental mental health on cognitive skills sharply falls with age, the effect on child mental health remains pronounced. This is in line with the psychiatric literature that emphasizes the importance of parental mental health for children's mental health (Section 2). As illustrated by the estimated α_{3st}, parental cognitive skills have an increasing importance on child cognitive skill, while having negligible impact on the mental health development. The impact of parental investments on child development as captured by α_{5st} decreases with age.

Fourth, row 13 of Table 2 illustrates the estimated complementarity between inputs in the production functions. For mental health, the corresponding elasticity of substitution $1/(1-\rho_{mt})$ is below unity for both stages and decreases with age.[29] This implies that parental investments are a complement in child mental health development. For cognitive skill development, the elasticity of substitution in the first stage exceeds one, and decreases to a value below one with age.

Mapping Results to Model Parameters. The model is set at a four year frequency. To map the estimated parameters of the biannual version of the development function to the model, we iterate once on the two biannual development functions. We assume that the shocks υ_{mt} and $\upsilon_{\theta t}$ only realize in the second half of the four-year period and that investment ι_t is constant within the four-year period. Since there is no cross-productivity between child mental health and child cognitive skills (Table 2), each of the two biannual development functions is iterated over independently.[30]

Elasticity of Substitution for Parental Investments. The elasticity of substitution between parental time and financial investments is governed by γ (Equation 12). We use the estimated elasticity of substitution of 0.25 in Caucutt, Lochner, Mullins, and Park (2026) to set $\gamma = 1 - \frac{1}{0.25} = -3$.

Initial Child Mental Health and Cognitive Skills. The psychiatric literature highlights the role of biological factors in shaping a child's mental health. In the model, these factors are captured by the dependence of a child's initial mental health on her parent's mental health. In particular, we assume that the initial draws of mental health and cognitive skills for the child probabilistically depend on the mental health and cognitive skills of their parent, namely:

$$\log \theta_{k1} = \beta_{1\theta} \log \theta + \beta_{2\theta} \log m + \varepsilon_\theta \qquad \text{and} \qquad \log m_{k1} = \beta_{1m} \log \theta + \beta_{2m} \log m + \varepsilon_m, \tag{18}$$

where ε_θ and ε_m are both mean-zero error terms and independent of the parental characteristics. Appendix E.5 discusses the calibration of $\beta_{1\theta}$, $\beta_{2\theta}$, β_{1m} and β_{2m}, the two variances of the error terms, and

[29] Cunha, Heckman, and Schennach (2010) report a constant elasticity of substitution for non-cognitive skills across the two development stages. Our estimates for the development of cognitive skills align with those reported in Cunha, Heckman, and Schennach (2010).

[30] The resulting self-productivity parameters for the four-year development technologies are their biannual analog squared. The other parameters inside the CES development function are equal to their biannual analog multiplied by 1 plus the self-productivity parameter. We assume that the variance of both shocks, $\sigma^2_{t\upsilon_m}$ and $\sigma^2_{t\upsilon_\theta}$ are twice their biannual analog.

Table 3: Intergenerational Transmission of Mental Health and Cognitive Skill: Initial Condition

	Child cognitive skills θ_k	Child mental health m_k
Parental cognitive skills $\theta - \beta_1$	0.030	0.085
Parental mental health $m - \beta_2$	0.043	0.057
Residual variance σ_ε^2	0.165	0.064
Residual covariance	-0.0043	

Table 3 presents the calibrated parameters for the system of equations that governs the initial draw of child mental health and cognitive skills (Equation 18).

the covariance of the two error terms.

To calibrate the initial draw for the child, we need to parameterize four coefficients $\beta_{1\theta}$, $\beta_{2\theta}$, β_{1m} and β_{2m}, the two variances of the error terms, and the covariance of the two error terms. In Appendix E.5, we show how all these parameters are identified from the covariances between parent cognitive skills θ, child cognitive skills θ_k, parent mental health m, and child mental health m_k. These covariances are obtained as part of the estimation of the child development technology.

Table 3 presents the calibrated parameters. The main takeaway is that children born to parents with worse mental health and lower cognitive skills tend to have worse initial mental health and lower cognitive skills.

Child Mental Health Treatment. The efficacy of child mental health treatment is given by $\mu(m_{kt})$ (11). We assume that $\mu(m_{k0}) = 0$, that is treatment does not improve mental health for healthy children. This assumption is motivated by the finding that healthy individuals rarely receive treatment Cronin, Forsstrom, and Papageorge (2025). We calibrate $\mu(m_{k1})$, the efficacy of treatment for children experiencing mental illness, using estimates on the efficacy of child mental health treatment from the medical literature. The medical literature reports estimates for the efficacy of treatment in the terms of standard mean differences (SMD). For child psychotherapy, a meta-analysis by Weisz et al. (2017) reports an average SMD of 0.46.[31] We set $\mu(m_{k1})$ so that the model-implied SMD equals 0.46. Given the childhood mental health technology (11), the model-implied SMD of treatment is $\mu(m_{k1})/\sigma_{tv_m}$ and we set $\mu(m_{k1})$

[31] Psychotherapy is the main form of child mental health treatment and is recommended as the primary choice of treatment by the medical literature (see, e.g., Birmaher and Brent (2007)) and government agencies (see, e.g., www.cdc.gov and Weisz, Sandler, Durlak, and Anton (2005)).

such that the SMD equals the empirical estimate of 0.46.[32]

Initial Adult Mental Health. At the of age 16, we transform child mental health into adult mental health in a monotonic fashion. In the CDS data, 29 percent of 16 and 17 year-olds experience mental illness, while in the TAS 18 percent of adolescents experience mild mental illness and 6.5 percent experience serious illness. The contemporaneous transition matrix from binary child mental health to three adolescent mental health states is designed to be consistent with this. All healthy 16 and 17 year-olds transition to healthy adult mental health. For the 29 percent experiencing mental illness, we assume 63 percent transition into mild mental illness (which yields 18 percent mild mental illness) and 22.2 percent transition into serious illness (which yields 6.5 percent serious mental illness).

4.3 Internally Estimated Parameters

We estimate remaining model parameters so that the model matches data moments related to child mental health, treatment and consumption data for adults, intergenerational income mobility and college choice, parent investments in child development, and tax policy. We estimate 18 parameters using a simulated method of moments. We use a Sobol sequence to solve and simulate the model in a 18-dimensional hypercube in which the parameters are distributed uniformly. We present evidence on the identification of each parameter in Appendix E.11.

Discount Factor and Altruism. We calibrate the discount factor β to match average household wealth in our PSID sample, which is 166 thousand dollars. The altruism factor δ is calibrated to match the intergenerational mobility of income. Increased levels of altruism induce more transfers and investments for children, especially for high-income children. Increased investments and transfers induce higher skills and capital income, overall leading to higher intergenerational mobility of income. We target an intergenerational rank coefficient of 0.34 following Chetty, Hendren, Kline, and Saez (2014).

Child Development. The parameters of the parental investment technology (12) are informed by data moments on parental financial and time investments. We calibrate the share parameter α_ι to match the average local financial spending on children. We target an average expenditure on education of

[32]Consider all children experiencing mental illness and compare the differences in the evolution of mental health between children that do undertake treatment and children who do not, that is $\mathbb{E}[\log m_{kt+1}|\tau_k = 1] - \mathbb{E}[\log m_{kt+1}|\tau_k = 0] = \mathbb{E}[\log \vartheta_t|\tau_k = 1] - \mathbb{E}[\log \vartheta_t|\tau_k = 0] + \mathbb{E}[\upsilon_{mt}|\tau_k = 1] - \mathbb{E}[\upsilon_{mt}|\tau_k = 0]$, where the equality follows by the mental health technology (11) and $\log \vartheta_t$ denotes the deterministic component. Provided random allocation to treatment and control group in the experiment, or $\mathbb{E}[\log \vartheta_t|\tau = 1] = \mathbb{E}[\log \vartheta_t|\tau = 0]$, the standardized mean difference is given by the differences in the means of the child mental health shock scaled by the standard error of the shock.

5,758 dollars for every child.[33] Investment productivity A_ι is normalized such that the average level of log cognitive skills in the stationary equilibrium equals zero at the end of childhood, following the normalization convention in Cunha, Heckman, and Schennach (2010). We calibrate $\nu \in [0, 1]$, which captures the extent to which time spent which children counts as leisure, so that the model matches the average amount of time spent on children in the data. When $\nu = 0$, time spent with kids is enjoyable as leisure, while $\nu = 1$ makes time spent with children equivalent to time spent at work. In the TAS data, parents spend an average of 17.7 hours per week with their child.

Child Mental Health Treatment. We calibrate the stigma cost for child mental health treatment, ξ^k_τ, so that the model matches the share of children with mental illness between the ages of 12 and 15 who receive mental health treatment, which is 0.232 in the CDS data. We estimate the utility cost of having a child with mental illness, ξ^k, to match the share of children with mental illness between the ages of 8 and 11 who receive mental health treatment, which is 0.186 in the CDS data. Since this disutility is incurred by parents during each period the child experiences mental illness, ξ^k has a disproportionate effect on treatment decisions earlier in childhood.

To compute the share of children with mental illness in the model in a manner that is compatible with the data, we first transform m_k in the model into our BPI measurement variable in the data. To do so, we use the measurement equation that was obtained as part of the estimation of the child development technology and that relates m_k, the latent child mental health variable, to BPI, our measurement variable for child mental health. We simulate the noise associated with the measurement variable from its estimated distribution.[34] We use the simulated BPI in the model whenever we compare child mental health moments in the model and in the data.

Adult Mental Health Treatment. We calibrate the stigma cost of treatment for adults $\xi_{\tau,t}$, such that the model matches the share of individuals experiencing serious mental illness who receive treatment in the National Survey on Drug Use and Health (NSDUH) analyzed by Substance Abuse and Mental Health Services Administration (2023). We allow the stigma cost of treatment to decline with age. Specifically, the stigma cost at age t is given by $\xi_{\tau,t} = \xi_\tau \exp\big(-\big(\frac{t-t_0}{T_R-1-t_0}\big)^{\rho_\tau} \ln(1000)\big)$, where t_0 is the first working-age period. The parameter ρ_τ governs the speed of decay of stigma with age. We calibrate ξ_τ to match the treatment share among young adults aged 16-27, which is 0.491 in the NSDUH, and ρ_τ to match the age

[33]See Appendix E.7.

[34]Since the age variation is only relevant within the two-year period for which the measurement equation is estimated, and given that the model period is 4 years, we do not use any the regressors when simulating the measurement variable.

gradient in treatment uptake, which we measure as the difference in treatment shares between ages 28-47 and ages 16-27.

Negative Thinking. Following Abramson, Boerma, and Tsyvinski (2024), we assume healthy individuals do not exhibit negative thinking (i.e. $\kappa(m_0) = 0$) and calibrate negative thinking among individuals with mild and serious mental illness so that the model matches observed differences in consumption by mental health status. The idea is that negative thinking induces precautionary savings and thus lowers consumption. In the PSID data, individuals experiencing mild mental (serious) illness consume 3.3 (5.5) percent less than healthy individuals, controlling for wealth, income, race, household composition, educational attainment as well as year and age fixed effects. Our model closely matches these observed differences with relative entropy constraints of $\kappa(m_1) = 0.376$ and $\kappa(m_2) = 1.446$.

College. The parameter ϖ determines the weight of non-college labor in the CES aggregate of labor efficiency units, $N = \left(\varpi\, N_0^{\Omega} + (1-\varpi)\, N_1^{\Omega}\right)^{1/\Omega}$. A higher ϖ increases the relative marginal product of non-college labor, compressing the college wage premium. We set $\varpi = 0.545$ to match the observed college wage premium of 1.775 in the PSID data.

We quantify the college preference parameters using TAS data. The intercept of the direct flow utility from college (Equation 17), α_0, is calibrated to match the unconditional college enrollment share of 0.426 in the data. The parameter α_1 is calibrated to match the relationship between parent education and child likelihood of college enrollment. In particular, we regress a dummy for child college enrollment on a dummy for whether the parent is a college graduate, controlling for the child's cognitive skill and mental health. The estimated coefficient is 0.242. The estimated coefficient on (standardized) cognitive skill in the same regression, which is 0.268, informs the relationship between cognitive skill and college taste in the model, α_θ. The root mean square error of the regression, which is 0.419, determines the scale parameter, χ_s.

Taxes and Redistribution. The tax rate τ_0 is chosen so that the model matches the ratio of government expenditures to aggregate output. We target a government expenditure to output share of 0.189, which is the ratio of government purchases to aggregate output in the U.S. between 2000 and 2019 according to the BEA's national income and product accounts. We set the transfer value ω such that the ratio of the variance of disposable income to the variance of pre-tax income is 0.61 following Heathcote, Perri, and Violante (2010).

Table 4: Endogenous Parameters

Parameter	Value	Moment	Data	Model
		Preferences		
Discount factor β	0.971	Average wealth (in thousands)	166	160
Altruism δ	0.480	Rank correlation income	0.340	0.330
		Child Mental Health		
Stigma cost of treatment ξ_τ^k	0.292	Treatment share, 12−15 years	0.232	0.191
Utility cost of illness ξ^k	1.815	Treatment share, 8−11 years	0.186	0.125
		Adult Mental Health		
Negative thinking, mild $\kappa(m_1)$	0.376	$\triangle$ Consumption, mild	−0.033	−0.040
Negative thinking, serious $\kappa(m_2)$	1.446	$\triangle$ Consumption, serious	−0.055	−0.062
Stigma cost of treatment ξ_τ	1.730	Treatment share, 16−27 years	0.491	0.335
Decay of stigma with age ρ_τ	2.724	Treatment share, age growth	0.009	0.046
		College		
Baseline college taste α_0	0.765	Enrollment share	0.426	0.350
College taste, college parents α_1	−1.340	$\triangle$ Enrollment share, e_p	0.242	0.263
College taste, cognitive skill α_θ	−1.299	Enrollment share, θ	0.268	0.263
Scale college taste shock χ_s	11.838	Enrollment regression RSME	0.419	0.421
Share non-college labor ϖ	0.545	Wage premium	1.775	1.904
		Parental Investment Technology		
Parental resources share α_ι	0.187	Mean expenditures	5,758	5,297
Investment productivity A_ι	7.710	Normalized cognitive skill	0.000	−0.033
Leisure cost of parental time ν	0.565	Mean time investment	0.177	0.171
		Fiscal Policy		
Labor earnings tax level τ_0	0.201	Government spending	0.189	0.162
Transfer ω	0.086	Disposable income variation	0.61	0.59

Table 4 presents the parameter values set to match model-generated moments to their data analog. The first two columns present the parameters and their values. The third column describes a moment that informs the parameter value. The fourth and fifth column present the data moment and the model-equivalent.

4.4 Validation

We next evaluate the model's ability to replicate empirical patterns that are important for child mental health but that were not directly targeted in the estimation. We show that the model accurately predicts the relationship between parental mental health and time invested in children, the treatment share among young children, and the relationship between subjective graduation expectations and adolescent mental health.

Time Investments. The psychiatric literature emphasizes that parental time investment is a critical determinant of child mental health and that it in turn depends on parents own mental health (Section 2). Table 5 reports the relationship between parental mental health and time investments in the CDS data and in the model. In particular, we regress log time invested on a mental illness indicator which is equal to one if the parent experiences mild or serious mental illness. When no controls are included, the data yields a coefficient of -23.1. The model equivalent is -27.6, in line with the fact that parents experiencing mental illness invest substantially less time in their children. When we control for age, income, and parent fixed effects, the empirical estimate is -15.0 versus -24.3 in the model.[35] Overall, our model is in line with this important elasticity in the data.

Intergenerational Persistence of Mental Health. A central mechanism in the model is that parental mental health shapes child mental health outcomes. Panel B of Table 5 reports the intergenerational persistence of mental health. The first row measures persistence as the coefficient from regressing an indicator for child mental illness on lagged parental mental illness, controlling for parental income and age. In the CDS data, the estimated coefficient is 0.25 (standard error 0.02), indicating that children of parents experiencing mental illness are 25 percentage points more likely to experience mental health illness. The model generates an identical coefficient of 0.25. We also examine how exposure to parental mental illness during childhood affects the child's mental health in young adulthood. Specifically, we regress an indicator for the child being unhealthy at ages 16 and 20 on the share of child-rearing periods during which the parent is unhealthy. The model generates coefficients of 0.24 at age 16 and 0.20 at age 20, indicating persistent but declining effects of parental mental health exposure. These moments are not targeted in the estimation, so the match provides an independent validation of the model's intergenerational transmission mechanism.

[35] This gap could reflect measurement error since the PSID time diary captures a single day, introducing noise that attenuates within-person estimates, whereas the model's time investment lacks such noise.

College and Mental Health. The psychiatric literature emphasizes that negative thinking is a key driver of the adverse outcomes associated with mental illness (Beck, 1967a, 2002, 2008) and that children experiencing mental illness are less likely to enroll in and graduate from college (Fergusson and Woodward, 2002; Breslau, Lane, Sampson, and Kessler, 2008). Panel C of Table 5 evaluates the relationship between adolescent mental health and college outcomes in both the model and the data.

We first examine how mental health affects college enrollment and graduation. In both the model and the data, we regress a binary indicator for college enrollment (or graduation) on an indicator for mental illness. Without controls, adolescents experiencing mental illness are 13.5 percentage points less likely to enroll in college in the data, compared to 5.2 percentage points in the model. Conditional on enrollment, mental illness reduces graduation by 7.3 percentage points in the data and 2.8 percentage points in the model. Adding income and parental education as controls attenuates the enrollment coefficient to 10 percentage points in the data and 3.5 percentage points in the model, while the graduation coefficient becomes 5.4 and 2.4 percentage points, respectively.

We next evaluate the relationship between mental health and subjective expectations of college graduation. In the TAS data, college students report whether or not they expect to graduate. In the model, we compute each student's subjective graduation likelihood when they decide to enroll and classify students whose subjective graduation likelihood is greater (lower or equal) than 50 percent as (not) expecting to graduate. In both the model and data, we regress the graduation expectation indicator on the student's mental health indicator. The empirical estimate is -3.7 percentage points, while the model generates -6.7 percentage points, both indicating that mental health substantially shifts perceived graduation probabilities.

5 Quantitative Results

We use the quantitative model to evaluate the mechanisms through which mental illness affects economic outcomes and to analyze the effects of policies that expand access to mental health services.

5.1 Evaluating the Model Mechanisms

We evaluate how mental health affects individual-level choices and aggregate economic outcomes through five distinct channels: negative thinking, rumination, income, the child development technology, and the children's initial cognitive skill and mental health. In all exercises in this section we hold wages and taxes at their baseline values. This allows us to isolate the direct effect of different forces in the model from

Table 5: Validation: Non-targeted Moments

	Data	Model
A. Parental Time Investments and Mental Health		
No controls	−23.1	−27.6
	(3.7)	
With controls	−15.0	−24.3
	(5.7)	
B. Intergenerational Persistence of Mental Health		
	0.25	0.25
	(0.02)	
C. College Enrollment		
No controls	−13.5	−5.2
	(3.3)	
With controls	−10.0	−3.5
	(3.0)	
D. College Graduation		
No controls	−7.3	−2.8
	(4.2)	
With controls	−5.4	−2.4
	(4.0)	
E. College Graduation Expectations		
	−3.7	−6.7
	(1.2)	

Panel A reports the estimated coefficient from a regression of (log) parental time investment on a mental illness indicator and different sets of controls in the model and in the data. Coefficients are reported in percentage points. Panel B reports the intergenerational persistence of mental health in the model and in the data. Panel C (D) reports the estimated coefficient from a regression of a college enrollment (graduation) indicator on a mental illness indicator and different sets of controls in the model and in the data. Panel E reports the estimated coefficient from a regression of a college graduation expectation indicator on a mental illness indicator in the model and in the data. All coefficients are reported in percentage points.

equilibrium changes in prices.

