THE UNIVERSITY OF CHICAGO

DEPRESSION AS A TRANSMISSION MECHANISM IMPACTING INTERGENERATIONAL MOBILITY

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

BY
CULLEN ALEXANDER ROBERTS

CHICAGO, ILLINOIS
JUNE 2018

To the late Robert LaLonde, who changed research and researchers – even in his final hour

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am grateful to many people who have supported me throughout my graduate education. My committee – James Heckman, Steven Durlauf, and Robert LaLonde – provided both inspiration and critical guidance throughout the research process. To anyone familiar with my advisers, their influence on my work will be obvious. Robert LaLonde deserves special mention, as he provided consistent feedback even in what would be the final months of his life. He taught me more than how to research – his determination in the face of deteriorating health has taught me how to live. I am also grateful to Gary Becker for his teaching and guidance; his universal approach to economics now permeates my thinking. I am thankful for Dan Black, who stepped in to fill LaLonde's role in my thesis committee on such short notice.

I would also like to thank several members of the Social Science Genetic Association Consortium, including: Dalton Conley, Daniel Benjamin, David Cessarini, Jonathan Beauchamp, and Patrick Turley. Some of these gave advice specifically relevant for my thesis; all of these helped me to better understand how biology can inform economics.

While I had the good fortune of finding many friends and colleagues in the ranks of this program, I would like to specifically thank James Best, Bocar Ba, and Daniel Hedblom, for providing stimulating conversation and critical feedback relevant for my thesis.

Finally, I would like to thank my wife and parents for their love and encouragement throughout this arduous process.

I acknowledge financial support from the National Science Foundation through their Graduate Research Fellowship. I also express gratitude to the University of Bristol for providing me with data from the Avon Longitudinal Study of Adults and Children. Decades of work by many unsung researchers have produced a truly unique source of knowledge.

ABSTRACT

Major depression is common, deleterious, and persistent across generations, thus making depression a crucial link between the economic fortunes of parents and children – albeit one that economists have largely neglected. A biological literature indicates that stressors causing depression in parents also cause depression in offspring through two mechanisms: parenting behaviors caused by parental depression and, potentially, epigenetic endowments. In this thesis, I explore these mechanisms and their relevance for the young adult outcomes of offspring. I utilize the Avon Longitudinal Study of Adults and Children (ALSPAC), which uniquely contains epigenetic data at birth of currently adult children. Furthermore, I develop an economic model of intergenerational mobility featuring depression as a facet of human capital, which aids interpretation of my findings and clarifies potential implications. My empirical findings indicate that parental depression, depression-associated parenting behaviors, and biological endowments play a key role in determining the young-adult outcomes of children. The child's adolescent depression mediates this relationship. Conditional on depression and other covariates, parental household income plays either a more modest role or no role, depending on the outcome variable considered. Meanwhile, child epigenetic data at birth explains twice the variance in the child's adolescent depression as does maternal depression in the subsample for which I have epigenetic data, providing suggestive evidence that biological endowments very early in life may establish a component of depression risk.

To interpret my findings, I crystalize ideas from the biological literature in a simple economic framework. I treat the mechanisms generating depression as a technology of preference formation. Stress, parenting behaviors, and direct biological endowments form *affective capital*, where low affective capital manifests as depression. Affective capital enters directly into the utility function to reduce disutility of effortful actions, including labor market effort and parenting effort. Thus, high affective capital both increases earnings and increases parental investment in the affective capital of children. In this way, my model resembles a Becker-Tomes/habit formation hybrid. This model highlights how investments in the child's human capital may be income inelastic. I discuss additional implications of this model.

CHAPTER 1

INTRODUCTION

Major depression is common, deleterious, and persistent across generations, thus making depression a crucial link between the economic fortunes of parents and children – albeit one that economists have largely neglected. A biological literature implicates several mechanisms behind this persistence: genetic endowments, (potentially) epigenetic endowments[1], and – crucially – parenting behaviors caused by parental depression. These parenting behaviors might be inelastic to household income, providing one mechanism through which parents matter – even when parental income does not. In turn, depression entails a reduced willingness to exert effort for rewards, which can help to explain both why and how depression is relevant for labor market outcomes.

I explore these mechanisms in the Avon Longitudinal Study of Adults and Children (ALSPAC) data, which provide rich longitudinal measures of parenting behaviors, depression inventories of parents and children, socioeconomic variables, and – uniquely – epigenetic data at birth of currently adult children. Furthermore, I crystalize ideas from the biology of depression into a simple intergenerational model of human capital formation, which helps to interpret my findings.

My empirical findings permit a sizeable role for parental depression in child young-adult outcomes but a more limited role for parental income conditional on depression. I focus on two young-adult outcomes: (1) being employed or in school at age 20; and (2) having attained an A-levels certification. Parental depression predicts both, and the child's adolescent depression mediates this relationship. In contrast, after conditioning on either parental or childd depression, parental income does not significantly predict being

---

[1] "Epigenetics" refer to mechanisms generating persistent patterns of genetic expression. I provide a more thorough description in the subsequent section.

employed or in school at age 20. Similarly, parental income explains substantially less of educational attainment after conditioning on parental depression.

Parenting behaviors peculiar to parental depression play a key role. Parental depression predicts a suite of affectively-laden parenting behaviors (cuddling with the child, physically playing with the child, etc...). Neither parental income nor parental education significantly associate with these parenting behaviors after conditioning on parental depression. In turn, these depression-specific parenting behaviors predict child depression – to some extent even after conditioning on parental depression – and also predict child young-adult outcomes. Income-specific parenting behaviors do not.

I find evidence that parental stress predicts adverse child outcomes and that this is largely mediated through parental depression. Several indices measuring parental neighborhood stressors, social support, and relationship stress strongly predict parental depression, explaining about a third of the variance for both mothers and fathers. These stressors predict depression-associated parenting behaviors. And these parental stressors also predict the child's young adult outcomes – but to a lesser extent after conditioning on parental depression or the child's adolescent depression.

These adverse young-adult outcomes may anticipate adverse middle-aged outcomes. For parents, depression explains a large portion of parental household income – about one third as much as parental education – and predicts parental unemployment. Depression in years past predicts current unemployment about as well as current depression does, so this relationship is unlikely to be driven by unemployment causing depression.

Altogether, these results are consistent with depression playing a key role in intergenerational mobility: the intergenerational transmission of depression generates intergenerational associations in income – not so much the reverse.

I also exploit epigenetic data to better understand the intergenerational transmission of depression. Two questions motivate the use of epigenetic data. What role do epigenetic endowments play in the intergenerational propagation of depression? Can epigenetic data be utilized to measure depression-relevant biological stocks when depression inventories are either too noisy or non-existent – in particular for younger children? Due to thorny identification problems I describe, neither question can be directly attacked in the ALSPAC without better motivated epigenetic markers. Instead, I attack a more modest question – to what extent can epigenetic data at birth forecast the child's adolescent depression?

I find that a leave-one-out forecast of the child's adolescent depression using epigenetic data at birth explains twice as much variance in the child's depression as does his mother's depression. This provides suggestive evidence that biological endowments may establish a component of depression risk very early in life.

I capture the above findings in an intergenerational model of human capital formation. This model crystalizes ideas from the biology literature on depression, yielding powerful theoretical implications.

To an economist, the biology underlying depression can only be described as a technology of preference formation. Social stressors, biological endowments, and certain parenting behaviors form biological stock. This stock modifies brain activity to reduce sensitivity to rewards and increase sensitivity to costs. Thus, rodents exhibiting depression-like behavior socialize less, lose interest in sex, are less willing to contest desirable locations in cages, spend less time exploring novel environments, groom (nurture) offspring less, and are less willing to expend effort for a variety of rewards. These behaviors reflect core symptoms of depression in humans – reduced interest in ordinarily rewarding activities and increased fatigue. These symptoms explain why humans with depression are less willing to expend effort to obtain things that they would otherwise want.

Naturally, this technology operates within the family. Parents provide parenting, and parents provide biological endowments. Through these mechanisms, stressors can propagate adverse outcomes across generations.

Crystalizing these ideas within an economic framework yields what looks like a Becker-Tomes/habit formation hybrid. Stressors, biological endowments, and parenting inputs combine to form a type of human capital I term *affective capital.* Affective capital enters exclusively into the utility function, reducing the marginal disutility of effortful actions (or, equivalently, increasing the marginal utility of things attained through effortful action). Low levels of affective capital correspond with depression. In principal, affective capital alters the marginal utility of many potential arguments, but two are relevant – parenting effort and labor market effort. Because affective capital reduces the marginal disutility of parenting effort, the parent's affective capital increases parental investments in the child's affective capital. And, because affective capital reduces the marginal disutility of labor market effort, affective capital increases income. Thus, affective capital generates intergenerational associations in income.

In contrast to Becker-Tomes, parents in my model can never abstract away from their utility functions when choosing investments in the child's human capital. In Becker-Tomes, credit constraints determine the materiality of parental utility. With binding credit constraints, parents must sacrifice consumption to make investments, so these investments come with a utility cost. But with relaxed credit constraints, parents first act like a firm, choosing the profit-maximizing investments. Only afterwards do they allocate income across generations in accordance with the utility function. In this case, optimal investments depend upon the child's ability but do not depend upon the parent. Fundamentally, parents do not matter. This central result is difficult to reconcile with both a limited causal role of income and also a behavioral genetics literature indicating that parents do matter.

My model implies that parental affective capital drives parenting inputs into the child's affective capital but parental income does not. Parents can never abstract away from their utility function. In this way my model explains my empirical findings.

My model offers additional predictions that merit investigation. Foremost among these, affective capital determines labor market outcomes through the utility function. Thus, depression should predict occupational choice and migratory decisions where the relevant margin is psychic, linking the theory of affective capital to compensating differentials. My model also highlights a worrisome possibility: to the extent that epigenetic endowments propogate depression across generations, equalizing household environment may not be sufficient to fully mitigate the intergenerational consequences of parental depression in one generation.

1.1 Thesis outline

The thesis contents are as follows. First, I motivate my research with an overeview of the biology of depression – how it affects behavior and how it transmits across generations. Second, I explore the intergenerational transmission of depression and its relevance for young adult outcomes in the ALSPAC data. Third, I explore the predictive content of epigenetic data at birth for the child's young adult outcomes. Fourth, I develop the model of affective capital to interpret results. Fifth, I conclude. I include a glossary of terminology in the appendix.

