# The Macroeconomics of Intergenerational Mental Health Transmission*

Boaz Abramson
Columbia GSB

Job Boerma
University of Wisconsin-Madison

Diego Daruich
University of Southern California

Aleh Tsyvinski
Yale University

March 2026

Link to Latest Version

**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 microdata on mental health of children and parents. Intergenerational transmission of mental illness arises due to parental time and monetary investments and correlated initial conditions. Adults experiencing mental illness have negative expectations and lose time due to rumination. As a result, parents experiencing mental illness invest fewer resources in their child's mental health. We use the model to evaluate the consequences of policies that have been proposed to improve mental health of children. 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 Mobility

*We thank Anmol Bhandari, Matteo Camboni and Anson Zhou for useful discussions, and Nicholas Garvey for excellence research assistance.

# 1 Introduction

The U.S. Surgeon General declared child mental health as "the defining public health crisis of our time" (Murthy, 2021). More than one in five adolescents in the U.S. are diagnosed with a mental health condition (Sappenfield et al., 2023). Approximately half of all adolescents experience a mental disorder at some point in their life (Merikangas et al., 2010).[1] The costs of mental illness in childhood and adolescence are substantial. Children experiencing mental health problems 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 first quantitative macroeconomic model of child mental health and uses it to study the child mental health crisis. Our model integrates the psychiatric literature on child mental health with quantitative macroeconomic models of child development. The psychiatric literature identifies five key features of child mental health. First, both biological factors and environmental factors are important in shaping child mental health. Second, parent choices plays a key role in shaping child mental health. Third, parent behavior depends on parents' own mental health. Fourth, mental health treatment is effective but taken up by a relatively small share of children experiencing mental illness. Fifth, child mental health is important for later in life outcomes. We use our model to evaluate policies that are often proposed for improving child mental health. In particular, we study the expansion of treatment availability for children in school ages and the subsidization of parental mental health during the perinatal period.

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 as well as environmental factors. Children are born with a mental health state that probabilistically depends on parents' mental health, capturing biological factors that shape mental health. Parents can invest time and money to improve child mental health, capturing that parental choices play a central role in the development of child mental health. Parents also choose whether to seek mental health treatment for their children. Treatment increases the probability of transitioning into

[1]Globally, the World Health Organization estimates that one in seven individuals between the age of 10 and 19 experiences a mental health disorder (see www.who.int). The most common disorders are anxiety disorders, behavior disorders and mood disorders (including depressive disorders.

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 outcomes later in life. They impact their education choices as adolescents and their income as adults.

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 parental investments and child mental health treatment. Rumination implies parents experiencing mental illness have less quality time to spend with their children. As a result of negative thinking and rumination, parents experiencing mental illness may invest less in their child's mental health. This captures the idea that parent behavior depends on parents' own mental health. Overall, in line with child psychiatry, mental illness in our model persists across generations through both exogenous biological factors as well as environmental factors, namely parental investment decisions.

We quantify the model using micro data on mental health of children and parents. We first estimate a child development technology for mental health and cognitive skills. Specifically, 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 Cunha, Heckman, and Schennach (2010) using the National Longitudinal Survey of Youth (NLSY), which measures parents and children mental health and cognitive skills, as well as parental investments, across time. The formation of initial child mental health and child cognitive skill is quantified using the covariance matrix of initial conditions in the estimation of the child development function.

The estimation of the child mental health development function delivers the following new insights. 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 developments are complementary at each development stage and the complementarity increases with age.

Second, we quantify the impact of treatment on the development of child mental health using estimates on the impact of psychotherapy treatment from the medical literature. We set the mean of the treatment efficacy for children experiencing illness and undertaking treatment such that the average model-implied

SMD equals 0.46 in every childhood period in line with the estimates of Weisz et al. (2017).

Third, we parameterize the extent of negative thinking, rumination and labor productivity for each adult mental health state building on Abramson, Boerma, and Tsyvinski (2024). To calibrate negative thinking, we use empirical differences in consumption by mental health status. In the model, negative thinking induces individuals to increase savings and thus decrease consumption conditional on income. We can thus discipline the extent of negative thinking by targeting consumption differences conditional on income. We calibrate the extent of rumination such that the model rationalizes working hours by mental health status. Using fixed effects regressions, we find that individuals experiencing mild (serious) mental illness work 2.7 (15.9) percent fewer hours in the data. 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 internally estimate the remaining model parameters using simulated method of moments. These include parameters governing preferences, the college decision, the parental investment technology, and fiscal policy. Specific to mental health, the utility costs of experiencing mental illness and the stigma costs associated with seeking treatment − for both adults and children − are identified from takeup rates of mental health treatment over the lifecycle.

We validate the model by evaluating model predictions against non-targeted moments that describe the relationship between mental health and economic outcomes. First, we evaluate the model predictions for time investments by parent mental health. In the PSID data, regressing log hours on a depression indicator as well as individual fixed effects yields a decrease of 15 log points, which lines up with the model where hours decrease 24 with parent mental illness. Second, we show that the model-implied take-up for child mental health treatment at early ages aligns with the data. The model correctly predicts near-zero treatment at age 0 and low uptake at age 4. Both model and data exhibit a monotonically increasing age profile, with treatment rising from near zero at birth to roughly one-fifth of unhealthy children by age 12. Third, we compare college graduation expectations. In the data, we find that students experiencing mental illness have 3.6 percentage point lower graduation probabilities, compared to 6.7 percentage point lower graduation probabilities in the model. The model successfully reproduces the qualitative prediction that college students experiencing mental illness hold more pessimistic beliefs about their graduation prospects.

We use the quantitative model to evaluate the consequences of mental illness and child mental health policies. We first decompose how each channel of mental illness – productivity effects, rumination, negative thinking, and the child development channel – affect economic outcomes and welfare. Eliminating all channels of serious mental illness for a single individual generates a welfare gain equivalent to 40 percent of annual consumption. Negative thinking is the dominant channel, accounting for almost three quarters of this gain: individuals with negative expectations invest less time and monetary resources in their children. The remaining channels – rumination and the wage effect – contribute comparatively modest gains. When we evaluate the aggregate consequences of eliminating each channel from the economy, the child development channel emerges as the most important driver of output, accounting for 7.5 out of the 8.5 percent total output increase. This channel operates through intergenerational propagation: parental mental illness reduces the cognitive and mental health development of children, who then enter adulthood as less productive workers and less effective parents. The contrast between individual-level and aggregate results reflects the distinct mechanisms at work: negative thinking affects individual decision-making, generating large welfare losses for affected individuals, while the child development channel shapes the long-run composition of the population, generating large aggregate output effects that compound across generations. These patterns are induced by differences in parental investments and human capital formation: eliminating the child development channel increases parental time and monetary investments by approximately 7 percent each, raises cognitive skills by 14 percent, and lifts college enrollment by nearly 5 percentage points.

We then evaluate prominent child mental health policy proposals in general equilibrium. Subsidizing child mental health treatment raises child treatment uptake from 11 to 36 percent and yields welfare gains of 0.8 percent of annual consumption. Subsidizing parent mental health treatment during the early parenting years generates a smaller, complementary gain of 0.1 percent. A combined policy subsidizing treatment for both children and all parents generates the maximum welfare gain among all policies we consider: 1.0 percent of annual consumption, close to the sum of the individual policies, reflecting that the two interventions operate through distinct channels. Decomposing welfare into compositional and direct effects reveals that the child subsidy and parental subsidies operate through different mechanisms. The child subsidy's fixed-distribution welfare – which evaluates welfare gains holding the population composition at its baseline level – is negative so the entire welfare gain arises from shifts in the long-run population composition as treated children become healthier adults who are themselves more effective parents. In contrast, parental subsidies generate positive fixed-distribution welfare, indicating that their direct benefits to current adults – through reduced negative thinking, rumination and productivity – exceed the fiscal cost even before accounting for intergenerational gains.

**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 parent 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. 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. 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 Literature on Child Psychiatry

The psychiatric, psychological, and epidemiological 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). That is, a portion of the variation in the risk of developing mental illness is attributable to genetic factors.[2] Estimates of heritability tend to be highest for schizophrenia and lowest for mood disorders (Sullivan and Geschwind, 2019).

**Maternal Prenatal Mental Illness**. Maternal mental illness during pregnancy also plays an important role in shaping child mental health. A large body of literature documents that prenatal maternal depression and anxiety are adversely associated with subsequent child mental health indicators (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 post-natal 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.[3] Motivated by this body of work and by psychiatric genetics,

[2]Through classical twin and adoption studies, as well as 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).

[3]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

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 key 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 and adolescent 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 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 model – a set of cognitive psychological structures that govern the way children perceive their self-worth and their 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 investment as a key driver of child mental health.

Parent behavior depends on parent 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.[4] 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.

development. Prenatal depression and anxiety can also lead to epigenetic dysregulation in serotonin transmission that delays offspring cognitive development (Oberlander et al., 2008).

[4] 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).

Overall, and consistent with the psychiatric literature, parents in our model shape child mental health through both biological factors and parenting behavior.

**Other Factors** Other environmental factors can also 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), 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 evidence, mental health in our framework evolves stochastically and is subject to shocks that are independent of parental behavior.

**Treatment**. The main pathway to improving child mental health is treatment. A vast medical literature estimates the effects of psychotherapy and antidepressant medication on child mental health using randomized 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.[5] 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 40.6 percent of 12-17 year-olds experiencing major depression received treatment.[6] 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 devel-

[5]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.

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

opment 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.

# 3 A Simple Model

Before presenting a rich quantitative model, we analyze a special two-period case of this model to highlight key intergenerational transmission channels of mental health. Mental illness is characterized by negative thinking, rumination, and an additional loss of income due to diminished productivity.

The model shows that each feature of mental illness in parents reduces the mental health of children through both biological and environmental factors. Negative thinking about the efficacy of mental health treatment reduces the propensity to seek treatment as an adult, before having children. Since mental health treatment is effective, such negative thinking reinforces mental illness and elevates the chances of experiencing mental illness as a parent. Through biological factors, a child's mental health is correlated with that of a parent, thereby reducing the child's mental health.

In terms of nurture, parents experiencing mental illness make fewer investments in the mental health of their children. Rumination induces parents to spend less time with their child, while negative thinking and lower income reduce the propensity for treatment for children, reinforcing their mental illness.

**Model**. We formalize the intergenerational mental health transmission channels described above in a two-period decision problem. The first period corresponds to adulthood before having a child, while the second period corresponds to parenthood. We first characterize the solution to the decision problem in parenthood, and then for adulthood without a child.

*Parenthood.* Individuals have one child with mental health $m_{k1}$. Parents choose consumption $c$ and time investments $n$ for child development, as well as treatment for the child $\tau_{k1}$ to solve:

$$v_1(m_1, m_{k1}) = \max_{c_1, \tau_{k1}} \left\{ \log c_1 + \min_p \mathbb{E}_p v_2(m_{k2}) \right\}$$

subject to the budget constraint $c_1 + \varphi_\tau \tau_{k1} \leqslant y_1(m_1)$, where household income varies with parents' mental health. The child continuation value is determined by the mental health $m_{k2}$ of the child.

Mental health is formed through a Cobb-Douglas technology:

$$\log m_{k2} = \log m_{k1} + \log m_1 + \log n_1 + \upsilon_\tau \tau_{k1}. \tag{1}$$

The mental health of a child when they become an adult, $m_{k2}$, depends on their mental health as a child $m_{k1}$, the mental health $m_1$ and time investments $n_1$ of parents, and the choice of treatment $\tau_{k1}$. Treatment efficacy follows a symmetrically truncated normal distribution $\upsilon_\tau \sim N(\mu(m_{k1}), \sigma^2_{\upsilon_\tau})$. We assume that $\mu(m_{k1}) = \max\{-\log m_{k1}, 0\}$ such that treatment for children $\tau_{k1}$ improves mental health for children experiencing mental illness, $\log m_{k1} < 0$. In the simple model, treatment thus neutralizes the impact of a child's mental illness on mental health as an adult.

The value function of a child with mental health $m_{k2}$ depends on their consumption as an adult. They consume labor income, which is log-linear with respect to mental health, so that $v_2(m_{k2}) = \log c_{k2}(m_{k2}) = \lambda_m \log m_{k2}$, where $\lambda_m$ is the earnings elasticity with respect to mental health. Using child mental health technology (1), negative thinking can now be formalized as:

$$\min_p \mathbb{E}_p v_2(m_{k2}) = \lambda_m \Big[ \log m_{k1} + \log m_1 + \log n_1 + \tau_{k1} \min_p \sum p(\upsilon_\tau) \log \upsilon_\tau \Big].$$

Negative thinking subject to a total variation constraint implies that mass $\kappa$ is moved from the best treatment efficacy shock realizations to place an additional $\kappa$ point mass at the bottom end of the support, lowering the expected value of treatment. Negative thinking, the expected value of treatment, is represented by $\mathcal{C}(m_1, m_{k1})$, which increases with parental mental health $m_1$.

Having characterized the negative thinking of parents, we remark that parents invest all disposable time with their children. Since disposable time is constrained by rumination and working hours, the time spent with children is $n_1 = 1 - n - n_r(m_1)$, where $n$ are working hours and $n_r(m_1)$ is the rumination time, which decreases with parent mental health. As a result, rumination decreases the time spent with children.