Individual-Level Choices. We first decompose how adult mental health affects individual-level choices. For an individual experiencing mental illness, we evaluate the contribution of each channel by setting the parameter associated with it to its corresponding value for healthy individuals. We evaluate the consequences of each channel for parental investments, consumption labor supply choices, and welfare using a Shapley decomposition. The negative thinking channel isolates the role of negative thinking, the rumination channel isolates the role of loss of available time, the labor productivity channel isolates the role of labor productivity losses associated with mental illness, and the child development technology channel captures the direct role of parent mental health in the development of cognitive skills and mental health of children.[36][37]

Table 6 reports the results, comparing adults experiencing serious mental illness to their healthy counterparts. The average compensating equivalent variation (CEV) gain from becoming healthy amounts to 40 percent, corresponding to a one-time wealth gain of 429 thousand dollars. Negative thinking accounts for about two thirds of the welfare gains, reflecting the significant impact of negative thinking on economic choices. Absent negative thinking individuals experiencing serious illness invest 34 percent more time and 17 percent more monetary resources in their children. By increasing labor income by 24 percent, the income effect accounts for 6 percentage points of the CEV. Rumination, through increased leisure, accounts for 4 percentage points of the CEV. The child development channel accounts for 4 percentage points of the CEV. It operates through parental investments in children, increasing time investment by 45 percent and monetary investment by 27 percent. These large investment responses reflect the complementarity in the child development technology: when parent mental health improves,

[36]For the income channel, we set income to its healthy counterpart; for rumination, we set the amount of time available for leisure, child investments and mental health treatment to its healthy counterpart; for negative thinking, the relative entropy constraint $\kappa(m)$ is set to zero; for the child development channel, the parental mental health value in the child skill production function, m_t, is set to m_0; and for the initial draw channel, the parental mental health value in the initial draw stochastic function, m_t, is set to m_0. Since the channels interact, we use the Shapley value decomposition to assign each channel a contribution equal to its weighted average marginal effect across all possible orderings in which channels are activated, ensuring that contributions sum exactly to the total effect (Shapley, 1953).

[37]The standard Shapley value assigns equal probability to all orderings of channels, which can attribute substantial interaction effects to channels with negligible direct influence. We instead adopt the weighted Shapley value (Kalai and Samet, 1987), where each channel's weight equals the absolute value of its isolated marginal contribution—the effect of activating that channel alone. This reflects the view that when two channels interact, the one with the stronger direct mechanism is the primary driver. For channel-outcome pairs where the isolated effect is exactly zero—for instance, child skill development has no direct mechanism affecting intergenerational transfers—the channel is treated as part of the baseline environment rather than as a separate contributor, and its interaction effects are attributed to active channels in proportion to their weights. Because the projection is outcome-specific, a given channel may be active for some columns of the table and held fixed for others.

Table 6: Shapley Decomposition of Mental Health Channels: Severe → Healthy

	Consumption	Savings	Labor		Parental Investments		Welfare	
			Income	Hours	Time	Money	Consumption	Wealth
Baseline	32,494	161,425	43,739	28.8	9.4	3,462	—	—
Negative Thinking	8.4	−11.2	0.0	0.0	34.3	17.1	24.8	320,000
Rumination	0.1	−0.2	0.0	0.0	4.2	1.1	3.7	29,400
Income Effect	4.1	10.0	23.7	17.1	−6.7	1.0	6.1	36,300
Child Development	0.5	−1.2	0.0	0.0	44.9	27.0	3.6	24,600
Initial Draw	−0.4	0.2	0.0	0.0	0.0	0.0	1.5	18,900
Total	12.8	−2.4	23.7	17.1	76.8	46.2	39.6	429,200

Table 6 decomposes the gain from transitioning an individual experiencing serious mental illness to perfect health using the weighted Shapley value (Kalai and Samet, 1987) on a projected game. Channel weights are $\omega_i = |\Delta_i(\emptyset)|$. Channels with zero isolated marginal contribution on a given outcome are held fixed (always "on") and their interaction effects are attributed to active channels. Baseline row reports levels. All values are reported in percentages, except for wealth welfare which is reported in dollars. Time and Money Investment refer to parental investments in children. The Child Development channel sets parental mental health to its healthy value in the child skill and mental health production functions. The Initial Draw channel sets parental mental health to its healthy value in the initial distribution of child mental health and cognitive skills. Table E.2 in the Appendix presents the analogous decomposition for mild mental illness.

Table 7: Welfare Effects of Curing Child Mental Health

	Consumption	Savings	Labor		Parental Investments		Welfare	
			Income	Hours	Time	Money	Consumption	Wealth
Baseline	30,442	169,691	50,218	33.4	16.4	5,101	—	—
Effect	0.6	−0.2	0.0	0.0	0.9	−0.6	20.8	292

Table 7 reports the welfare gain from curing a child's mental health from unhealthy to healthy, evaluated from the parent's perspective during early childhood development ages. For each parent state, the table compares parental policy functions when the child has unhealthy versus healthy mental health, holding cognitive type fixed, and weights by the ergodic distribution. The Baseline row reports average levels for parents of unhealthy children: Consumption, Labor Income, and Money Investment are reported in dollars per year; Savings are reported in dollars (stock); Hours and Time Investment are reported in hours per week. The Effect row reports percentage changes. Consumption welfare is the compensating equivalent variation (CEV) in percent. Wealth welfare is the one-time transfer in thousands of dollars that equates the parent's value function across the two child health states. Labor Income and Hours are zero because child mental health does not enter the parent's wage equation.

the marginal product of financial and time investments rises, inducing parents to invest substantially more in their child. The initial draw channel, which sets the child's initial mental health and cognitive skill distribution to be the one associated with healthy parents, accounts for 1.5 percentage points of the CEV. The importance of negative thinking for welfare is consistent with the psychiatric literature emphasizing that mental illness is primarily a cognitive disorder characterized by negative thinking (Section 2).[38]

Child Mental Health. Table 7 quantifies the value parents place on curing their child's mental health during the parenting ages. The CEV gain is 20.8 percent, corresponding to a one-time wealth equivalent of 292 thousand dollars. Labor income and hours are unaffected because child mental health does not directly enter the parent's wage equation. However, curing the child increases parental consumption by 0.6 percent and time investment by 0.9 percent, but reduces savings by 0.2 percent and monetary investment by 0.6 percent. The savings reduction is consistent with a precautionary motive: parents of unhealthy children accumulate additional assets to buffer against worse expected child outcomes.

Aggregate Outcomes. We next decompose how mental health affects aggregate equilibrium outcomes. We compare two steady states: the baseline economy and a steady state in which one or more mental health channels are eliminated. In both steady states wages and taxes are held fixed at their baseline levels. We use the weighted Shapley value (Kalai and Samet, 1987) to assign contributions, with weights proportional to each channel's isolated marginal effect.

[38] The analogous decomposition for mild mental illness (Table E.2 in the Appendix) shows attenuated but qualitatively similar results, with the relative importance of channels preserved.

Table 8 presents the contribution of each channel to output, labor income, consumption, hours worked, and savings. Eliminating all mental health mechanisms increases aggregate output by 8.3 percent. The child development channel contributes the largest share. This contribution reflects the channel's operation through children's cognitive skill and mental health formation: parental mental illness reduces both the cognitive skills and mental health of children, which in turn determines adult labor productivity.

Elimination of negative thinking reduces output by 0.5 percent and savings by 18 percent. Eliminating negative thinking causes individuals experiencing mental illness to reduce precautionary savings, as they perceive future outcomes more positively. The decline in household savings translates into reduced parental monetary investments. Despite the negative output effect, eliminating negative thinking generates the largest welfare gains, as shown in Table 10: a 36 percentage point CEV gain, equivalent to a one-time wealth transfer of about 354 thousand dollars. The income effect contributes 1.4 percent to output. Rumination and the initial draw channel have small aggregate output effects but contribute 0.5 and 4.1 percentage points of the CEV respectively. Together, the five channels generate a consumption equivalent welfare gain of 71.7 percent (Table 10), corresponding to a one-time wealth gain of 696,700 dollars. This aggregate welfare gain is substantially larger than the individual-level CEV of 40 percent for two reasons. First, the aggregate comparison is between two stationary distributions: eliminating mental illness permanently alters the composition of the population across generations, compounding gains as healthier parents raise healthier children who are themselves more productive. Second, the aggregate CEV averages welfare gains across the entire population under a veil of ignorance − including individuals who are currently healthy but whose descendants would have experienced mental illness − rather than conditioning on current mental health status.

Table 9 decomposes the contribution of the mental health channels to aggregate education outcomes, parental investments, and the intergenerational persistence of mental health. Time investment in children increases by 6.5 percent if mental illness is eliminated, with the child development channel accounting for most of the aggregate effect. College enrollment rises by 5.0 percentage points, driven by the child development channel. Parental monetary investments increase by 6.2 percent. Two channels contribute negatively to parental investments. The initial draw channel produces healthier children who require less treatment. Because treatment is cheaper, parents reduce both time and monetary investments, which lowers cognitive skills by 0.6 percent and college enrollment by 0.3 percentage points. The income effect channel removes the productivity penalty associated with mental illness, raising labor income. Higher income increases the desire for both consumption and leisure. Since time is finite, parents reduce time investments in children (by 0.6 percent) to enjoy additional leisure, and monetary investments decline as well (by 0.3 percent). The intergenerational persistence of mental health, measured by a rank-rank regression of child mental health on lagged parent mental health, drops to zero when all five channels are eliminated. The child development channel accounts for nearly all of the reduction (-0.22 out of -0.25), with negative thinking contributing a modest reduction of (-0.05). The initial draw channel has no effect on intergenerational persistence, consistent with the fact that it changes the distribution of initial conditions without altering the mechanisms that transmit mental health across generations.

It is useful to distinguish between the CEV and the fixed-distribution CEV (denoted by CEV FD). The fixed-distribution CEV evaluates the counterfactual value functions at the baseline stationary distribution rather than the endogenous counterfactual distribution. Under the fixed baseline distribution, eliminating the child development channel leads to only a 4.6 consumption equivalent variation (relative to 29.7 consumption equivalent variation under the endogenous distribution). This decline indicates that the welfare gains operate primarily through compositional changes: a larger proportion of children growing up with healthy parents which increases cognitive skills and improves mental health of children. The children enter adulthood as healthier individuals, and are themselves more likely to be healthy parents, propagating the gains to subsequent generations. By contrast, eliminating negative thinking generates a CEV FD and CEV that are similar in magnitude. This indicates that eliminating negative thinking benefits individuals directly. Eliminating the initial draw channel also generates a CEV FD and CEV that are similar in magnitude. Holding the baseline distribution fixed, making children's initial mental health independent of parental mental health generates a sizable welfare gain for children born to unhealthy parents. In the new steady state, the improved composition of the population − fewer unhealthy parents − reduces the mass of individuals who benefit from this channel, partially offsetting the gain.

5.2 Mental Health Policies

We now use our model to evaluate three policies that aim to improve mental health of children and parents. In Appendix F, we provide an overview of the main public policies that are commonly proposed to address child mental health. We assume that labor income taxes adjust so that government spending in the new equilibrium is unaffected relative to the baseline (see Appendix B).

I. Subsidizing Treatment for Children. Policymakers are considering various policies to expand the take-up of mental health services among children, for example by providing school-based mental health services or by providing access to treatment through community health clinics. We assess the consequences of subsidizing mental health treatment for children. In the model, this corresponds to

Table 8: Shapley Decomposition of Aggregate Economic Effects

Channel	Output	Labor Income	Consumption	Hours	Savings
Negative Thinking	−0.5	−0.4	−2.5	−0.0	−17.7
Rumination	−0.1	−0.1	−0.0	−0.0	0.5
Income Effect	1.4	1.4	0.6	0.9	−2.8
Child Development	7.8	8.3	5.4	−0.0	4.1
Initial Draw	−0.3	−0.3	−0.1	−0.0	0.8
Total	8.3	8.8	3.4	0.9	−15.0

Table 8 presents a Shapley value decomposition of the aggregate effects from eliminating each mental health mechanism. All the effects are expressed as percentage changes from the initial steady state. The Income Effect channel removes the productivity penalty of mental illness, affecting both the wage per hour and the number of hours worked, without affecting leisure hours. Rumination changes both free hours and leisure hours. Negative Thinking eliminates the pessimistic bias in expectations about future mental health states. Child Development sets parental mental health to its healthy value in the cognitive skill and child mental health production functions. Initial Draw sets parental mental health to its healthy value in the initial distribution of child mental health and cognitive skills.

treatment being freely available for all children and paid for by the government.

II. Subsidizing Treatment for Young Parents. Policymakers recognize the importance of parents' mental health for children's mental health. For example, the U.S. Surgeon General specifically phrased parental mental health as a child healthcare issue (Murthy, 2021). Legislation targeting parents' mental health primarily focuses on maternal mental health in the perinatal period − from pregnancy through 12 months postpartum. These policies typically aim to expand the take-up of mental health services by lowering their out of pocket cost. In the model, we consider a subsidy that subsidizes adult mental health treatment during ages 24 through 31, covering the pre-birth period and the period of birth.

III. Subsidizing Treatment for Parents and Children. The third set of policies that is considered involves subsidizing treatment for both children and parents. In the model, we consider subsidizing mental health treatment for all parents between age 24 and 43 and simultaneously for all children.

Treatment. Table 11 reports treatment rates among individuals experiencing mental illness under the three policies. Subsidizing treatment for children generates a large take-up response: treatment of children experiencing mental illness increases from 11 percent to 36 percent. Subsidizing treatment for young parents increases take-up for young adults from 51 to 54 percent. Simultaneously subsidizing treatment for all parents and children increases parent take-up from 43 to 48 percent and child treatment

Table 9: Shapley Decomposition of Education, Investment, and Intergenerational Persistence

	Parental Investments			Cognitive	College	IGM
Channel	Time	Money	Transfers	Skills		
Baseline	17.1 h/wk	5,291	50,290	0.97	43.3%	0.25
Negative Thinking	1.3	0.7	−26.8	−0.0	0.6	−0.05
Rumination	0.0	−0.1	0.4	−0.1	−0.1	0.00
Income Effect	−0.6	−0.3	−10.2	−0.9	−0.4	0.01
Child Development	6.5	6.8	8.9	14.5	5.2	−0.22
Initial Draw	−0.7	−0.8	1.5	−0.6	−0.3	0.00
Sum	6.5	6.2	−26.2	12.9	5.0	−0.25

Table 9 presents a Shapley value decomposition of effects on parental time investment, money investment, transfers to adult children, cognitive skills, college enrollment (percentage points), and the intergenerational persistence of mental health (level changes). All effects as percentage changes from initial steady state, except for college which is in percentage points and IGE MH which is in level changes. The Child Development channel sets parental mental health to its healthy value in the child skill and mental health production functions. The Initial Draw channel sets parental mental health to its healthy value in the initial distribution of child mental health and cognitive skills.

rates to 36 percent.

Prevalence. Table 12 reports the impact of the different policies on the prevalence of mental illness. Subsidizing treatment for children generates a large improvement in child and adolescent mental health. The fraction of children experiencing mental illness declines by 0.7 percentage points, while mild and serious illness among adolescents declines by 2.2 and 0.8 percentage points respectively. Adult mental health also improves modestly as children enter adulthood healthier. Subsidizing treatment for young parents primarily affects adult mental health at young parenting ages, with limited spillovers to child mental health. Subsidizing treatment simultaneously for all parents and children yields the largest improvements. It reduces mental illness among children by 0.7 percentage points, mild and serious illness among parents by 0.2 and 0.1 percentage points, and mild and serious illness among adolescents by 2.3 and 0.8 percentage points.

Education and Parental Investments. Table 13 establishes a sharp contrast in terms of how the different policies impact parental investment. Subsidizing treatment for children reduces parental time and monetary investments as well as transfers. Parents reallocate resources away from child investments and

Table 10: Shapley Values: Wealth Equivalence and Welfare Effects

	Welfare		Welfare: Fixed Distribution	
Channel	Wealth	Consumption	Wealth	Consumption
Negative Thinking	353,600	35.9	330,000	33.8
Rumination	3,500	0.5	6,700	0.9
Income Effect	13,200	1.6	34,100	4.0
Child Development	290,800	29.7	42,100	4.6
Initial Draw	35,600	4.1	49,100	5.4
Sum	696,700	71.7	462,000	48.7

Table 10 presents a Shapley decomposition of welfare effects from eliminating each mental health channel. Wealth reports the wealth-equivalent welfare gain in dollars. Consumption reports the compensating equivalent variation as a percentage of baseline consumption. Fixed Distribution columns hold the stationary distribution (over savings, cognitive skills and mental health) at its baseline level.

Table 11: Effects of Treatment Subsidies on Treatment Uptake

Policy	All Adults	Young Parents	All Parents	Children
Baseline	30.4	51.2	43.3	10.6
Children	30.8	51.4	43.6	36.4
Young Parents	30.9	54.4	44.7	10.6
Young Parents and Children	32.2	54.0	48.4	36.4

Table 11 presents treatment rates among individuals experiencing mental illness for different age groups. All Adults corresponds to ages 16 and above. Parents corresponds to ages 24–43. Children refers to ages 0–12. Adolescents refers to ages 16-19.

toward their own consumption. College enrollment also declines modestly, consistent with the reduction in parental transfers. Cognitive skills decline by 0.4 percent under the child subsidy. When the treatment is free, parents reduce their own investments in children, and the resulting decline in time and monetary inputs lowers the marginal product of parental effort in the child skill production function, which more than offsets the direct mental health improvement from subsidized treatment. Subsidizing treatment for young parents produces the opposite pattern. Healthier parents face lower costs of engaging with their children as negative thinking and rumination are attenuated. As a result, they invest more. Subsidizing treatment simultaneously for all parents and children shows that the crowding-out effect of subsidizing child treatment dominates the crowding-in effect of subsidizing parental treatment.

Table 12: Effects of Treatment Subsidies on Prevalence of Mental Illness

	All Adults		Parents		Children	Adolescents		IGM
Policy	Mild	Serious	Mild	Serious	Unhealthy	Mild	Serious	
Baseline	10.4	4.7	9.3	4.2	28.9	16.6	5.9	0.25
Children	−0.18	−0.09	−0.05	−0.06	−0.66	−2.19	−0.78	−0.02
Young Parents	−0.01	−0.01	−0.04	−0.03	−0.03	−0.01	−0.00	−0.00
Young Parents and Children	−0.24	−0.13	−0.17	−0.14	−0.71	−2.27	−0.81	−0.02

Table 12 presents the distribution of mental health for different age groups and the intergenerational persistence of mental health. The baseline row reports the prevalence of mental illness in percent and IGM in levels in the initial steady state. Policy rows report steady-state changes in percentage points relative to the baseline for the prevalence of mental health, and level changes for IGM. All Adults corresponds to ages 16 and above. Parents corresponds to ages 24–43. Children refers to ages 0–12. Adolescents refers to ages 16-19.

Table 13: Effects of Treatment Subsidies on Education and Parental Investments

	Parental Investments			Cognitive	College
Policy	Time	Monetary	Transfers	Skills	
Children	−0.52	−0.77	−1.94	−0.36	−0.18
Young Parents	0.07	0.07	0.15	0.09	0.01
Young Parents and Children	−0.41	−0.66	−2.38	−0.20	−0.12

Table 13 presents the effects of policies on parental time investment, money investment, transfers to children, cognitive skills, and college enrollment. All effects are reported in terms of percentage changes relative to the baseline steady-state except for the effect on college enrollment which is reported in percentage points relative to the baseline.

Table 14: Effects of Treatment Subsidies on Aggregate Economic Outcomes

	Tax	Output	Labor Income	Consumption	Hours	Savings
Children	0.13	−0.16	−0.13	−0.32	0.02	−0.68
Young Parents	0.02	0.04	0.04	0.05	0.00	0.06
Young Parents and Children	0.20	−0.07	−0.05	−0.27	0.03	−0.75

Table 14 presents the effects of mental health policies on aggregate economic outcomes. All effects are expressed as percentage differences relative to the baseline steady state except for the effect on taxes is reported in terms of percentage points difference.

Table 15: Effects of Treatment Subsidies on Wealth Equivalence and Welfare

	Welfare		Welfare: Fixed Distribution	
Policy	Wealth	Consumption	Wealth	Consumption
Children	5,300	0.78	−1,100	−0.17
Young Parents	900	0.13	200	0.04
Young Parents and Children	7,200	1.04	−800	−0.12

Table 15 presents the welfare effects of different policies. Wealth refers to the wealth-equivalent welfare gain in dollars. Consumption refers to the compensating-equivalent variation as a percentage of baseline consumption. Fixed Distribution columns refer to cases where the distribution of individuals over the state space is the baseline steady state distribution.

Aggregate Effects. Table 14 reports the macroeconomic consequences of the different policies. Subsidizing treatment for children produces a small negative output effect, driven by both the increase in income taxes associated with the subsidy and the negative thinking, rumination, and initial draw channels (Table 8). Subsidizing treatment for young parent generates a modest output increase as healthier parents are more productive and invest more in their child's human capital. Tax changes across the different policies are small, ranging from 0.02 to 0.20 percentage points.