1.2 Biological Motivation

1.2.1 A Very Brief Epigenetics Primer

Broadly, "epigenetics" refers to mechanisms generating highly persistent patterns of genetic expression (see Berger et al., 2009, for a discussion of terminology).[2] These mech-

---

[2] Biologists disagree about how "epigenetics" should be defined. Waddington (1942) coined the term to explain tissue differentiation. How is it that skin cells are so different from liver cells when both contain

anisms underlie tissue differentiation, cellular memory of past environmental exposures, and potentially the transmission of certain phenotypes across generations. Epigenetics may be of interest to economists for two broad reasons. First, the intergenerational transmission of phenotypes may represent a channel through which certain dimensions of parental human capital directly seed the child's human capital at conception. Interventions prior to conception might impact these epigenetic endowments, but interventions after conception would not. Second, epigenetic markers may proxy for corresponding dimensions of human capital.

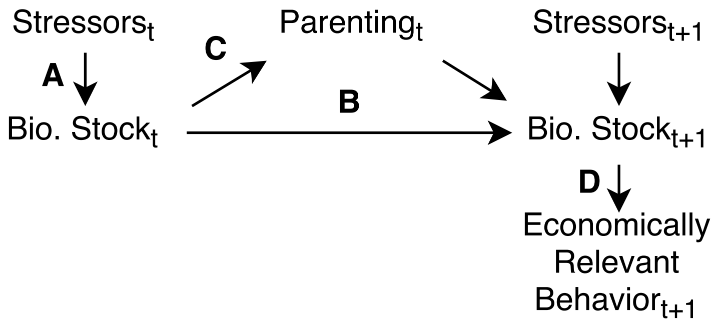
As a concrete example of an epigenetic mechanism, DNA *methylation* adds a chemical group to DNA, which changes how that region of DNA is read. In some circumstances, addition of this chemical group can prevent cellular machinery from binding to the DNA, ultimately reducing the quantity of the specific protein produced from this region of DNA. Because methylation can be highly persistent, methylation can cause long-lasting changes in the patterns of gene expression. To the extent certain changes in protein expression also change behavior, past experiences that alter methylation can cause persistent changes in behavior.

I provide a discussion of epigenetic data in the section preceding analysis of epigenetic data. I review terminology in Appendix A and further describe some epigenetic mechanisms in Appendix B.

---

identical genetics? How do the descendents of skin cells "remember" to be skin cells? Epigenetic mechanisms underly this cellular memory. Subsequent research into these mechanisms indicated that they serve two additional roles. First, epigenetic mechanisms can mediate long lasting changes to gene expression following transient environmental exposures. For example, epigenetic mechanisms appear to underlying immune cell tolerance to antigens. Second, epigenetic mechanisms can transmit information from parents to children through the germ line to govern expression of the child's genes. *Imprinting* is the most well-established example of this – certain genes will be expressed differently depending on whether these genes are inherited from the mother or from the father.

See Holliday (1994) and Felsenfeld (2014) for discussion of how terminology has evolved.



A Formation of biological stock mediating depression
B Transmission of depression: genetic and (potentially) epigenetic endowments
C Transmission of depression: depression-specific parenting behaviors
D Reduced sensitivity to rewards/increased sensitivity to costs

Figure 1.1. Representation of channels through which depression links economic outcomes of parents and children

1.2.2 Biological insights: stressors cause depression in descendants; depression impedes economically relevant behavior

Biological research provides compelling evidence that stressors in one generation increase depression in the next. The literature implicates both parenting behaviors caused by parental depression as well as epigenetic endowments. The biological mechanisms underlying depression involve alterations to the brain's reward pathway, which reduce anticipation of reward and increase perception of costs. These mechanisms explain why and how depression is relevant for economic outcomes. In this biological primer, I highlight key papers describing the formation, intergenerational transmission, and behavioral ramifications of depression (see Figure 1.1).

Biological research on depression involves both rodent experiments and human studies. Rodent experiments identify causal parameters but may lack external validity. In contrast, the human studies rarely identify causal parameters but do help to shore up the external validity of rodent experiments when human associations mirror effects seen in rodents.

Taken together, rodent experiments and human associations elucidate mechanisms relevant for intergenerational mobility.

Certain stressors cause biological modifications that mediate mediate changes in economically relevant behavior – both in the generation exposed to these stressors and in descendants. In rodent experiments used to simulated depression, both "chronic intermitent stress" (random cage reassigment, random changes in cage lighting or temperature, etc.) and "social defeat stress" (loss of social status simulated by forcing a small rodent to contest the territory of a large rodent) cause these biological modification, but stressors simulating loss of social status cause the most profound and durable modifications. Key modifications inclue alterations to the brain's reward pathway (e.g. blunted dopamine signaling in the shell of the nucleus accumbens (Lemos et al., 2012; Reed et al., 2012)). These and other modifications reduce the willingness to exert effort for rewards (Salamone et al., 2006), underlying the depression symptom of "anticipatory anhedonia": a reduced ability to anticipate reward from ordinarily rewarding activities.

Apropos, severe stressors in humans can cause depression, generating symptoms like anhedonia, fatigue, and social withdrawal. From an economic perspective, these changes indicate a greater psychic cost of effortful actions (or, equivalently, a decrease in the utility of things attained through effortful action), thereby resulting in reduced effort. Concordantly, those with depression are considerably more likely to become unemployed (Lerner et al., 2004), attain less education (Peyrot et al., 2015), and earn less income (Zimmerman and Katon, 2005). Likewise, single mothers and impoverished mothers are more likely to experience depression (Brown and Moran, 1997). And consonant with the experimental evidence on reward pathways in mice, children of depressed mothers demonstrate a blunted reward sensitivity and an increased loss sensitivity in corresponding regions of the brain

(Luking et al., 2016), and depressed patients are less willing to exert effort for things they enjoy (Sherdell et al., 2012)[3]

Both epigenetic endowments and parenting behaviors caused by parental depression may transmit the effects of parental stress to subsequent generations. For example, male mice subject to social defeat stress and subsequently bred with normal female mice sired pups exhibiting higher stress hormone levels and more depression-like behaviors – even when the female mice were impregnated via in vitro fertilization and had no contact with the male mice (Dietz et al., 2011). Subsequent research implicates epigenetic mechanisms present at conception (Rodgers et al., 2015, 2013), although this research requires further replication.

Alternatively, stressing maternal rats induces parenting behaviors that ultimately cause depression in offspring. For example, chronic intermittent stress in lactating maternal rodents disrupts grooming behavior – a behavior that relieves stress in pups. In adulthood, the offspring demonstrate elevetated depression. Furthermore, the offspring exhibit disrupted

---

[3] It merits speculating why natural selection might preserve mechanisms that convert a parent's stressors into both parental depression and the child's predisposition towards depression. Indeed, loss-of-function mutations could prevent many of the behavioral effects of chronic stress. As one example, murine models indicate the migration of peripheral monocytes into the brain mediates the development of anxiety-like behavior following social-defeat stress. Loss-of-function mutations in monocyte chemokine receptors prevent such migration and so prevent social stressors from causing anxiety-like behavior (Wohleb et al., 2013). The apparent fragility of this pathway would be difficult to reconcile with the existence of depression-like behavior in a wide variety of species unless depression-like behavior serves some adaptive purpose in certain environments. Biologists offer several explanations why depression-like behavior might be an adaptive response to psychosocial stress (for a review see Gilbert, 2005) and why parents might endow offspring with a predisposition towards depression-like behavior. For instance, Sheriff et al. (2010) find evidence that snowshoe hares impregnated during the peak of a predator cycle endow their offspring with a more sensitive HPA-axis. Although the offspring eat less and reproduce less, such offspring also suffer less predation, thus netting a reproductive advantage. Similarly, in social animals prone to conspecific violence, depression-like behavior might increase reproductive fitness by reducing the willingness to contest resources if contesting these resources risks a violent altercation (Gilbert, 2005). Given the possibly high rates of conspecific violence in the human evolutionary past (Gómez et al., 2016; McCall and Shields, 2008), depression-like behaviors may have prevented humans from contesting food, territory, or mates when doing so carried too great a risk of death or injury . As a caveat, an evolutionary explanation for depression-like behavior in no way implies that such behavior would improve income, education, welfare, or even reproductive fitness for humans in the modern environment. Rather, such an explanation lends credibility to, and aids interpretation of, the aforementioned biological evidence.

grooming behavior towards their own pups, thus transmitting depression to the third generation (Champagne and Meaney, 2006). The long-run effects of disrupted grooming may be transmitted in part through epigenetic modifications of a glucocorticoid receptor gene in the HPA-axis (that is the feedback circuit regulating the response to stressors) (Weaver et al. 2004; see also Roth and Sweatt 2011), although epigenetic changes throughout the genome may be relevant. See Maccari et al. (2014) for a review.

Associations in humans mirror the experimental results in rodents. Depression in mothers corresponds with reduced physical affection towards children (Lovejoy et al., 2000), and physical affection towards children reduces stress hormones in children (Feldman et al., 2010; Haley and Stansbury, 2003). In addition, children of mothers scoring high in depression in the third trimester show increased methylation of a glucocorticoid receptor gene, paralleling changes seen in rodents (Oberlander et al., 2008). Furthermore, those who report experiencing less parental warmth in childhood show increased depression risk as well as signs of a dysregulated HPA-axis in adulthood (Hanson and Chen, 2010).

Altogether, the experimental murine models and human associations support the proposition that stressors in one generation change behavior in the next through both epigenetic endowments and epigenetic alterations mediated through parenting; among other things, these epigenetic modifications reduce willingness to exert effort for reward, which naturally translates into worse socio-economic outcomes. Thus, the intergenerational transfer of the biological residue of stressors may play a substantial role impeding intergenerational mobility. This proposition motivates my research.

1.3 Related Research

1.3.1 Economic research related to depression

Economists have given relatively little attention to depression.

Some economists have considered the effects of various exogenous stressors on child outcomes; some of these stressors resemble those that may cause depression. This literature generally fits within the prenatal origins literature. For example, Black et al. (2016) consider whether the bereavement of a pregnant woman causes lasting harm to the woman's child. They find short-term harm to certain health parameters, mirroring other studies finding negative effects of parental stress on neonates. But they find no adverse effects persisting into adulthood.