We next analyze the parent treatment decision for their child. The cost of treatment is a reduction in consumption, while the benefit is that treatment improves a child's mental health development. Parents choose to get treatment for their child if and only if the benefit exceeds the costs:

$$\log(y_1(m_1)) - \log(y_1(m_1) - \varphi_\tau) \leqslant \lambda_m \mathcal{C}(m_1, m_{k1}).$$

In order to highlight the intergenerational transmission of mental health, we next consider the impact of negative thinking and reduced income on treatment choice. As the expected benefit of treatment decreases with the severity of adult mental illness due to negative thinking, negative thinking reduces the propensity to seek treatment. Lower income increases the utility cost of treatment, further reducing the propensity to seek treatment. Parents experiencing mental illness thus make time investments and treatment choices that tend to sustain mental illness in the family.

The simple model captures the two broad channels identified in Section 2: biological factors—encompassing genetic heritability and prenatal conditions—and environmental factors, primarily parenting behavior.

**Claim 1.** *Environmental Factors.* Environmental factors sustain mental illness within the family. Negative thinking and income reductions decrease the propensity to seek treatment for children, reinforcing mental illness. Rumination reduces children's mental health by diminishing parental time investments.

*Adulthood.* Agents have mental health $m_0$ which generates income as $y(m_0)$. The decision they make is to get treatment $\tau_0$ to solve $\max_{\tau_0} \big\{ \log\big(y(m_0) - \varphi_\tau \tau_0\big) + \min_p \mathbb{E}_p v_1(m_1, m_{k1}) \big\}$. We assume that initial child mental health is monotonically increasing with parental mental health. The mental health of parents during parenthood is related to their mental health in adulthood as

$$\log m_1 = \rho_m \log m_0 + \upsilon_\tau \tau_0$$

where treatment efficacy follows a symmetrically truncated normal distribution $\upsilon_\tau \sim N(\mu(m_0), \sigma^2_{\upsilon_\tau})$.

Adults choose to get treatment if and only if:

$$\log y(m_0) - \log\big(y(m_0) - \varphi_\tau\big) \leqslant \min_p \mathbb{E}_p v_1(m_1, m_{k1} | \tau_0 = 1) - v_1(m_1, m_{k1} | \tau_0 = 0)$$

The treatment decisions of adults lead to the reinforcement of mental illness through behavior. Consider the impact of mental illness, in terms of negative thinking and lower income, on treatment choice. Negative thinking reduces the expected benefit of treatment, while lower income increases the costs. Hence, both lead to reinforcement through behavior.

**Claim 2.** *Biological Factors.* Negative thinking and the reduction in income decrease the propensity to seek treatment for adults before becoming parents, thus reinforcing mental illness into adulthood. Through biological factors, this reduces the initial mental health of children.

# 4 Quantitative Model

We formalize our theory of child mental health in an overlapping generations model with heterogeneous agents. We consider an infinite horizon economy populated by overlapping generations, where each generation has mass one. Individuals live for 20 periods $t = 1, 2, 3, \ldots, 20$ and die deterministically thereafter.[7] 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 $\ell$. Individuals have preferences that are separable in time and discount the future with a discount factor $\beta$. Total time each period is normalized to one. Agents also care about the utility of their children, modeled following Barro and Becker (1989).

**Adult Mental Health**. Adult mental health status is denoted $m \in \mathcal{M}$, where $\mathcal{M}$ is a finite set. We consider three mental health states: healthy $m_0$, mild illness $m_1$, and serious illness $m_2$. In 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 treatment choice $\tau_t$ and on idiosyncratic labor productivity $z_t$. Negative labor market shocks can thus affect mental health. Adult mental health governs negative thinking, rumination, treatment efficacy, labor productivity, hours worked, flow utility, and the development of child mental health and cognitive skills.

**Negative Thinking**. The first distinctive feature of mental illness is negative thinking. Building on the cognitive model of mental illness (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 $\chi$ in a finite set of realizations $\Omega_\chi$. Let $q(\chi)$ be the objective probability of the outcome. Negative thinking is represented by individuals forming expectations over the random outcome according to:

$$\min_{p} \ \mathbb{E}_p w(\chi) = \min_{p(\chi)} \sum_{\chi \in \Omega_\chi} p(\chi) w(\chi). \tag{2}$$

That is, individuals form expectations based on the subjective probability distribution $p(\chi)$ that minimizes the expected value. Minimization is subject to the relative entropy constraint that limits the choice of

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

subjective probabilities to those that are close to the objective probabilities:

$$\mathcal{R}(p\|q) = \sum_{\chi\in\Omega_\chi} p(\chi)\log\frac{p(\chi)}{q(\chi)} \leqslant \kappa(m). \tag{3}$$

The extent to which subjective probabilities can differ from objective probabilities is governed by $\kappa$, 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 $\chi$ is:

$$p(\chi) = \frac{q(\chi)\exp(-\lambda(m)w(\chi))}{\sum q(\chi)\exp(-\lambda(m)w(\chi))}, \tag{4}$$

where $\lambda$ is the inverse of the Lagrange multiplier with respect to the relative entropy constraint. Adverse outcomes are states with low values $w(\chi)$, and $\lambda$ controls the size of the subjective probabilities distortion. An increase in $\lambda$ corresponds to more negative thinking. The value $\lambda = 0$ implies $p = q$, in which case the subjective probabilities coincide with the objective probabilities. The relationship between $\lambda$ and the extent of negative thinking $\kappa$ is monotonic, implying that the larger $\kappa$, 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, $\kappa$ increases with the severity of mental illness.[8]

**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.

**Productivity**. Mental health affects labor productivity and hours worked. Working-age households receive an hourly wage rate $w_t$. Hourly wages depend on a deterministic age component $\zeta_t^e$ that depends on their education level $e$, mental health $m$, cognitive ability $\theta$, and an idiosyncratic productivity component $\log z_t$:

$$\log w_t(m, z_t, e, \theta) = \log\zeta_t^e + \lambda_m + \lambda_\theta^e \log\theta + \log z_t, \tag{5}$$

[8] 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.

where $\lambda_m$ captures the effect of mental health on wages, and $\lambda_\theta^e$ captures the impact of cognitive skills on (log) wages, which varies by educational attainment. Idiosyncratic productivity log $z_t$ follows a first-order autoregressive process:

$$\log z_t = \rho_t^e \log z_{t-1} + \upsilon_t^e, \tag{6}$$

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

**Income**. Labor income for working-age households is the product of the hourly wage rate $w_t$ and working hours $n_t$, which vary by mental health. Household labor income before taxes is denoted $y_t(m, z_t, e, \theta) = w_t(m_t, z_t, e, \theta)n_t(m_t)$. In retirement, households receive pension income $y_t^p(\theta, e)$ which depends on cognitive skills and educational attainment.

**Mental Health Treatment**. Adults decide whether to seek mental health treatment. Let $\tau_t = 0$ denote that the adult does not undertake treatment, and let $\tau_t = 1$ denote that the adult undertakes treatment. Treatment increases the probability of transitioning into better mental health states. Treatment is associated with a time cost $n_\tau$, a financial cost $\varphi_\tau$, and a utility cost $\xi_\tau$. We introduce a utility cost of treatment to capture the stigma associated with mental health treatment. The psychiatric literature identifies stigma as an important factor contributing to low mental health treatment rates, despite the efficacy of treatment (see, for example, Corrigan (2004) and Clement et al. (2015)).

Adults also decide whether to seek mental health treatment for their child. We denote by $\tau_t^k = 0$ if the child does not undertake treatment and by $\tau_t^k = 1$ if the child does undertake treatment. As with adults, treatment probabilistically improves child mental health. When a child undertakes treatment, the parent incurs a time cost, $n_\tau^k$, a financial cost, $\varphi_\tau^k$, and a utility cost, $\xi_\tau^k$.

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

$$\ell(m, e, \tau, \tau_k) = 1 - n_w(m) - n_r(m) - n_\tau \tau - n_\tau^k \tau_k - n_e e, \tag{7}$$

where $n_e$ is the time individuals spend studying if they choose to enroll in college.

**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 $R_b \geqslant R$. Borrowing

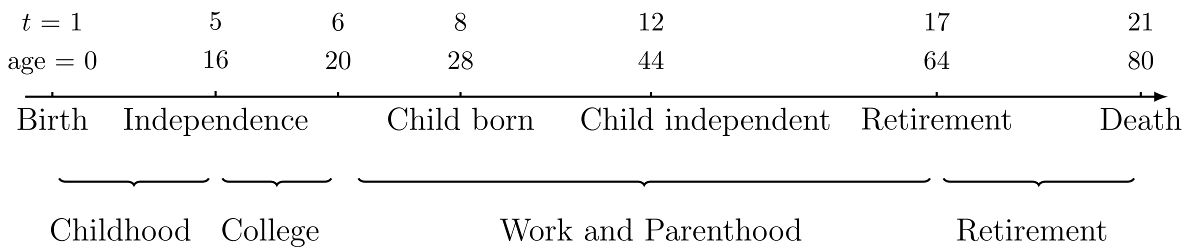



Figure 1: Lifecycle in the Model with Four Stages

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

or negative net asset positions are considered positions of unsecured debt, and the difference between interest rates captures the cost of unsecured borrowing. Borrowing limits vary by age, that is,

$$a_{t+1} \geqslant \underline{a}_{t+1}, \tag{8}$$

where borrowing limits $\underline{a}_{t+1}$ are equal to the natural borrowing limit (see Section 5.2). Student loans are explained in detail below.

## 4.1 Life-Cycle Problem

The lifecycle is divided into four stages: childhood, college, parenthood, and retirement. Figure 1 presents the timing of these stages. Each period corresponds to four calendar years, that is, $t = 1$ refers to ages $0{-}3$, $t = 2$ to ages $4{-}7$, and so on. Childhood spans period 1 through 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 starts in period 5. The first decision individuals make is whether to attend college. After the college stage, the individual works and becomes a parent in period 8. Individuals retire at age 64, or period 17, and die deterministically at age 80, at the beginning of period 21.

## 4.2 Education Stage and College Decision

In period 5, individuals choose between going to college for one period ($e_t = 1$) or directly entering the labor market as non-college workers ($e_t = 0$). Individuals make this decision based on five state variables: assets $a_t \geqslant 0$, mental health $m_t$, cognitive ability $\theta$, an indicator for whether their parents are college graduates $e_p$, and school taste shocks $\chi$. College education is associated with a time cost $n_e$ and a monetary cost $\tau_e$. It can be financed by parental transfers or with subsidized loans. Subsidized loans come with 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. Note that individuals experiencing mental illness think negatively about their probability of completing college, which can affect their enrollment decisions.

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

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

subject to the entropy constraint (3), the borrowing constraint (8), and the budget constraint:

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

where $y_t = y_t(m_t, z_t, e_t, \theta)$ is labor income during college. Students receive labor income while in college, but their hours worked are reduced by studying hours $n_e$ and $\log z_t = 0$. The initial wage shock $\log z_{t+1} \sim N(0, \sigma^e_{t+1})$ depends on the education achieved $e_{t+1}$ (see Section 5.2 for details). College students can access subsidized loans at a rate $R_e$, and these loans are subject to the borrowing constraint (8). 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 (i.e., 20 years) following graduation and an interest rate $R_b$, where $\tilde{a}_s(a_{t+1})$ is the function performing this transformation (see Appendix E.1 for details). Income and consumption taxes are captured by the function $T$, which depends on labor income, capital income, and consumption.

Individuals can decide not to enroll in college ($e_t = 0$), in which case they directly enter the labor market. Upon entering the labor market, individuals draw an idiosyncratic productivity shock $z_t$ where $\log z_t \sim N(0, \sigma^0_t)$. Individuals choose consumption $c_t$, savings $a_{t+1}$, and treatment $\tau_t$. The value of choosing not to enroll in college is given by:

$$v_t(a_t, m_t, z_t, e_t, \theta) = \max_{c_t, a_{t+1}, \tau_t} \Big\{ u(c_t, \ell_t(m_t, e_t, \tau_t, 0)) - \xi_\tau \tau_t + \beta \min_p \mathbb{E}_p v_{t+1}\big(a_{t+1}, m_{t+1}, z_{t+1}, e_t, \theta\big) \Big\} \tag{9}$$

where $e_t = 0$ and subject to the entropy constraint (3), the borrowing constraint (8), and the budget constraint:

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

where $y_t = y_t(m_t, z_t, e_t, \theta)$.

At the beginning of period 5, the individual chooses between going to college and directly entering the labor market. 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. Their value function is given by:

$$v_5^e(a_t, m_t, \theta, e_p, \chi) = \max_{e_t} \Big\{ \min_{p} \mathbb{E}_p v_t(a_t, m_t, z_t, 0, \theta) + \chi_0, v_t(a_t, m_t, e_t = 1, \theta) + \Upsilon(\theta, e_p) + \chi_1 \Big\} \tag{10}$$

subject to the entropy constraint (3). $\Upsilon$ is a direct flow utility from college and depends on cognitive ability $\theta$, parental education $e_p$, and an additive school taste $\chi$.[9]

## 4.3 Work and Parenthood

An individual is a childless adult in periods 6 and 7. The optimization problem in these periods is given by (9). Adults deterministically have one child at the beginning of period 8. Children are born with cognitive skills $\theta_{kt}$ and mental health $m_{kt}$, which are drawn from a joint distribution that depends on parental cognitive ability $\theta_t$ and parental mental health $m_t$. For periods 8 to 11, which correspond to ages 0 to 15 for the child, parents can invest time $n_k$ and financial resources $x$ into developing their child's skills and mental health. While the child is not yet independent, parents purchase consumption goods for their child $c_k$. Children have the same utility function $u$ as adults, but they do not work. Parents care about the utility of their child with an altruism factor $\delta$.