Welfare. Table 15 reports the welfare effects of the different policies. Subsidizing child treatment generates substantially larger gains than subsidizing parental treatment. The child subsidy generates a CEV of 0.8 percent of annual consumption and a wealth-equivalent of 5,300 dollars, compared to 0.13 percent for the young parent subsidy. Subsidizing both parents and children generates the largest welfare gains: a CEV of 1.0 percent and wealth-equivalent of 7,200 dollars. Note that subsidizing treatment for both parents and children yields a welfare gain that is roughly the sum of the welfare gains associated with subsidizing treatment for children only and subsidizing treatment for parents only. This additivity reflects separate mechanisms at work: parental subsidies improve parent mental health, reducing negative thinking and rumination and thereby increasing private parental investments in children; child subsidies directly treat children through a separate channel.

It is informative to distinguish between the overall welfare and the welfare gain holding the baseline distribution over state variables fixed. Under the baseline distribution, the welfare effects of subsidizing child treatment are negative. Evaluated at the baseline distribution over individual state variables, the fiscal cost exceeds its benefits. In other words, the entire welfare gain arises from shifts in the long-run composition of the population. By incentivizing child mental health treatment early on, the policy improves mental health of adults later on. This translates to higher adult earnings and more parental investment, which in turn improves mental health of future generations. This compositional channel is relatively muted when subsidizing treatment for young adults. The welfare effects of subsidizing treatment for young adults under the baseline distribution are fairly similar to the overall welfare effects. This indicates that parental subsidies generate a larger share of their gains by improving outcomes for individuals at their current states rather than through long-run distributional shifts.

6 Conclusion

We develop a first quantitative macroeconomic theory of child mental health. The theory incorporates the key features of child psychiatry in a life-cycle heterogeneous agent framework of child development. In line with the psychiatric literature, both biological and environmental factors shape child mental health, and parental behavior plays a key role in the intergenerational transmission of mental illness. We quantify the model using U.S. micro data on the mental health of children and parents, and use it to study the drivers of the child mental health crisis and the policies designed to address it.

References

Abbott, B. et al. (2019): "Education Policy and Intergenerational Transfers in Equilibrium," *Journal of Political Economy*, 127(6), 2569–2624.

Abramson, B., Boerma, J. and Tsyvinski, A. (2024): "Macroeconomics of Mental Health," Discussion paper, NBER Working Paper No. 32354.

Abramson, L.Y., Seligman, M.E. and Teasdale, J.D. (1978): "Learned Helplessness in Humans: Critique and Reformulation.," *Journal of Abnormal Psychology*, 87(1), 49.

Achenbach, T.M. (1983): "Manual for the Child Behavior Checklist and Revised Child Behavior Profile," *University of Vermont*.

Achenbach, T.M. and Edelbrock, C.S. (1981): "Behavioral Problems and Competencies Reported by Parents of Normal and Disturbed Children Aged Four Through Sixteen," *Monographs of the Society for Research in Child Development*, pp. 1–82.

Adamopoulou, E. et al. (2026): "The Game of Life: An Economic Model of Gaming, Employment and Family Formation," Discussion paper, University of Pennsylvania Working Paper.

Agostinelli, F. et al. (2026): "It Takes a Village: The Economics of Parenting with Neighborhood and Peer Effects," *Journal of Political Economy*, 134(1), 313–365.

Ainsworth, M.D. et al. (1978): *Patterns of Attachment: A Psychological Study of the Strange Situation*. Lawrence Erlbaum.

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

Andreassen, O.A. et al. (2023): "New Insights from the Last Decade of Research in Psychiatric Genetics: Discoveries, Challenges and Clinical Implications," *World Psychiatry*, 22(1), 4–24.

Arseneault, L. (2018): "Annual Research Review: The Persistent and Pervasive Impact of Being Bullied in Childhood and Adolescence: Implications for Policy and Practice," *Journal of Child Psychology and Psychiatry*, 59(4), 405–421.

Attanasio, O., Boneva, T. and Rauh, C. (2022): "Parental Beliefs about Returns to Different Types of Investments in School Children," *Journal of Human Resources*, 57(6), 1789–1825.

Attanasio, O., Cunha, F. and Jervis, P. (2019): "Subjective Parental Beliefs: Their Measurement and Role," Discussion paper, NBER Working Paper No. 26516.

Aughinbaugh, A., Pierret, C.R. and Rothstein, D.S. (2005): "The Impact of Family Structure Transitions on Youth Achievement: Evidence from the Children of the NLSY79," *Demography*, 42, 447–468.

Beck, A.T. (1967a): *Depression: Clinical, Experimental and Therapeutic Aspects*. New York: Harper & Row.

——— (1967b): *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.

Becker, G.S. and Tomes, N. (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87(6), 1153–1189.

——— (1986): "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 4(3, Part 2), S1–S39.

Benabou, R. (2002): "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?," *Econometrica*, 70(2), 481–517.

Bhandari, A., Borovička, J. and Ho, P. (2025): "Survey Data and Subjective Beliefs in Business Cycle Models," *Review of Economic Studies*, 92(3), 1375–1437.

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

Birmaher, B. and Brent, D. (2007): "Practice Parameter for the Assessment and Treatment of Children and Adolescents with Depressive Disorders," *Journal of the American Academy of Child & Adolescent Psychiatry*, 46(11), 1503–1526.

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

Bowlby, J. (1969): *Attachment and Loss*. Random House.

Brand, A.E. and Brinich, P.M. (1999): "Behavior Problems and Mental Health Contacts in Adopted, Foster, and Nonadopted Children," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 40(8), 1221–1229.

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.

Breslau, J. et al. (2008): "Mental Disorders and Subsequent Educational Attainment in a US National Sample," *Journal of Psychiatric Research*, 42(9), 708–716.

Breuer, J. and Freud, S. (1895): *Studien uber Hysterie*. Franz Deuticke.

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. (2020): "Longitudinal Assessment of Mental Health disorders and comorbidities across 4 decades among participants in the Dunedin birth cohort study," *JAMA network open*, 3(4), e203221–e203221.

Caucutt, E. et al. (2026): "Child Skill Production: Accounting for Parental and Market-Based Time and Goods Investments," *Journal of Political Economy*, 134(1), 150–209.

Caucutt, E.M. and Lochner, L. (2020): "Early and Late Human Capital Investments, Borrowing Constraints, and the Family," *Journal of Political Economy*, 128(3), 1065–1147.

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.

Chetty, R. et al. (2014): "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *Quarterly Journal of Economics*, 129(4), 1553–1623.

Christakis, D.A. et al. (2004): "Early Television Exposure and Subsequent Attentional Problems in Children," *Pediatrics*, 113(4), 708–713.

Cipriani, A. et al. (2016): "Comparative Efficacy and Tolerability of Antidepressants for Major Depressive Disorder in Children and Adolescents: A Network Meta-Analysis," *Lancet*, 388(10047), 881–890.

Clayborne, Z.M., Varin, M. and Colman, I. (2019): "Systematic Review and Meta-Analysis: Adolescent Depression and Long-Term Psychosocial Outcomes," *Journal of the American Academy of Child & Adolescent Psychiatry*, 58(1), 72–79.

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.

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.

College Board (2015): "Trends in College Pricing 2015," accessed via `https://research.collegeboard.org/media/pdf/trends-college-pricing-2015-full-report.pdf`.

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. (2025): "What Good are Treatment Effects Without Treatment? Mental Health and the Reluctance to Use Talk Therapy," *Review of Economic Studies*, 92(3), 1699–1737.

Cummings, E.M. and Davies, P.T. (1994): "Maternal Depression and Child Development," *Journal of Child Psychology and Psychiatry*, 35(1), 73–112.

Cunha, F. and Heckman, J. (2007): "The Technology of Skill Formation," *American Economic Review*, 97(2), 31–47.

Cunha, F. and Heckman, J.J. (2008): "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Journal of Human Resources*, 43(4), 738–782.

Cunha, F., Heckman, J.J. and Schennach, S.M. (2010): "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 78(3), 883–931.

Currie, J. (2025): "Presidential Address: Investing in Children to Address the Child Mental Health Crisis," *American Economic Review*, 115(5), 1369–1399.

Currie, J. and Stabile, M. (2006): "Child Mental Health and Human Capital Accumulation: The Case of ADHD," *Journal of health economics*, 25(6), 1094–1118.

——— (2007): "Mental Health in Childhood and Human Capital," in *The Problems of Disadvantaged Youth: An Economic Perspective*, pp. 115–148.

Danese, A. et al. (2009): "Adverse Childhood Experiences and Adult Risk Factors for Age-Related Disease: Depression, Inflammation, and Clustering of Metabolic Risk Markers," *Archives of Pediatrics & Adolescent Medicine*, 163(12), 1135–1143.

Daruich, D. (2024): "The Macroeconomic Consequences of Early Childhood Development Policies," Discussion paper, University of Southern California Working Paper.

Daruich, D. and Kozlowski, J. (2020): "Explaining Intergenerational Mobility: The Role of Fertility and Family Transfers," *Review of Economic Dynamics*, 36, 220–245.

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.

DeJong, H., Fox, E. and Stein, A. (2016): "Rumination and Postnatal Depression: A Systematic Review and a Cognitive Model," *Behaviour Research and Therapy*, 82, 38–49.

Del Boca, D., Flinn, C. and Wiswall, M. (2014): "Household Choices and Child Development," *Review of Economic Studies*, 81(1), 137–185.

Dix, T. and Meunier, L.N. (2009): "Depressive Symptoms and Parenting Competence: An Analysis of 13 Regulatory Processes," *Developmental Review*, 29(1), 45–68.

Doepke, M., Sorrenti, G. and Zilibotti, F. (2019): "The Economics of Parenting," *Annual Review of Economics*, 11(1), 55–84.

Doepke, M. and Zilibotti, F. (2017): "Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission," *Econometrica*, 85(5), 1331–1371.

Dowdney, L. (2000): "Annotation: Childhood Bereavement Following Parental Death," *Journal of Child Psychology and Psychiatry*, 41(7), 819–830.

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.

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

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.

Fearon, R.P. et al. (2010): "The Significance of Insecure Attachment and Disorganization in the Development of Children's Externalizing Behavior: A Meta-Analytic Study," *Child Development*, 81(2), 435–456.

Feldstein, M.S. (1969): "The Effects of Taxation on Risk Taking," *Journal of Political Economy*, 77(5), 755–764.

Fergusson, D.M. and Woodward, L.J. (2002): "Mental Health, Educational, and Social Role Outcomes of Adolescents with Depression," *Archives of general psychiatry*, 59(3), 225–231.

Fernandez, R. and Rogerson, R. (1996): "Income Distribution, Communities, and the Quality of Public Education," *Quarterly Journal of Economics*, 111(1), 135–164.

——— (1998): "Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education-Finance Reform," *American Economic Review*, pp. 813–833.

Field, T. et al. (1990): "Behavior-State Matching and Synchrony in Mother-Infant Interactions of Nondepressed versus Depressed Dyads," *Developmental Psychology*, 26(1), 7.

Fowler, P.J. et al. (2009): "Community Violence: A Meta-Analysis on the Effect of Exposure and Mental Health Outcomes of Children and Adolescents," *Development and Psychopathology*, 21(1), 227–259.

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.

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

Goldin, C.D. and Katz, L.F. (2008): *The Race between Education and Technology*. Harvard University Press.

Goodman, A., Joyce, R. and Smith, J.P. (2011): "The Long Shadow Cast by Childhood Physical and Mental Problems on Adult Life," *Proceedings of the National Academy of Sciences*, 108(15), 6032–6037.

Gortmaker, S.L. et al. (1990): "Chronic Conditions, Socioeconomic Risks, and Behavioral Problems in Children and Adolescents," *Pediatrics*, 85(3), 267–276.

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

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

Groh, A.M. et al. (2012): "The Significance of Insecure and Disorganized Attachment for Children's Internalizing Symptoms: A Meta-Analytic Study," *Child Development*, 83(2), 591–610.

Gross, D.B. and Souleles, N.S. (2002): "Do Liquidity Constraints and Interest Rates Matter for Consumer Behavior? Evidence from Credit Card Data," *Quarterly Journal of Economics*, 117(1), 149–185.

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

Grotzinger, A.D. et al. (2025): "Mapping the Genetic Landscape Across 14 Psychiatric Disorders," *Nature*, pp. 1–15.

Gu, S. and Zhang, L. (2026): "Educational Competition, Mental Health, and the Allocation of Human Capital: A Macroeconomic Analysis," Discussion paper, Hong Kong University Working Paper.

Gulliver, A., Griffiths, K.M. and Christensen, H. (2010): "Perceived Barriers and Facilitators to Mental Health Help-Seeking in Young People: A Systematic Review," *BMC Psychiatry*, 10(1), 113.

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.

Heathcote, J., Perri, F. and Violante, G.L. (2010): "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967–2006," *Review of Economic Dynamics*, 13, 15–51.

Heathcote, J., Storesletten, K. and Violante, G.L. (2014): "Consumption and Labor Supply with Partial Insurance: An Analytical Framework," *American Economic Review*, 104(7), 2075–2126.

Heckman, J.J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, pp. 153–161.

Heckman, J.J., Lochner, L. and Taber, C. (1998): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1(1), 1–58.

Heckman, J.J., Lochner, L.J. and Todd, P.E. (2006): *Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond.*pp. 307–458, Handbook of the Economics of Education.

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.

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.

James, W. (1890): *The Principles of Psychology.* Henry Holt and Company.

Jensen, M.F. and Zhang, N. (2026): "Effects of Parental Death on Labor Market Outcomes," *American Economic Review*, 116(5), 1811–36.

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.

Kalai, E. and Samet, D. (1987): "On Weighted Shapley Values," *International Journal of Game Theory*, 16(3), 205–222.

Katz, L. and Murphy, K.M. (1992): "Changes in Relative Wages, 1963–1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 107(1), 35–78.

Kessler, R.C. et al. (2005): "Lifetime Prevalence and Age-Of-Onset Distributions of DSM-IV disorders in the National Comorbidity Survey Replication," *Archives of General Psychiatry*, 62(6), 593–602.

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.

Kim-Cohen, J. et al. (2003): "Prior Juvenile Diagnoses in Adults with Mental Disorder: Developmental Follow-Back of a Prospective-Longitudinal Cohort," *Archives of general psychiatry*, 60(7), 709–717.

Kinsler, J. and Pavan, R. (2021): "Local Distortions in Parental Beliefs over Child Skill," *Journal of Political Economy*, 129(1), 81–100.

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.

Korenman, S., Miller, J.E. and Sjaastad, J.E. (1995): "Long-term Poverty and Child Development in the United States: Results from the NLSY," *Children and Youth Services Review*, 17(1-2), 127–155.

Krueger, D. and Ludwig, A. (2013): "Optimal Progressive Labor Income Taxation and Education Subsidies when Education Decisions and Intergenerational Transfers are Endogenous," *American Economic Review*, 103(3), 496–501.

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

Lee, S.Y. and Seshadri, A. (2019): "On the Intergenerational Transmission of Economic Status," *Journal of Political Economy*, 127(2), 855–921.

Lovejoy, M.C. et al. (2000): "Maternal Depression and Parenting Behavior: A Meta-Analytic Review," *Clinical Psychology Review*, 20(5), 561–592.

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.

Madigan, S. et al. (2018): "A Meta-Analysis of Maternal Prenatal Depression and Anxiety on Child Socioemotional Development," *Journal of the American Academy of Child & Adolescent Psychiatry*, 57(9), 645–657.

McCormick, M.C. et al. (2006): "Early Intervention in Low Birth Weight Premature Infants: Results at 18 Years of Age for the Infant Health and Development Program," *Pediatrics*, 117(3), 771–780.

McLeod, J.D. and Kaiser, K. (2004): "Childhood Emotional and Behavioral Problems and Educational Attainment," *American Sociological Review*, 69(5), 636–658.

Merikangas, K.R. et al. (2010a): "Prevalence and Treatment of Mental Disorders Among US Children in the 2001−2004 NHANES," *Pediatrics*, 125(1), 75–81.

Merikangas, K.R. et al. (2010b): "Lifetime Prevalence of Mental Disorders in US Adolescents: Results From the National Comorbidity Survey Replication − Adolescent Supplement (NCS-A)," *Journal of the American Academy of Child & Adolescent Psychiatry*, 49(10), 980–989.

Merikangas, K.R. et al. (2011): "Service Utilization for Lifetime Mental Disorders in US Adolescents: Results of the National Comorbidity Survey−Adolescent Supplement (NCS−A)," *Journal of the American Academy of Child & Adolescent Psychiatry*, 50(1), 32–45.

Moore, S.E. et al. (2017): "Consequences of Bullying Victimization in Childhood and Adolescence: A Systematic Review and Meta-Analysis," *World Journal of Psychiatry*, 7(1), 60.

Murthy, V. (2021): "Protecting Youth Mental Health: The U.S. Surgeon General's Advisory," US Surgeon General's Office. https://www.hhs.gov/sites/default/files/surgeon-general-youth-mental-health-advisory.pdf (accessed March 23, 2026).

National Center for Education Statistics (2015): "Digest of Education Statistics," accessed via https://nces.ed.gov/programs/digest/.

——— (2017): "Digest of Education Statistics," accessed via https://nces.ed.gov/programs/digest/.

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.

Oberlander, T.F. et al. (2008): "Prenatal Exposure to Maternal Depression, Neonatal Methylation of Human Glucocorticoid Receptor Gene (NR3C1) and Infant Cortisol Stress Responses," *Epigenetics*, 3(2), 97–106.

O'Connor, T.G. et al. (2002): "Maternal Antenatal Anxiety and Children's Behavioural/Emotional Problems at 4 Years: Report from the Avon Longitudinal Study of Parents and Children," *British Journal of Psychiatry*, 180(6), 502–508.

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.

Pavuluri, M.N., Luk, S.L. and McGee, R. (1996): "Help-Seeking for Behavior Problems by Parents of Preschool Children: A Community Study," *Journal of the American Academy of Child & Adolescent Psychiatry*, 35(2), 215–222.

Peterson, J.L. and Zill, N. (1986): "Marital Disruption, Parent-Child Relationships, and Behavior Problems in Children," *Journal of Marriage and the Family*, 48(2), 295–307.

Pettit, G.S. et al. (2001): "Antecedents and Behavior-Problem Outcomes of Parental Monitoring and Psychological Control in Early Adolescence," *Child Development*, 72(2), 583–598.

Pinquart, M. (2017): "Associations of Parenting Dimensions and Styles with Internalizing Symptoms in Children and Adolescents: A Meta-Analysis," *Marriage & Family Review*, 53(7), 613–640.

Polderman, T.J. et al. (2015): "Meta-Analysis of the Heritability of Human Traits Based on Fifty Years of Twin Studies," *Nature Genetics*, 47(7), 702–709.

Radloff, L.S. (1977): "The CES-D Scale: A Self-Report Depression Scale for Research in the General Population," *Applied psychological measurement*, 1(3), 385–401.

Rajyaguru, P. et al. (2021): "Maternal and Paternal Depression and Child Mental Health Trajectories: Evidence from the Avon Longitudinal Study of Parents and Children," *British Journal of Psychiatry Open*, 7(5), e166.

Restuccia, D. and Urrutia, C. (2004): "Intergenerational Persistence of Earnings: The Role of Early and College Education," *American Economic Review*, 94(5), 1354–1378.

Rice, F. et al. (2010): "The Links Between Prenatal Stress and Offspring Development and Psychopathology: Disentangling Environmental and Inherited Influences," *Psychological Medicine*, 40(2), 335–345.

Rogers, A. et al. (2020): "Association Between Maternal Perinatal Depression and Anxiety and Child and Adolescent Development: A Meta-Analysis," *JAMA Pediatrics*, 174(11), 1082–1092.

Sappenfield, O. et al. (2023): "Adolescent Mental and Behavioral Health, 2023," Health Resources and Services Administration. https://www.ncbi.nlm.nih.gov/books/NBK608531/ (accessed March 23, 2026).

Seckl, J.R. and Meaney, M.J. (2004): "Glucocorticoid Programming," *Annals of the New York Academy of Sciences*, 1032(1), 63–84.

Shapley, L.S. (1953): "A Value for n-Person Games," II(28), 307–317.

Sherman, J.D. et al. (2003): "School District Expenditures for Elementary and Secondary Education: 1997–98," Statistical Analysis Report NCES 2004–311, National Center for Education Statistics.

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.

Spruit, A. et al. (2020): "The Relation Between Attachment and Depression in Children and Adolescents: A Multilevel Meta-Analysis," *Clinical Child and Family Psychology Review*, 23(1), 54–69.

Stein, A. et al. (2012): "Maternal Cognitions and Mother−Infant Interaction in Postnatal Depression and Generalized Anxiety Disorder," *Journal of Abnormal Psychology*, 121(4), 795.