The theoretical literature is as sparse as the empirical literature. To my knowledge, the only economic model of depression is that of de Quidt et al. (2016). In this model, individuals play what amounts to a one-armed bandit game to learn about the returns to effort; depression is the pattern of behavior that follows from learning, perhaps erroneously, that effort is not well-rewarded. Because those with depression exert less effort, they enjoy fewer opportunities to update their beliefs, thus generating persistence in depression. While this model can capture some of the features of depression, the model implies four untenable counterfactuals. First the model implies that information and information alone can cause or remediate depression. It is not possible for a person to both be depressed and believe, as a matter of fact, that effort is well-rewarded. Furthermore, stressors that do not indicate the returns to effort cannot cause depression. Second, because the model is one of Bayesian updating and yields efficient decisions, the model implies that remedying depression would benefit only those who received erroneously low signals of the return to effort. Conversely, causing depression would benefit those who received erroneously high signals of the return to effort. Third, the model omits stressors as a cause of depression unless these stressors signal the return to effort. Consequently, their model has nothing to say about abuse, privation, and other social stressors causing depression. Fourth, the model predicts that depression would decrease dramatically with higher wages. While within-country variation of depression supports this prediction, cross-country variation does not –

an unskilled laborer in a developed country earns more than an order of magnitude higher wage rate than a skilled worker in an impoverished country. Thus, by implication of the Quidt and Haushofer model, most workers in poor countries should be depressed in comparison to even the poorest workers in developed countries. International comparisons in rates of depression rates do not support such a proposition.

In contrast, the human capital model of depression I present sidesteps all four problematic predictions. In my model, depression manifests with low levels of the affective capital stock; an individual might cognitively recognize high returns to effort but recognition of this fact does not remediate low affective capital. Likewise, my model permits interpretation of low affective capital as an unambiguous bad, which cannot be said of the optimal depression-posterior from the one-armed bandit model. Furthermore, my model permits social stressors and parenting inputs to cause depression even when these stressors do not provide information about the returns to effort. Finally, my model explains reduced willingness to exert effort among the depressed without real incomes determining the willingness to exert effort, thus sidestepping the problematic cross-country comparisons.

1.3.2 Skill formation literature

The skill formation literature explores the evolution of various forms of human capital over the life cycle with a focus on the early years. While in some ways this literature finds its roots in Becker-Tomes, the skill formation literature considers more dimensions of human capital (see Borghans et al. 2008, Cunha and Heckman 2008, Biroli 2014, and Attanasio et al. 2016, for example) and emphasizes the technology of human capital formation. Heckman defines skills to include all personal characteristics that provide the capacity to act (Heckman and Mosso, 2014), and so can be thought of as incorporating a broad interpretation of human capital. The emphasized skills include non-cognitive skills (such as conscientiousness) in addition to more conventional skills, such as cognitive ability.

The possibility of dynamic complementarity motivates much of this research. Dynamic complementarity refers to when the production of a stock $H_{t+1} = f_t\left(H_t, I_t\right)$ features a positive cross-partial between current investment and previous investment via current levels of the stock (see Cunha and Heckman, 2007). If there were only one period of investment, then this positive cross-partial would imply that investments in those initially more able H_0 would produce more human capital H_1, suggesting that efficient investments increase inequality. However, this intuition does not necessarily apply when investments accrue over multiple periods; if the production function in later periods is sufficiently concave in human capital, then early investments in the least able may in fact be more productive than early investments in the most able. Intuitively, the declining marginal yproduct of investing in human capital in later periods offsets the lower marginal product of investment in less able students in earlier periods. This possibility motivates research on early-childhood education as well as skill formation within the household. Such early-childhood investments promise to both improve efficiency and reduce inequality. In this vein, researchers often estimate CES production functions or nested CES production functions (exe. Heckman and Mosso (2014)) to assess the complementarity between stocks and investments. Many papers explicitly or implicitly allow for sensitive periods, so that investments in children may be more productive than investments in adults.

My paper fits squarely within the skill-formation literature. My paper explores an unconventional form of human capital, directs attention to the technology of its formation, particularly in the early years, and extends the concept of dynamic complementarity to an intergenerational setting. Although the epigenetic data and the specific type of human capital considered are non-standard, this paper is thematically conventional.

CHAPTER 2

LINKAGES BETWEEN PARENTAL DEPRESSION AND CHILD OUTCOMES IN THE ALSPAC

2.1 Introduction

My empirical findings permit a sizeable role for parental depression in child young-adult outcomes but a more limited role for parental income. Child depression mediates the predictive content of parental depression. Meanwhile, parenting behaviors peculiar to parental depression predict child outcomes and, again, child depression mediates the predictive content of parenting behaviors peculiar to parental depression. Furthermore, parental stressors strongly predict both parental depression and adverse child outcomes. But, conditioning on parental depression or child adolescent depression obliterates the predictive content of parental stressors for child depression, thus indicating that depression represents a key channel through which parental stress adversely effects child outcomes. In turn, associations between the child's adolescent depression and young-adult outcomes anticipate adverse middle-aged outcomes. Parental depression strongly predicts worse socioeconomic outcomes on several dimensions, and some evidence suggests that causality does not flow from parental economic outcomes to parental depression.

In Appendix C, I provide a description of the ALSPAC survey variables utilized in this section (Boyd et al., 2013; Fraser et al., 2013) as well as a discussion of measurement error and selective non-response. Of note, this appendix section describes depression inventories and their potentially severe measurement error, especially when applied to men or children. This appendix also describes the intra- and inter-generational correlation structure of depression, interpreting this structure to draw inferences about the persistence of depression and about measurement error in these inventories.

2.2 Estimation of affective inputs using observed parenting behaviors and depression factors

2.2.1 Parenting behaviors peculiar to parental depression

A psychology literature suggests that relatively subtle interactions parents have with children – and not time spent with children per se – predispose or protect children from depression. For example, in the still-face paradigm (an experimental procedure that originated to simulate observed behavior of depressed mothers), researchers instruct a mother to stare impassively at her infant (see Haley and Stansbury, 2003, for example). The procedure robustly increases stress hormone levels in infants, which mirrors the depressogenic effects of maternal depression on rodent offspring. From this perspective, I investigate which of the parenting behaviors measured in the ALSPAC are peculiar to parental depression and the extent to which these behaviors predict depression in offspring.

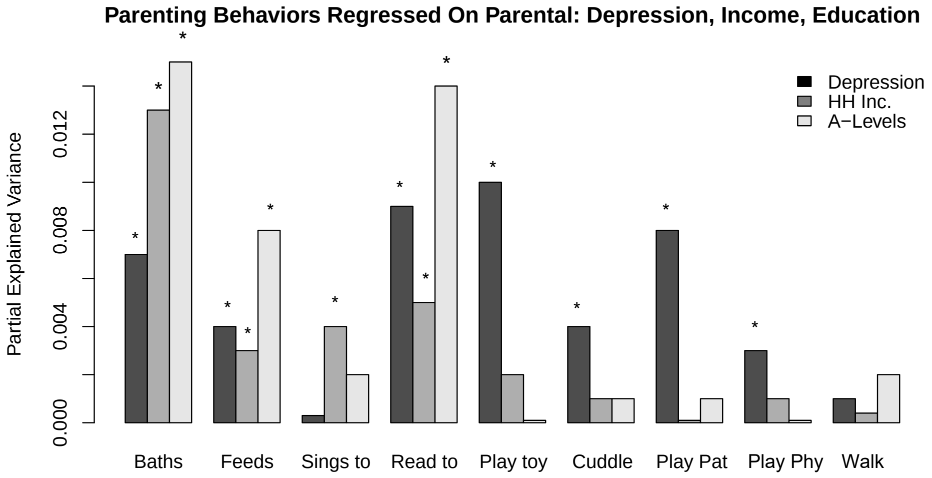
I treat a vector of observed parenting behavior variables P_i as noisy measurements of depression-associated latent parenting behaviors $\theta_{dep,i}$ as well as income-associated $\theta_{inc,i}$ and education- associated $\theta_{educ,i}$ parenting behaviors:

$$P_i = \Gamma_{dep}\theta_{dep,i} + \Gamma_{inc}\theta_{inc,i} + \Gamma_{educ}\theta_{educ,i} + \varepsilon_i$$

Parental depression can serve as a proxy for latent depression-associated pareneting behaviors[1]. Specifically, if parental depression d_i as well as unobserved variabls ε_i cause these latent pareneting behaviors, then $\Gamma_{dep}d_i$ can be substituted into the measurement sytem and

---

[1] The destinction between depression-associated parenting behaviors and depression will be fleshed out in the chapter on the theory of affective capital. In short, depression reduces the extent to which parents are willing to invest in the mental health of the child. However, variation in other parental attributes – like altruism – can either offset or compound the effects of depression on these parental investments. In this sense, depression-associated parenting behaviors and depression are distinct – they imperfectly covary.



Each set of three columns comes from a multivarate regression of the corresponding parenting behavior on parental household income at timepoint nineteen, parental educational certification, and the paternal depression inventory factors averaged over all periods in which the father completed the depression inventory. The height of the bars indicates partial explaine variance for ease of comparison. Asterisks denote that the associated regression coefficients are significant at the 99% significance level for a two-tailed t-test.

Figure 2.1. Affectively laden parenting behaviors (father) associate with depression and not income

$\Gamma_{dep}\varepsilon_i$ absorbed into the residual. Regressing each parenting behavior on parental depression, income, and education will identify the coefficient vectors $\begin{pmatrix}\Gamma_{dep} & \Gamma_{inc} & \Gamma_{educ}\end{pmatrix}$ for the associated latent parenting behaviors (I view the income- and education-associated parenting behaviors as nuissance variables, but do investigate whether these variables predict child outcomes).

With this framework in mind, I regress each of the father's observed parenting behaviors from timepoint 1.5 on the father's average depression scores and two controls: household income and the father's education. I repeat these regressions for the mother using her respective variables. Tables 2.1 and 2.2 show results. Figure 2.1 visualizes results for the father by comparing the partial explained variance of each covariate for each parenting behavior.