**Child Development**. The development of a child's mental health and cognitive ability is modeled using a CES specification. Specifically, the evolution of a child's mental health, $m_{kt}$, depends on their cognitive ability and mental health, their parent's cognitive ability and mental health, as well as parental investments and mental health treatment:

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

where $\upsilon_{mt} \sim N(0, \sigma^2_{t\upsilon_m})$ is an i.i.d. mental health shock. The efficacy of treatment for children is given by $\upsilon_\tau$ and follows a normal distribution $\upsilon_\tau \sim N(\mu_t(\tau_{kt}, m_{kt}), \sigma^2_{t\upsilon_\tau}(\tau_{kt}, m_{kt}))$. The investment aggregate $\iota_t$ is given by:

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

where $x$ are parental monetary investments and $n_{kt}$ are parental time investments. Mental health treatment for children enters the development function multiplicatively through the term $e^{\upsilon_\tau \tau_{kt}}$, where the

[9] We emphasize that the college decision is made by the young adult. Negative thinking regarding the value of not going to college entails negative thinking by the young adult regarding the first productivity realization $z_t$. School taste $\chi$ is realized prior to the college decision and is introduced to account for observed variation in educational patterns.

treatment effect $\upsilon_\tau$ is drawn from a distribution with a positive mean. This implies that treatment probabilistically improves child mental health.

The evolution of a child's cognitive ability follows an analogous specification to mental health, but there is no direct impact of mental health treatment. The evolution depends on their own cognitive ability and mental health, their parent's cognitive ability and mental health, and parental investments:

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

where $\upsilon_{\theta t} \sim N(0, \sigma^2_{\upsilon_\theta})$ 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 $x_t$, and time investments $n_t$ in child development, as well as mental health treatment for the parent $\tau_t$ and the child $\tau_{kt}$ to solve[10]

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

subject to the entropy constraint (3), the borrowing constraint (8), the childhood skill development production function, and the budget constraint:

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

The parameter $\nu \in [0, 1]$ captures the extent to which childcare time counts as work versus leisure. Parents experience a direct utility cost $\xi_k$ when their children experience mental illness. This parameter allows the model to match the age pattern of child mental health treatment in the data: because the continuation value of improving a child's mental health – which operates through better adult outcomes – is discounted more heavily in early childhood, the direct utility cost provides an additional incentive for parents to seek treatment early on.

**Inter Vivos Transfer**. The transfer decision is the final channel through which parents affect their child's outcomes, determining the child's initial wealth at independence. Shortly before period 12 (i.e., before the child becomes independent), parents can transfer $\hat{a} \geqslant 0$ to their child.[11] Parents transfer at most all their wealth and borrow up to the borrowing constraint, that is, $\hat{a} \leqslant a_t - \underline{a}_t$. After receiving the transfer, the child becomes independent and decides whether to go to college. The transfer is made in

[10] Period t=11 is slightly different in that the continuation value is the transfer subperiod $v^{\text{iv}}$ rather than $v_{t+1}$.

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

a subperiod where the parent knows the state $(a_t, m_t, z_t, e, \theta, m_{kt}, \theta_k)$ but before the realization of their child's college preference shocks $\chi$. The inter vivos transfer problem is:

$$v^{\text{iv}}(a_t, m_t, z_t, e, \theta, m_{kt}, \theta_k) = \max_{\hat{a} \geqslant 0} \left\{ v_{12}(a_t - \hat{a}, m_t, z_t, e, \theta) + \delta \min_p \mathbb{E}_p v_5^e(\hat{a}, m_{kt}, \theta_k, e, \chi) \right\},$$

subject to the entropy constraint (3).

The min operator before the child's continuation value reflects the parent's negative thinking: a parent experiencing mental illness evaluates the child's education-specific taste shocks $\chi$ under distorted probabilities. However, because these shocks follow an EV1 distribution, the entropy-constrained distortion preserves the child's choice probabilities, and its level effect cancels out in the first-order condition for the optimal transfer (see details in Appendix E.10). The child's expected continuation value is therefore unaffected by negative thinking. Instead, negative thinking affects transfers through the parent's continuation value $v_{12}$: the parent's own problem at age 12 is solved under distorted expectations of future income and mental health outcomes, which inflates the marginal utility of their own future consumption at the time of the transfer decision. This raises the value of retaining wealth and tilts the first-order condition away from transfers and toward own savings.

After making the transfer, the individual reverts to the standard consumption and savings problem without a child (9) in the periods leading up to retirement.

## 4.4 Retirement Stage

Individuals are retired from period 17 to period 21. They supply no labor during retirement; their only sources of income are asset income and pension income $y_t^p(e, \theta)$. The consumption and savings problem is:

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

subject to the entropy constraint (3), the borrowing constraint(8), and the budget constraint:

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

where we allow pension income to be a function of education and cognitive skills.

## 4.5 Aggregate Production

Firms produce using a Cobb-Douglas technology with a capital share $\alpha$. Aggregate output $Y = AK^\alpha N^{1-\alpha}$, where $A$ is total factor productivity and $K$ denotes capital. Aggregate labor is represented by 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 effective labor input of non-college graduates, and $N_1$ is the effective labor input of college graduates. Firms rent capital and hire labor in competitive factor markets. In the small open economy, the capital-labor ratio equalizes the marginal product of capital to the world interest rate $R$ plus depreciation $\delta_K$. Education-specific wages $w_e$ are determined in general equilibrium by the marginal product of each labor type. The government budget constraint and equilibrium definition are presented in Appendix B. When a policy is introduced, exogenous government expenditures are held constant at their initial steady-state level, and the government budget is rebalanced through the labor-income tax rate $\tau_0$.

# 5 Model Quantification and Validation

We quantify the model using U.S. data. All monetary values are expressed in 2015 dollars. Economic variables − including consumption, earnings, time, and wealth − are measured at the per-adult level within the household to align with the model's unit of analysis. We first describe the data sources used for quantification, then present the externally estimated parameters, the estimated child development technology, and the internally calibrated parameters.

## 5.1 Data on Child Mental Health and Economic Outcomes

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). A key feature of the PSID is that it records the mental health of respondents, which allows quantifying the relations between mental health and economic outcomes such as consumption, labor supply and wealth. A key feature of the CDS and TAS is that they record child mental health of PSID respondents throughout childhood and early adulthood. By linking parents in the PSID to their children in the CDS and TAS, we measure the relation between parent and child mental health as well as the persistence of mental health from childhood to adulthood. The CDS also records time investments of parents in children and whether a child receives mental health treatment, which allows us to measure how these investments vary with parent and child mental health. Finally, the TAS records young adults' expectations regarding college graduation as well as their realized college enrollment and graduation outcomes. A key feature of the NLSY is that it measures mothers' and children's cognitive abilities 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.

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 the primary mental health measure used in U.S. government administered health surveys as well as the WHO World Mental Health Surveys. [12] The K6 scale has been extensively validated against clinical mental health diagnoses and has been shown to consistently predict clinical diagnoses of mood and anxiety disorders (Kessler et al., 2002, 2003; Furukawa, Kessler, Slade, and Andrews, 2003; Cairney et al., 2007). We classify adults into the three mental health states (healthy, mild, serious) based on the K6 scale following Kessler et al. (2008).[13]

We measure child mental health in the CDS using the Behavioral Problem Index (BPI). The BPI measures the frequency and types of childhood 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.[14] 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

[12]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.

[13]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.

[14]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.

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 according to the K6 scale in the TAS.

Our measures for parental investments, maternal cognitive skills, and child cognitive skills in the NLSY are constructed as in Cunha, Heckman, and Schennach (2010). Our measures for child mental health are equivalent to the measures of non-cognitive skills in Cunha, Heckman, and Schennach (2010). Our measures for maternal mental health are various subscales of the Center for Epidemiologic Studies Depression Scale (CES-D).

## 5.2 Externally Estimated Parameters

*Taxes and Transfers.* Taxes are the sum of taxes on labor earnings, asset income and consumption. Labor earnings are taxed according to the nonlinear function used by Feldstein (1969); Benabou (2002); Heathcote, Storesletten, and Violante (2014), where after-tax earnings $\tilde{y}$ are a power function of pre-tax earnings $y$, that is $\tilde{y} = (1-\tau_0)y^{1-\tau_1}$. Consumption is taxed at a linear rate $\tau_c = 0.05$, while asset income is taxed at a rate $\tau_a = 0.36$ when positive following Trabandt and Uhlig (2011). The progressivity parameter $\tau_1 = 0.185$ is directly taken from the estimates of Heathcote, Storesletten, and Violante (2014). $\tau_0$ is estimated internally to match the size of government spending relative to GDP. Finally, all households receive a lump-sum transfer $\omega$ which is estimated internally to match the amount redistribution in the data.

*Time Allocation.* Individuals have a unit time endowment, which corresponds to 100 hours per week. We calibrate working hours, rumination, and time costs of treatment following Abramson, Boerma, and Tsyvinski (2024) using average per-adult measures at the household level. Work takes 33.75 hours per week when individuals are healthy, or $n(m_0) = 0.3375$. To determine the working hours of individuals experiencing mental illness, we use PSID data and estimate what happens to working hours for a given individual as they transition between mental health states. In particular, we regress (log) working hours on individual fixed effects. Individuals experiencing mild mental illness work on average 2.6 percent fewer hours, while individuals experiencing serious mental health problems work on average 15.8 percent fewer hours, which determines $n(m)$. Healthy individuals do not ruminate, while individuals who experience mild and serious illness ruminate 1.3 and 3 hours per week, so that $n_r(m_0) = 0$, $n_r(m_1) = 0.013$, and $n_r(m_2) = 0.03$. Based on Abramson, Boerma, and Tsyvinski (2024), we set the time cost of mental health treatment for adults and for children to two hours per week $n_\tau = 0.02$. Lastly, if individuals enroll in college, they spend $n_e$ hours per week in school. We estimate $n_e = 0.3375 - 0.104$ to match the average working hours of college students.[15] Leisure is determined residually from the time constraint described in the model section.

*Earnings.* Earnings are the product of hours worked $n(m)$ and the wage per hour $w_t(m_t, z_t, e, \theta)n_t(m_t)$. We estimate the components of this equation using PSID and NLSY79 data. Using PSID data, we regress (log) wages on a third-order age polynomial by education level and on indicator variables for mild and serious mental illness. We include individual-level fixed effects that control time invariant characteristics at the individual level like cognitive skills as in our model. The third-order polynomial in age informs the deterministic life-cycle component, $\log \zeta_t^e$, while the coefficients on the indicator variables for mental illness capture how hourly wages vary with mental health. We estimate a decrease of 2.2 percent in hourly wages associated with mild mental illness, $\lambda_m(m_1) = -0.022$, and 5.5 percent decrease associated with serious mental illness, $\lambda_m(m_2) = -0.055$.[16] We estimate the second two components of the wage process using NLSY79 data.

*Assets.* We set the annual real gross return on assets $r = 0.0524$, corresponding to the annual real returns on safe assets in Jordà, Knoll, Kuvshinov, Schularick, and Taylor (2019) between 2001 and 2020. The net interest rate on borrowing $r^b = r + \iota^b$ with $\iota^b = 0.12$ to reflect the average credit card borrowing rate reported by Gross and Souleles (2002). College students have access to subsidized loans at rate $r^e = r + \iota^e$. We focus on Stafford loans, which are the dominant source of college borrowing. We set $\iota^e = 0.009$ using the weighted average interest rate on different types of Stafford loans (Daruich and Kozlowski, 2020). Gross returns are equal to the net returns plus one, $R = 1 + r$.

Individuals can borrow up to the natural borrowing limit. Thus, individuals are not allowed to borrow in the final period, $\underline{a}_{21} = 0$. In period 20, individuals are allowed to borrow up to the lump-sum government transfer $\omega/R_b$, and generally for $t \leqslant 19$:

$$\underline{a}_t = -\omega \sum_{j=t}^{20} R_b^{j-20}. \tag{14}$$

[15] 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).

[16] 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 of $0.4 \times 2.6 = 1.1$ 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).

The borrowing limit does not depend on education. This ensures that we do not get unreasonable behavior around the college decision when there is a probability of not passing college.

*Preferences.* Individuals have flow utility over consumption $c$ and leisure $\ell$ given by:

$$u(c,\ell) = \log c + \psi \frac{\ell^{1-\frac{1}{\eta}} - 1}{1-\frac{1}{\eta}}\ , \tag{15}$$

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 $\delta$.[17] We choose the parameter $\eta$ so that the Frisch elasticity of labor supply for an average person who does not experience mental illness and works the average amount of hours of healthy workers ($\bar{n} = 0.3375$) equals 0.55, following Chetty, Guren, Manoli, and Weber (2012). To align with the Frisch elasticity of labor supply for these workers in the model, we require $\eta = \frac{\bar{n}}{1-\bar{n}} \frac{0.55}{1-\tau_1(1+0.55)} = 0.521$.

*College.* We take the following specification for the direct utility from going to college:

$$\Upsilon(\theta, e_p) = \chi_s \big[\alpha_0 + \alpha_1 \mathbf{1}_{\{e_p=1\}} + \alpha_\theta \log \theta\big], \tag{16}$$

The college taste shock $\chi$ is distributed according to a Type I extreme value distribution with scale $\chi_s$. The direct utility from college is scaled by $\chi_s$, the scale parameter of the taste shock distribution. This normalization ensures that the parameters $\alpha_0$, $\alpha_1$, and $\alpha_\theta$ are estimated in units comparable to the taste shock, simplifying identification.