Substance Abuse and Mental Health Services Administration (2023): "Key Substance Use and Mental Health Indicators in the United States: Results from the 2022 National Survey on Drug Use and Health," Annual report, Center for Behavioral Health Statistics and Quality, SAMHSA, Rockville, MD, Among adults with any mental illness (AMI), treatment receipt by age: 49.1% (ages 18–25), 50.0% (ages 26–49), 52.7% (ages 50+). For serious mental illness (SMI): 61.4% (ages 18–25), 67.4% (ages 26–49), 71.0% (ages 50+). Older adults report lower stigma as a barrier to treatment (Chapter 7, Tables 7.1–7.3).

Sullivan, P.F., Daly, M.J. and O'Donovan, M. (2012): "Genetic Architectures of Psychiatric Disorders: The Emerging Picture and Its Implications," *Nature Reviews Genetics*, 13(8), 537–551.

Sullivan, P.F. and Geschwind, D.H. (2019): "Defining the Genetic, Genomic, Cellular, and Diagnostic Architectures of Psychiatric Disorders," *Cell*, 177(1), 162–183.

Swing, E.L. et al. (2010): "Television and Video Game Exposure and the Development of Attention Problems," *Pediatrics*, 126(2), 214–221.

Talge, N.M., Neal, C. and Glover, V. (2007): "Antenatal Maternal Stress and Long-Term Effects on Child Neurodevelopment: How and Why?," *Journal of Child Psychology and Psychiatry*, 48(3-4), 245–261.

Teixeira, J.M., Fisk, N.M. and Glover, V. (1999): "Association Between Maternal Anxiety in Pregnancy and Increased Uterine Artery Resistance Index: Cohort Based Study," *Obstetrical & Gynecological Survey*, 54(9), 545–546.

Todd, P.E. and Wolpin, K.I. (2003): "On the Specification and Estimation of the Production Function for Cognitive Achievement," *Economic Journal*, 113(485), F3–F33.

Trabandt, M. and Uhlig, H. (2011): "The Laffer Curve Revisited," *Journal of Monetary Economics*, 58(4), 305–327.

Visscher, P.M. et al. (2021): "Discovery and Implications of Polygenicity of Common Diseases," *Science*, 373(6562), 1468–1473.

Weissman, M.M. and Paykel, E.S. (1974): *The Depressed Woman: A Study of Social Relationships.* University of Chicago Press.

Weissman, M.M. et al. (1999): "Depressed Adolescents Grown Up," *Jama*, 281(18), 1707–1713.

Weisz, J.R. et al. (2017): "What Five Decades of Research Tells Us About the Effects of Youth Psychological Therapy: A Multilevel Meta-Analysis and Implications for Science and Practice," *American Psychologist*, 72(2), 79.

Weisz, J.R. et al. (2005): "Promoting and Protecting Youth Mental Health Through Evidence-Based Prevention and Treatment," *American Psychologist*, 60(6), 628–648.

Widom, C.S., DuMont, K. and Czaja, S.J. (2007): "A Prospective Investigation of Major Depressive Disorder and Comorbidity in Abused and Neglected Children Grown Up," *Archives of General Psychiatry*, 64(1), 49–56.

Wilson, S. and Durbin, C.E. (2010): "Effects of Paternal Depression on Fathers' Parenting Behaviors: A Meta-Analytic Review," *Clinical Psychology Review*, 30(2), 167–180.

Yum, M. (2023): “Parental Time Investment and Intergenerational Mobility,” *International Economic Review*, 64(1), 187–223.

Zheng, A. and Graham, J. (2022): “Public Education Inequality and Intergenerational Mobility,” *American Economic Journal: Macroeconomics*, 14(3), 250–282.

The Macroeconomics of Intergenerational Mental Health Dynamics

Online Appendix

Boaz Abramson, Job Boerma, Diego Daruich, and Aleh Tsyvinski

May 2026

A Negative Thinking Problem

In this appendix, we first characterize the negative thinking problem. Second, we show how to use of the endogenous grid method in dynamic problems with negative thinking.

A.1 Solving the Negative Thinking Problem

The negative thinking problem is to choose a subjective probability distribution p to minimize (1) subject to the relative entropy constraint (2). The Lagrangian for the negative thinking problem is written as:

$$\sum p(\chi)w(\chi) - \frac{1}{\lambda}\Big(\kappa(m) - \sum p(\chi)\log p(\chi) + \sum p(\chi)\log q(\chi)\Big) - \frac{\phi}{\lambda}\Big(1 - \sum p(\chi)\Big).$$

The optimality conditions are given by:

$$\log p(\chi) = \log q(\chi) - \lambda w(\chi) - \phi - 1 \qquad \implies \qquad p(\chi) = q(\chi)\exp\big(-\lambda w(\chi) - \phi - 1\big),$$

which shows that states χ with lower values are assigned increased subjective probabilities. By integration, $1 = \sum q(\chi)\exp(-\lambda w(\chi) - \phi - 1)$. Dividing the optimality condition by this constraint:

$$p(\chi) = \frac{q(\chi)\exp(-\lambda w(\chi))}{\sum q(\chi)\exp(-\lambda w(\chi))}. \tag{3}$$

We next show that there is a monotonic relation between the entropy budget κ and the inverse of the Lagrange multiplier λ. In order to do so, we evaluate the relative entropy constraint using the subjective probabilities (3):

$$\kappa(\lambda) = \sum p(\chi)\log\frac{p(\chi)}{q(\chi)} = -\lambda\sum\frac{q(\chi)\exp(-\lambda w(\chi))}{\sum q(\chi)\exp(-\lambda w(\chi))}w(\chi) - \log\sum q(\chi)\exp(-\lambda w(\chi)),$$

We differentiate this expression with respect to the inverse of the Lagrange multiplier λ to show that the derivative of the relative entropy with respect to the inverse of Lagrange multiplier is equal to the variance of the values under the subjective probabilities:

$$\kappa'(\lambda) = \lambda \sum \frac{q(\chi)\exp(-\lambda w(\chi))}{\sum q(\chi)\exp(-\lambda w(\chi))} w(\chi)^2 - \lambda \frac{\left(\sum q(\chi)\exp(-\lambda w(\chi)) w(\chi)\right)^2}{\left(\sum q(\chi)\exp(-\lambda w(\chi))\right)^2}$$
$$= \lambda \sum p(\chi) w(\chi)^2 - \lambda \left(\sum p(\chi) w(\chi)\right)^2 = \lambda \mathrm{Var}(w(\chi)).$$

Since the variance of values $w(\chi)$ under the subjective probabilities is greater than zero, relative entropy budget κ is indeed monotonically increasing with respect to λ.

A.2 Using the Endogenous Grid Method

We illustrate how we use the endogenous grid method in our framework for the consumption and savings problem as well as the inter vivos transfer problem.

Consumption and Savings Problem. In order to illustrate how we use the endogenous grid method for the consumption and savings problem, we consider a two period problem, where idiosyncratic productivity in the final period is stochastic. The problem is written as:

$$v_0(a_0) = \max_{c_0, a_1} u(c_0) + \beta \min_p \sum p_i v_1(a_1, z_i)$$

subject to the constraint that $c_0 + a_1 \leqslant a_0$. By the analysis of the negative thinking problem in Section A.2, the solution to the negative thinking problem for some choice of savings a_1 is:

$$p(a_1, z_i) = \frac{q_i \exp(-v_1(a_1, z_i)\lambda(a_1))}{\sum q_i \exp(-v_1(a_1, z_i)\lambda(a_1))} .$$

The key observation is that the multiplier depends on the choice variable, that is, $\lambda(a_1)$. We substitute the subjective probabilities into the entropy constraint to find $\lambda(a_1)$ such that the relative entropy constraint binds:

$$\kappa = \sum p(a_1, z_i)\left(-\lambda(a_1) v_1(a_1, z_i)\right) - \log\Big(\sum q_i \exp\big(-\lambda(a_1) v_1(a_1, z_i)\big)\Big). \tag{A.1}$$

The optimality condition with respect to savings a_1 is:

$$u_c(a_0 - a_1) = \beta \sum \left(p_1(a_1, z_i) v_1(a_1, z_i) + p(a_1, z_i) v_1'(a_1, z_i)\right)$$

The key difference compared to a standard Euler equation is that the individual takes into account that the savings choice affects the subjective probabilities. To use the endogenous grid method, we thus need to characterize the derivative of the subjective probabilities for each a_1.

We characterize the derivative of the subjective probabilities given $\lambda(a_1)$. Differentiating subjective probabilities with respect to asset choice a_1 gives:

$$\begin{aligned}p_1(a_1,z_i) &= p(a_1,z_i)\left[\sum p(a_1,z_j)\big(v_1'(a_1,z_j)\lambda(a_1)+v_1(a_1,z_j)\lambda'(a_1)\big)-v_1'(a_1,z_i)\lambda(a_1)-v_1(a_1,z_i)\lambda'(a_1)\right]\\ &= p(a_1,z_i)\left[\lambda(a_1)\sum p(a_1,z_j)\big(v_1'(a_1,z_j)-v_1'(a_1,z_i)\big)+\lambda'(a_1)\sum p(a_1,z_j)\big(v_1(a_1,z_j)-v_1(a_1,z_i)\big)\right]\end{aligned}$$

From the perspective of our numerical analysis, the derivative of the inverse Lagrange multiplier is a new unknown.

We characterize the derivative of the Lagrange multiplier analytically to avoid having to numerically differentiate the function $\lambda(a_1)$. To do so, we first write the entropy constraint as:

$$-\frac{\kappa}{\lambda(a_1)}-\frac{1}{\lambda(a_1)}\log\Big(\sum q_i\exp\big(-\lambda(a_1)v_1(a_1,z_i)\big)\Big)=\sum p(a_1,z_i)v_1(a_1,z_i). \tag{A.2}$$

Differentiating the left-hand side with respect to a_1 yields:

$$\frac{\lambda'(a_1)}{\lambda(a_1)^2}\left[\kappa+\log\big(\sum q_i\exp\big(-\lambda(a_1)v_1(a_1,z_i)\big)\big)\right]+\frac{\lambda'(a_1)}{\lambda(a_1)}\sum p(a_1,z_i)v_1(a_1,z_i)+\sum p(a_1,z_i)v_1'(a_1,z_i)$$

and differentiating the right-hand side gives:

$$\sum\big(p_1(a_1,z_i)v_1(a_1,z_i)+p(a_1,z_i)v_1'(a_1,z_i)\big).$$

Using the expression for the derivative of the subjective probabilities $p_1(a_1,z_i)$, we make further analytical progress. We note that we only need to do this for first term on the right hand side $\sum p_1(a_1,z_i)v_1(a_1,z_i)$ since the second term cancels:

$$\begin{aligned}&\sum_i p(a_1,z_i)v_1(a_1,z_i)\left[\lambda(a_1)\sum_j p(a_1,z_j)\big(v_1'(a_1,z_j)-v_1'(a_1,z_i)\big)+\lambda'(a_1)\sum_j p(a_1,z_j)\big(v_1(a_1,z_j)-v_1(a_1,z_i)\big)\right]\\ =&\lambda(a_1)\sum_{i,j}v_1(a_1,z_i)p_ip_j\big(v_1'(a_1,z_j)-v_1'(a_1,z_i)\big)+\lambda'(a_1)\sum_{i,j}v_1(a_1,z_i)p_ip_j\big(v_1(a_1,z_j)-v_1(a_1,z_i)\big),\end{aligned}$$

which is to be equated against the remainder of the left-hand side:

$$\frac{\lambda'(a_1)}{\lambda(a_1)^2}\left[\kappa+\log\big(\sum q_i\exp\big(-\lambda(a_1)v_1(a_1,z_i)\big)\big)\right]+\frac{\lambda'(a_1)}{\lambda(a_1)}\sum p(a_1,z_i)v_1(a_1,z_i)$$

to identify $\lambda'(a_1)$. The last two expressions are linear in the derivative of the inverse Lagrange multiplier $\lambda'(a_1)$ and therefore can be used to describe $\lambda'(a_1)$ for any asset level a_1 in closed-form given the set of subjective probabilities $p(a_1,z_i)$ and $\lambda(a_1)$. The characterization of the derivative of the inverse Lagrange multiplier $\lambda'(a_1)$ in turn characterizes the derivative of the subjective probabilities with respect to assets $p_1(a_1,z_i)$, which is what we wanted.

Intervivos Transfer Problem. We next show how we use the endogenous grid method in combination for the intervivos transfer problem with negative thinking. In the intervivos transfer problem, the parent faces uncertainty with respect to both the college taste χ and the mental health of their child becoming independent $\hat{m}$. By the analysis in Section A.2, the negative thinking problem is solved as:

$$p(\hat{a}_t, \chi_i, m_j) = \frac{q_{ij} \exp(-v_1(\hat{a}_t, \chi_i, m_j)\lambda(\hat{a}_t))}{\sum q_{ij} \exp(-v_1(\hat{a}_t, \chi_i, m_j)\lambda(\hat{a}_t))}, \tag{A.3}$$

where we emphasize the dependence of the choice variable $\hat{a}_t$ and on the sources of uncertainty χ and $\hat{m}$. The first-order optimality condition for the intervivos transfer $\hat{a}$ is:

$$v_{t1}(a_t - \hat{a}_t, m_t, z_t, e, \theta) = \delta \sum p(\hat{a}_t, \chi_i) v_{5,1}(\hat{a}_t, m_{kt}, \theta_{kt}, \chi_i) + \delta \sum p_1(\hat{a}_t, \chi_i) v_5(\hat{a}_t, m_{kt}, \theta_{kt}, \chi_i) + \underline{\lambda}_{\hat{a}} - \bar{\lambda}_{\hat{a}},$$

where $\underline{\lambda}_{\hat{a}} \geqslant 0$ is the Lagrange multiplier on the non-negativity constraint for the transfer and $\bar{\lambda}_{\hat{a}} \geqslant 0$ is the Lagrange multiplier on the maximum transfer amount. In this case, we have to characterize $p_1(\hat{a}_t, \chi_i)$ also. This characterization follows the same steps as for the consumption-savings problem above.

A.3 Discretizing the Outcomes

While the child development shocks υ are continuously distributed, our computational implementation discretizes the outcomes of child mental health and cognitive ability onto a finite grid. In our numerical approach, we solve the negative thinking problem over discretized outcomes rather than the underlying continuous shock. In this appendix, we show that these formulations are equivalent.

Suppose child mental health is discretized onto M grid points $\{m_1, \ldots, m_M\}$ whereas cognitive skill is discretized onto N grid points $\{\theta_1, \ldots, \theta_N\}$ with thresholds $-\infty < \upsilon_1^m < \cdots < \upsilon_M^m = \infty$ and $-\infty < \upsilon_1^\theta < \cdots < \upsilon_N^\theta = \infty$ such that mental health equals m_i if $\upsilon_m \in (\upsilon_{i-1}^m, \upsilon_i^m]$ and such that cognitive skill equals θ_j if $\upsilon_\theta \in (\upsilon_{j-1}^\theta, \upsilon_j^\theta]$. This produces $M \times N$ outcome states (m_i, θ_j) with objective probabilities $q_{ij} = \mathbb{P}(\upsilon \in R_{ij})$ with rectangle $R_{ij} = (\upsilon_{i-1}^m, \upsilon_i^m] \times (\upsilon_{j-1}^\theta, \upsilon_j^\theta]$.

Negative Thinking over Continuous Shocks. The individual subjective probabilities are the solution to the negative thinkign problem. The negative thinking problem over continuous shock υ is given by:

$$\min_p \ \mathbb{E}_p[w(\upsilon)]$$

subject to the entropy constraint $\mathcal{R}(p\|q) \leqslant \kappa(m)$. Negative thinking is given by $p(\upsilon) \propto q(\upsilon)\exp(-\lambda w(\upsilon))$. Since $w(\upsilon)$ is constant within each region R_{ij}, this reduces to:

$$\frac{p(\upsilon)}{q(\upsilon)} = \frac{\exp(-\lambda w_{ij})}{Z(\lambda)}$$

where $Z(\lambda) = \sum q_{ij} \exp(-\lambda w_{ij})$. Integrating over each region, we obtain $p_{ij} = q_{ij} \exp(-\lambda w_{ij})/Z(\lambda)$.

Negative Thinking over Discrete Outcomes. We next show that we obtain the same solution by directly analyzing the discrete problem. The negative thinking problem over discretized outcomes is:

$$\min_p \sum p_{ij} w_{ij}$$

subject to the entropy constraint $\sum p_{ij} \log \frac{p_{ij}}{q_{ij}} \leqslant \kappa(m)$. The subjective probabilities are given by (3) as $p_{ij} = q_{ij} \exp(-\lambda w_{ij})/Z(\lambda)$, showing that the subjective probabilities are indeed identical.

Moreover, the entropy in the problem with continuous shocks is equal to the entropy in the problem with discrete shocks:

$$\mathcal{R}(p\|q) = \int p(\upsilon) \log \frac{p(\upsilon)}{q(\upsilon)} d\upsilon = \sum_{ij} \int_{R_{ij}} p(\upsilon) \log \frac{p(\upsilon)}{q(\upsilon)} d\upsilon = \sum p_{ij} \log \frac{p_{ij}}{q_{ij}},$$

where the third equality uses that $p(\upsilon)/q(\upsilon)$ is constant in each region. The extent of negative thinking $\kappa(m)$ therefore implies the same restriction in both formulations, the Lagrange multiplier takes the same value. The two negative thinking problems are equivalent. The equivalence holds because the value w is constant within each region $-$ it depends on υ only through the discrete outcome that is realized.

B Equilibrium Definition

We focus on a stationary equilibrium. Let Ω_t denote the set of possible state of individuals of age t. Let Θ_t denote the distribution over the state space at age t.

Definition 1. *Stationary Recursive Competitive Equilibrium.* A stationary recursive competitive equilibrium is (i) a collection of policy functions for college enrollment s, consumption c, treatment τ, savings a, consumption c_k and mental health treatment τ_k for children, parental time n_k and financial investments x, and inter vivos transfers $\hat{a}$; (ii) a collection of value functions $(v_1, ..., v_{21}, v^{\text{iv}})$; (iii) aggregate capital and labor efficiency units (K, N_0, N_1); (iv) wages (w_0, w_1); (v) a tax code T, pension income y_t^p, mental health treatment subsidies $\{g_{\tau t}, g_{\tau k}\}$ and other government expenditure G; ; (vi) net exports NX; and (vii) probability distributions Θ_t such that:

1. Given wages, the set of policy functions and value functions solve the individual Bellman equations in Section 3.1.
2. Given wages, aggregate capital and labor efficiency units solve the representative firm optimization problem.

3. Labor market for each education level clears. That is, for each education group the aggregate labor efficiency units supplied equals the aggregate labor efficiency units demanded. For non-college labor, the labor market clearing condition reads as:

$$N_0 = \frac{1}{w_0}\left[\sum_{t \geqslant 6} \int_{\Omega_t^i} y_t^i \mathbf{1}\{e^i = 0\} d\Theta_t + \sum_{t=5} \int_{\Omega_t^i} y_t^i d\Theta_t\right]$$

where the first term corresponds to the supply of labor efficiency units of non-college graduates post the college period and the second term corresponds to supply of labor efficiency units during the college period. For college graduates, the labor market clearing condition reads as:

$$N_1 = \frac{1}{w_1} \sum_{t \geqslant 6} \int_{\Omega_t^i} y_t^i \mathbf{1}\{e^i = 1\} d\Theta_t.$$

4. Goods market clears:

$$\sum_{t \geqslant 5} \int_{\Omega_t^i} \left(c_t^i + c_{kt}^i + x_t^i + c_\tau(\tau_t^i + \tau_{kt}^i)\right) d\Theta_t + \tau_s \int_{\Omega_5^i} \mathbf{1}\left\{s^i = 1\right\} d\Theta_5 + \delta K + G + NX = F\left(K, H\right).$$

5. The government budget holds:

$$g_{\tau t} \sum \int_{\Omega_t^i} \tau_t^i d\Theta_t + g_{\tau k} \sum \int_{\Omega_t^i} \tau_{kt}^i d\Theta_t + \sum \int_{\Omega_t^i} y_t^p\left(e^i, \theta^i\right) d\Theta_t + G = \sum \int_{\Omega_t^i} T\left(y_t^i, a_t^i, c_t^i\right) d\Theta_t.$$

The government subsidy for mental health services is the difference between the overall costs of treatment, denoted by c_τ, and the out of pocket cost paid by individuals, φ_τ. The subsidy can vary by whether the treated person is an adult or child as well as by the age of the adult. φ_τ. When a policy is introduced, G is held constant at its baseline steady-state level and the budget is rebalanced through a change in the income tax rate τ_0.