LHS var (father), t=1.5	RHS vars (father)				
Parenting behavior (larger numbers indicate lower frequency)	mean depression	HH income in 10k lbs. t=19	attained A-levels	obs	R2
Baths child	0.109	-0.085	-0.307	2917	0.044
	0.007	*0.013*	*0.015*		
	(0.025)	(0.014)	(0.046)		
	{0.000}	{0.000}	{0.000}		
Feeds child	0.078	-0.037	**-0.211**	2891	0.018
	0.004	*0.003*	*0.008*		
	(0.024)	(0.013)	(0.044)		
	{0.001}	{0.006}	{0.000}		
Sings to child	0.029	-0.054	**-0.126**	2910	0.008
	0.000	*0.004*	*0.002*		
	(0.030)	(0.017)	(0.056)		
	{0.333}	{0.001}	{0.025}		
Reads to child	**0.126**	**-0.050**	**-0.277**	2915	0.033
	0.009	*0.005*	*0.014*		
	(0.024)	(0.014)	(0.045)		
	{0.000}	{0.000}	{0.000}		
Plays w/ toys with child	**0.103**	-0.027	**-0.017**	2917	0.014
	0.010	*0.002*	*0.000*		
	(0.019)	(0.011)	(0.035)		
	{0.000}	{0.010}	{0.638}		
Cuddles child	0.031	-0.008	**-0.024**	2918	0.007
	0.004	*0.001*	*0.001*		
	(0.009)	(0.005)	(0.016)		
	{0.001}	{0.101}	{0.142}		
Plays pat-a-cake etc. with child	0.121	-0.010	-0.075	2913	0.009
	0.008	*0.000*	*0.001*		
	(0.025)	(0.014)	(0.047)		
	{0.000}	{0.500}	{0.113}		
Physically plays with child	**0.051**	-0.015	**0.013**	2918	0.004
	0.003	*0.001*	*0.000*		
	(0.018)	(0.010)	(0.033)		
	{0.004}	{0.126}	{0.686}		
Takes child on walks	0.031	**-0.011**	**-0.099**	2914	0.004
	0.001	*0.000*	*0.002*		
	(0.020)	(0.012)	(0.038)		
	{0.122}	{0.323}	{0.009}		
Other activity with child	0.179	0.014	**-0.377**	571	0.020
	0.009	*0.000*	*0.011*		
	(0.078)	(0.047)	(0.157)		
	{0.022}	{0.762}	{0.016}		

*Variable on left is regressed on three variables on right. First line gives the regression coefficients. Bolded estimates are significant at the 99% significance level. Second line (in italix) gives the partial explained sum of squares for this variables -- the coefficient squared times the variance of that RHS variable (mean depression scores, household income, father attaining A-levels credentials) divided by the variance of the LHS variabls (parenting behavior frequency). Standard errors in parenthesis. p-values in braces.

*Parenting behaviors are scored by frequency: 1 = "daily" ; 2= "3-5 times per week"; 3 = "1-2 times per week" ; 4 = "less than once a week" ; 5= "never". Father's depression is the standardized average of depression inventory factors for all periods in which he took the depression inventory. A-levels attined is a dummy indicating that the father has at least an A-levels certification.

Table 2.1: Regressions of father's parenting behaviors on father's depression, household income, and educational attainment

LHS var (mother), t=1.5	RHS vars (mother)				
Parenting behavior (larger numbers indicate lower frequency)	mean depression	HH income in 10k lbs. t=19	attained A-levels	obs	R2
Baths child	0.022 *0.000* (0.019) {0.246}	0.002 *0.000* (0.010) {0.812}	-0.023 *0.000* (0.033) {0.498}	3375	0.001
Feeds child	-0.001 *0.000* (0.015) {0.926}	-0.004 *0.000* (0.008) {0.659}	**-0.120** *0.007* (0.026) {0.000}	3344	0.007
Sings to child	0.021 *0.000* (0.019) {0.271}	-0.008 *0.000* (0.010) {0.462}	**-0.118** *0.004* (0.033) {0.000}	3373	0.006
Reads to child	**0.061** *0.006* (0.014) {0.000}	**-0.021** *0.003* (0.007) {0.005}	**-0.166** *0.015* (0.024) {0.000}	3376	0.031
Plays w/ toys with child	**0.031** *0.003* (0.009) {0.001}	0.002 *0.000* (0.005) {0.671}	**-0.058** *0.004* (0.016) {0.000}	3377	0.008
Cuddles child	-0.001 *0.000* (0.001) {0.515}	-0.001 *0.001* (0.001) {0.155}	**0.004** *0.001* (0.002) {0.065}	3373	0.001
Plays pat-a-cake etc. with child	0.028 *0.001* (0.014) {0.039}	0.002 *0.000* (0.007) {0.796}	0.010 *0.000* (0.024) {0.682}	3369	0.001
Physically plays with child	**0.057** *0.003* (0.018) {0.001}	-0.001 *0.000* (0.010) {0.889}	**-0.082** *0.002* (0.031) {0.008}	3368	0.006
Takes child on walks	0.035 *0.001* (0.017) {0.037}	**0.042** *0.007* (0.009) {0.000}	**0.079** *0.002* (0.030) {0.008}	3368	0.012
Other activity with child	-0.054 *0.001* (0.088) {0.537}	-0.032 *0.001* (0.049) {0.511}	**-0.909** *0.066* (0.159) {0.000}	556	0.072

*Variable on left is regressed on three variables on right. First line gives the regression coefficients. Bolded estimates are significant at the 99% significance level. Second line (in italix) gives the partial explained sum of squares for this variables -- the coefficient squared times the variance of that RHS variable (mean depression scores, household income, father attaining A-levels credentials) divided by the variance of the LHS variabls (parenting behavior frequency). Standard errors in parenthesis. p-values in braces.

*Parenting behaviors are scored by frequency: 1 = "daily" ; 2= "3-5 times per week"; 3 = "1-2 times per week" ; 4 = "less than once a week" ; 5= "never". Father's depression is the standardized average of depression inventory factors for all periods in which he took the depression inventory. A-levels attined is a dummy indicating that the father has at least an A-levels certification.

Table 2.2: Regressions of father's parenting behaviors on father's depression, household income, and educational attainment

Parenting behaviors related to affective engagement (such as physical play with the child, playing games with the child, and playing with toys with the child) all significantly associate with lower paternal depression and with lower maternal depression but do not associate with household income or paternal education when conditioned on depression. In contrast, less affectively-laden parenting behaviors, such as bathing the child and feeding the child, significantly associate with all three right-hand-side variables. These results are consistent with the existence of a suite of parenting behaviors that are peculiar to parents with depression.

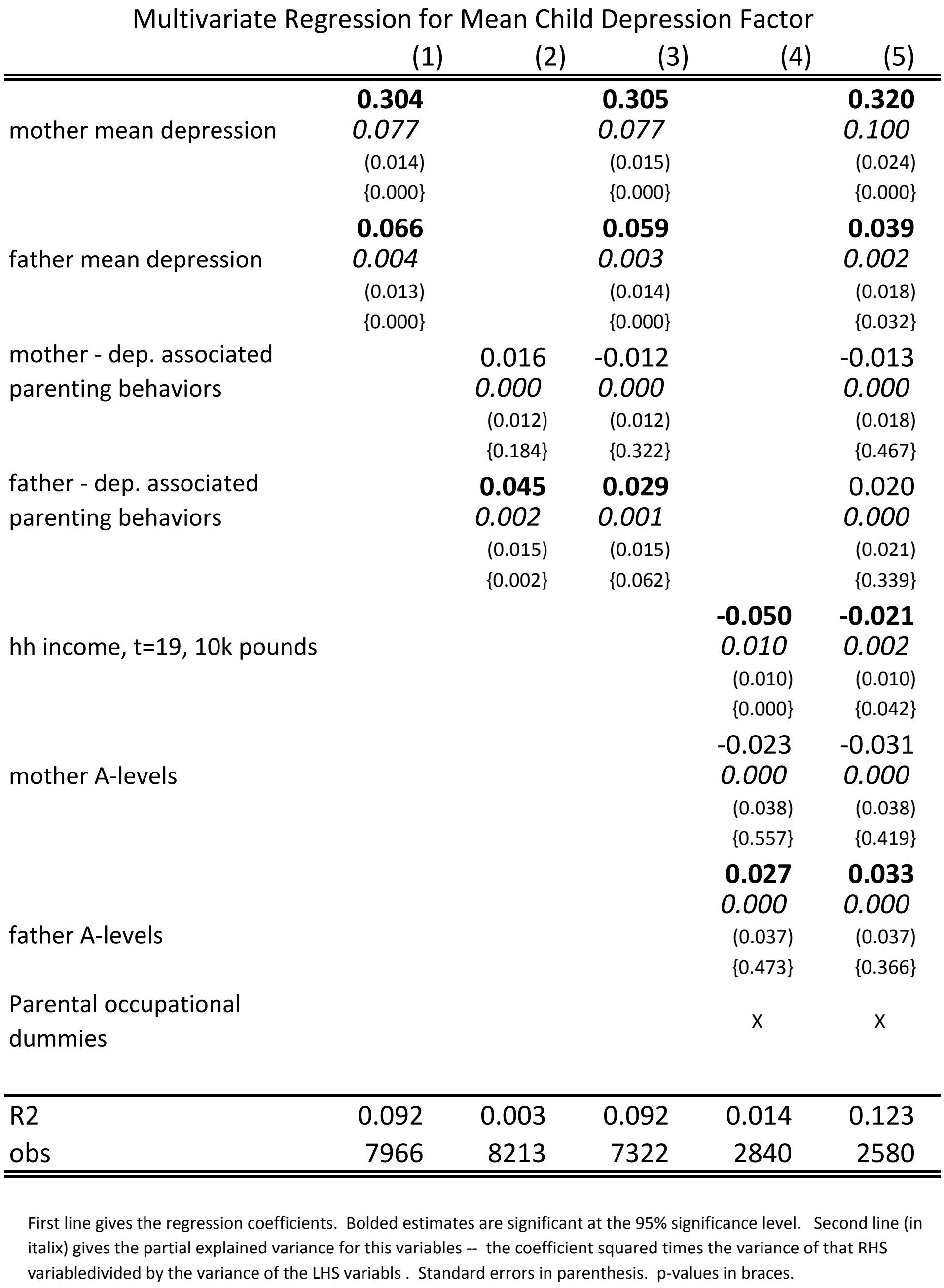
2.2.2 Parenting behaviors and the intergenerational transmission of depression

The measurement system of parenting behaviors can be inverted to estimate the latent parenting behaviors. Define the matrices $\Gamma \equiv \begin{pmatrix} \Gamma_{dep} & \Gamma_{inc} & \Gamma_{educ} \end{pmatrix}$ and $\theta_i \equiv \begin{pmatrix} \theta_{dep} & \theta_{inc} & \theta_{educ} \end{pmatrix}'$. Then an estimate of the vector of latent parenting behaviors θ_i is given by x

$$\hat{\theta}_i = (\Gamma'\Gamma)^{-1} \Gamma' P_i$$

Accordingly, I estimate these latent parenting behaviors by utilizing the regression estimates of Γ provided in the previous section.