In the TAS data, we estimate a linear probability model for the graduation probability conditional on enrollment. We quantify a baseline graduation probability of 50 percent, which increases with 23.6 percentage points per standard deviation of cognitive skills (see Appendix C.2). In the model, the graduation probability $p_e(\theta)$ is a linear function of cognitive skills, clipped to the unit interval, with a baseline probability of 50 percent that increases by 23.6 percentage points per standard deviation of cognitive skills, matching the linear probability model estimated from the TAS data.

*Adult Mental Health Treatment.* We quantify the mental health transition matrix for adults $\Gamma_m(\tau, z)$ following Abramson, Boerma, and Tsyvinski (2024). We report the mental health transition matrix as a function of idiosyncratic labor market shocks and treatment in Table 1. 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

[17] 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).

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 two years ahead $m'$.

efficacy of psychotherapy as estimated by the psychiatric literature. For our calibration, we use an average SMD of $-0.66$ from the meta-analysis by Cuijpers, Andersson, Donker, and Van Straten (2011). The second takeaway is that bad labor market shocks increase the likelihood to experience mental illness in the future consistent with the unconditional transitional probabilities from the PSID. For example, the likelihood to transition from the healthy state into serious (mild) illness is 2.2 (7.7) percent in normal productivity states, while it is 2.8 (7.8) percent in low productivity states.[18]

The financial cost of treatment is set using data from the Medical Expenditure Panel Survey (MEPS) documented in Cronin, Forsstrom, and Papageorge (2025), who 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 $\varphi_\tau$ of 1,250 dollars, and insurer cost of $g_\tau$ of 5,300 dollars.

*Production.* We set the labor share as the compensation of employees in the divided by national income net of proprietors income between 2000 and 2020 which gives $\alpha = 0.692$. 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 annual rate of depreciation $\delta_K = 0.065$ equals the average annual depreciation rate of private fixed assets in the national income and product accounts.

[18] The cutoff distinguishes low and normal productivity states for the mental health transition matrix. We set $\underline{z}$ at the 25th percentile of the stationary distribution of the idiosyncratic productivity process. In the model, we define low as the lowest value of $z_t$.

Table 2: Child Development Function

| | Mental Health ($s = m$) | | Cognitive Skill ($s = \theta$) | |
|---|---|---|---|---|
| | 1st Stage | 2nd Stage | 1st Stage | 2nd Stage |
| Cognitive skill ($\alpha_{1st}$) | 0.000 | 0.000 | 0.491 | 0.857 |
| | (0.028) | (0.091) | (0.027) | (0.012) |
| Mental health ($\alpha_{2st}$) | 0.531 | 0.797 | 0.000 | 0.000 |
| | (0.035) | (0.012) | (0.026) | (0.005) |
| Parent cognitive skill ($\alpha_{3st}$) | 0.020 | 0.000 | 0.044 | 0.062 |
| | (0.013) | (0.007) | (0.013) | (0.008) |
| Parent mental health ($\alpha_{4st}$) | 0.348 | 0.166 | 0.297 | 0.040 |
| | (0.032) | (0.015) | (0.027) | (0.011) |
| Investments ($\alpha_{5st}$) | 0.102 | 0.037 | 0.168 | 0.041 |
| | (0.023) | (0.008) | (0.015) | (0.007) |
| Complementarity ($\rho_{st}$) | −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 child development technology for mental health (second and third column) and for cognitive skill (fourth and fifth column).

## 5.3 Child Mental Health and Cognitive Skill Development

We next estimate the child skill development technology using the framework of Cunha, Heckman, and Schennach (2010). Like the parameters in the previous subsection, these are estimated externally — that is, without simulating the structural model — but we present them separately given the importance of the child development technology for our results. Specifically, we estimate the development of child mental health and cognitive skill as a function of parental investments as well as the current mental health and cognitive skill of both parent and child. We estimate CES development functions for mental health and cognitive skill, with parameters varying with the age of the child, following equations (11) and (13).

Table 2 shows the parameter estimates and the corresponding standard errors. First, the results show that self-productivity of child mental health is large and increasing with age. This is shown by the weight on child mental health ($\alpha_{2mt}$) in the first two 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 column.

Second, the results display no evidence that cognitive skill affects child mental health development as seen by the weight on cognitive skill ($\alpha_{1mt}$) in the first two rows and columns of Table 2. Child mental health does also not directly improve cognitive skill development in rows 3 and 4. In sum, there appears to be no cross-productivity between child mental health and cognitive skill.

Third, row 7 shows that 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. Row 5 displays that parental cognitive skills ($\alpha_{3st}$) have an increasing importance on child cognitive skill of the child, while having negligible impact on the mental health development. Row 9 displays that the impact of parental investments on child development as captured by $\alpha_{5st}$ decreases with age.

Fourth, row 13 displays 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.[19] 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.

A difference between child mental health and adult mental health is that we consider mental health of children as a latent variable. Unlike for adults, where K6 measures mental health, we do not use a direct measurement for child mental health. To compare the model to data, we simulate in the model the measurement variable for child mental health using the measurement equation with the latent model child mental health state and noise.[20] We classify a child as experiencing mental illness if their measured BPI is below a threshold. This threshold is set such that the share of 18 to 19 year-olds classified as experiencing mental illness according to the K6 index in TAS aligns with the share of 16 to 17 year-olds classified as experiencing mental illness according to BPI in the CDS.

We transform the estimated parameters from 2-year periods to 4-year periods. To go from 2-year

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

[20] Following Cunha, Heckman, and Schennach (2010), we use five regressors for each measurement equation: child sex, cohort dummy (1 if child is born after 1987), dummy if 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. 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 of the regressors when simulating the measurement variable.

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

| | Child cognitive skills $\theta_k$ | 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 $\sigma^2_\varepsilon$ | 0.165 | 0.064 |
| Residual covariance | −0.0043 | |

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

periods to 4-year periods, we iterate on the child production function and assume that the shock $\nu$ only takes place in the last iteration. The self-productivity parameter is squared, while other parameters inside the CES function need to be multiplied by 1 plus the persistence parameter. For the shock, in the baseline exercise we assume that the variance of the shock in the 4-year model is twice the one in the 2-year model.

*Elasticity of Parental Investments* The elasticity of substitution between parental time and financial investments in the investment aggregate $\iota_t$ is governed by $\gamma$. 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 initial draw of mental health and cognitive skills for the child are a linear function of the mental health and cognitive skills of their parent:

$$\log \theta_{k1} = \beta_{1\theta} \log \theta + \beta_{2\theta} \log m + \varepsilon_\theta \qquad \text{and} \qquad \log m_{k1} = \beta_{1m} \log \theta + \beta_{2m} \log m + \varepsilon_m \tag{17}$$

where $\varepsilon_\theta$ and $\varepsilon_m$ are both mean-zero error terms independent of the parental characteristics. To calibrate the initial draw for the child, we need to parameterize four coefficients $\beta$, two variances for the error terms, and their covariance. In Appendix E.5, we show how all parameters are identified of the covariation between cognitive skills for the parent and child and mental health for the parent and child that we obtain from the estimation of the child development technology.

The intergenerational transmission operates primarily through the cognitive channel for child mental health. Parent cognitive skills are a strong predictor for child mental health ($\beta_{1m} = 0.085$), while parent mental health $m$ contributes modestly ($\beta_{2m} = 0.057$). For child cognitive skills, both parent characteristics have smaller effects. The residuals have weak negative covariation (−0.0043), indicating that the child's initial cognitive and mental health draws are nearly independent.

*Child Mental Health Treatment* We calibrate the efficacy of treatment for children using estimates from the medical literature. Estimates of the efficacy of psychological therapy are reported in the terms of standard mean differences (SMD), which for children is estimated to be 0.46 (Weisz et al., 2017).[21] We normalize the mean of the treatment efficacy $\mu_t(\tau_{kt}, m_{kt})$ to zero for individuals who do not take treatment, that is, $\mu(0, m_{kt}) = 0$. The standard deviation of the mental health shock by age $\sigma_{t\upsilon_m}$ is set to align with our estimated error variance of the mental health factor by each period.[22] Next, we set the mean of the treatment efficacy for children experiencing mental illness and undertaking treatment $\mu(1,1)$ such that the average model-implied SMD equals 0.46 in every childhood period. The standard deviation of the treatment efficacy is set to correspond to the 95 percent confidence interval of the SMD from 0.41 to 0.51 in Weisz et al. (2017). Given childhood mental health technology (11), and since $\mu(0,1) = 0$, the model-implied SMD of treatment is $\mu(1,1)/\sigma_{t\upsilon_m}$ and we set $\mu(1,1)$ such that the SMD equals the empirical estimate of 0.46.[23] We set $\sigma^2_{t\upsilon_m}(1,1)$ such that the 95 percent confidence interval of the SMD equals the empirical estimate of 0.10, which is $4\sigma_{t\upsilon_\tau}(1,1)/\sigma_{t\upsilon_m}$.

## 5.4 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 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 $\beta$ to calibrate average household wealth in our PSID sample, which is 166 thousand dollars. The altruism factor $\delta$ is calibrated to match the

[21]Psychological therapy is the main form of child mental health treatment and is recommended as the treatment of choice for young people within practice guidelines of medical disciplines (see, e.g., Birmaher and Brent (2007)) and government policy documents (see, e.g., www.cdc.gov and Weisz, Sandler, Durlak, and Anton (2005)).

[22]In order to obtain the estimated error variance of the mental health factor for each period, we linearly interpolate between the estimate for the two stages in Section 5.3.

[23]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 $\mathbb{E}[\log m_{kt+1}|\tau = 1] - \mathbb{E}[\log m_{kt+1}|\tau = 0] = \mathbb{E}[\log \vartheta_t|\tau = 1] - \mathbb{E}[\log \vartheta_t|\tau = 0] + \mathbb{E}[\upsilon_{mt}|\tau = 1] - \mathbb{E}[\upsilon_{mt}|\tau = 0]$, 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.

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 wealth, 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 parent investment technology (12) is informed by moments on financial and time investments. We calibrate share parameter $\alpha_\iota$ to match the average amount of financial resources spent on children. We target an average local expenditure on education of 5,758 for every child.[24] Investment productivity $A_\iota$ 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 parameter $\nu \in [0,1]$, which captures the extent to which time spent which children counts as leisure, to align the average amount of time spent on children. 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, parents spent about 17.7 hours per week of quality time with their child. For $\nu = 0.565$, the model generates a value of 17.1 hours per week.

*Child Mental Health Treatment*. We calibrate the stigma cost for child mental health treatment, $\xi^k_\tau$, to match the share of children classified as unhealthy (based on BPI scores) who receive treatment between ages 12 and 15, which is 0.232 in the CDS of the PSID.[25] The utility cost of having a child experiencing mental illness, $\xi^k$, enters the parent's flow utility in each period of childhood. Because parents of younger children face more remaining periods of this disutility, $\xi^k$ has a disproportionate effect on treatment decisions early in childhood, generating variation in conditional treatment rates across age groups. We target the conditional treatment share among unhealthy children aged 8–11, which is 0.186 in the CDS.

*Adult Mental Health Treatment*. We calibrate the utility cost of treatment for adults $\xi_\tau$, such that the model matches the share of individuals experiencing serious mental illness who receive treatment. According to the National Institute of Mental Health's 2021 National Survey on Drug Use and Health, 65.6 percent of those who experience serious illness receive treatment during the year. 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 $\rho_\tau$ governs the

[24] See Appendix E.7.

[25] Merikangas et al. (2010) estimates that only about half of 8 to 15 year-olds experiencing mental illness receive treatment. Lu (2019) finds that 39 percent of 12 to 17 year-olds experiencing depression received mental health treatment in 2016. Ghandour et al. (2019) establishes that, among 3 to 17 year-olds, 80 percent of those with depression received treatment, compared to 59.3 percent of those with anxiety disorders and 53.5 percent of those with behavioral disorders.

speed of decay: higher values concentrate the decline toward later ages, while lower values produce rapid early decay. We calibrate $\xi_\tau$ to match the treatment share among young adults aged 16–27, which is 0.491 in the NSDUH, and $\rho_\tau$ to match the age gradient in treatment uptake, which we measure as a coefficient of 0.009 from a linear regression of treatment on age.

*Negative Thinking.* Following Abramson, Boerma, and Tsyvinski (2024), we quantify negative thinking for adults experiencing mild and serious mental illness by targeting the differences in consumption by mental health status, controlling for income and wealth. In the data, individuals experiencing mild mental illness consume 3.3 percent less than healthy individuals, and those experiencing serious mental illness consume 5.5 percent less. The estimated entropy constraint parameters are $\kappa(m_1) = 0.376$ for mild and $\kappa(m_2) = 1.446$ for serious mental illness, generating model consumption gaps of 4.0 and 6.2 percent, respectively.

*College.* There are two main sets of college parameters. First, the parameter $\varpi$ determines the weight of non-college labor in the CES aggregate of effective labor, $H = \left(\varpi\, H_{nc}^{\rho} + (1-\varpi)\, H_{c}^{\rho}\right)^{1/\rho}$. A higher $\varpi$ 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. Second, college taste parameters are quantified using TAS. For the college taste specification (16), the intercept for the case where parents did not attend college, $\alpha_0$, is calibrated to match the college entry share of 0.426. The intercept for the case where parents did attend college, $\alpha_1$, 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 informs the relationship between cognitive skill and college taste $\alpha_\theta$. The root mean square error of the regression, which is 0.419, determines the scale parameter $\chi_s$ for the Type I extreme value college taste shock.