6. The distributions Ω_t are time-invariant.

Firm Problem. The firm chooses capital as well as college and non-college labor to maximize profits:

$$\max \; AK^\alpha N^{1-\alpha} - w_0 N_0 - w_1 N_1 - (r + \delta_K)K,$$

where A represents total factor productivity and δ_K is the depreciation rate of capital. The first-order condition with respect to capital gives:

$$r + \delta_K = \alpha A \left(\frac{N}{K}\right)^{1-\alpha}, \tag{A.4}$$

while the first-order condition with respect to non-college labor N_0 is:

$$\begin{aligned}
w_0 &= (1-\alpha)AK^{\alpha}\left(\varpi N_0^{\Omega} + (1-\varpi)N_1^{\Omega}\right)^{\frac{1-\alpha}{\Omega}-1}\varpi N_0^{\Omega-1} \\
&= (1-\alpha)\varpi A^{\frac{1}{1-\alpha}}\left(\frac{\alpha}{r+\delta_K}\right)^{\frac{\alpha}{1-\alpha}}\left(\varpi + (1-\varpi)\left(\frac{N_1}{N_0}\right)^{\Omega}\right)^{\frac{1-\Omega}{\Omega}}
\end{aligned}$$

where we observe that the second equality follows from the first-order condition for capital and yields an expression that is independent of K. The first-order condition with respect to college labor N_1 is:

$$w_1 = (1-\alpha)AK^{\alpha}\left(\varpi N_0^{\Omega} + (1-\varpi)N_1^{\Omega}\right)^{\frac{1-\alpha}{\Omega}-1}(1-\varpi)N_1^{\Omega-1},$$

which together with the first-order condition for non-college labor yields the relative wage:

$$\frac{w_1}{w_0} = \frac{1-\varpi}{\varpi}\left(\frac{N_1}{N_0}\right)^{\Omega-1}. \tag{A.5}$$

Equilibrium Construction. Fix interest rate R, and normalize wages for non-college workers $w_0 = 1$. Guess relative wage w_1. Using the optimality condition for non-college labor, set A so that the wage for the low skill worker is one, given that the relative wage implies a relative demand for college labor (A.5). Given prices, solve the individual problems and aggregate to obtain the relative supply of college labor. Update the implied relative wage by evaluating equation (A.5) at the relative supply for college update.

C Empirical Evidence

We present empirical evidence from the Panel Study of Income Dynamics (PSID) on the intergenerational persistence of mental health, parental time investments, college enrollment, and college graduation.

Panel Study of Income Dynamics. The Panel Study of Income Dynamics collects information on a nationally representative sample of U.S. families. We incorporate data from all waves from 2000 to 2020 since earlier waves do not contain information on mental health. Our analysis focuses on households with heads of households between 24 and 63 years of age. Our measure of mental health is based on the K6 scale. All dollar values are reported in 2015 values.

Our measure of income is household labor income per adult over the past calendar year. Hours worked are measured as total hours worked per adult including overtime. Hourly wage rates are constructed as the ratio of income and hours worked per adult. Our benchmark measure of consumption is annual non-durable expenditures which include expenditures on food, utilities, child care, clothing, home insurance, telecommunications, home maintenance, variable transportation costs, education, and recreation.[39] For all analyses, we use the sample weights provided by the surveys. Our measure of wealth is total wealth, the sum of all assets net of liabilities.[40]

Child Development Supplement. The Child Development Supplement (CDS) of the PSID records information for a nationally representative sample of U.S. children between age 0 and age 18. The first cohort of the CDS was launched in 1997 and included up to two children between age 0 and age 12 from each PSID household. Children and their parents were interviewed in 1997, 2002 and 2007. The second cohort of the CDS was launched in 2014 and included all children in PSID households. Children and parents were interviewed in 2014, 2019, and 2021.

Mental Health. The CDS provides information on child mental health, child mental health treatment, the amount of quality time parents spend with children, and parent mental health. We measure child mental health in the CDS using the Behavioral Problem Index (BPI) as explained in the main text. We measure whether a child receives mental health treatment as an indicator that is equal to one if the child has seen a mental health professional in the past year.

Parental Time Investments. The CDS collects detailed time diaries for each child, recording hour by hour which activity is being performed, whether a parent is doing the activity together with the child, and whether a parent is present but not actively engaged. Time diaries are available for both a weekday and a weekend day, which we use to construct weekly hours. We define quality time as the weekly hours a parent spends actively engaged with the child in social activities, entertainment, sports, eating, studying, home computer activities and games (excluding video games), and going to the museum. We exclude time spent watching television or playing video games, since these activities involve limited parent-child interaction and are typically not associated with positive developmental outcomes (Christakis, Zimmerman, DiGiuseppe, and McCarty, 2004; Swing, Gentile, Anderson, and Walsh, 2010). If both parents are performing the activity together with the child, we count this as double the hours since time constraints must hold for each parent individually. This is the measure of parental time investment n_{kt}

[39] Our consumption measure is closest to the consumption measures used by Krueger and Perri (2006) 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.

[40] We drop observations where the head of the household or the partner 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 rate below 3 dollars or above 300 dollars while working less than 10 hours per week, and observations where households report working more than 92 hours on average per week.

Table C.1: Parental Time Investments and Mental Health

Depression κ	-23.1	-25.3	-12.9	-11.8	-15.0
	(3.7)	(3.5)	(3.5)	(3.6)	(5.7)
Controls	None	+ Age	+ Income	+ Child Mental	+ Fixed Effects
R^2	0.01	0.12	0.16	0.16	0.42

Table C.1 displays the regression coefficient κ estimated from equation (A.6) and the corresponding standard errors (in row 2). The control variables include age, (log) income, child mental health, and a constant. From the first to the final column, we incorporate additional control variables. The number of observations is equal to 5,747.

used in the estimation and validation of the model.

Transition into Adulthood Supplement (TAS). The TAS of the PSID follows children of PSID respondents as they transition into adulthood, that is, after age 18. It is conducted biannually in conjunction with the main PSID since 2005. Between 2005 and 2015, only children who were part of the CDS were included in the TAS. Starting in 2017, all children of PSID respondents were surveyed as part of the TAS. The TAS provides information on young adults' mental health (measured using the K6 scale) as well as their education expectations and realized outcomes. Individuals are defined as being ever-enrolled in college if the maximum number of schooling years they report across the TAS waves is at least 13 and they reported at least once that they are seeking a four-year college degree. Individuals are defined as ever-graduated from college if the maximum number of schooling years they report across TAS waves is at least 16. Individuals are classified as currently enrolled in college if they report to be enrolled in a four-year college. Individuals are classified as expecting to graduate from college in the future if they think they will graduate from a four-year college or get more than four years of college.

C.1 Parent Time Investments and Mental Health

Table C.1 shows the effect of parent mental health on quality time they spent with their child. Quality time includes social activities, entertainment, sports, eating, studying, home computer activities and games (excluding video games), and going to the museum. We assess how parent quality time investments vary with parental mental illness, by estimating the following regression:

$$\log n_{ki} = \kappa D_i + \kappa_x X_i + \varepsilon_i, \tag{A.6}$$

Table C.2: Mental Health and College

	Enrollment	Graduation		
Depression γ_1	-7.6	-5.2	-4.8	-3.1
	(2.9)	(4.1)	(4.0)	(4.3)
Cognitive Skill γ_2	26.8	23.6	20.8	15.9
	(1.9)	(2.7)	(2.7)	(4.2)
Education Mother γ_3	24.1		18.7	16.0
	(2.6)		(3.1)	(3.5)
Observations	1,218	766	766	681
R^2	0.25	0.09	0.13	0.13

Table C.2 displays regression estimates for college enrollment (column 2) and graduation (columns 3 to 5). The corresponding standard errors are in the row below the estimated values. The additional control variables in the final column include the mental health and income of the parents as well as the letter word score of the child.

where $\log n_{ki}$ is the log of quality time investment by parent i, and D_i is a dummy variable which takes the value one when parent i is classified as experiencing mental illness. Control variables X_i include age, (log) income, child mental health, and a constant.

Table C.1 demonstrates how quality time investments varies with parental mental health. Each column corresponds to a regression that differs in the control variables that are included. From the first to the final column, we add control variables. For example, the first column shows that without controls, we find that parents experiencing mental illness spent 23.1 percent less time with their children relative to healthy individuals (first row). The penultimate column shows that this finding is robust to the inclusion of other control variables. Individuals experiencing mental illness spent roughly 11.8 percent less quality time with their children. In the final column, we show that parents spent 15.0 percent less quality time with their children when controlling for individual fixed effects.

C.2 College

We next assess the relation between college enrollment, college graduation and subjective college graduation probabilities with mental health. We find that adolescents experiencing mental illness are less likely to enroll into college, are equally likely to graduate, but have lower subjective graduation probabilities.

Table C.3: Expectations on College Graduation and Mental Health

Mental illness κ	-3.7	-3.7	-3.5	-3.7	-3.6
	(1.2)	(1.2)	(1.2)	(1.2)	(1.2)
Controls	None	+ Age, Sex	+ Parental Education	+ Income	+ Cognitive Skills

Table C.3 displays the regression coefficient κ estimated from equation (A.6) and the corresponding standard errors (in row 2). The control variables include age, (log) income, child mental health, and a constant. From the first to the final column, we incorporate additional control variables. The number of observations is equal to 1,683. All regressions control for year fixed effects.

Enrollment. We first analyze how college enrollment varies with child mental health, cognitive skills, and parental education. The results are presented in the second column of Table C.2.

We find that children who experience mental illness are 7.6 percent less likely to enroll in college. The probability of college enrollment is strongly increasing with cognitive skills: the enrollment probability increases by 26.8 percent for each standard deviation increase in cognitive skills. When the mother has completed college, the enrollment probability increases by 24.1 percent.

College Graduation. In columns 3 to 5 of Table C.2, we analyze the statistical relation between graduation probability and student mental health. We find no economic and statistically significant impact of mental health on college graduation conditional on enrollment in the second row of the table. Table C.2 does show that cognitive skill and parental educational attainment increase the likelihood of graduation as shown in rows 3 to 7. We use the regression coefficient in the second column to calibrate the dependence of the college graduation probability on cognitive skill in the quantitative model.

Expectations of Graduation. Finally, we evaluate the subjective graduation probabilities by mental health. Table C.3 shows how subjective graduation probabilities vary by mental health status. Each column corresponds to a regression that differs in the control variabels that are included. From the first to the final column, we add control variables. For example, the first column in Table C.3 shows that with only year fixed effects, we find that individuals experiencing mental illness have a subjective graduation probability that is 3.7 percentage point lower relative to healthy college students (first row). The final column shows that this finding is robust to the inclusion of all control variables. In sum, college students experiencing mental illness have lower subjective graduation probabilities.

Transition From Adolescence to Adult Mental Health. We use the CDS and the TAS to estimate the transition from the binary classification of child mental health into the three-way classification of adult mental health. In the CDS, we find that 29.25 percent of adolescents between age 16 and 17 experience mental health problems. In the TAS, we find that at age 18, 18.27 percent experiences mild mental illness while 6.55 percent experiences serious mental illness. The model-implied mental health transitions align with this data by setting the transition probability from experiencing mental illness as an adolescent to experiencing serious mental illness as an adult at $\frac{6.55}{29.25} = 0.224$, and to experiencing mild mental illness at $\frac{18.27}{29.25} = 0.625$. A remaining fraction of 0.153 adolescents that experience mental illness are healthy as adults. All children that are healthy remain healthy when they become adults.

C.3 Persistence

In order to evaluate the persistence of child mental health problems, we regress child mental health on lagged child mental health and parent mental health controlling for income, age, and year fixed effects. When a parent experiences mild mental illness, the child is 14.9 percent more likely to experience mental illness. When a parent experiences serious illness, this percentage jumps to 24.0 percent. When a child currently experiences mental illness, they are 39.2 percent more likely to experience mental illness five years later.

C.4 Consumption and Hours Worked

Following Abramson, Boerma, and Tsyvinski (2024), we quantify the relationship between consumption, hours worked, and mental health using the Panel Study of Income Dynamics (PSID). We incorporate data from all waves from 2000 to 2020 since earlier waves do not contain information on mental health. Our analysis focuses on households with heads of households between 24 and 63 years of age. All dollar values are reported in 2015 values.

Our income measure is household labor income per adult over the past calendar year. Hours worked are measured as total hours worked per adult including overtime. Hourly wage rates are constructed as the ratio of income and hours worked per adult. Our benchmark measure of consumption is annual non-durable expenditures which include expenditures on food, utilities, child care, clothing, home insurance, telecommunications, home maintenance, variable transportation costs, education, and recreation.[41] For all analyses, we use the sample weights provided by the surveys. Our measure of wealth is total wealth,

[41] Our consumption measure is closest to the consumption measures used by Krueger and Perri (2006) 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.

Table C.4: Consumption and Mental Health

Variable (in logs)	Non-durables	Durables
Mild γ_1	-2.2	-3.2
	(0.9)	(1.0)
Serious γ_2	-4.8	-5.5
	(1.6)	(1.7)
Observations	27,301	27,301
R^2	0.28	0.28
Mean (in levels)	19,400	22,200

Table C.4 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 sum of all assets net of liabilities.[42]

The PSID reports mental health 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. The K6 scale is included in all PSID waves conducted between 2000 and 2020 expect for 2004.

D Numerical Solution of Negative Thinking Problem

We describe the numerical approach to solving the negative thinking problem that arises in the individual decision problem. The challenge is that for each state and dynamic choice, households solve a constrained minimization problem to form subjective expectations. We develop an approximation method to reduce computational time by several orders of magnitude while maintaining accuracy.

D.1 The Computational Challenge

Individuals with mental health m form subjective beliefs by choosing a subjective probability distribution p to minimize (1) subject to the relative entropy constraint (2). The computational difficulty is that: (i)

[42]We drop observations where the head of the household or the partner 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 rate below 3 dollars or above 300 dollars while working less than 10 hours per week, and observations where households report working more than 92 hours on average per week.

this problem must be solved at every point in the state space, (*ii*) the value function iteration requires solving this problem repeatedly until convergence, and (*iii*) the Lagrange multiplier λ corresponding to a given value of κ does not have a closed-form solution and must be found numerically for each configuration of continuation values and objective probabilities.

In our model, the state space includes assets, mental health, idiosyncratic productivity, education, cognitive ability, together with child states. With a fine grid for the states, this implies solving millions of entropy problems per iteration of the value function. Standard approaches − grid search over candidate values for the inverse Lagrange multiplier λ values or numerical root-finding of this multiplier − become computationally prohibitive.

Problem Statement. We solve the negative thinking problem: Specifically, agents choose a subjective probability distribution p to minimize the expected value (1) subject to the relative entropy constraint (2). In this appendix, we write this shorthand as:

$$\min_p \sum_{i=1}^{N} p_i w_i$$

subject to $\mathcal{R}(p\|q) \leqslant \kappa$ and where p is required to be a probability mass function.

The vector of continuation values for N states is written as $w = (w_1, \ldots, w_N)^\top$, while the vector of objective probabilities is $q = (q_1, \ldots, q_N)^\top$. The parameter κ governs the maximum permissible entropy of the subjective probability distribution. The solution to the negative thinking problem yields subjective probabilities (3) where the inverse Lagrange multiplier λ has to satisfy the entropy constraint (2):

$$\kappa = \sum p_i \log \frac{p_i}{q_i} = -\lambda \sum \frac{q_i \exp(-\lambda w_i)}{\sum q_i \exp(-\lambda w_i)} w_i - \log\left(\sum q_i \exp(-\lambda w_i)\right). \tag{A.7}$$

Solving for λ using root-finding methods is computationally expensive. We next develop an approximation method. Observe that the solution depends on the distribution of the continuation values w.

Approximation Method. We develop a fast approximation using Padé approximants derived from a Puiseux series expansion. This approximation speeds up the code by a factor 4 while preserving accuracy through a hybrid strategy.

To develop the approximation method, we first define the standardized continuation values:

$$\hat{w}_i = \frac{w_i - \mu}{\sigma}$$

with mean $\mu = \sum q_i w_i$ and variance $\sigma^2 = \sum q_i (w_i - \mu)^2$.

Claim 1. *Scale Invariance.* Consider some $\kappa \geqslant 0$ and fix a probability vector q. Suppose $\hat{\lambda}$ satisfies the entropy constraint (A.7) for continuation vector $\hat{w}$, and suppose λ satisfies the entropy constraint for w. Then, $\lambda\sigma = \hat{\lambda}$.

Proof. Substituting $w_i = \sigma\hat{w}_i + \mu$ into the subjective probabilities (3):

$$p_i(\lambda, w) = \frac{q_i \exp(-\lambda(\sigma\hat{w}_i + \mu))}{\sum q_i \exp(-\lambda(\sigma\hat{w}_i + \mu))} = \frac{q_i \exp(-\lambda\sigma\hat{w}_i)}{\sum q_i \exp(-\lambda\sigma\hat{w}_i)}.$$

Defining $\hat{\lambda} = \lambda\sigma$, $p_i(\lambda, w) = p_i(\hat{\lambda}; \hat{w})$. Since the relative entropy constraint depends only on the probability vectors, both subjective probability distributions have the same entropy when $\hat{\lambda} = \lambda\sigma$. □

This claim implies that when $\hat{\lambda}$ satisfies the entropy constraint for the normalized values, $\lambda = \hat{\lambda}/\sigma$ satisfies the entropy constraint with non-normalized values. Recall that the entropy constraint (A.7) depends only on the distribution of continuation values and κ. Since the continuation values are bounded, their distribution is completely characterized by its moments or, equivalently, its cumulants. As the first two moments are normalized, $\hat{\lambda}$ satisfies the entropy constraint for the normalized values only depends on the cumulants k_n for $n \geqslant 3$.

The cumulants of a random variable are defined using the cumulant-generating function, which is the logarithm of the moment-generating function. The cumulant generating function for the distribution of $\hat{w}$ is:

$$K(\hat{\lambda}) := \log \mathbb{E}_q\left[e^{\hat{\lambda}\hat{w}}\right] = \sum_{n=1}^{\infty} \frac{k_n}{n!}\hat{\lambda}^n,$$

where k_n is the n-th cumulant of $\hat{w}$. For a normalized distribution, the first two cumulants are $k_1 = 0$ and $k_2 = 1$. The cumulant-generating function of $\hat{w}$ and the entropy constraint (A.7) are related as:

$$\kappa = -\hat{\lambda}K'(-\hat{\lambda}) - K(-\hat{\lambda}).$$

Substituting the cumulant-generating function, the coefficient for $\hat{\lambda}^n$ in the entropy constraint is:

$$\kappa = \sum_{n=2}^{\infty} a_n \hat{\lambda}^n, \tag{A.8}$$

with coefficients $a_n = (-1)^n \frac{n-1}{n!} k_n$. Given the cumulants, this gives κ as a function of $\hat{\lambda}$. However, rather than finding κ as a function of $\hat{\lambda}$, we need $\hat{\lambda}$ as a function of κ. Thus, we seek $\hat{\lambda}(\kappa)$ as a Puiseux series in $\kappa^{1/2}$:

$$\hat{\lambda} = \sum_{n=1}^{\infty} c_n \kappa^{n/2} = c_1\kappa^{1/2} + c_2\kappa + c_3\kappa^{3/2} + \dots \tag{A.9}$$

To determine the coefficients c_n, we substitute (A.9) into (A.8) − the Taylor expansion of $\kappa(\hat{\lambda})$ − and match powers of $\kappa^{1/2}$. This yields a system that can be solved sequentially.

Since the Puiseux series (A.9) will diverge for moderate-to-large values for the relative entropy κ, we convert it using the Padé $[m/n]$ approximant. The Padé $[m/n]$ approximant is an approximation of a function using rational polynomials:

$$\hat{\lambda}(\kappa) \approx \frac{P_m(t)}{Q_n(t)},$$

where $t = \kappa^{\frac{1}{2}}$, $P_m(t) = p_0 + p_1 t + p_2 t^2 + \cdots + p_m t^m$, and $Q_n(t) = 1 + q_1 t + q_2 t^2 + \cdots + q_n t^n$. We exactly match the first $(m+n)$ Puiseux coefficients to determine $(m+n)$ unknown Padé coefficients.

D.2 Comparing Different Methods

We compare three approaches for computing the solution: a direct numerical solver used as a baseline, a truncated Puiseux series augmented with a fallback mechanism, and the Padé approximant combined with the same fallback mechanism.