Using these estimates of latent parenting behaviors, Table 2.3 explores the role of depression-associated parenting behaviors in the intergenerational transmission of depression. Broadly: depression is intergenerationally persistent; depression-associated parenting behaviors help to explain this persistence; neither income, education, income- nor education-associated parenting behaviors help to explain this persistence.



Multivariate Regression for Mean Child Depression Factor

	(1)	(2)	(3)	(4)	(5)
mother mean depression	**0.304** *0.077* (0.014) {0.000}		**0.305** *0.077* (0.015) {0.000}		**0.320** *0.100* (0.024) {0.000}
father mean depression	**0.066** *0.004* (0.013) {0.000}		**0.059** *0.003* (0.014) {0.000}		**0.039** *0.002* (0.018) {0.032}
mother - dep. associated parenting behaviors		0.016 *0.000* (0.012) {0.184}	-0.012 *0.000* (0.012) {0.322}		-0.013 *0.000* (0.018) {0.467}
father - dep. associated parenting behaviors		**0.045** *0.002* (0.015) {0.002}	**0.029** *0.001* (0.015) {0.062}		0.020 *0.000* (0.021) {0.339}
hh income, t=19, 10k pounds				**-0.050** *0.010* (0.010) {0.000}	**-0.021** *0.002* (0.010) {0.042}
mother A-levels				-0.023 *0.000* (0.038) {0.557}	-0.031 *0.000* (0.038) {0.419}
father A-levels				**0.027** *0.000* (0.037) {0.473}	**0.033** *0.000* (0.037) {0.366}
Parental occupational dummies				x	x
R2	0.092	0.003	0.092	0.014	0.123
obs	7966	8213	7322	2840	2580

First line gives the regression coefficients. Bolded estimates are significant at the 95% significance level. Second line (in italix) gives the partial explained variance for this variables -- the coefficient squared times the variance of that RHS variabledivided by the variance of the LHS variabls . Standard errors in parenthesis. p-values in braces.

Table 2.3: Regressions of child depression on parental parenting behaviors and covariates

Column one indicates that parental depression strongly predicts child depression, which should be unsurprising. Despite severe measurement error in depression inventories for adults and children, maternal depression explains about 10 percent of the variation in child depression, indicating strong intergenerational persistence

Column two indicates that depression-associated parenting behaviors may play a role in the child's depression. The explained variance is substantially smaller, which can partly be accounted for by noise in the estimation of the latent depression-associated parenting behaviors. In seperate regressions not included in this table, neither income nor educational associated latent parenting behaviors significantly associated with the child's depression after conditioning on depression associated parenting behaviors. Column three indicates that the father's depression-associated parenting behaviors significantly associate with child's depression even after controlling for parental depression, suggesting that depression-associated parenting behaviors play a causal role in the intergenerational transmission of depression.

Columns four and five indicate that parental household income and education explain little of the child's depression after controlling for parental depression and depression-associated parenting behaviors. Column four illustrates that parental houshold income significantly predicts the child's depression in a trivariate regression. This result is necessary for the intergenerational transmission of depression to play a role in intergenerational immobility. In contrast, parental education does not significantly associate with the child's depression. This result is consistent with depression capturing form of human capital distinct from educational human capital. Column five demonstrates that the association between parental household income and child depression largely evaporates after conditioning on parental depression and depression-associated parenting behaviors.

Altogether, these results indicate that neither parental education nor parental income play much role in the intergenerational transmission of depression – except to the extent that parenetal income and education affect depression in the parents.

2.3 The intergenerational transmission of depression and child young adult outcomes

In this section, I explore the role of the intergenerational transmission of depression in explaining child young-adult outcomes. I focus on two young-adult outcomes: (1) being employed or in school at age 20; and (2) having attained an A-levels certification. Failure to be enrolled or employed at timepoint 20 represents one of the worst potential outcomes for children. About 92 percent of children are enrolled or employed at timepoint 20. Meanwhile, the A-levels certification is the single most predictive educational outcome for a variety of SES outcomes among parents. Because A-levels certifcation occur prior to timepoint 20, these certifications adequately summarize the child's educational attainment.

Table 2.4 and display regressions of being employed or in school at timepoint 20 on various combinations of parental depression, child depression, depression-associated parenting behaviors, parental income, parental education, and controls. As a whole, these regressions indicate that depression – of the parents or the child – better explain young adult employment/enrollment status than any single other variable; moreover, parental income explains little to nothing after conditioning on depression – of either the parent or the child. In turn, parental depression explains child young-adult outcomes but not after conditioning on the child's depression, indicating that parental depression proxies for the transmission of depression to offspring.

The effect size of adolescent depression on young adudlt employment/enrollment status is both substantial and robust to which covariates are included in the model. Indeed, Multivariate Regression for Child: Enrolled or Employed t=20

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Child mean depression	**-0.037**				**-0.032**					**-0.033**
	0.015				*0.010*					*0.013*
	(0.006)				(0.007)					(0.009)
	{0.000}				{0.000}					{0.000}
mother mean depression		**-0.022**			-0.005			**-0.019**	-0.014	
		0.005			*0.000*			*0.004*	*0.002*	
		(0.007)			(0.006)			(0.007)	(0.010)	
		{0.001}			{0.464}			{0.005}	{0.143}	
father mean depression		-0.003			0.000			0.000	0.000	
		0.000			*0.000*			*0.000*	*0.000*	
		(0.005)			(0.006)			(0.006)	(0.007)	
		{0.595}			{0.988}			{0.996}	{0.967}	
mother - dep. assoc. par. Behaviors					-0.008		**-0.016**	-0.012	-0.001	
					0.001		*0.003*	*0.002*	*0.000*	
					(0.006)		(0.006)	(0.006)	(0.008)	
					{0.172}		{0.008}	{0.070}	{0.916}	
father - dep. assoc. par. behaviors					-0.010		-0.018	-0.013	-0.003	
					0.001		*0.004*	*0.002*	*0.000*	
					(0.011)		(0.011)	(0.011)	(0.015)	
					{0.367}		{0.087}	{0.265}	{0.821}	
mother - educ assoc. par. behaviors					-0.005		**-0.012**	-0.008	-0.002	
					0.000		*0.002*	*0.001*	*0.000*	
					(0.006)		(0.006)	(0.006)	(0.009)	
					{0.400}		{0.049}	{0.230}	{0.792}	
father - educ. assoc. par. behaviors					0.018		**0.027**	0.022	-0.004	
					0.004		*0.009*	*0.005*	*0.000*	
					(0.012)		(0.011)	(0.012)	(0.017)	
					{0.129}		{0.018}	{0.070}	{0.813}	
hh income, t=19, 10k pounds			**0.010**			0.004			0.003	0.002
			0.006			*0.001*			*0.000*	*0.000*
			(0.003)			(0.004)			(0.004)	(0.004)
			{0.000}			{0.264}			{0.470}	{0.633}
mother A-levels				**0.035**		0.017			0.012	0.015
				0.003		*0.001*			*0.001*	*0.001*
				(0.009)		(0.014)			(0.015)	(0.014)
				{0.000}		{0.239}			{0.414}	{0.291}
father A-levels				0.014		-0.007			-0.004	-0.006
				0.001		*0.000*			*0.000*	*0.000*
				(0.009)		(0.013)			(0.014)	(0.013)
				{0.138}		{0.607}			{0.748}	{0.604}
Parental occupational dummies						X			X	X
R2	0.015	0.006	0.006	0.006	0.013	0.018	0.003	0.006	0.023	0.030
obs	4005	3613	2176	3571	3296	1739	3666	3375	1598	1734

First line gives the regression coefficients. Bolded estimates are significant at the 95% significance level. Second line (in italix) gives the partial explained variance for this variables -- the coefficient squared times the variance of that RHS variabledivided by the variance of the LHS variabls . Standard errors in parenthesis. p-values from two-tailed t-test in braces.

Table 2.4: Regressions of child being employed or in school at age 20 on depression of child, depression of parents, parenting behaviors, and covariates

a single standard deviation increase in the child's adolescent depression predicts about a 4 percentage point increase in the probability of being neither enrolled nor employed at timepoint 20, indicating that children in highest fifth of the depression distribution in prior years are three times as likely to be neither employed nor enrolled in education than children in the bottom fifth of the depression distribution.

Parental depression also significantly associates with worse enrollment/employment for the child, but not when conditioning on the child's depression, indicating that the child's depression mediates this relationship. Likewise, depression and education-associated parenting behaviors predict worse enrollment/employment of the child, but not when conditioning on the child's depression. Including a large set of parental variables, including parental education, income, and occupation, does little to attenuate the relationship between child's depression and the child's employment/enrollment outcome.

Table 2.5 displays comparable regressions using A-levels certification as the dependent variable. The patterns are broadly similar, with child depression predicting large changes in certification depressions what other variables are included as controls. Two key differences stand out. First, depression and education-associated parenting behaviors remain significantly associated with the child's A-levels certification after including the child's depression. Second, household income and parental education remain significantly associated with the child's A-levels certification after including the child's depression. These results indicate that educational attainment and labor market participation may depend upon different dimensions of human capital.