*Taxes and Redistribution.* The level of taxes $\tau_0$ is chosen to match government expenditures to aggregate output. We target a government expenditure share in output of 0.189, which is the ratio of government purchases to aggregate output in the U.S. between 2000 and 2019. Finally, we set the transfer value $\omega$ such that the variance of disposable income to the variance of pre-government income is 0.61 following Heathcote, Perri, and Violante (2010), which in the model corresponds to $\text{Var}(\log(\tilde{y} + (1 + (1-\tau_a)r)a + \omega))/\text{Var}(\log(y + (1+r)a))$.

Table 4: Endogenous Parameters

| Parameter | Value | Moment | Data | Model |
|---|---|---|---|---|
| | | Preferences | | |
| Discount factor $\beta$ | 0.971 | Average wealth (in thousands) | 166 | 160 |
| Altruism $\delta$ | 0.480 | Rank correlation income | 0.340 | 0.330 |
| | | Child Mental Health | | |
| Stigma cost of treatment $\xi_\tau^k$ | 0.292 | Treatment share, 12−15 years | 0.232 | 0.191 |
| Utility cost of illness $\xi^k$ | 1.815 | Treatment share, 8−11 years | 0.186 | 0.125 |
| | | Adult Mental Health | | |
| Negative thinking, mild $\kappa(m_1)$ | 0.376 | Consumption, mild | −0.033 | −0.040 |
| Negative thinking, serious $\kappa(m_2)$ | 1.446 | Consumption, serious | −0.055 | −0.062 |
| Stigma cost of treatment $\xi_\tau$ | 1.730 | Treatment share, 16−27 years | 0.491 | 0.335 |
| Decay of stigma with age $\rho_\tau$ | 2.724 | Treatment share, age growth | 0.009 | 0.046 |
| | | College | | |
| Baseline college taste $\alpha_0$ | 0.765 | Entry share | 0.426 | 0.350 |
| College taste, college parents $\alpha_1$ | −1.340 | Entry regression, parents | 0.242 | 0.263 |
| College taste, cognitive skill $\alpha_\theta$ | −1.299 | Entry regression, cognitive skill | 0.268 | 0.263 |
| Scale college taste shock $\chi_s$ | 11.838 | Entry regression RSME | 0.419 | 0.421 |
| Share non-college labor $\varpi$ | 0.545 | Wage premium | 1.775 | 1.904 |
| | | Parental Investment Technology | | |
| Parental resources share $\alpha_\iota$ | 0.187 | Mean expenditures | 5,758 | 5,297 |
| Investment productivity $A_\iota$ | 7.710 | Normalize cognitive skill | 0.000 | −0.033 |
| Leisure cost of parental time $\nu$ | 0.565 | Mean time investment | 0.177 | 0.171 |
| | | Fiscal Policy | | |
| Labor earnings tax level $\tau_0$ | 0.201 | Government spending | 0.189 | 0.162 |
| Transfer $\omega$ | 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 three columns present the parameters and their values. The fourth column describes a moment that informs the parameter value. The fifth and sixth column present the data moment and the model-equivalent.

## 5.5 Validation

We next evaluate the model's ability to replicate empirical patterns that were not directly targeted in the estimation. These exercises cover three dimensions: parental time investments, mental health treatment for children, and college graduation expectations.

**Time Investments**. We compare predictions for the effect of parental mental illness on time investments in children to our empirical estimates reported in Table A.1. Regressing log hours on a depression indicator yields a coefficient of −23.1 (in 100× log points) without controls. The model generates −27.6, capturing that parents experiencing mental illness invest substantially less time in their children. The sign, order of magnitude, and the fact that the raw gap is large are all consistent with the data. When income controls are added, our empirical coefficient attenuates sharply to −12.9. The model coefficient attenuates less, to −25.2. In the fixed effects specification, the empirical estimate is −15.0 versus −24.3 in the model.[26] Overall, model and data show that parental mental illness substantially reduces time investments.

**Child Mental Health Treatment**. Figure 2 displays the share of children classified as having behavioral problems (using the BPI measure) who receive treatment, by age. The estimation targets treatment rates at ages 8 and 12 only, so ages 0 and 4 serve as out-of-sample validation. The model correctly predicts near-zero treatment at age 0 and low uptake at age 4. Model and data exhibit a monotonically increasing age profile, with treatment rising from near zero at birth to roughly one-fifth of unhealthy children by age 12. The model understates the steepness of this age gradient: treatment increases more slowly with age compared to the data, particularly at age 4 where empirical uptake is substantially higher than what the model generates. This age gradient emerges endogenously from parental choices, which takes into account both the current utility from living with a healthier child during the remaining childhood periods and the value of the child's mental health upon leaving the household.

**College Graduation Expectations**. We construct a model analog of college graduation expectations. For each college student, we record the expected graduation probability. In the survey, students report whether or not they expect to graduate − a binary response. In order to match this, we convert every student's perceived graduation probability to one if it exceeds 51 percent, that is, if the student perceives graduation as more likely than not, and regresses the expectation on a mental health indicator.

The empirical estimate is −3.6 percentage points, while the model generates −6.7 percentage points,

[26] 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.

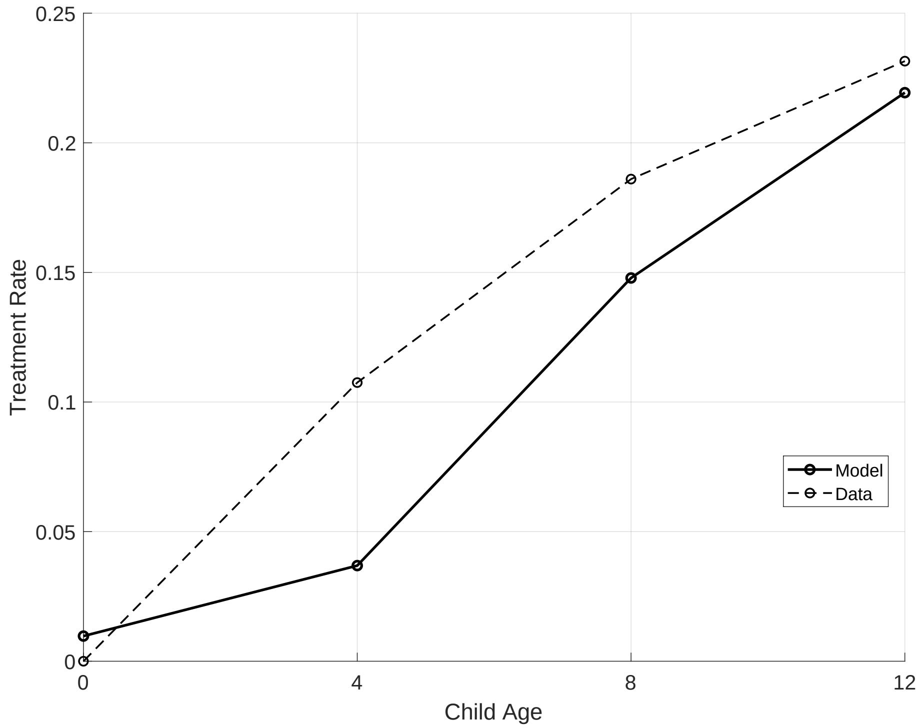



Figure 2: Child Mental Health Treatment Rate Conditional on BPI Unhealthy, by Age

both indicating that mental health substantially shifts perceived graduation probabilities. The model successfully reproduces the qualitative prediction that college enrollees experiencing mental illness hold more negative expectations about their graduation prospects, a pattern that was not targeted in estimation and that validates the entropy-based distortion mechanism.

# 6 Quantitative Results

We use the quantitative model to evaluate the consequences of mental illness and the effects of policies that expand access to treatment services intended to improve child mental health. First, we decompose how four channels of mental illness – negative thinking, rumination, productivity effects, child development – affect outcomes and welfare. Second, we evaluate prominent child mental health policy proposals.

## 6.1 Evaluating the Model Mechanisms

We next evaluate the mental health channels by decomposing the effects of each of the four channels on (a) individual-level choices and (b) aggregate economic outcomes.

Table 5: Validation: Untargeted Moments

| | Data | Model |
|---|---|---|
| *Parental Time Investments and Mental Health* | | |
| No controls | −23.1 | −27.6 |
| + Age | −25.3 | −27.6 |
| + Income | −12.9 | −25.2 |
| + Child Mental Health | −11.8 | −24.7 |
| + Fixed Effects | −15.0 | −24.3 |

Table 5 reports moments not directly targeted in the estimation. Parental time investment coefficients are 100× the OLS coefficient on a depression indicator with log hours as the dependent variable, estimated with progressive controls as in Table A.1.

**Individual-Level Choices**. We first decompose how mental health affects individual outcomes through four distinct channels. For an individual experiencing mental illness, we evaluate each channel by setting its parameters to the values that a healthy individual would have, while holding the remaining channels at their mental illness values. This isolates the consequences of each mechanism for parental investments, consumption, labor supply, and welfare.[27]

We consider four channels through which mental health operates. The negative thinking channel isolates the implications of negative thinking, the rumination channel reduces available time, while the productivity effect reduces labor productivity for individuals experiencing mental illness. Finally, the child development channel captures the role of parent mental health in the development of cognitive skills and mental health of children. Parent mental illness reduces both cognitive skill formation and mental health development in children, operating through the child development technology. This channel is amplified by the complementarity in the child development function: when inputs are complementary, the marginal product of monetary and time investments is higher when parent mental health is better.

In Panel A of Table 6, we first compare individuals experiencing serious mental illness to their healthy counterparts.The compensating equivalent variation (CEV) gain amounts to 31 percent, corresponding to a one-time wealth gain of 430 thousand dollars. Negative thinking accounts for about three quarters

[27]For each mental health channel we replace the corresponding parameter with its healthy counterpart. For the productivity effect, we set the productivity penalty to zero; for rumination, the time loss parameter is set to zero; for negative thinking, the subjective probabilities are set to align with the objective probabilities; and for the child development channel, the parental mental health values in the child skill production function are set to the healthy values. 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).

Table 6: Shapley Decomposition of Mental Health Channels

| | Consumption | Labor | | Parental Investments | | | Welfare | |
|---|---|---|---|---|---|---|---|---|
| | | Income | Hours | Time | Money | Transfers | Consumption | Wealth |
| *Panel A: Severe → Healthy* | | | | | | | | |
| Baseline | 32,500 | 43,800 | 28.8 | 9.4 | 3,461 | 27,515 | — | — |
| Wage Effect | 1.1 | 6.1 | 0.0 | −0.7 | 0.3 | 5.3 | 0.8 | 8,500 |
| Rumination | 3.1 | 17.6 | 17.1 | 2.2 | 2.0 | 12.3 | 3.6 | 42,000 |
| Negative Thinking | 8.5 | 0.0 | 0.0 | 38.7 | 25.1 | 78.6 | 23.8 | 340,000 |
| Child Development | 0.2 | 0.0 | 0.0 | 37.1 | 19.0 | −2.4 | 3.2 | 39,800 |
| Total | 12.8 | 23.7 | 17.1 | 77.2 | 46.4 | 93.8 | 31.4 | 430,300 |
| *Panel B: Mild → Healthy* | | | | | | | | |
| Baseline | 36,000 | 55,600 | 32.9 | 15.9 | 4,900 | 44,300 | — | — |
| Wage Effect | 0.4 | 2.2 | 0.0 | −0.1 | −0.1 | 1.0 | 0.3 | 3,700 |
| Rumination | 0.6 | 2.7 | 2.7 | −0.1 | −0.5 | 0.5 | 1.0 | 12,600 |
| Negative Thinking | 3.8 | −0.0 | 0.0 | 4.5 | 4.9 | 22.6 | 11.5 | 170,300 |
| Child Development | −0.0 | −0.0 | 0.0 | 3.4 | 2.2 | 0.4 | 2.0 | 27,500 |
| Total | 4.8 | 4.9 | 2.7 | 7.7 | 6.4 | 24.5 | 14.8 | 214,200 |

Table 6 decomposes the gain from transitioning an individual from a given mental health state to perfect health using Shapley values. Panel A presents results for serious mental illness; Panel B for mild. Baseline row reports levels. Al values are reported in percentages, except for wealth welfare which is reported in thousand of dollars. Time and Money Investment refer to parental investments in children.

of the welfare gains, reflecting the significant impact of negative thinking on economic choices. Negative thinking causes individuals experiencing serious illness to invest 39 percent less time and 25 percent less monetary resources in their children, and reduce transfers by 79 percent. Rumination accounts for 4 percent of the CEV, operating primarily through increased hours worked (17 percent). The productivity effect is small, generating a CEV of 0.8 percent and increasing labor income by 6 percent. Finally, the child development channel accounts for 3 percent of the CEV. It operates through parental investments in children, increasing time investment by 37 percent and monetary investment by 19 percent. These large investment responses reflect the complementarity in the child development technology: when parent mental health improves, the marginal product of financial and time investments rises, inducing parents to invest substantially more in their child.

In Panel B, we compare individuals experiencing mild mental illness to their healthy counterparts. The CEV gain is attenuated to 15 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 12 percent of the CEV, rumination 1 percent, the productivity effect is negligible, and child development accounts for 2 percent. The importance of negative thinking is consistent with the predominant theory of depression − Beck's cognitive model of depression − which posits that depression is primarily a cognitive disorder characterized by negative thinking (Beck, 1967b, 1976, 2002, 2008).

**Aggregate Outcomes**. We next evaluate the aggregate implications of each mental health mechanism. We use the Shapley decomposition to decompose the contribution of each mental health channel.