The fallback mechanism is designed to ensure numerical accuracy. When the relative entropy implied by the Puiseux or Padé approximation deviates from the target relative entropy value κ by more than a threshold − set here to one percent of κ − the approximation is rejected and the direct numerical solver is used instead. This approach ensures that approximation errors do not propagate into the final solution.

Figure D.1 summarizes the findings. Panel (a) shows that the approximation methods are substantially faster than relying exclusively on the direct numerical solver. The Puiseux-series method is about 2.7 faster than the direct solver-only benchmark, while the Padé approximation achieves a speedup of about 3.8. This difference is largely driven by the lower fallback rate of the Padé approximation, which allows the solver to be bypassed more frequently.[43]

Both Puiseux series and Padé approximations perform well when relative entropy κ is close to zero, as shown in panels (b) and (c), where accuracy is high and the fallback is negligible. As κ increases, however, their performance diverges. The Padé approximation remains accurate and stable for $\kappa > 0.05$, while the Puiseux series deteriorates rapidly, particularly for intermediate values of relative entropy between 0.05 and 0.5.

Implementation. Based on these findings, we adopt a Padé [4/3] approximation together with a fallback numerical solver to balance computational efficiency and accuracy across the range of entropy values. The

[43]We tested multiple truncation orders for both the Puiseux series and the Padé approximants and report results using the degrees that minimize fallback rates for each method.

Figure D.1: Methods Comparison

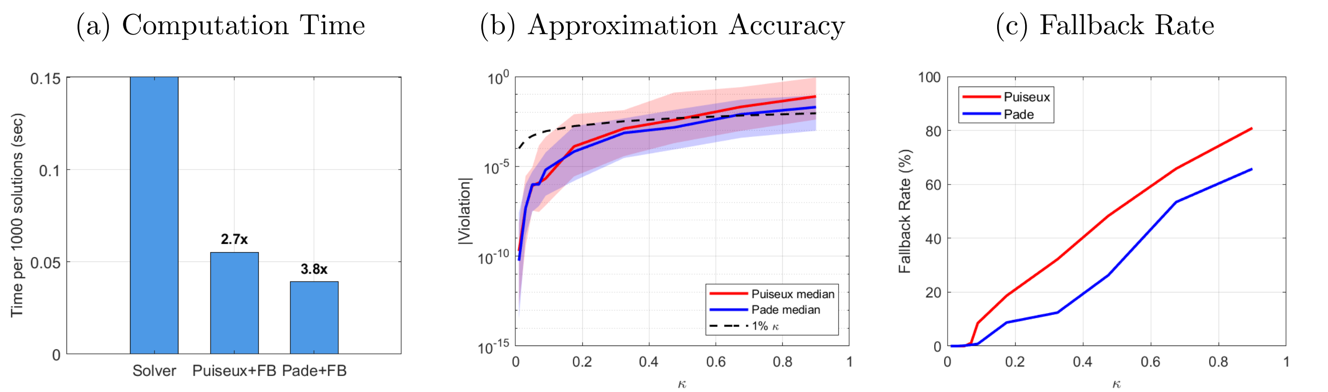


Figure D.1 presents results on the speed and accuracy of the Puiseux series and Padé approximations. Panel (a) shows computation time per 1,000 solutions with speedup factors. Panel (b) shows the approximation accuracy median |violation| of κ before fallback on log scale, with 1 percent threshold shown as dashed line. Shaded areas show the 90th to 10th percentile ranges. Panel (c) shows the fallback rate.

algorithm proceeds as follows:

1. *Standardize.* We first standardize the continuation values $\hat{w}_i = (w_i - \mu)/\sigma$.
2. *Compute cumulants.* Calculate k_n from moments $\mathbb{E}_q[\hat{w}^n]$ for $n = 3, \ldots, 12$ using standard formulas.
3. *Compute Puiseux coefficients*: $c_1, \ldots, c_{m+n}$ from cumulants.
4. *Solve Padé system*: Linear algebra to get $\{p_k, q_k\}$.
5. *Evaluate at* $t = \sqrt{\kappa}$: $\hat{\lambda} = P_m(t)/Q_n(t)$.
6. *Fallback if needed*: If implied entropy exceeds a threshold, use root-finder with initial guess $\sqrt{2\kappa}$.

E Quantifying the Model

In this appendix, we provide additional detail on how we quantify the structural framework.

E.1 College Borrowing

We transform college loans into market loans following Abbott, Gallipoli, Meghir, and Violante (2019). Individuals can choose to borrow an amount b for college. They borrow an amount b and repay the loan in five periods after college. If the individual obtains such a loan, akin to a mortgage, at the subsidized rate, the mortgage payment formula specifies that the annual payment b_p is given by:

$$b_p = \frac{R_e - 1}{1 - R_e^{-5}} b,$$

where b_p is the payment, b is the loan amount and R_e is the subsidized interest rate on borrowing.

Instead, if the individual was making the same fixed payments on a mortgage that with an unsubsidized rate $R_b \geqslant R_e$, the present value of this mortgage is instead lower at:

$$\tilde{b} = \frac{1 - R_b^{-5}}{R_b - 1} b_p.$$

This means the market value of debt of the fixed payments b_p after college can be written as:

$$\tilde{b} = \frac{1 - R_b^{-5}}{R_b - 1} \frac{R_e - 1}{1 - R_e^{-5}} b.$$

E.2 Consumption Choices

Following Lee and Seshadri (2019), we show that the consumption choice for the parent and their child can be written as choosing a single consumption level when consumption preferences follow a CRRA specification, that is, $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$. At optimum, the marginal utility of the parent and child must be proportional. In order to see this formally, note that for a given level of consumption expenditures C, parents choose parental consumption c, and child consumption c_k to maximize $\frac{c^{1-\gamma}}{1-\gamma} + \delta \frac{c_k^{1-\gamma}}{1-\gamma}$ subject to $c + c_k \leqslant C$. This problem is thus equivalent to maximizing $\frac{c^{1-\gamma}}{1-\gamma} + \delta \frac{(C-c)^{1-\gamma}}{1-\gamma}$, and yields the first-order condition:

$$c^{-\gamma} = \delta (C - c)^{-\gamma}.$$

Together with the expenditure constraint, this yields the consumption choice for parents $c = \frac{1}{1+\delta^{\frac{1}{\gamma}}} C$ and for children $c_k = \frac{\delta^{\frac{1}{\gamma}}}{1+\delta^{\frac{1}{\gamma}}} C$ and the indirect utility from total consumption expenditures C:

$$U(C) = \left[\left(\frac{1}{1 + \delta^{\frac{1}{\gamma}}} \right)^{1-\gamma} + \delta \left(\frac{\delta^{\frac{1}{\gamma}}}{1 + \delta^{\frac{1}{\gamma}}} \right)^{1-\gamma} \right] \frac{C^{1-\gamma}}{1-\gamma} = \left(1 + \delta^{\frac{1}{\gamma}}\right)^{\gamma} u(C) = \Psi(\delta, \gamma) u(C)$$

In the special case of logarithmic preferences for consumption utility, we obtain $c = \frac{1}{1+\delta} C$ and for children $c_k = \frac{\delta}{1+\delta} C$ and hence the indirect utility is:

$$U(C) = (1 + \delta) \log C + \delta \log \delta - (1 + \delta) \log(1 + \delta)$$

E.3 Pension Benefits: US Social Security System

The pension benefit is constructed using the Old Age Insurance of the US Social Security System. We use cognitive skills and education to estimate a proxy for average lifetime earnings, on which the replacement benefit is based. Average earnings at age t are estimated as $\hat{y}_t(e, \theta) = w_t(m, z_t, e, \theta) n_w(m, 0)$ evaluated at mean idiosyncratic productivity $\log z_t = 0$ and weighted across the mental health distribution. Averaging over t allows average lifetime income $\hat{y}(e,\theta)$ to be calculated.

Given average lifetime income $\hat{y}(e,\theta)$, the formula for the primary insurance amount is given by:

$$y^{\text{pi}}(e,\theta) = \begin{cases} 0.9 \times \hat{y}\left(e,\theta\right) & \text{if } \hat{y}\left(e,\theta\right) \leqslant 9,912 \\ 0.9 \times 9,912 + 0.32\left(\hat{y}\left(e,\theta\right) - 9,912\right) & \text{if } 9,912 \leqslant \hat{y}\left(e,\theta\right) \leqslant 59,760 \\ 0.9 \times 9,912 + 0.32\left(59,760 - 9,912\right) + 0.15\left(\hat{y}\left(e,\theta\right) - 59,760\right) & \text{if } 59,760 \leqslant \hat{y}\left(e,\theta\right) \end{cases}$$

Pension benefits are capped at $\bar{y}^p = 31,956$ dollars per year, or $y^p(e,\theta) = \min\left\{y^{\text{pi}}(e,\theta), \bar{y}^p\right\}$.

E.4 Utility from Leisure and Rumination

In order to parameterize the utility from leisure, we first set the level of leisure ψ as well as the curvature of the function which is governed by η. We set these parameters by ensuring that the first-order condition of labor supply are satisfied for individuals who are healthy, do not have children, and do not go to college, that is, for individuals who only have to spend their time on work and leisure.

We set η to align the Frisch elasticity of labor supply. Let λ capture the marginal utility of wealth, the first-order condition with respect to hours worked can be written as:

$$(1-\tau_0)(1-\tau_1)w^{1-\tau_1}n^{-\tau_1}\lambda = \psi(1-n)^{-\frac{1}{\eta}}.$$

We totally differentiate this expression with respect to w, holding constant the marginal utility of wealth. Let $A := (1-\tau_0)(1-\tau_1)\lambda$, then differentiating the left-hand side we obtain:

$$\frac{d}{dw}\left[Aw^{1-\tau_1}n^{-\tau_1}\right] = A\left[(1-\tau_1)w^{-\tau_1}n^{-\tau_1} - \tau_1 w^{1-\tau_1}n^{-\tau_1-1}\frac{dn}{dw}\right],$$

and for the right-hand side we obtain:

$$\frac{d}{dw}\left[\psi(1-n)^{-\frac{1}{\eta}}\right] = \frac{\psi}{\eta}(1-n)^{-\frac{1}{\eta}-1}\frac{dn}{dw}.$$

Setting both sides equal and solving for $\frac{dn}{dw}$:

$$\frac{dn}{dw} = \frac{A(1-\tau_1)w^{-\tau_1}n^{-\tau_1}}{\tau_1 Aw^{1-\tau_1}n^{-\tau_1-1} + \frac{\psi}{\eta}(1-n)^{-\frac{1}{\eta}-1}}$$

Multiplying both sides by $\frac{w}{n}$, we isolate $\frac{w}{n}\frac{dn}{dw}$:

$$\varepsilon_F = \frac{dn}{dw}\frac{w}{n} = \frac{A(1-\tau_1)}{\tau_1 A + \dfrac{\psi}{\eta}w^{\tau_1-1}n^{\tau_1+1}(1-n)^{-\frac{1}{\eta}-1}} = \frac{1-\tau_1}{\tau_1 + \frac{1}{\eta}\frac{n}{1-n}} \tag{A.10}$$

where the final equality uses that $Aw^{1-\tau_1}n^{-\tau_1} = \psi(1-n)^{-\frac{1}{\eta}}$. We use this expression to calibrate η using evidence on the Frisch elasticity of labor supply ε_F. Note that this expression is independent of the level shifter for the utility from leisure.

We set ψ to align the labor supply condition in the model with the data. The first-order condition for labor supply is:

$$(1-\tau_1)(1+\tau_c)\frac{\tilde{y}}{c} = \psi\frac{n}{(1-n)^{\frac{1}{\eta}}}.$$

We evaluate the labor supply condition for hours worked n and the ratio of consumption to after-tax labor income for healthy households without children in the data, together with the Frisch elasticity of labor supply η and tax code τ to obtain ψ. For healthy households without children, labor income before taxes is 49,195 dollars, consumption before taxes is 46,075 dollars, and average hours worked equal 0.338.

Given the calibration of ψ using the labor supply condition of healthy individuals without children, we next calibrate the extent of rumination using the first-order condition for labor supply of individuals without children experiencing mental illness to calibrate rumination. The first-order condition for labor supply for individuals experiencing mental illness can be written as:

$$\bar{n} = n + \left(\frac{\psi n\, c}{(1-\tau_1)(1+\tau_c)\,\tilde{y}}\right)^{\eta}.$$

In order to calibrate hours ruminating when experiencing mild and serious mental illness that are consistent with the labor supply decision of these individuals, we evaluate this labor supply condition using average hours worked, after-tax income, and consumption by mental health status. Using fixed effects regressions, we find that households experiencing mild (serious) mental illness work 2.7 (15.9) percent fewer hours, have 2,000 (7,000) dollars less income, and spent 725 (1,434) dollars less on consumption.

E.5 Initial Conditions for Child Mental Health and Cognitive Skills

This section describes the parameterization of the system of equations that governs the initial draw of the child mental health and cognitive skills. As described in Section 4.2.1, the initial conditions for the child are a linear function of the mental health and cognitive skills of their parent, namely:

$$\log\theta_k = \beta_{1\theta}\log\theta + \beta_{2\theta}\log m + \varepsilon_\theta \qquad \text{and} \qquad \log m_k = \beta_{1m}\log\theta + \beta_{2m}\log m + \varepsilon_m$$

where ε_θ and ε_m are both mean-zero error terms and independent of the parental characteristics. To calibrate the initial draw for the child, we need to parameterize the four coefficients $\beta_{1\theta}$, $\beta_{2\theta}$, β_{1m}, β_{2m}, the two variances of the error terms, and the covariance between the error terms. We first discuss the identification of the four coefficients and then discuss the identification of the covariance matrix of the error term.

Consider the two regression above in general form $y = \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, where u is independent and identically distributed. The regression coefficient vector $\beta = \{\beta_1, \beta_2\}$ minimizes the squared error term, $\beta = \arg\min \mathbb{E}[\varepsilon^2]$, where $\varepsilon = y - \beta_1 x_1 - \beta_2 x_2$. The first-order conditions imply $\mathbb{E}[x_1\varepsilon] = 0$ and $\mathbb{E}[x_2\varepsilon] = 0$. Expanding the first-order conditions yields a linear system with two unknowns. We invert this system to identify the coefficients:

$$\begin{bmatrix} \beta_1 \\ \\ \beta_2 \end{bmatrix} = \frac{1}{\mathbb{E}[x_1^2]\mathbb{E}[x_2^2] - \mathbb{E}[x_1 x_2]^2} \begin{bmatrix} \mathbb{E}[x_2^2]\mathbb{E}[x_1 y] - \mathbb{E}[x_1 x_2]\mathbb{E}[x_2 y] \\ \\ \mathbb{E}[x_1^2]\mathbb{E}[x_2 y] - \mathbb{E}[x_1 x_2]\mathbb{E}[x_1 y] \end{bmatrix}.$$

Next, we identify the covariance matrix for the error using the estimated coefficients. Given the linear structure of the initial conditions, it follows that:

$$\begin{aligned} \mathrm{Var}(\varepsilon_\theta) &= \mathrm{Var}(\log\theta_k) - \beta_{1\theta}^2 \mathrm{Var}(\log\theta) - \beta_{2\theta}^2 \mathrm{Var}(\log m) \\ \mathrm{Var}(\varepsilon_m) &= \mathrm{Var}(\log m_k) - \beta_{1m}^2 \mathrm{Var}(\log\theta) - \beta_{2m}^2 \mathrm{Var}(\log m) \end{aligned}$$

and that the covariance is given by:

$$\begin{aligned} \mathrm{Cov}(\varepsilon_\theta, \varepsilon_m) = \mathrm{Cov}(\log\theta_k, \log m_k) &- \beta_{1\theta}\beta_{1m}\mathrm{Var}(\log\theta) - \beta_{2\theta}\beta_{2m}\mathrm{Var}(\log m) \\ &- \big(\beta_{1\theta}\beta_{2m} + \beta_{1m}\beta_{2\theta}\big)\mathrm{Cov}(\log\theta, \log m). \end{aligned}$$

Covariance Matrix of Initial Conditions. Using the estimated covariance matrix of initial conditions, we operationalize the identification argument for the biological factors. Specifically, we use the estimated covariance matrix of initial conditions in Table E.1. The implied coefficient of determination is 0.07 for mental health and below 0.01 for child cognitive skills.

E.6 College Costs

We parameterize college costs τ_e by calculating the annual monetary cost faced by undergraduate students, excluding room and board. Annual college costs are defined as net tuition and mandatory fees after grant aid, plus books and supplies at four-year colleges.

We calculate the costs exclude housing, food, transportation, and other living expenses using data from the *Trends in College Pricing* (College Board, 2015) and the National Center for Education Statistics

Table E.1: Covariance Matrix of Initial Conditions

	Child Cognitive	Child Mental	Maternal Cognitive	Maternal Mental	Unobserved
Child cognitive	0.1657				
Child mental	−0.0024	0.0690			
Maternal cognitive	0.0193	0.0515	0.5811		
Maternal mental	0.0050	0.0084	0.0412	0.0869	
Unobserved heterogeneity	0.0000	0.0000	−0.0021	−0.0066	0.0057

Table E.1 displays the estimated covariance matrix of initial conditions for the cognitive skills and mental health of children, the cognitive skills and mental health of parents as well as unobserved heterogeneity.

(2015). In 2015 dollars, net tuition and fees (3,980) and books and supplies (1,298) sum to 5,278 dollars per year for in-state students at public colleges. For out-of-state students at public colleges, we assume they receive the same amount of grant aid as in-state students since net tuition and fees are not reported for out-of-state students. Given tuition and fees of 23,893 for out-of-state students and grant aid of 5,430, net tuition and fees are 18,463, which together with books and supplies (1,298) sum to 19,761 dollars. In order to calculate the average costs at public colleges, we use data from Figure 28 of College Board (2015), which indicates that about 80 percent of students at public institutions are in-state students. The average annual cost at public institutions is therefore 8,175 dollars ($0.80 \times 5{,}278 + 0.20 \times 19{,}761 \approx 8{,}175$). For students at private nonprofit colleges, net tuition and fees (14,890) and books and supplies (1,249) sum to 16,139 dollars.

In order to calculate the overall average college costs, we take a weighted average using enrollment data from the National Center for Education Statistics. Restricting attention to students enrolled in public and private nonprofit institutions, approximately 67 percent attend public institutions and 33 percent attend private nonprofit institutions (National Center for Education Statistics, 2015). The population-weighted average annual cost is thus 10,755 dollars per year ($0.676 \times 8{,}175 + 0.324 \times 16{,}139 \approx 10{,}755$).

E.7 Child Expenditures

The National Center for Education Statistics reports that for the 2014−15 school year the average total expenditure was 12,796 dollars per student (National Center for Education Statistics, 2017, Table 236.55). Federal and state sources together account for 55 percent of public school revenues, while local sources account for 45 percent (National Center for Education Statistics, 2017, Table 235.10). Following the education literature in economics − which emphasizes that households choose locations taking into account locally financed education expenditures funded through their taxes (e.g., Fernandez and Rogerson (1996, 1998) and Zheng and Graham (2022)) − we focus on the local component, yielding approximately 5,758 dollars per student. The coefficient of variation of cost-adjusted total expenditures per student between districts was 0.24.[44] We target average local expenditures as a moment, and set the low and high levels of monetary expenditures to be 48 percent (twice the coefficient of variation) away from the mean.

E.8 Mental Health Transition Matrix for Adults

We estimate transition rates between mental health states as a function of the agent's treatment decision and idiosyncratic productivity. We denote the transition probability from state m to m', conditional on the treatment decision τ and idiosyncratic productivity state ν, by $\Gamma_m(m'|m, \tau, \nu)$. 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 agents, or $\Gamma(m' \mid m_0, 1, \nu) = \Gamma(m' \mid m_0, 0, \nu)$ for all m' and ν. This is motivated by the fact that, in the data, healthy individuals rarely receive treatment (see, e.g., Cronin, Forsstrom, and Papageorge (2025)). 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. The last two assumptions allow us to capture, in a parsimonious way, the idea that negative income shocks deteriorate future mental health.

We next describe the data moments used for estimation. First, we compute the biannual transition probabilities between mental health states from the PSID sample. 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 households who have normal idiosyncratic productivity (i.e., $\nu_i \geqslant \underline{\nu}$) and for households who have low idiosyncratic productivity (i.e., $\nu_i < \underline{\nu}$). These transitions are denoted by $\Gamma^d(m' \mid m_0, \nu \geqslant \underline{\nu})$, and $\Gamma^d(m' \mid m_0, \nu < \underline{\nu})$ and. The empirical transition probabilities from the healthy state are independent of

[44]This figure is from Sherman, Gregory, Poirier, and Ye (2003), which reports dispersion measures for the 1997−98 school year using F33 School District Finance Survey data adjusted by the geographic cost of education index.

treatment. We multiply the transition matrix by itself to obtain the four-year transition matrix.