2.4 Parental stress and child outcomes

I find evidence that parental stress predicts adverse child outcomes and that this is mediated through parental depression. Several indices measuring parental neighborhood stressors, Multivariate Regression for Child: Attained A-levels Qualification

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Child mean depression	**-0.051**				**-0.048**			**-0.038**
	0.009				*0.007*			*0.005*
	(0.009)				(0.011)			(0.014)
	{0.000}				{0.000}			{0.006}
mother mean depression		**-0.044**		**-0.040**	-0.024		-0.012	
		0.007		*0.005*	*0.002*		*0.000*	
		(0.010)		(0.010)	(0.011)		(0.014)	
		{0.000}		{0.000}	{0.028}		{0.414}	
father mean depression		-0.004		0.001	0.006		-0.011	
		0.000		*0.000*	*0.000*		*0.000*	
		(0.009)		(0.010)	(0.010)		(0.013)	
		{0.638}		{0.916}	{0.545}		{0.395}	
mother - dep. assoc. par. Behaviors			**-0.043**	**-0.043**	**-0.042**		0.000	
			0.007	*0.007*	*0.007*		*0.000*	
			(0.010)	(0.010)	(0.010)		(0.014)	
			{0.000}	{0.000}	{0.000}		{0.994}	
father - dep. assoc. par. behaviors			**-0.088**	**-0.082**	**-0.079**		-0.019	
			0.030	*0.022*	*0.020*		*0.001*	
			(0.017)	(0.018)	(0.019)		(0.025)	
			{0.000}	{0.000}	{0.000}		{0.458}	
mother - educ assoc. par. behaviors			**-0.032**	**-0.032**	**-0.030**		-0.004	
			0.005	*0.005*	*0.004*		*0.000*	
			(0.010)	(0.011)	(0.011)		(0.014)	
			{0.002}	{0.003}	{0.007}		{0.774}	
father - educ. assoc. par. behaviors			**0.083**	**0.084**	**0.082**		0.000	
			0.028	*0.025*	*0.023*		*0.000*	
			(0.018)	(0.019)	(0.019)		(0.026)	
			{0.000}	{0.000}	{0.000}		{0.999}	
hh income, t=19, 10k pounds						0.023	0.017	0.021
						0.008	*0.004*	*0.006*
						(0.007)	(0.007)	(0.007)
						{0.001}	{0.017}	{0.003}
mother A-levels						**0.097**	**0.105**	**0.096**
						0.012	*0.014*	*0.012*
						(0.027)	(0.028)	(0.027)
						{0.000}	{0.000}	{0.000}
father A-levels						**0.080**	**0.078**	**0.082**
						0.008	*0.007*	*0.008*
						(0.025)	(0.026)	(0.025)
						{0.001}	{0.003}	{0.001}
Parental occupational dummies						X	X	X
R2	0.009	0.007	0.013	0.018	0.023	0.097	0.094	0.101
obs	3879	3499	3557	3273	3196	1708	1568	1703

First line gives the regression coefficients. Bolded estimates are significant at the 95% significance level. Second line (in italix) gives the partial explained variance for this variables -- the coefficient squared times the variance of that RHS variabledivided by the variance of the LHS variabls . Standard errors in parenthesis. p-values in braces.

Table 2.5: Regressions of child A-levels certification on depression of child, depression of parents, parenting behaviors, and other covariates

social support, and relationship stress strongly predict parental depression, explaining about a third of the variance for both mothers and fathers. These stressors predict depression-associated parenting behaviors. And these parental stressors also predict the child's young adult outcomes – but to a lesser extent after conditioning on parental depression or the child's adolescent depression

Measures of parental stress and stress buffers explain about a third of the variance in parental depression. These stressors/buffers all contain a strong social dimension – and are inherently endogenous. Plausibly exogenous stressors, such as bereavement shocks and criminal victimization (conditional on expected risk of victimization), predict little or no significant increase in parental depression. See Appendix D.

For each survey in which I have the relevant measures, I estimate a single latent factor of social support for the mother and father, neighborhood quality, and partner relationship quality. I then calculate the averages of each of these factors for surveys at all timepoints in which respondents completed the surveys.

Together, these four variables explain 21 percent of the variation in mean depression scores of women and 32 percent of the variation in mean depression scores of men. Given the putatively high measurement error in depression factors and probably high measurement error in observed endogenous stressors, this result indicates an extremely strong association between these endogenous stressors and unobserved depression. One source of endogeneity should be emphasized: depression causes social withdrawal as well as a more pessimistic outlook. As a consequence, many of these measures of stressors can be thought of as alternative measures of depression (rather than causes).

These parental stressors also predict child outcomes; parental depression partly mediates this relationship. For readability, I construct an aggregate stress statistic as a weighted average of the life-events index, social support factor, neighborhood quality factor, and partner relationship quality factor, where the weights are determined by the coefficients from a regression of maternal depression on these stressors. Table 2.6 displays regressions of child depression on parental depression and the aggregate stress statistics as well as regressions of the child being employed or enrolled at age 20 on parental depreession, the aggregate stress statistic, and covariates.

These results indicate that aggregate parental stress explains 4.4 percent of the variance in child adolescent depression, but conditioning on parental depression drops the partial explained variance to 1.1 percent. This result is consistent with parental stress impacting the child primarily through parental depression. In turn, parental stressors predict adverse young-adult outcomes for the child ("employed or in school at age 20"), but to a greatly reduced extent after controlling for the child's depression and not at all after controlling for the child's depression and other parental characteristics. In this regression with many controls, only the child's depression significantly associates with this young-adult outcome and the coefficient on child's depression is no smaller than in the univariate regression.

Regressions on Parental Social Stressors

	Child Mean Depression			Child: Job or School t=20				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(7)
summary stress index	**0.223**		**0.113**	**-0.022**			**-0.017**	0.000
	(0.014)		(0.015)	(0.006)			(0.006)	(0.008)
	{0.000}		{0.000}	{0.000}			{0.009}	{0.998}
depression		**0.054**	0.019		-0.001		0.005	0.001
		(0.013)	(0.014)		(0.005)		(0.005)	(0.007)
		{0.000}	{0.154}		{0.800}		{0.404}	{0.929}
mother mean depression		**0.279**	**0.242**		**-0.014**		-0.002	-0.003
		(0.014)	(0.015)		(0.006)		(0.006)	(0.007)
		{0.000}	{0.000}		{0.023}		{0.719}	{0.662}
child mean depression						**-0.033**	**-0.030**	**-0.042**
						(0.007)	(0.007)	(0.009)
						{0.000}	{0.000}	{0.000}
hh income, t=19, 10k pounds								0.004
								(0.004)
								{0.298}
mother A-levels								**0.025**
								(0.013)
								{0.049}
father A-levels								0.007
								(0.012)
								{0.553}
R2	0.044	0.078	0.086	0.005	0.002	0.006	0.012	0.029
obs	8,094	8,094	8,094	3,593	3,593	3,593	3,593	1,909

First line gives the regression coefficients. Bolded estimates are significant at the 95% significance level. Huber-White Standard errors in parenthesis. p-values in braces.

Table 2.6: Parental stressors independently predict child depression but not after conditioning on parental depression

CHAPTER 3

EPIGENETIC TRANSMISSION

3.1 Introduction

Two questions motivate the use of epigenetic data. What role do epigenetic endowments play in the intergenerational transmission of depression? Can epigenetic data be utilized to measure depression when depression inventories are either too noisy or non-existent – in particular for young children? In this section, I provide suggestive evidence that biological endowments predisposing depression may be present at birth. Specifically, a forecast utilizing epigenetic data at birth explains twice as much variance in the child's adolescent depression in this sample as does maternal depression. However, this forecast should be interpreted warily. I fail to robustly replicate associations between depression and epigenetic variables highlighted in the biology literature, which limits interpretation. While this epigenetic forecast may feature predictive content because it proxies epigenetic endowments, it may also simply proxy for comorbid health behaviors of parents or for genetic endowments.

In this section, I describe the epigenetic data available in ALSPAC, develop an interpretive framework for this data, forecast child depression using the child's epigenetic data at birth, attempt replication of prior research using epigenetic data, and discuss findings.

3.2 Epigenetic Data

3.2.1 Overview of epigenetic data utilized in the ALSPAC

The ALSPAC study assayed methylation using the Infinium 450k array, which measures methylation in about 480,000 *probes* (variables) broadly covering the genome. On average, the nucleotide sequence covered by each probe includes 17 *CpG sites*, which are the locations at which DNA is methylated. These probes measure the share of methylated CpG sites. Reported as such, the datum is called a *beta-value*. Logistic transformation of the beta value yields the *M-value.* Mechanical failure of the 450k array causes some measurement error. *Cross-hybridization* – when a DNA sequence matches with the wrong probe – is one potentially serious source of measurement error (Chen et al., 2013), and the full extent of this problem is unknown. I drop from my analysis all probes known to cross-hybridize.

Auxiliary *annotation* data describe features of the DNA covered by probes. Such annotation data includes gene names in or near the probe, proximity to different positions on or near the gene (e.g., within 200 base pairs of the transcription start site), certain regulatory features (promoter, enhancer), relative CpG density in the neighborhood of the probe (e.g. CpG island, shore, shelf...), and some indicators of the variability of methylation within the probe (e.g. differentially methylated regions).

The ALSPAC study applies the Infinium array to white blood cells in blood samples. Methylation systematically varies across tissues types, so methylation patterns in white blood cells do not necessarily reflect methylation patterns in other tissues. Experiments on rodents often explore the epigenetic underpinnings of depression through analysis of brain tissue; results from these studies may not replicate in blood tissue. However, methylation patterns in human blood tissue may be predictive of depression for two reasons. First, changes in hormones or immune regulation that parallel depression may be reflected in altered white blood cell methylation, so that monocytes can serve as a barometer for broader allostatic changes. In addition, some experimental evidence indicates that *monocytes* (a type of white blood cell) play a central role mediating the effects of psychosocial stress on anxiety and depression-like behavior (Wohleb et al., 2013), so monocytes may serve as more than just barometers.

Blood samples for newborn children were drawn from the umbilical cord (*cord blood)*. Some research indicates that cord blood samples can be contaminated with maternal blood, but the severity of this problem is unknown.

Shares of different types of white blood cells may vary across samples, and the ALSPAC methylation data includes estimates of these shares. In principal, methylation rates could vary across samples due to variation in white blood cell shares rather than due to changes in methylation rates within particular cell types[1]. In the ALSPAC data, estimated shares do not vary substantially between samples and are not predictive of depression, so I make no further mention of these shares.

3.2.2 Methylation data subsample

The ALSPAC methylation data is longitudinal, including methylation data from blood samples of mothers and their respective children. Samples from mothers were taken two months prior to the due date and about 15 years after birth of the study child. Samples from children were taken at birth (cord blood), at about seven years of age, and at about fifteen years of age. The subjects included in the methylation data are a random subsample of those ALSPAC participants who participated in all of five clinic studies. Problems with the quality of blood samples or the methylation arrays caused some of these samples to be dropped. Altogether, the methylation data includes 987 mothers prior to birth, 992 mothers when the child is over the age of fifteen, 914 children at birth, 980 children at age seven, and 981 children over the age of fifteen.