Table 7 presents the aggregate contribution of each channel to output, labor income, consumption, hours worked, and savings. Eliminating all mental health mechanisms increases aggregate output by 9 percent. The child development channel contributes the largest share to output growth at 7.5 percent. 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 0.2 percent and savings by 19 percent. Eliminating negative thinking causes individuals experiencing mentally illness to reduce precautionary savings, as they perceive future outcomes more positively. In this open economy, the reduction in savings does not lower the capital stock but it does reduce transfers to adult children and parent monetary investments. Despite the negative output effect, eliminating negative thinking generates the largest welfare gains, as shown in Table 9: a 36 percent CEV gain, equivalent to a one-time wealth transfer of about 351 thousand dollars.

Table 7: Shapley Decomposition of Aggregate Economic Effects

| Channel | Output | Labor Income | Consumption | Hours | Savings |
|---|---|---|---|---|---|
| Wage Effect | 0.4 | 0.4 | 0.3 | −0.0 | −0.0 |
| Rumination | 0.8 | 0.8 | 0.5 | 1.0 | −0.6 |
| Negative Thinking | −0.2 | −0.2 | −2.6 | −0.1 | −18.7 |
| Child Development | 7.5 | 8.0 | 5.3 | 0.0 | 4.3 |
| Total | 8.5 | 9.0 | 3.5 | 0.9 | −15.0 |

Table 7 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.

Rumination contributes 0.8 percent to output, while the productivity effect contributes 0.4 percent. Together, the four channels generate a consumption equivalent welfare gain of 67.7 percent (Table 9), corresponding to a one-time wealth gain of 654,900 dollars.

Table 8 decomposes the contributions of different mental health channels to education outcomes and parental investments. Parental time investment in children rises by 7 percent in aggregate, with the child development channel accounting for almost all of it. College enrollment rises by 4.8 percentage points, driven by the child development channel. Parental money investment increases by 7 percent. Transfers to adult children decline by 26 percent as healthier parents raise children who require less financial support later in life.

The contrast between CEV and fixed-distribution CEV (CEV FD) is informative to understand the mechanisms that deliver welfare gains. The fixed-distribution measure evaluates the counterfactual value functions at the baseline stationary distribution rather than the endogenous counterfactual distribution. The child development channel CEV drops from 29 to 5 percent under the fixed 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, the CEV of negative thinking remains large at 36 percent under fixed distribution, indicating that negative thinking correction benefits individuals directly.

Table 8: Shapley Decomposition of Education and Investment Effects

| | Parental Investments | | | Cognitive | College |
|---|---|---|---|---|---|
| Channel | Time | Money | Transfers | Skills | |
| Wage Effect | −0.1 | 0.0 | −1.0 | −0.1 | −0.0 |
| Rumination | −0.2 | −0.2 | −4.5 | −0.3 | −0.1 |
| Negative Thinking | 0.4 | −0.3 | −29.6 | −0.4 | 0.2 |
| Child Development | 7.1 | 7.2 | 8.6 | 14.1 | 4.7 |
| Sum | 7.2 | 6.7 | −26.4 | 13.3 | 4.8 |

Table 8 presents a Shapley value decomposition of effects on parental time investment, money investment, transfers to adult children, cognitive skills, and college enrollment (percentage points). All effects as percentage changes from initial steady state, except for college which is in percentage points.

## 6.2 Expanding Mental Health Treatment

We next evaluate the consequences of prominent mental health policy proposals targeted towards children and parents. In Appendix F, we provide more detail on specific policies. We evaluate all policies in general equilibrium with endogenous fiscal adjustment through tax rate on income.

**Child Mental Health Policies**. We next evaluate the economic consequences of the following prominent mental health policies.

**I. Expanding Treatment Services for Children**. Policymakers are considering various policies to expand mental health treatment services among children. A first policy is to provide school-based mental health services. A second policy is to expand access to treatment through community health clinics. We assess the consequences of a policy that makes treatment services available to all children. In the model, this corresponds to treatment being freely available for all children. That is, we consider a child subsidy that provides free mental health treatment to children by completely subsidizing out-of-pocket treatment costs.

**II. Expanding Treatment Services for Young Parents**. Policymakers recognize the important link between parental mental health and the development of cognitive skills and mental health for children. For example, the U.S. Surgeon General specifically phrased poor parental mental health as a child healthcare issue. Legislation targeting parent mental health primarily focuses on maternal mental health in the

Table 9: Shapley Values: Wealth Equivalence and Welfare Effects

| | Welfare | | Welfare: Fixed Distribution | |
|---|---|---|---|---|
| Channel | Wealth | Consumption | Wealth | Consumption |
| Wage Effect | 5,900 | 0.6 | 7,200 | 0.8 |
| Rumination | 19,700 | 2.1 | 24,400 | 2.8 |
| Negative Thinking | 350,600 | 36.2 | 344,200 | 36.1 |
| Child Development | 277,700 | 28.7 | 41,400 | 4.5 |
| Sum | 653,900 | 67.7 | 417,100 | 44.3 |

Table 9 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 10: 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 |
| Parents and Children | 32.2 | 54.0 | 48.4 | 36.4 |

Table 10 presents treatment uptake as percentage of non-healthy population.

perinatal window – from pregnancy through 12 months postpartum. These legislations increase access to mental health treatment for parents. 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. Expanding Treatment Services for Parents and Children**. The third policy combines subsidies for children and all parents. This policy extends free mental health treatments services for all ages between age 24 and 43, and simultaneously provides free treatment for children. We use our structural framework to explore the potential gains from extending such programs to a larger population.

**Treatment**. Table 10 reports treatment rates among the non-healthy population under the three policies. The subsidy for child mental health treatment services generates a large behavioral response: treatment of children experiencing mental illness increases from 11 percent to 36 percent. The young parents subsidy

Table 11: Effects of Treatment Subsidies on Prevalence of Mental Illness

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

Table 11 presents the distribution of mental health for different age groups. Baseline row reports prevalence in percent in the initial steady state. Policy rows report steady-state changes in percentage points relative to baseline. All Adults includes all adult ages. Parents covers ages 24–43. Children are ages 0–12. Adolescents are age 16 (the first adult period in the model).

operates through adult uptake, increasing treatment at young parenting ages from 51 to 54 percent. The combined parents-and-child policy increases treatment across all parenting ages from 44 to 48 percent and child treatment to 36 percent.

**Prevalence**. Table 11 reports the impact of policy on the prevalence of mental illness. The treatment subsidy generates a large improvement in child and adolescent mental health. The fraction of unhealthy children declines by 0.7 percentage points, while mild and serious illness among adolescents decline by 2.2 and 0.8 percentage points, respectively. Adult mental health improves modestly as children enter adulthood healthier. The young parent subsidy primarily affects adult mental health at young parenting ages, with limited spillovers to child mental health. The combined parents-and-child policy produces the largest mental health improvements, reducing mental illness among children by 0.7 percentage points, mild and serious illness by 0.2 and 0.1 percentage points among parents, and adolescent mild and serious illness by 2.3 and 0.8 percentage points, respectively.

**Education and Parental Investments**. Table 12 establishes a sharp contrast in parent investment behavior between subsidies for the treatment of children and subsidies for the treatment of young parents. Subsidies for child mental health treatment reduce parental time and monetary investments, as well as transfers. Parents reallocate resources away from child investments and toward own consumption. College enrollment also declines modestly, consistent with the reduction in parental transfers. Parental subsidies produce the opposite pattern: healthier parents face lower costs of engaging with their children − negative thinking and rumination are attenuated − and respond by investing more. The combined all-parents-

Table 12: 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 |
| Parents and Children | −0.41 | −0.66 | −2.38 | −0.20 | −0.12 |

Table 12 presents the policy effects on parental time investment, money investment, transfers to adult children, cognitive skills, and college enrollment. All effects as percentage changes from initial steady-state, except for college which is in percentage points.

Table 13: Effects of Treatment Subsidies on Aggregate Economic Outcomes

| | Tax Change | 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 |
| Parents and Children | 0.20 | −0.07 | −0.05 | −0.27 | 0.03 | −0.75 |

Table 13 presents the consequences of child mental health policies on aggregate economic outcomes. All effects are expressed as percentage differences from the initial steady state, except for tax change which refers to the change in the average tax rate in percentage points.

and-child policy shows that the crowding-out of the child subsidy dominates.

**Aggregate Effects**. Table 13 documents the macroeconomic consequences. The child subsidy produces a small negative output effect as the fiscal cost of the program slightly exceeds the productivity gains within the current generation. The young parent subsidy generates a modest output increase as healthier parents are directly 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**. The welfare results in Table 14 reveal that 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 equivalence of 5,300 dollars, compared to 0.13 percent for the young parent subsidy. The combined all-parents-and-child policy generates the largest welfare gain among all policies we consider: a CEV of 1.0 percent and wealth equivalence of 7,200 dollars, close to the sum of the child-only and all-parents-only policies. 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.

Table 14: 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 |
| Parents and Children | 7,200 | 1.04 | −800 | −0.12 |

Table 14 presents the welfare effects for each policy. 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 entire initial stationary distribution (over savings, cognitive skills and mental health) at its initial steady-state level.

The distinction between total welfare and fixed-distribution welfare is informative for understanding why child subsidies are so effective. The child subsidy's fixed-distribution wealth equivalence is negative. Evaluated at the baseline distribution over savings, permanent income, mental health, and education, the fiscal cost exceeds its benefits. The entire welfare gain arises from shifts in the long-run composition of the population. By treating child mental health problems early, the policy produces healthier adults in the next generation, who earn more, make better decisions, and are themselves healthier parents, propagating improvements across generations. This compositional channel is relatively muted for the parental subsidies, where the fixed-distribution welfare is positive (200 dollars for young parents), indicating 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.

# 7 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.

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.

Barro, R.J. and Becker, G.S. (1989): "Fertility Choice in a Model of Economic Growth," *Econometrica*, 57(2), 481–501.

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.

Cuijpers, P. et al. (2011): "Psychological Treatment of Depression: Results of a Series of Meta-Analyses," *Nordic Journal of Psychiatry*, 65(6), 354–364.

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. and Stabile, M. (2006): "Child mental health and human capital accumulation: the case of ADHD," *Journal of health economics*, 25(6), 1094–1118.

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.

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.

Ghandour, R.M. et al. (2019): "Prevalence and Treatment of Depression, Anxiety, and Conduct Oroblems in U.S. Children," *Journal of Pediatrics*, 206, 256–267.

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.

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., 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.

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.

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.

Lu, W. (2019): “Adolescent Depression: National Trends, Risk Factors, and Healthcare Disparities,” *American Journal of Health Behavior*, 43(1), 181–194.

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. (2010): “Prevalence and Treatment of Mental Disorders Among US Children in the 2001−2004 NHANES,” *Pediatrics*, 125(1), 75–81.

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.

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.

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 Transmission

## Online Appendix

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

March 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 (2) subject to the relative entropy constraint (3). 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:

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

which shows that states $\chi$ 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:

$$p(\chi) = \frac{q(\chi)\exp(-\lambda w(\chi))}{\sum q(\chi)\exp(-\lambda w(\chi))}. \tag{4}$$

We next show that there is a monotonic relation between the entropy budget $\kappa$ and the inverse of the Lagrange multiplier $\lambda$. In order to do so, we evaluate the relative entropy constraint using the subjective probabilities (4):

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

We differentiate this expression with respect to the inverse of the Lagrange multiplier $\lambda$ 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 $\kappa$ is indeed monotonically increasing with respect to $\lambda$.

## 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:

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

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:

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

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:

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

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:

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

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

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

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 $\chi$ 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 $\chi$ 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 $\upsilon$ 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 $\theta_j$ if $\upsilon_\theta \in (\upsilon_{j-1}^\theta, \upsilon_j^\theta]$. This produces $M \times N$ outcome states $(m_i, \theta_j)$ 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 $\upsilon$ 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 (4) 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 $\upsilon$ only through the discrete outcome that is realized.

# B Equilibrium Definition

We focus on a stationary equilibrium. In order to define a stationary equilibrium more easily, we introduce some notation. Let $s_t$ be the state of an individual of age $t$, as defined by the recursive representation of the individual's problems in Section 4.1. Let the distribution of individuals over the states at age $t$ be given by $\mu_t$.

**Definition 1.** *Stationary Recursive Competitive Equilibrium.* A stationary competitive equilibrium is $(i)$ a collection of policy functions for education $e$, consumption $c$, treatment $\tau$, savings $a$; consumption $c_k$ and mental health treatment $\tau_k$ for children, parental time $n$ and financial investments $x$ for child development; intervivos transfers $\hat{a}$; $(ii)$ a collection of value functions $(v_t, v_t^{\text{iv}})$; $(iii)$ aggregate capital and labor inputs $(K, N_0, N_1)$; $(iv)$ wages $(w_0, w_1)$; $(v)$ tax policy $T$ and $y_t^p$; and $(vi)$ probability distributions $\mu_t$ such that:

1. Given prices, policy functions and value functions solve the individual problem.
2. Given prices, aggregate capital and labor inputs solve the firm problem.
3. Labor market for each education level clears. For the non-college and college level $i \in \{0, 1\}$:

$$N_0 = \frac{1}{w_0} \left[ \sum_{t \geqslant 5} \int_{S_t} y_t \mathbf{1}\{e = 0\} d\mu_t + \sum_{t=5} \int_{S_t} y_t \mathbf{1}\{e = 1\} d\mu_t \right]$$

where the first summation is the supply of non-college graduates while the second is that labor supply of college students. For college level:

$$N_1 = \frac{1}{w_1} \sum_{t \geqslant 6} \int_{S_t} y_t \mathbf{1}\{e = 1\} d\mu_t.$$

4. Goods market clears:

$$\sum_{t \geqslant 5} \int_{S_t} \big(c_t + c_{kt} + x_t + c_\tau(\tau_t + \tau_{kt})\big) d\mu_t + \tau_e \int_{S_5} \mathbf{1}\{e = 1\}\, d\mu_5 + \delta K + G + NX = F(K, H)$$

where $NX$ denotes net exports of goods.