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 experiencing mild illness. 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 Use and Health of the Substance Abuse. 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.[45] Fourth, we obtain the share of healthy individuals who have an idiosyncratic productivity above $\underline{\nu}$ in our PSID sample, which we denote by π_d^{ν}.

Finally, we use estimates on the efficacy of 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 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.

Given the data, we estimate the mental health transition matrix following Appendix D in Abramson, Boerma, and Tsyvinski (2024).

E.9 Mapping Adult Mental Health to Child Development Input

In order to map parental mental health into the child skill production technology we assume that parental mental health and $m \sim \mathcal{N}(0, \sigma^2)$ are ordered in an identical fashion. For example, the highest 80 percent in the distribution of mental health factor m are healthy, the bottom 5 percent experience serious mental health problems. We discretize the normal distribution so that our grid reflects this.

Let $x \sim \mathcal{N}(\mu, \sigma^2)$ be a normal random variable. We partition its support into three intervals capturing proportions such that mass p_1 is in the first bin; mass p_2 is in the second bin, and mass $p_3 = 1 - p_1 - p_2$ is in the third bin. Each bin is assigned a value equal to the conditional mean of x in that bin.

[45] 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 $\frac{5.5}{22.8} \times 65.4 + (1 - \frac{5.5}{22.8}) \times 0.41 = 47.2$.

Given μ, σ^2, and bin shares $p_1, p_2 \in (0,1)$ such that $0 < p_3 < 1$, we set cutoffs (q_1, q_2) and generate representative means m_1, m_2, m_3. In order to determine the cutoffs, we use the inverse cumulative distribution function for the standard normal distribution to yield $\alpha_i = \Phi^{-1}(p_i)$ for $i = \{1, 2\}$. In order to compute the conditional means, we evaluate the means of the truncated normal distribution, that is:

$$m_1 = \mu - \sigma \frac{\varphi(\alpha_1)}{\Phi(\alpha_1)} \qquad m_2 = \mu + \sigma \frac{\varphi(\alpha_1) - \varphi(\alpha_2)}{\Phi(\alpha_2) - \Phi(\alpha_1)} \qquad m_3 = \mu + \sigma \frac{\varphi(\alpha_2)}{1 - \Phi(\alpha_2)}.$$

This gives the three values for the lognormal distribution. These values are exponentiated to give the cutoffs and conditional means in levels.

E.10 Negative Thinking and Discrete Choice Probabilities

We prove that negative thinking of parents preserves the choice probabilities of the child going to college. To prove this result, we characterize negative thinking over the child's discrete choice problem when the parent evaluates outcomes pessimistically subject to a relative entropy constraint.

Setup. The child chooses among $j \in \{1, \ldots, J\}$ options with an associated value:

$$U_j = V_j + \varepsilon_j,$$

where ε_j are independent and identically distributed Type 1 extreme value shocks with location μ and scale parameter σ. Let f be the objective density of the shock $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_J)$. Let $j^*(\varepsilon) = \arg\max_k \{V_k + \varepsilon_k\}$ be the child's optimal choice given shocks ε.

Under the objective probability distribution, the choice probabilities are given by:

$$q(j) = \frac{\exp(V_j/\sigma)}{I(\sigma)}, \tag{A.11}$$

where $I(\sigma) = \sum \exp(V_k/\sigma)$ and the expected utility of the realized choice is:

$$\mathbb{E}_f[V_{j^*} + \varepsilon_{j^*}] = \mu + \sigma(\gamma + \log I(\sigma)), \tag{A.12}$$

where γ is the Euler-Mascheroni constant. We set the location parameter μ such that the expected value is identical whenever the subjective probabilities across the different states j are identical.

Negative Thinking. Parents that experience mental illness think negatively with respect to the random outcome of the college taste shock ε for their child where $M(\varepsilon) = V_{j^*(\varepsilon)} + \varepsilon_{j^*(\varepsilon)}$ is the value associated with taste shock ε. The subjective probability distribution $\pi(\varepsilon)$ solves:

$$\min_{\pi} \ \mathbb{E}_\pi \big[M(\varepsilon)\big] \tag{A.13}$$

subject to the relative entropy constraint $\mathcal{R}(\pi\|f) = \int \pi(\varepsilon)\log\frac{\pi(\varepsilon)}{f(\varepsilon)}d\varepsilon \leqslant \kappa$.

The subjective probability distribution is given by:

$$\pi(\varepsilon) = \frac{f(\varepsilon)\exp\bigl(-\lambda M(\varepsilon)\bigr)}{Z(\lambda)} \tag{A.14}$$

where $j^*(\varepsilon)$ is independent of the distribution of ε and λ is the inverse Lagrange multiplier on the entropy constraint and $Z(\lambda) = \mathbb{E}_f\bigl[\exp(-\lambda M)\bigr]$.

Claim 2. Parental subjective probabilities with respect to their child going to college are independent of parental negative thinking, or $p(j) = q(j)$ for every j and all κ. The continuation value is $\mathbb{E}_\pi[M(\varepsilon)] = \mu + \sigma\log I(\sigma) - \sigma\psi(1+\lambda\sigma)$, where λ is the inverse multiplier on the negative thinking constraint.

We prove this claim below. For any strictly positive inverse Lagrange multiplier λ, $\psi(1+\lambda\sigma) > \psi(1) = -\gamma$, which implies that the expected continuation value decreases with the extent of negative thinking.

Negative thinking of parents with respect to the college choice of their child has the striking feature that the subjective probabilities of going to college are equivalent to the objective probability that the child goes to college. In other words, parental negative thinking does not affect the parental expectation that their child goes to college, but that the child will experience worse realizations of taste shocks regardless of which option is chosen. The continuation value differs from value (A.12) only in the constant term: $-\sigma\psi(1+\lambda\sigma)$ replaces $\sigma\gamma = -\sigma\psi(1)$. As pessimism increases (κ increases and hence λ increases), we have $\psi(1+\lambda\sigma) \to \infty$, and the continuation value tends to negative infinity.

Proof to Claim 2. We next prove Claim 2. In this section, we adopt the following notation: Γ denotes the gamma function, $\psi := \Gamma'/\Gamma$ is the digamma function.

Marginal Choice Probabilities. The subjective choice probability for option j under subjective probability distribution π is:

$$p(j) = \int_{A_j}\pi(\varepsilon)d\varepsilon = \frac{1}{Z(\lambda)}\int_{A_j}\exp\bigl(-\lambda(V_j+\varepsilon_j)\bigr)f(\varepsilon)d\varepsilon = \frac{\exp(-\lambda V_j)}{Z(\lambda)}\int_{A_j}\exp(-\lambda\varepsilon_j)f(\varepsilon)d\varepsilon,$$

where A_j denotes the support of the college taste shock ε where the optimal choice is given by option j, that is, $A_j = \{\varepsilon : V_j + \varepsilon_j > V_k + \varepsilon_k \text{ for all } k \neq j\}$.

The integral can equivalently be written as

$$\int_{A_j}\exp(-\lambda\varepsilon_j)f(\varepsilon)d\varepsilon = \int_{-\infty}^{\infty}\exp(-\lambda\varepsilon_j)f_j(\varepsilon_j)\prod_{k\neq j}F_k(\varepsilon_j + V_j - V_k)d\varepsilon_j,$$

where $f_j(\varepsilon_j) = \frac{1}{\sigma}\exp(-\varepsilon_j/\sigma)\exp(-\exp(-\varepsilon_j/\sigma))$ is the probability density function for the Type I extreme value distribution and $F_k(x) = \exp(-\exp(-x/\sigma))$ is the corresponding cumulative distribution function. In order to evaluate this integral, we substitute $u = \exp(-\varepsilon_j/\sigma)$ so that $\varepsilon_j = -\sigma \log u$, and differentiating gives $d\varepsilon_j = -\frac{\sigma}{u}du$. As ε_j tends to $-\infty$, u tends to ∞, and when ε_j tends to ∞, u tends to 0. The bounds of integration reverse and we obtain:

$$\int_{-\infty}^{\infty} \exp(-\lambda\varepsilon_j)\frac{1}{\sigma}\exp(-\varepsilon_j/\sigma)\exp(-\exp(-\varepsilon_j/\sigma))\prod_{k\neq j}\exp(-\exp(-(\varepsilon_j + V_j - V_k)/\sigma))d\varepsilon_j$$
$$= \int_0^{\infty} u^{\lambda\sigma}\exp\Big(-u\sum\exp\big(-(V_j - V_k)/\sigma\big)\Big)du = \int_0^{\infty} u^{\lambda\sigma}\exp\big(-u\exp(-V_j/\sigma)I(\sigma)\big)du.$$

We substitute $t = u\exp(-V_j/\sigma)I(\sigma)$. Then, $u = t\exp(V_j/\sigma)/I(\sigma)$ and $du = \exp(V_j/\sigma)/I(\sigma)dt$. The bounds remain 0 to ∞ and hence,

$$\int_{A_j}\exp(-\lambda\varepsilon_j)f(\varepsilon)d\varepsilon = \int_0^{\infty}\left(\frac{t\exp(V_j/\sigma)}{I(\sigma)}\right)^{\lambda\sigma}\exp(-t)\frac{\exp(V_j/\sigma)}{I(\sigma)}dt = \frac{\exp(V_j(\lambda\sigma+1)/\sigma)}{I(\sigma)^{\lambda\sigma+1}}\Gamma(\lambda\sigma+1),$$

where $\Gamma(z) = \int_0^{\infty} t^{z-1}\exp(-t)\,dt$ is the gamma function.

Substituting this expression back into the subjective choice probability for option j, we obtain:

$$p(j) = \frac{\exp(-\lambda V_j)}{Z(\lambda)}\frac{\exp(V_j(\lambda\sigma+1)/\sigma)}{I(\sigma)^{\lambda\sigma+1}}\Gamma(\lambda\sigma+1) = \frac{\exp(V_j/\sigma)}{Z(\lambda)I(\sigma)^{\lambda\sigma+1}}\Gamma(\lambda\sigma+1).$$

Summing the choice probabilities across alternative options:

$$1 = \sum p(j) = \frac{\Gamma(\lambda\sigma+1)}{Z(\lambda)I(\sigma)^{\lambda\sigma}} \qquad \Longrightarrow \qquad Z(\lambda) = \frac{\Gamma(\lambda\sigma+1)}{I(\sigma)^{\lambda\sigma}}. \tag{A.15}$$

Substituting $Z(\lambda)$ back into the expression for the subjective probability distribution p, we conclude

$$p(j) = \frac{\exp(V_j/\sigma)}{I(\sigma)} = q(j),$$

where the last equality follows from (A.11).

Expected Continuation Value. We use the characterization of the subjective probability distribution with respect to the college taste shock to derive the continuation value:

$$\mathbb{E}_\pi\big[M(\varepsilon)\big] = \frac{\mathbb{E}_f[M(\varepsilon)\exp(-\lambda M(\varepsilon))]}{Z(\lambda)}.$$

A standard result in extreme value theory states that if ε_k are independent and identically distributed Type I extreme value shocks with location $\hat{\mu}$ and scale σ, then the maximum $M = \max_k\{V_k + \varepsilon_k\}$ also follows a Type I extreme value distribution with location $\hat{\mu} + \sigma\log I(\sigma)$ and scale σ.[46]

[46]Let all ε_k be independent and identically distributed Type I extreme value random variables with cumulative distribution function $\mathbb{P}(\varepsilon_k \leqslant t) = \exp(-\exp(-(t-\hat{\mu})/\sigma))$. The maximum value is below m with probability $\mathbb{P}(M \leqslant m) = \prod\mathbb{P}(\varepsilon_k \leqslant m - V_k) = \prod\exp\big(-\exp(-(m - V_k - \hat{\mu})/\sigma)\big) = \exp(-\exp(-(m-\hat{\mu})/\sigma)I(\sigma)) = \exp(-\exp(-(m-\hat{\mu}-\sigma\log I(\sigma)))/\sigma)$, which is the cumulative distribution function for a Type I distribution with scale σ and location parameter $\hat{\mu} + \sigma\log I(\sigma)$.

For a Type I extreme value distribution with location μ and scale s, the moment generating function is $\mathbb{E}[\exp(tX)] = \exp(t\mu)\Gamma(1-ts)$ for $ts < 1$. Setting location $\mu = \hat{\mu} + \sigma \log I(\sigma)$, $s = \sigma$ and using $t = -\lambda$, we use the moment generating function in order to write:

$$\mathbb{E}_f[\exp(-\lambda M(\varepsilon))] = \exp(-\lambda(\hat{\mu} + \sigma \log I(\sigma)))\Gamma(1+\lambda\sigma) = \exp(-\lambda\hat{\mu})I(\sigma)^{-\lambda\sigma}\Gamma(1+\lambda\sigma) = Z(\lambda).$$

We next rewrite the numerator of the entropy constraint above. We observe that $\mathbb{E}_f[M(\varepsilon)\exp(-\lambda M(\varepsilon))] = -\frac{d}{d\lambda}\mathbb{E}_f[\exp(-\lambda M(\varepsilon))] = -\frac{d}{d\lambda}Z(\lambda)$. Differentiating $Z(\lambda) = \exp(-\lambda\hat{\mu})\Gamma(1+\lambda\sigma)I(\sigma)^{-\lambda\sigma}$:

$$\begin{aligned}\frac{dZ}{d\lambda} &= -(\hat{\mu} + \sigma \log I(\sigma))\exp(-\lambda\hat{\mu})I(\sigma)^{-\lambda\sigma}\Gamma(1+\lambda\sigma) + \sigma \exp(-\lambda\hat{\mu})\Gamma'(1+\lambda\sigma)I(\sigma)^{-\lambda\sigma} \\ &= \exp(-\lambda\hat{\mu})I(\sigma)^{-\lambda\sigma}\Gamma(1+\lambda\sigma)\left[-(\hat{\mu} + \sigma \log I(\sigma)) + \sigma\psi(1+\lambda\sigma)\right].\end{aligned}$$

Dividing the numerator expression by $Z(\lambda) = \exp(-\lambda\hat{\mu})I(\sigma)^{-\lambda\sigma}\Gamma(1+\lambda\sigma)$, we finally obtain the continuation value of the Lemma.

Entropy Constraint. The entropy is given by:

$$\mathcal{R}(\pi\|f) = \mathbb{E}_\pi\left[\log\frac{\pi(\varepsilon)}{f(\varepsilon)}\right] = \mathbb{E}_\pi\big[-\lambda M(\varepsilon) - \log Z(\lambda)\big] = -\lambda\big[\sigma \log I(\sigma) - \sigma\psi(1+\lambda\sigma)\big] - \log Z(\lambda),$$

where the second equality follows from the characterization of the subjective probabilities (A.14) and the third equality follows from the continuation value. Using the characterization of $Z(\lambda)$ in equation (A.15):

$$\mathcal{R}(\pi\|f) = \lambda\sigma\psi(1+\lambda\sigma) - \log\Gamma(1+\lambda\sigma),$$

We find λ by setting the entropy equal to κ. We note that λ is only a function of κ and σ.

Derivative of Continuation Value with Respect to Assets. We construct the derivative of the expected value of the child. The expected continuation value is $\sigma \log I(\sigma) - \sigma\psi(1+\lambda\sigma)$ with $I(\sigma) = \sum \exp(V_j(a)/\sigma)$ and where λ is determined by setting the entropy to κ. We can differentiate this continuation value with respect to assets:

$$\sum \frac{\exp(V_j(a)/\sigma)}{I(\sigma)}V_{j,a} - \sigma\psi'(1+\lambda\sigma)\sigma\frac{d\lambda}{da} = \sum q(j)V_{j,a} - \sigma\psi'(1+\lambda\sigma)\sigma\frac{d\lambda}{da} = \sum q(j)V_{j,a},$$

where the second equality follows by equation (A.11). The final equality follows since $d\lambda/da = 0$. Relative entropy does not depend on assets since both κ and σ are primitive parameters that are invariant to the transfer a.

E.11 Parameter Identification

We internally estimate $P = 18$ parameters to match $P = 18$ moments following two steps. In the first step, we estimate the model globally. Given a hypercube of the parameter space, we draw approximately 100,000 candidate parameter vectors from uniform Sobol (quasi-random) points, solve and simulate the model, and compute the implied moments in steady state. In the second step, we use the best 400 parameter sets as initial points for a Nelder-Mead local search algorithm. The resulting estimation is shown in Table 4.

We use the approach developed in Daruich (2024) to show which parameters are closely related to each moment, providing a transparent mapping between parameters and moments. Since our approach is exactly identified (equal number of moments and parameters), we can show this mapping one-by-one for each parameter.

Although the model is highly nonlinear, so that almost all parameters affect all outcomes, the identification of each parameter relies on particular moments in the data. Figures E.1 show the result of the following exercise. For each parameter, we associate a target moment and divide the Sobol draws of that parameter into 50 quantiles. For each quantile, we compute the 25th, 50th, and 75th percentiles of the associated moment across all Sobol draws in that quantile. Importantly, the remaining $P - 1$ parameters vary freely within each quantile, so they are generally far from their estimated values.

A moment is informative for a parameter's identification if, as we move across quantiles, the percentiles of the associated moment change monotonically and cross the horizontal dashed line (the value of that moment in the data). The slope of each curve reflects how sensitive the moment is to the parameter: a steeper curve implies the moment is more informative. The gap between the 25th and 75th percentiles indicates the relative importance of other parameters. When these lines are close together, the focal parameter is the dominant determinant of the moment. When they are far apart, at least one other parameter has a quantitatively relevant effect.

Consider some examples. Panel (a) of Figure E.1 shows that average wealth increases with the discount factor β: more patient households accumulate more assets. The narrow interquartile band indicates that other parameters have limited influence on this moment. Panel (b) shows that the intergenerational rank correlation of income increases with altruism δ: higher altruism induces larger transfers and investments, raising intergenerational persistence. Panels (c)–(d) show that child treatment shares identify the child mental health parameters ξ_τ^k and ξ^k. Higher stigma costs reduce treatment uptake among older children, while the utility cost of child illness primarily affects treatment decisions for younger children.

As expected, the treatment moments show wider interquartile bands, reflecting the interaction between stigma and utility costs.

The adult mental health parameters (panels (e)–(h) of Figure E.1) are identified from consumption differences and treatment patterns. The entropy constraint parameters $\kappa(m_1)$ and $\kappa(m_2)$ generate consumption gaps through precautionary savings distortions, while the stigma cost ξ_τ and its age decay ρ_τ are identified from treatment uptake levels and age gradients.

The college parameters (panels (i)–(m) of Figure E.1–E.1) are identified from enrollment regressions. The baseline college taste α_0 governs the overall entry share, while α_1 and α_θ are identified from the coefficients on parental education and cognitive skill in enrollment regressions. The scale of the taste shock χ_s determines the residual variation in enrollment decisions, and the production weight ϖ is pinned down by the college wage premium.

The investment technology parameters (panels (n)–(p) of Figure E.1) are identified from expenditure and time investment data: α_ι from mean financial expenditures, A_ι from the normalization of cognitive skills, and ν from mean time investments. Finally, the fiscal parameters τ_0 and ω (panels (q)–(r)) are identified from the government spending share and the degree of income redistribution, respectively.