---

[1] Variation in assay methylation rates caused by variation in shares of different cell types may be more of a problem for biologists and medical researchers than for social scientists. Such contamination may impede the ability of biologists to understand changes at the cellular level. However, variation in cell type may contain useful information in itself. If a researcher uses methylation data to proxy for past exposures or biological stocks, it may be immaterial whether in cell type shares or changes in methylation within cell types drives the variation in methylation.

Mother-child pairs participating in all five clinic studies are a selected subsample. Prenatal EPDS depression factor scores are about a tenth of a standard deviation lower in this subsample than in the full sample. However, the share of mothers scoring above the 11-point EPDS threshold for depression is no lower in this subsample.

Lower depression scores in the clinic study should not be surprising in light of the model of affective capital. Participating in the clinics required traveling to the clinic and spending time under evaluation. ALSPAC compensated participants with about 40 british pounds. Implicitly, those who participated felt that the monetary compensation (as well as non-monetary compensation, such as satisfying feelings of responsibility) was worth the effort and the opportunity cost of time. Those with less affective capital would overweight the costs, biasing selection against them.

3.2.3 Preparation of methylation data

Prior to delivery of the data, ALSPAC regressed out 10 principal components from a set of control probes and regressed out slide fixed effects. This 'functional normalization' can be thought of as a way to calibrate the assays. The cleaned methylation data include beta-values of 481,764 probes. 70,890 of these probes are known to cross-hybridize (that is, match with the wrong DNA) and 29,234 contain polymorphic CpG's (that is, will fail to match with DNA of some test subjects), both of which cause measurement error (see Chen et. al., 2013); I drop these probes from my analysis. I also drop probes on the Y-chromosome. I converted the beta-values (rates of methylation within probes, bounded between 0 and 1) to M-values (logistic transformation).

3.3 Interpretive framework and forecasting methodology

I interpret the methylation data from the perspective of a linear measurement system. Latent variables describe unobserved biological stocks; these biological stocks feature unique epigenetic patterns so that variation in these stocks generates variation in the epigenetic data. Analysis of the epigenetic data aims to extract information about the biological stocks that generate the data. This perspective yields insights for identification and forecasting. In this section, I describe these insights as well as the forecasting approach I apply to my data.

$$\underbrace{M_i}_{m\times 1} = \underbrace{\Gamma}_{m\times k}\underbrace{\theta_i}_{k\times 1} + \underbrace{\varepsilon_i}_{m\times 1} \quad \text{linear measurement system}$$

M_i : column vector of epigenetic data for person i

θ_i : row vector of latent variables; 1st element $\theta_{1,i}$ represents depression associated stock

Γ : parameter matrix; each column is an epigenetic pattern for corresponding stock

ε_i : idiosyncratic errors (e.g. measurement error)

3.3.1 Insights into identification

Identification of a depression-associated latent stock requires strong assumptions which much be justified from the biology (I provide greater depth and formally discuss identification in appendix F). Although exogeneous variation in a variable causing only the depression-associated stock $\theta_{1,i}$ can identify the epigenetic pattern for this stock Γ_1 if this stock does not in turn cause other stocks, identification of this stock's methylation pattern Γ_1 is not sufficient to identify the stock $\theta_{1,i}$. This is because another stock, say $\theta_{2,i}$, may feature a similar epigenetic pattern $\Gamma_1'\Gamma_2 \neq 0$. In this case, variation in this other stock would be confused for variation in $\theta_{1,i}$. Two approaches can circumvent this identification problem. First, exogenous variation in a cause of $\theta_{1,i}$ can be paired with proxies for stocks that feature epigenetically similar patterns. My data lacks useful exogenous variation, so I discuss this approach no further. Second, attention can be restricted to those epigenetic variables that only $\theta_{1,i}$ is purported to affect. Doing so requires a strong prior informed by the biology of depression. Given current knowledge, such a strong prior is unwarranted. Nevertheles, In section 3.4.2, I consider specific epigenetic variables emphasized by experiments on rodents (e.g. methylation near the promoter of the NR3C1 gene). Because I find no robustly significant associations between these epigenetic variables and maternal or child depression after adjusting for multiple hypothesis testing, I cannot use these epigenetic variables to identify $\theta_{1,i}$ even if it were reasonable to assume that only $\theta_{1,i}$ affected these epigenetic variables.

Given that identification of depression-associate stock is so fraught, I redirect attention to simply forecasting depression with epigenetic data.

3.3.2 Forecasting

In this section I describe my forecasting approach and some intuition behind it. I motivate this forecasting algorithm from a latent variable measurement system and provide more rigorous justification of claims in the appendix.

I utilize epigenetic data to calculate a leave-one-out forecast of depression $\hat{\theta}_{1,i}$ for each person i with the formula

$$\hat{\theta}_{1,i} = \dot{\theta}_{1,-i} M_{-i}' M_i$$

$\hat{\theta}_{1,i}$	forecast of depression for person i
$\dot{\theta}_{1,-i}$	row vector of standardized depression inventory scores for individuals other than i
M_{-i}	row-demeaned matrix of methylation data for individuals other than i...
	... each column is the vector of methylation data for a different individual.
M_i	column vector of methylation data for person i

and rescale the forecasts to be mean zero and unit variance across the N observations, where $\dot{\theta}_{1,-i}$ is the $1 \times (n-1)$ row vector of standardized depression inventory scores of all persons other than person i, and M_{-i} is the $m \times (n-1)$ matrix of methylation column vectors for each person other than i.

This leave-one-out forecast is equivalent to a normalized ridge regression forecast with arbitrarily large ridge penalty[2]. It is interesting that an estimate of depression derived from performing regression algebra on a measurement system would a limiting case of yield ridge regression – a canonical forecasting method. Moreover, if such a measurement system generates the data, then ridge regression is preferable to other forecasting methods that orient around model selection such as lasso or stepwise regression. In such a measurement system, the predictive content of predictors for a latent variable increases with their collinearity. It is impossible for two different predictors to be both predictive and orthogonal in this data generating process. However, model selection methods like lasso and step-

---

[2] A leave-one-out ridge regression forecast of y_i for person i using covariates x is given by $\hat{y}_i = x_i \left(X'X_{-i} + \lambda I\right)^{-1} X'_{-i} y_{-i}$ where I is the identity matrix, λ is the ridge regression penalty parameter, and the $-i$ index indicates that person i has been removed from the variable matrix. Because I am forecasting depression inventory factors, which have no natural scale, the forecast can be multiplied by a constant without loss of content. Multiply both sides of the equation by the ridge penalty. Then $\hat{y}_i \lambda = x_i \left(X'X_{-i} \lambda^{-1} + I\right)^{-1} X'_{-i} y_{-i}$. In the limit as λ approaches infinity, we have $\hat{y}_i \lambda = x_i X'_{-i} y_{-i}$. Renormalizing this forecast to be unit variance renders multiplication by λ irrelevant. This formula provides intuition about how ridge regression operates with a large penalty, linking ridge regression forecasts to estimates of latent variables derived from the latent variables simultaneous equations framework.

wise regression tend to drop the less predictive of collinear variables even if these dropped variables are highly predictive, instead retaining those predictors that are less collinear with other predictors. For this reason, a prior that covariates predict an outcome only because they load on that single outcome in a measurement system should lead a researcher to favor ridge regression over model selection methods.

This forecasting equation provides additional intuition for ridge regression with a large ridge penalty. That is, the term $M'_{-i}M_i$ describes the epigenetic similarity of person i with each other person, providing each epigenetic variable identical weight. Thus the forecast$\dot{\theta}_{-i}\left(M'_{-i}M_i\right)$ predicts person i's depression from the depression levels of epigenetically similar persons without giving consideration to how relevant any particular epigenetic variable is to depression. Such uniform weighting is appropriate under a strong prior that the effect size for each epigenetic variable is close to zero.

This forecast is computationally efficient when the number of predictors is large (after appropriately grouping matrix multiplication, computation time is $O\left(N^2K\right)$ where N is the number of observations and K is the number of predictors).

A forecast of depression using methylation data exploits the predictive content of epigenetic stocks correlated with depression. This is a double-edged sword, improving the forecast but reducing interpretability. For example, smoking and depression are highly comorbid. Because smoking associates with epigenetic changes, researchers should strongly anticipate that epigenetic data can predict depression if for no other reason than that that epigenetic data can predict smoking and smoking predicts depression. On the other hand, such confounds prevent interpreting the forecast of depression using epigenetic data as indicative of any one epigenetic stock (this is true for forecasts of smoking as well).

3.3.3 Forecasting procedure

My epigenetic data include about 1000 observations for mothers at two time points and about 1000 observations of children at three time points. To increase power, I would prefer to utilize all observations. However, methylation patterns predictive of depression in adults might not be identical to those predictive of depression in children.