5. Government budget holds with equality

$$g_\tau \sum \int_{S_t} (\tau_t + \tau_{kt}) d\mu_t + \sum \int_{S_t} y_t^p(e, \theta)\, d\mu_t + G = \sum \int_{S_t} T(y_t, a_t, c_t)\, d\mu_t.$$

where summation is with respect to age groups. The technological costs of treatment is $c_\tau$, and the effective government subsidy for mental health treatment is denoted by $g_\tau := c_\tau - \varphi_\tau$. The net tax function $T(y_t, a_t, c_t)$ combines labor income taxes $\tilde{y} = (1 - \tau_0) y^{1-\tau_1}$, capital income taxes at rate $\tau_a$, consumption taxes at rate $\tau_c$, and the lump-sum transfer $\omega$, so that

$$T(y_t, a_t, c_t) = y_t - (1 - \tau_0) y_t^{1-\tau_1} + \tau_a \max\{r \cdot a_t, 0\} + \tau_c\, c_t - \omega.$$

Government expenditures on mental health treatment, retirement benefits, and residual expenditures $G$ equal net revenues from taxes. When a policy is introduced, $G$ is held constant at its initial steady-state level and the budget is rebalanced through $\tau_0$.

6. The distributions $\mu_t$ are time-invariant and consistent with the policy functions.

**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. 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:

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

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:

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

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 presents 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, as discussed in the main text. 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.[28] 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.[29]

[28] 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.

[29] We drop observations where the head of the household or the partner is a student; where reported consumption

**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, whether they receive mental health treatment, the amount of quality time parents spend with children, parental expenditures, 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. We measure parent mental health using the K6 Scale.

*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 not typically 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 $n_{kt}$ used in the estimation and validation of the model.

**Transition into Adulthood Supplement**. The Transition into Adulthood Supplement (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

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 A.1: Parental Time Investments and Mental Health

| | | | | | |
|---|---|---|---|---|---|
| Depression $\kappa$ | $-23.1$ | $-25.3$ | $-12.9$ | $-11.8$ | $-15.0$ |
| | (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 A.1 displays the regression coefficient $\kappa$ 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.

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 A.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:

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

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 A.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,

Table A.2: Mental Health and College

| | Enrollment | | Graduation | |
|---|---|---|---|---|
| Depression $\gamma_1$ | −7.6 | −5.2 | −4.8 | −3.1 |
| | (2.9) | (4.1) | (4.0) | (4.3) |
| Cognitive Skill $\gamma_2$ | 26.8 | 23.6 | 20.8 | 15.9 |
| | (1.9) | (2.7) | (2.7) | (4.2) |
| Education Mother $\gamma_3$ | 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 A.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.

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.

**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 A.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.

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

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

Table A.3 displays the regression coefficient $\kappa$ 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.

**College Graduation**. In columns 3 to 5 of Table A.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 A.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 A.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 A.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.[30] 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.[31]

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

[30] 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.

[31] 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 A.4: Consumption and Mental Health

| Variable (in logs) | Non-durables | Durables |
|---|---|---|
| Mild $\gamma_1$ | $-2.2$ | $-3.2$ |
| | (0.9) | (1.0) |
| Serious $\gamma_2$ | $-4.8$ | $-5.5$ |
| | (1.6) | (1.7) |
| Observations | 27,301 | 27,301 |
| $R^2$ | 0.28 | 0.28 |
| Mean (in levels) | 19,400 | 22,200 |

Table A.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 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 (2) subject to the relative entropy constraint (3). The computational difficulty is that: $(i)$ 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 $\lambda$ corresponding to a given value of $\kappa$ 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 $\lambda$ 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 (2) subject to the relative entropy constraint (3). 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 $\kappa$ governs the maximum permissible entropy of the subjective probability distribution. The solution to the negative thinking problem yields subjective probabilities (4) where the inverse Lagrange multiplier $\lambda$ has to satisfy the entropy constraint (3):

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

Solving for $\lambda$ 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 3.** *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 $\lambda$ satisfies the entropy constraint for $w$. Then, $\lambda\sigma = \hat{\lambda}$.

*Proof.* Substituting $w_i = \sigma\hat{w}_i + \mu$ into the subjective probabilities (4):

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

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 $\kappa$. 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 $\kappa$ as a function of $\hat{\lambda}$. However, rather than finding $\kappa$ as a function of $\hat{\lambda}$, we need $\hat{\lambda}$ as a function of $\kappa$. 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 $\kappa$, 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.

Figure A.1: Methods Comparison

(a) Computation Time (b) Approximation Accuracy (c) Fallback Rate

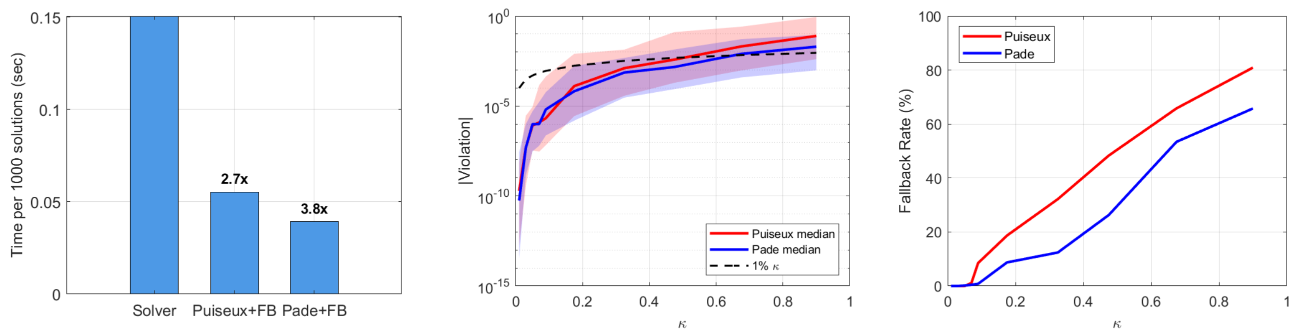



Figure A.1 presents results on the speed and accuracy of the Puiseux series and Padé approximations. Panel (a) shows computation time per 1,000 solutions with speedup factors. Panel (b) shows the approximation accuracy median |violation| of $\kappa$ before fallback on log scale, with 1 percent threshold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

## 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 $\kappa$ by more than a threshold – set here to one percent of $\kappa$ – 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 A.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.[32]

Both Puiseux series and Padé approximations perform well when relative entropy $\kappa$ is close to zero, as shown in panels (b) and (c), where accuracy is high and the fallback is negligible. As $\kappa$ 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.

[32]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.

**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 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 $\hat{y}_t(e,\theta) = w_t(m, z_t, e, \theta)n(m)$ 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:

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

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 $\psi$ as well as the curvature of the function which is governed by $\eta$. 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 $\eta$ to align the Frisch elasticity of labor supply. Let $\lambda$ 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 $\frac{dn}{dw}$:

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

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 $\eta$ using evidence on the Frisch elasticity of labor supply $\varepsilon_F$. Note that this expression is independent of the level shifter for the utility from leisure.

We set $\psi$ 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 $\eta$ and tax code $\tau$ to obtain $\psi$. 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 $\psi$ 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

We describe the parameterization for the initial draw of the child mental health and cognitive skills.

We first describe the statistical system for the initial conditions of child mental health and cognitive skills. Let $m_k$ and $\theta_k$ denote the initial cognitive skill and mental health for the child, and let $m$ and $\theta$ be the corresponding parental variables. The initial conditions for the child are a linear function of the current mental health and cognitive skills of their parent:

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

where $\varepsilon_\theta$ and $\varepsilon_m$ are both mean-zero error terms independent of the parental characteristics. To calibrate the initial draw for the child, we need to parameterize four coefficients $\beta$, two variances for the error terms, and their covariance. We proceed to first identify the coefficients, and then continue to parameterize the covariance matrix of the error term.

In the first step, we identify the coefficients $\beta$. We consider the regression $y = \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, where $u$ is independent and identically distributed. Regression coefficient $\beta$ 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 $m_{11}\beta_1 + m_{12}\beta_2 = m_{1y}$ and $m_{12}\beta_1 + m_{22}\beta_2 = m_{2y}$, where $m_{11} = \mathbb{E}[x_1^2]$, $m_{12} = \mathbb{E}[x_1 x_2]$, $m_{22} = \mathbb{E}[x_2^2]$, $m_{1y} = \mathbb{E}[x_1 y]$, and $m_{2y} = \mathbb{E}[x_2 y]$.

We invert this system to write:

$$\begin{bmatrix} \beta_1 \\ \\ \beta_2 \end{bmatrix} = \frac{1}{D} \begin{bmatrix} m_{22} & -m_{12} \\ \\ -m_{12} & m_{11} \end{bmatrix} \begin{bmatrix} m_{1y} \\ \\ m_{2y} \end{bmatrix} = \frac{1}{D} \begin{bmatrix} m_{22}m_{1y} - m_{12}m_{2y} \\ \\ m_{11}m_{2y} - m_{12}m_{1y} \end{bmatrix},$$

where $D = m_{11}m_{22} - m_{12}^2$.

In the second step, we parameterize the covariance matrix for the error using the estimated coefficients. Given the linear structure of the initial conditions, it follows that:

$$\mathrm{Var}(\varepsilon_\theta) = \mathrm{Var}(\log \theta_k) - \beta_{1\theta}^2 \mathrm{Var}(\log \theta) - \beta_{2\theta}^2 \mathrm{Var}(\log m)$$

$$\mathrm{Var}(\varepsilon_m) = \mathrm{Var}(\log m_k) - \beta_{1m}^2 \mathrm{Var}(\log \theta) - \beta_{2m}^2 \mathrm{Var}(\log m)$$

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 of the biological factors. Specifically, we use the estimated covariance matrix of initial conditions in Table A.5.

Table A.5: Covariance Matrix of Initial Conditions

| | Child Cognitive | Child Mental | Parent Cognitive | Parent Mental | Unobserved |
|---|---|---|---|---|---|
| Child cognitive | 0.1657 | | | | |
| Child non-cognitive | −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 A.5 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.

## E.6 College Costs

We parameterize college costs $\tau_e$ 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 (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.[33] 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.

[33] 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.

## 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 $\tau$ and idiosyncratic productivity state $\nu$, by $\Gamma_m(m'|m, \tau, \nu)$. In this appendix, we drop the subscript $m$ on $\Gamma_m$ 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 $\nu$. 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 $m = \{m_1, m_2\}$, $\tau$ and $(\nu, \nu')$. 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 $\rho_\nu$ and $\sigma_\nu^2$. 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 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.[34] Fourth, we obtain the share of healthy individuals who have an idiosyncratic productivity above $\underline{\nu}$ in our PSID sample, which we denote by $\pi_d^{\underline{\nu}}$.

Finally, we use estimates on the efficacy of 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 $\mu_T$ is the average outcome in the treatment group, $\mu_C$ is the average outcome in the control group, $\sigma_T^2$ is the variance of the outcome in the treatment group, and $\sigma_C^2$ 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.

Given $\mu$, $\sigma^2$, 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.

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

## 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 $\varepsilon_j$ are independent and identically distributed Type 1 extreme value shocks with location $\mu$ and scale parameter $\sigma$. 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 $\varepsilon$.

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 $\gamma$ is the Euler-Mascheroni constant. We set the location parameter $\mu$ 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 $\varepsilon$ for their child where $M(\varepsilon) = V_{j^*(\varepsilon)} + \varepsilon_{j^*(\varepsilon)}$ is the value associated with taste shock $\varepsilon$. 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:

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

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

**Claim 4.** 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 $\kappa$. The continuation value is $\mathbb{E}_\pi[M(\varepsilon)] = \mu + \sigma \log I(\sigma) - \sigma\psi(1 + \lambda\sigma)$, where $\lambda$ is the inverse multiplier on the negative thinking constraint.

We prove this claim below. For any strictly positive inverse Lagrange multiplier $\lambda$, $\psi(1+\lambda\sigma) > \psi(1) = -\gamma$, 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 ($\kappa$ increases and hence $\lambda$ increases), we have $\psi(1 + \lambda\sigma) \to \infty$, and the continuation value tends to negative infinity.

**Proof to Claim 4**. We next prove Claim 4. In this section, we adopt the following notation: $\Gamma$ 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 $\pi$ is:

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

where $A_j$ denotes the support of the college taste shock $\varepsilon$ 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 $\varepsilon_j$ tends to $-\infty$, $u$ tends to $\infty$, and when $\varepsilon_j$ tends to $\infty$, $u$ 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 $\infty$ 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 $\varepsilon_k$ are independent and identically distributed Type I extreme value shocks with location $\hat{\mu}$ and scale $\sigma$, 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 $\sigma$.[35]

For a Type I extreme value distribution with location $\mu$ 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:

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

[35] Let all $\varepsilon_k$ 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 $\sigma$ and location parameter $\hat{\mu} + \sigma\log I(\sigma)$.

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 $\lambda$ by setting the entropy equal to $\kappa$. We note that $\lambda$ is only a function of $\kappa$ and $\sigma$.

*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 $\lambda$ is determined by setting the entropy to $\kappa$. 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 $\kappa$ and $\sigma$ 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 G.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 G.1 shows that average wealth increases with the discount factor $\beta$: 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 $\delta$: higher altruism induces larger transfers and investments, raising intergenerational persistence. Panels (c)–(d) show that child treatment shares identify the child mental health parameters $\xi_\tau^k$ and $\xi^k$. 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 G.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 $\xi_\tau$ and its age decay $\rho_\tau$ are identified from treatment uptake levels and age gradients.