Figure E.1: Parameter Identification: Global Results

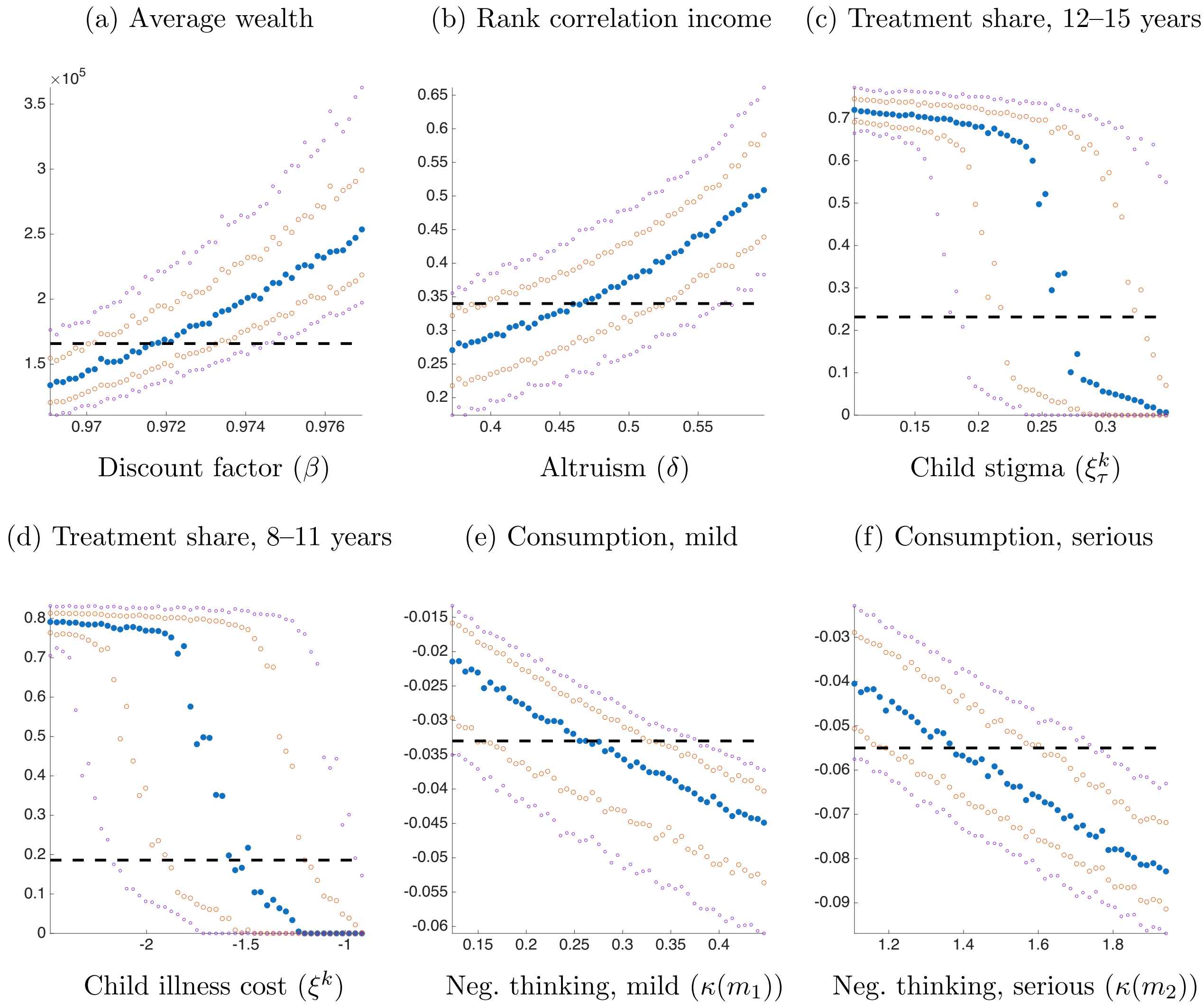
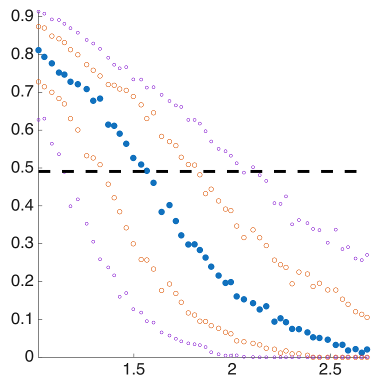


Figure E.1 provides results from the global identification exercise for the parameters in Table 4. The vertical axis of each panel is a target moment (denoted in the panel label) while the horizontal axis is a focal parameter (labeled below the panel). For each parameter quantile, the blue dots show the median of the moment across Sobol draws. The orange squares show the 25th and 75th percentiles, and the purple dots show the 10th and 90th percentiles. The black dashed line shows the value of the moment in the data. See Appendix E.11 for details on the methodology.

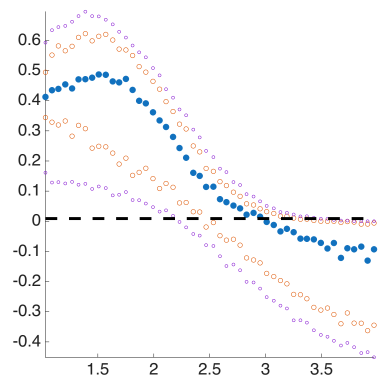
Figure E.1: Parameter Identification: Global Results (continued)

(g) Treatment share, 16–27 years



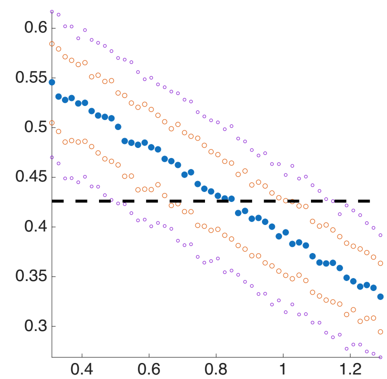
Adult stigma (ξ_τ)

(h) Treatment share, age growth



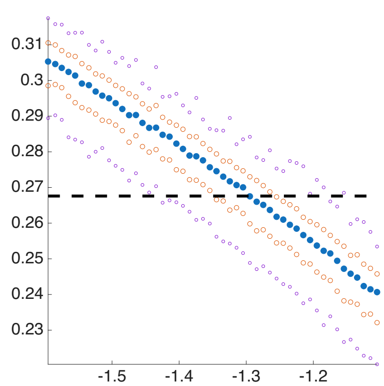
Stigma decay (ρ_τ)

(i) Entry share



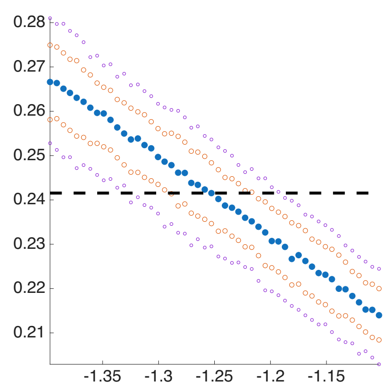
Baseline college taste (α_0)

(j) Entry regression, parents



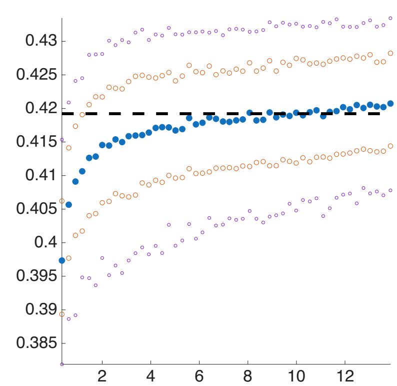
College taste, parents (α_1)

(k) Entry regression, cognitive skill



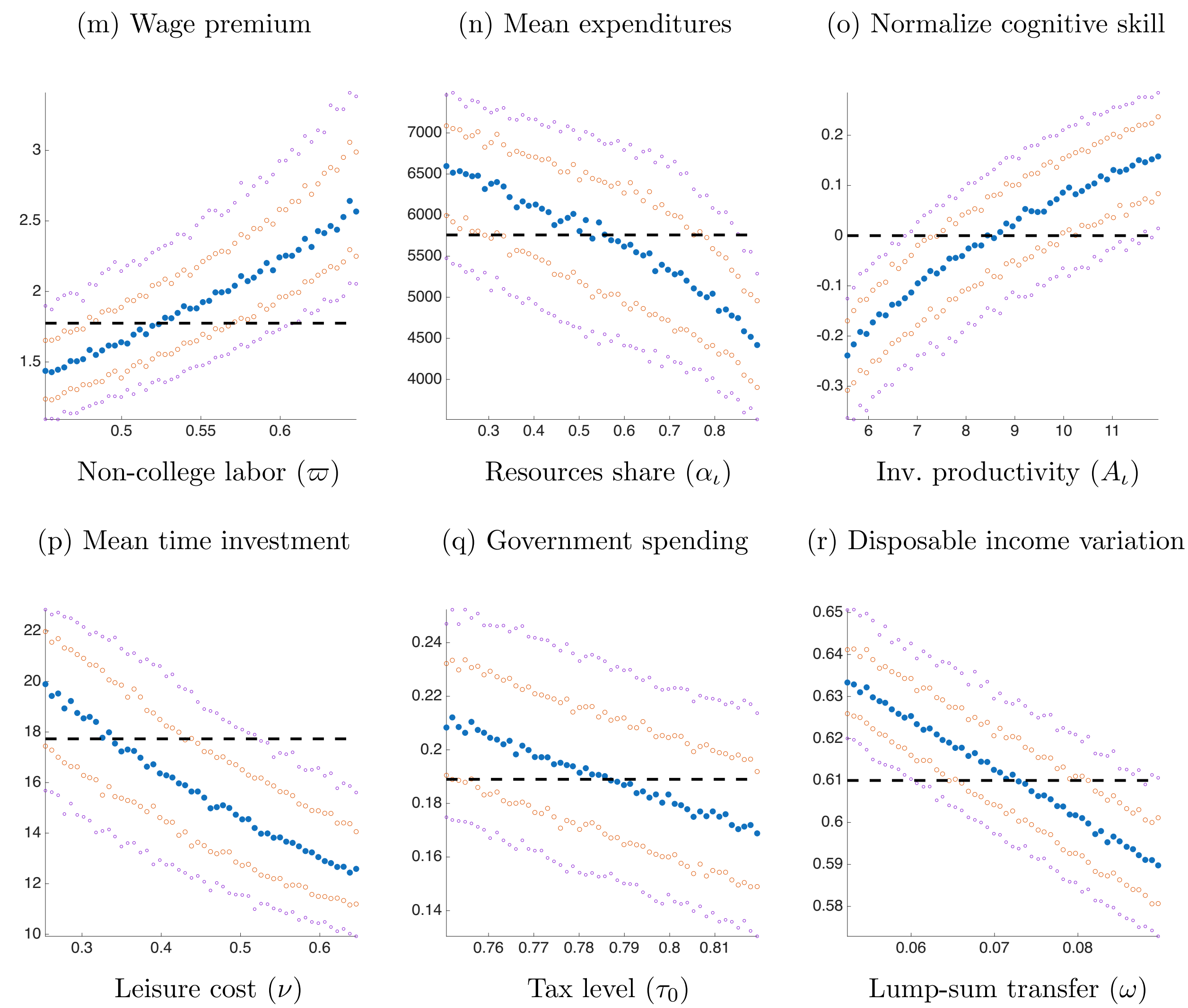
College taste, skill (α_θ)

(l) Entry regression RMSE



Taste shock scale (χ_s)

Figure E.1: Parameter Identification: Global Results (continued)



F Mental Health Policies Examples

We next provide examples of government policies focused on child mental health.

Expanding Mental Health Treatment Services for Children. These policies expand mental health services for children via schools, community health centers, and remote healthcare options. A first policy is to provide school-based treatment services. A second policy is to expand access to treatment through community health clinics.

Federal. Prior to the Covid-era expansion, the most significant federal program dedicated to early childhood mental health was Project LAUNCH, administered by SAMHSA:

- Project LAUNCH (Linking Actions for Unmet Needs in Children's Health)

  Project LAUNCH provides five-year grants to states and tribal communities to promote the social, emotional, cognitive, and behavioral health for children up to age 8. The program implements five core strategies: developmental screening and assessment of children across child-serving settings, integration of behavioral health into pediatric primary care, mental health consultation in early care and education environments, enhanced home visiting focused on social and emotional well-being, and family strengthening and parent skills training. While the primary focus is child mental health, the latter two strategies involve a broader family engagement component, training parents on child social-emotional development and strengthening the home environment.[47]

The most significant Covid-era legislation regarding child mental health at the federal level is the 2021 American Rescue Plan Act (ARPA). While the primary purpose was pandemic recovery, ARPA contained significant investments in children's mental health. For example:[48]

- School based mental health services
  - Elementary and Secondary School Emergency Relief Fund (ESSER) provided state education agencies 122.8 billion dollars in grants to support schools. One of the central allowable uses was to support school mental health systems.
  - ARPA provided 80 million for the Pediatric Mental Health Care Access (PMHCA) program, which promotes integrating care for mental health into pediatric primary care settings. This program trains pediatricians and other primary care providers, including school counselors, to recognize and respond to mental health conditions.
  - ARPA provided 30 million for Project AWARE, a grant program for state education agencies to advance school based mental health services, for example by training school professionals to help them identify and respond to mental health issues and by connecting school-aged youth to services.
- Community health clinics
  - ARPA invested 420 million dollars in Certified Community Behavioral Health Clinics (CCBHC). These clinics are primarily funded through federal grants and agencies and have to provide a

[47] See https://www.samhsa.gov/mental-health/children-and-families/early-childhood/project-launch.

[48] ARPA also provided 150 million for the MIECHV program that we discussed above, and 20 million to support youth suicide prevention programs.

comprehensive range of services including crisis care, outpatient mental health and substance use treatment, and case management. They must serve anyone who requests care regardless of their ability to pay, age, or place of residence. CCBHCs use a sliding fee schedule based on income and accept all insurance types, including Medicaid, Medicare, and private insurance.[49]

The most significant post-Covid legislation on child mental health is the Bipartisan Safer Communities Act (BSCA) enacted in June 2022. The BSCA child mental health provisions included:

- School based mental health services
  - BSCA funds the expansion of school based mental health services through the School-Based Mental Health Services Grant Program and Mental Health Service Professional Demonstration Grant. provided 500 million dollars for training professionals providing mental health services in schools, and 285 million dollars for schools to hire and train mental health counselors. The BSCA also appropriated 1 billion dollars in Stronger Connections Grants for high-need local education agencies to fund positive school climate initiatives and direct school-based mental health services.
  - PMHCA was reauthorized for five years under the BSCA.
- Community health clinics
  - BSCA expanded the CCBHC program nationally, allowing all states to apply to participate beginning in 2024, with up to ten states added every two years.

State. At the state level, many states enacted laws aimed at supporting schools in the delivery of school-based mental health services and at expanding mental telehealth services. Some notable examples are:

- California enacted the Children and Youth Behavioral Health Initiative (CYBHI, 2021), a 4.4 billion dollar statewide effort to support the mental health of young people. The CYBHI expands school-based services, for example by helping public schools, colleges, and school-linked sites get reimbursed for mental health services. Students under age 26 can get these services if they attend a public school or college in California and have private health insurance, disability insurance, or are under California's Medicaid program. Services are free of charge. CYBHI also provides capacity and infrastructure grants to support implementation of behavioral health services in schools and school-linked settings.

[49] 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).

- Florida budgeted 160 million for the 2023-24 school year to assist school districts in establishing or expanding school based mental health care, train educators in detecting and responding to mental health issues, and connect children and families with behavioral health services (CS/SB 1340).
- New Jersey enabled Medicaid to pay for mental health services provided in schools (A 3334, 2023).
- Colorado established a school-based program to administer mental health screenings for students (HB 23-1003, 2023).
- Illinois established a school-based mental health screening program (SB 0724, 2023).
- Massachusetts allocated funds for a pilot program for telemental health services through schools (H 4002, Chapter 24).

Subsidizing Mental Health Treatment for Parents. Poor parental mental health is tied to adverse outcomes for their children. Legislation targeting parent mental health primarily focuses on parent mental health in the perinatal window – from pregnancy through 12 months postpartum. Examples include:

- Into the Light for Maternal Mental Health and Substance Use Disorder Act (2022)
  As part of the Restoring Hope for Mental Health and Wellbeing Act, this law reauthorized and increased federal funding to support and expand maternal mental health screening and treatment programs at the state level.
- National Maternal Mental Health Hotline
  The Health Resources and Services Administration (HRSA) established a free mental health hotline to pregnant women, new mothers, and parents.
- Medicaid Postpartum Coverage Extension
  The American Rescue Plan Act of 2021 gave states the option to extend postpartum Medicaid coverage from the previously mandated 60 days to up to 12 months after delivery.
- Postpartum Mental Health Screening
  Several states have enacted laws which require new parents to be screened for postpartum depression and/or reimburse postpartum mental health screening.
- The Maternal, Infant, and Early Childhood Home Visiting (MIECHV) Program
  This federally-funded program, which is administered by the states, sends nurses, social workers, or parent educators to the homes of at-risk families to provide mental health support, parenting skill-building, and referrals for treatment. These programs explicitly target parental mental health as a mechanism for improving child development. The American Rescue Plan Act of 2021 included significant additional MIECHV funding.

- MOMS Act (2024)

  Passed as part of the National Defense Authorization Act and signed into law in December 2024, this law supports military mothers by establishing a program in the military health care system to provide clinical and non-medical resources to prevent and treat maternal mental health conditions.

Joint Parent-Child Mental Health Programs. A smaller and more recent class of policies treats the parent-child dyad as the unit of intervention, providing mental health services to parents and children simultaneously rather than targeting each separately. These programs remain relatively novel in the policy landscape.

Federal. Two recent federal initiatives have introduced dyadic or multigenerational approaches to mental health service delivery:

- Family First Prevention Services Act (2018)

  Signed into law as part of the Bipartisan Budget Act, this act allows states to use federal Title IV-E matching funds for time-limited prevention services—up to 12 months—in mental health, substance abuse treatment, and in-home parent skill-based programs, specifically for families where a child is at imminent risk of foster care placement. Services are available to both children and their parents or kin caregivers simultaneously. To qualify for reimbursement, interventions must be rated by the Title IV-E Prevention Services Clearinghouse. Approved programs include dyadic therapies such as Child-Parent Psychotherapy and Parent-Child Interaction Therapy, which treat the parent-child relationship as the clinical unit, as well as depression-focused parenting programs and multisystemic therapy for families. The act does not establish a general subsidy for these treatments; rather, it redirects existing child welfare funding toward prevention for at-risk families.[50]

- Infant and Early Childhood Mental Health (IECMH) Grant Program (2018)

  Authorized by Section 10006 of the 21st Century Cures Act (2016) and first funded in 2018, this SAMHSA-administered program funds states and communities to develop mental health promotion, intervention, and treatment services for young children (birth to 12) and their caregivers. Services include multigenerational therapy and mental health screening for both children and parents. Program data indicate that over 26,000 young children and caregivers have received services, and over 38,000 have been screened.[51]

[50] See www.acf.gov/p1. The Title IV-E Prevention Services Clearinghouse maintains a list of approved evidence-based programs at www.acf.gov/p2.

[51] See www.samhsa.gov.

State. At the state level, dyadic service models have begun to enter Medicaid coverage:

- California Medicaid Dyadic Services Benefit (2023)

  Effective January 2023, California became the first state to establish a dedicated Medicaid benefit covering dyadic behavioral health services for children ages 0–20 and their caregivers within a single clinical encounter. Services include mental health screening for both child and caregiver, family counseling, and psychoeducation. The benefit explicitly treats the parent-child dyad as the clinical unit.[52]

G Quantitative Results

G.1 Shapley Value Decomposition

All Shapley decompositions hold prices and taxes fixed at their baseline equilibrium values. For the individual-level decomposition (Tables 6 and E.2), the exercise conditions on a specific mental health state (serious or mild) and evaluates the gains from switching each channel to the healthy benchmark while holding the stationary distribution fixed at the conditional distribution for that mental health state. For the aggregate decomposition (Tables 8–10), each coalition is solved holding wages, interest rates, and the tax rate at their baseline steady-state levels, so that the Shapley values isolate the direct and compositional effects of each channel without confounding from price adjustments.

G.2 Individual-Level Decomposition: Mild Mental Illness

Table E.2 presents the individual-level Shapley decomposition for mild mental illness, analogous to Table 6 in the main text. The CEV gain from becoming healthy is attenuated to 17 percent, corresponding to a one-time wealth gain of about 214 thousand dollars. The relative importance of the different channels is preserved. Negative thinking accounts for 11 percentage points of the CEV, the income effect for 2 percentage points, the child development channel for 2 percentage points, rumination for 1 percentage point, and the initial draw is negligible at 0.3 percentage points.

[52] See www.first5center.org.

Table E.2: Shapley Decomposition of Mental Health Channels: Mild $\rightarrow$ Healthy

	Consumption	Savings	Labor		Parental Investments		Welfare	
			Income	Hours	Time	Money	Consumption	Wealth
Baseline	36,048	163,618	55,577	32.9	15.9	4,925	—	—
Negative Thinking	4.0	−4.1	0.0	0.0	5.9	5.0	11.4	159,600
Rumination	−0.1	0.2	0.0	0.0	0.2	−0.3	1.1	9,700
Income Effect	1.2	2.0	4.9	2.7	−1.1	−1.4	1.7	14,200
Child Development	−0.1	0.2	0.0	0.0	2.6	2.9	2.2	25,300
Initial Draw	−0.2	0.3	0.0	0.0	0.0	0.0	0.3	5,200
Total	4.8	−1.5	4.9	2.7	7.6	6.3	16.8	214,000

See Table 6 for details on the decomposition methodology.