The college parameters (panels (i)–(m) of Figure G.1–G.1) are identified from enrollment regressions. The baseline college taste $\alpha_0$ governs the overall entry share, while $\alpha_1$ and $\alpha_\theta$ are identified from the coefficients on parental education and cognitive skill in enrollment regressions. The scale of the taste shock $\chi_s$ determines the residual variation in enrollment decisions, and the production weight $\varpi$ is pinned down by the college wage premium.

The investment technology parameters (panels (n)–(p) of Figure G.1) are identified from expenditure and time investment data: $\alpha_\iota$ from mean financial expenditures, $A_\iota$ from the normalization of cognitive skills, and $\nu$ from mean time investments. Finally, the fiscal parameters $\tau_0$ and $\omega$ (panels (q)–(r)) are identified from the government spending share and the degree of income redistribution, respectively.

Figure G.1: Parameter Identification: Global Results

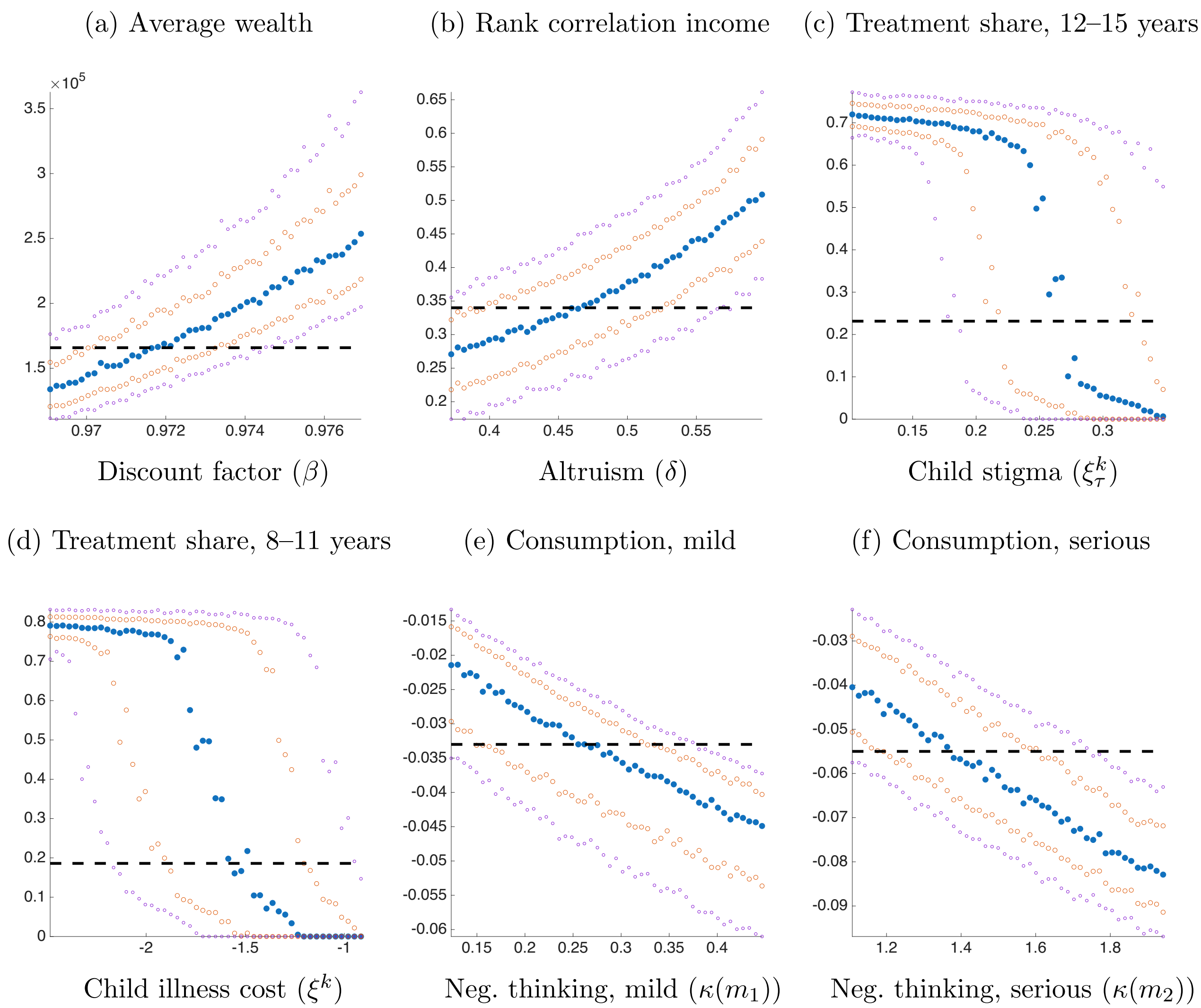



Figure G.1 provides results from the global identification exercise for the parameters in Table 4. The vertical axis of each panel is a target moment (denoted in the panel label) while the horizontal axis is a focal parameter (labeled below the panel). For each parameter quantile, the blue dots show 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 G.1: Parameter Identification: Global Results (continued)

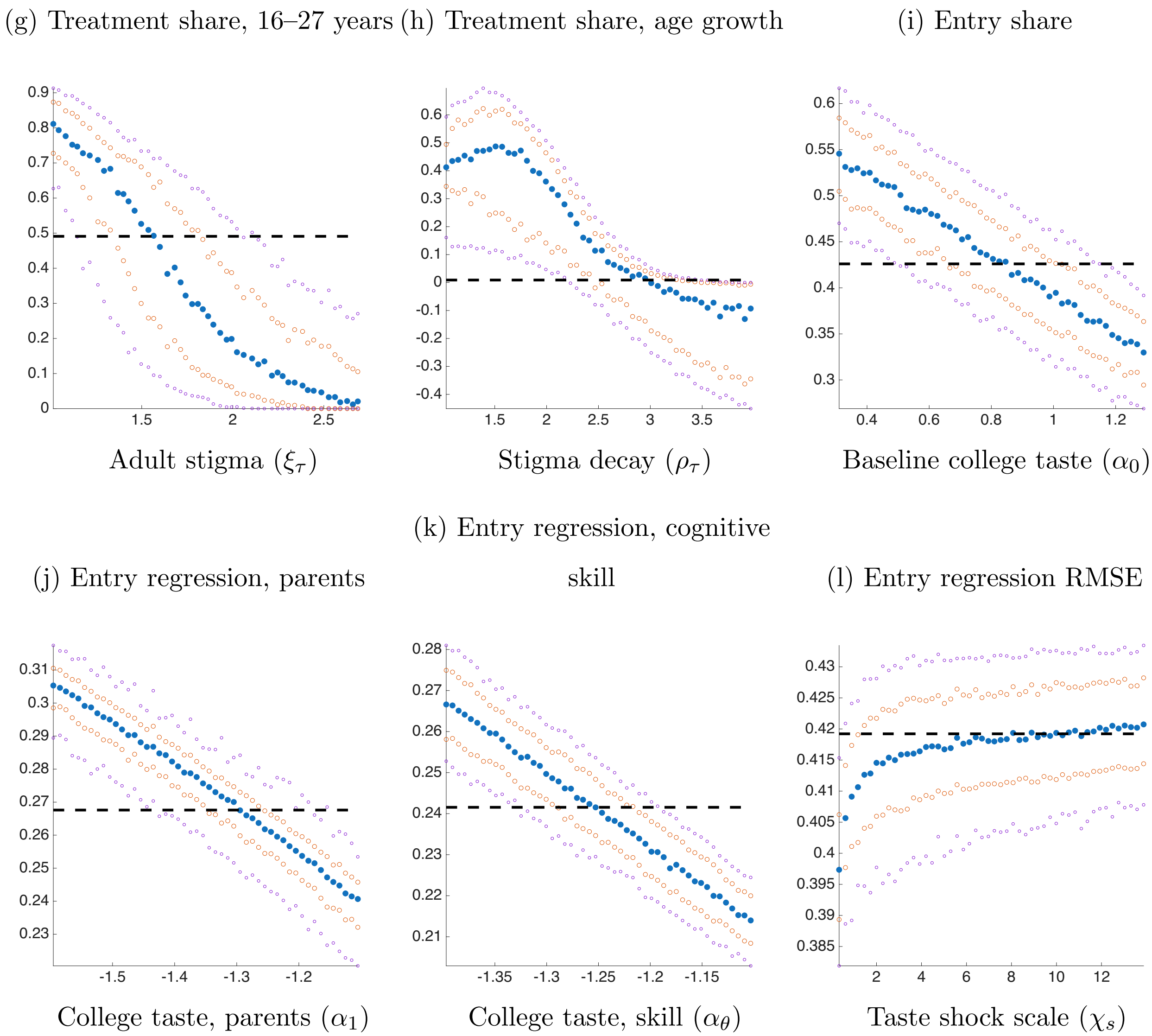
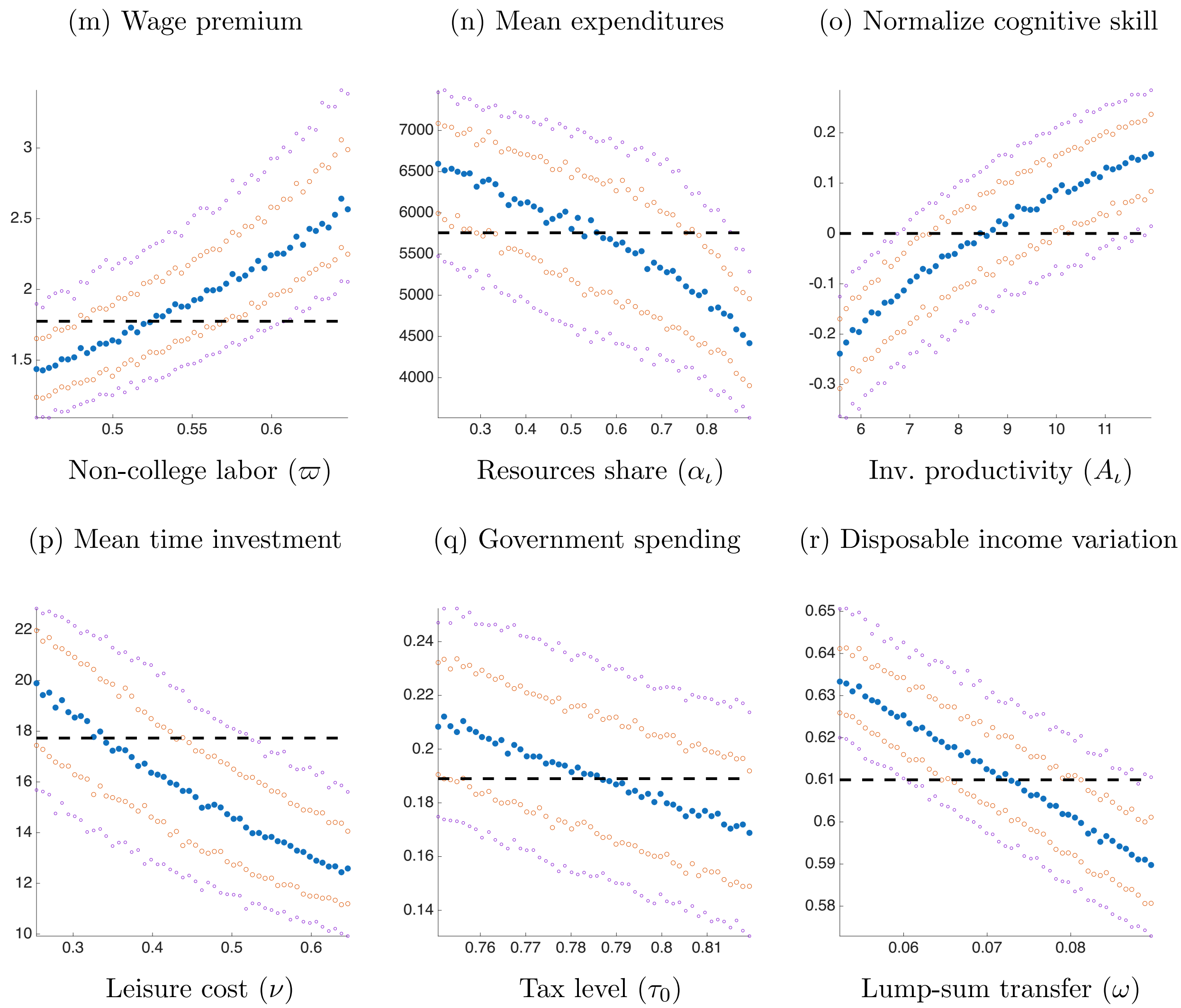


Figure G.1: Parameter Identification: Global Results (continued)




# F Mental Health Policies Examples

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

**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)
  Enacted 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, assessment, 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.

**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)

  Administered by SAMHSA since 2008, Project LAUNCH provides five-year grants to states and tribal communities to promote the social, emotional, cognitive, and behavioral health of children from birth 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 program's 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. Since its inception, the program has funded 55 sites across states, tribal communities, territories, and the District of Columbia.[36]

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:[37]

- 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 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

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

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

income and accept all insurance types, including Medicaid, Medicare, and private insurance.[38]

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.
- 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).

[38] 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).

- 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).

**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 (P.L. 115-123), 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.[39]
- 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.

[39] 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.

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.[40]

*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.[41]

[40] See www.samhsa.gov.

[41] See www.first5center.org.