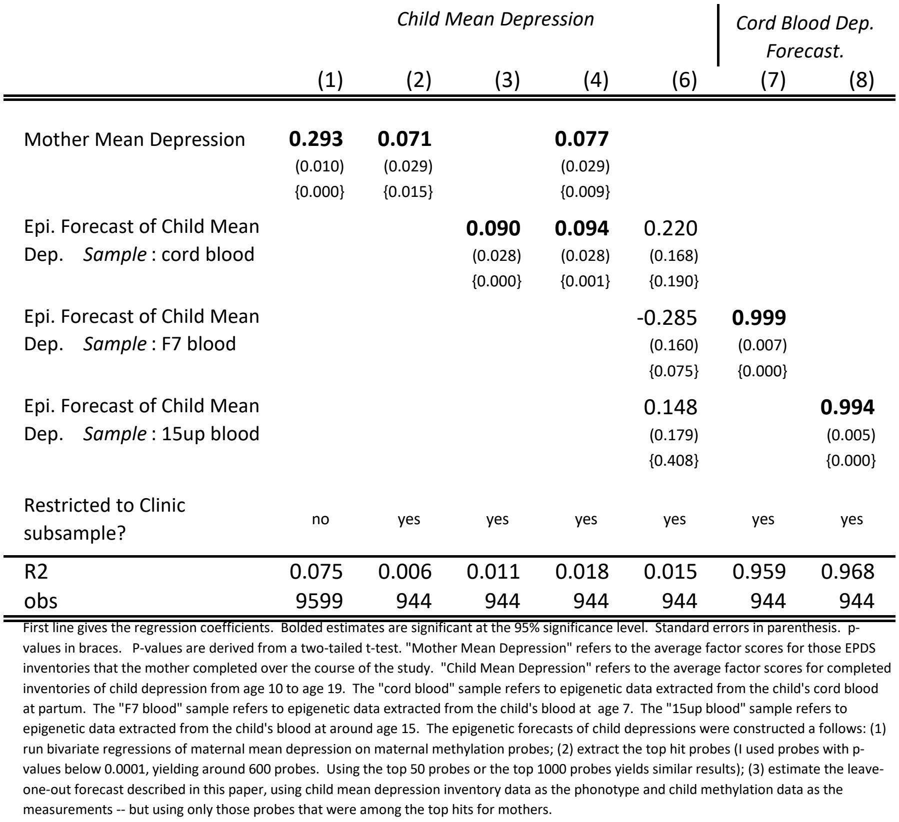
To construct forecasts for the child's depression using child epigenetic data, I:

1. I estimate bivariate regressions of maternal timepoint 15 methylation data on mean maternal depression.
    (a) I ascertain which methylation variables featured p-values below 0.001, yielding around 500 methylation variables – about twice as many as would be expected by chance under the null. Results using this procedure are robust to the threshold I utilize.
2. I estimate the leave-one-out forecast of child mean depression in adolescence using the formula described above:

$$\hat{\theta}_{1,i} = \dot{\theta}_{1,-i} M'_{-i} M_i$$

    (a) I construct a different forecast for each of the child's methylation samples: cord blood, age 7 blood, and age 15 blood.

To forecast maternal depression, I swap the above steps. That is, I take the top hits from bivariate regression of child depression on child methylation data and use these top hits in the leave-one-out forecast of maternal depression using maternal methylation data.



	Child Mean Depression					*Cord Blood Dep. Forecast.*	
	(1)	(2)	(3)	(4)	(6)	(7)	(8)
Mother Mean Depression	**0.293** (0.010) {0.000}	**0.071** (0.029) {0.015}		**0.077** (0.029) {0.009}			
Epi. Forecast of Child Mean Dep. *Sample* : cord blood			**0.090** (0.028) {0.000}	**0.094** (0.028) {0.001}	0.220 (0.168) {0.190}		
Epi. Forecast of Child Mean Dep. *Sample* : F7 blood					-0.285 (0.160) {0.075}	**0.999** (0.007) {0.000}	
Epi. Forecast of Child Mean Dep. *Sample* : 15up blood					0.148 (0.179) {0.408}		**0.994** (0.005) {0.000}
Restricted to Clinic subsample?	no	yes	yes	yes	yes	yes	yes
R2	0.075	0.006	0.011	0.018	0.015	0.959	0.968
obs	9599	944	944	944	944	944	944

First line gives the regression coefficients. Bolded estimates are significant at the 95% significance level. Standard errors in parenthesis. p-values in braces. P-values are derived from a two-tailed t-test. "Mother Mean Depression" refers to the average factor scores for those EPDS inventories that the mother completed over the course of the study. "Child Mean Depression" refers to the average factor scores for completed inventories of child depression from age 10 to age 19. The "cord blood" sample refers to epigenetic data extracted from the child's cord blood at partum. The "F7 blood" sample refers to epigenetic data extracted from the child's blood at age 7. The "15up blood" sample refers to epigenetic data extracted from the child's blood at around age 15. The epigenetic forecasts of child depressions were constructed a follows: (1) run bivariate regressions of maternal mean depression on maternal methylation probes; (2) extract the top hit probes (I used probes with p-values below 0.0001, yielding around 600 probes. Using the top 50 probes or the top 1000 probes yields similar results); (3) estimate the leave-one-out forecast described in this paper, using child mean depression inventory data as the phonotype and child methylation data as the measurements -- but using only those probes that were among the top hits for mothers.

Table 3.1: Forecast of Child Depression Using Epigenetic Data

3.4 Results of Epigenetic Analysis

3.4.1 Forecasting results

Leave-one-out forecasts of maternal depression using maternal methylation data and forecasts of child depression using child methylation data each explain a little over one percent of the variation in depression. In all five cases, these forecasts are highly significant.

Table 3.1 illustrates performance of the leave-one-out forecast of the child's adolescent depression. Regression results in this table compare the predictive content of the epigenetic forecast of child depression against maternal mean depression. Maternal depression is a key predictor of child depression and so should serve as a benchmark for the predictive content of the epigenetic forecast. By this standard, the epigenetic forecast performs well. Four features of this table deserve mention.

First, regressions one and two indicate that sample selection obliterates intergenerational associations in depression. Regressions one and two both regress child mean depression on maternal depression, but regression two restricts the sample to observations for which I have methylation data. The methylation data were drawn from those mother-child pairs who completed all five clinics; mothers in this sample scored lower on depression inventory scores, indicating that those with depression were less likely to participate in the clinics. While maternal depression is highly significant in both regressions, the explained variance drops by more than an order of magnitude. In my interpretation, selection into the methylation data reduces useful variation – membership in the epigenetic sample required participation in all five clinic visits, which required that the participants transport themselves to the hospital to undergo examination. In return, participants were compensated with around 40 pounds. Thus, participation in the clinic visit amounted to an effortful task that those with depression would be less likely to complete.

Second, regression three indicates that the epigenetic data at birth contains substantial predictive power. The leave-one-out forecast is quite significant and explains more than twice the variance in child depression as maternal depression does in this subsample.

Third, regression four indicates that the performance of the forecast is not explained by the child's epigenetics proxying for maternal depression. That is, the fit of the forecast is not a statistical artifact of the forecasting approach nor of contamination of child cord blood by maternal blood. In addition, the coefficients in this multivariate regression are virtually unchanged from those of the bivariate regressions. The results is consistent with large amounts of measurement error in both the forecast and the maternal measure of depression.

Fourth, regressions five, six, and seven indicate that epigenetic forecasts of adolescent depression are quite stable using different blood samples. These forecasts are extremely correlated. Including forecasts from multiple samples does little to improve the fit. This suggests that the epigenetics predictive of adolescent depression change very little in the years following birth.

3.4.2 Replication of prior research on methylation and depression

Experimental evidence indicates that epigenetic changes near particular genes may serve as key markers for depression (e.g., methylation in the promoter to the NR3C1 gene). Replicating such results in my sample and using these methylation variables in my forecast would improve interpretability of the forecast. However, I do not find robustly significant associations between these epigenetic variables and depression after multiple hypothesis testing (see Appendix G and Appendix H).

I also attempt replication of results from methylome-wide association studies. Prior research has associated depression and depression-related phenotypes with methylation changes in human blood tissue. I test if these methylation variables are significantly associated with depression in the mothers or children of the ALSPAC sample. Rarely do I find robustly significant associations after multiple hypothesis testing (see Appendix H).

3.4.3 Discussion of epigenetic results

It is puzzling that I do not robustly replicate prior results linking depression to epigenetic modifications. Four possibilities may account for limited replication. First, the epigenetic subsample is highly selected. The subsample features lower depression scores and obliterates intergenerational associations in depression. Intuitively, participation in the clinics selects against depression, amplifying measurement error in depression inventories in this subsample. In turn, larger measurement error reduces power to replicate prior resutls. Second, methylation data for white blood cells might not be the most useful epigenetic data.

For example, rodent experiments often utilize brain tissue for epigenetic analysis; but epigenetic changes in brain tissue might not be indicative of epigenetic changes in blood tissue (even if some human associations have explored these epigenetic changes in blood tissue). Third, human studies associating methylation with depression often feature small sample sizes and marginally significant results. In combination with the second reason, previously researched associations may be false positivese driven by publication bias.

Several of these issues could be rectified in future expansions of the ALSPAC data set. First, depression inventories of children in future waves of the ALSPAC – when children are in their mid twenties and depression inventories are more reliable – might reduce measurement error. Likewise, linkages with medical records concerning depression diagnoses or antidepressant usage in adulthood might serve as more accurate proxies for the child's adult depression. Second, other epigenetic data, such as histone modification data, may be more useful than methylation data. For instance, recent research indicates that, in monocytes (a type of white blood cell), histone acetylation near the nuclear factor kappa beta gene may play a key role in depression. Third, ALSPAC could mitigate sample selection in the epigenetic data by drawing a random sample from only the first clinic visit. That is, ALAPAC collected cord blood samples in hospitals where mothers gave birth. Presumably, depression would not be as predictive of participating in this clinic.

Failure to replicate prior epigenetic associations with depression reduces interprettability of my epigenetic forecast; nevertheless, compared to maternal depression in the epigenetic sample, the epigenetic forecast explains a large share of child depression. This result provides suggestive evidence that biological stocks predisposing depression are established early in life. Future research needs to further explore this possibility.

CHAPTER 4

A THEORY OF AFFECTIVE CAPITAL

4.1 Introduction

In my interpretation, chronic stress reduces a type of human capital I term affective capital. Low levels of affective capital increase the psychic cost of effort, thereby impairing labor market outcomes as well as reducing parenting inputs. In turn, parents produce a child's affective capital partly through endowments at birth and partly through investments over childhood, including goods investments and unpriced parenting inputs. Thus, affective capital transmits across generations and generates intergenerational associations in income. In the spirit of Becker-Tomes (Becker and Tomes, 1979) and subsequent iterations (e.g., Solon, 2014), I formalize these ideas and specialize the model so as to produce comparable log-linear intergenerational processes. The model highlights the economic nature of affective capital—both the maximization decision governing its accrual and its role in modulating the utility function—even though the biological nature of affective disorders might make them appear to reside outside economics. In addition, this model illustrates how investments that incur a direct psychic cost may be less sensitive to shocks to the parental budget constraint, even in the case that credit constraints bind.

The neologism "affective capital" might seem redundant given how heavily my research relies on both depression inventories and prior biological research into depression, but it serves two purposes. First, "affective capital" imposes a human capital interpretation, focusing on specific symptoms likely to be economically relevant (e.g. anticipatory anhedonia). The use of "affective capital" allows me to distinguish my economic interpretation from psychological research as such. Second, in my interpretation, depression manifests when affective capital drops below some threshold; however, variation in affective capital above this threshold may still be economically relevant. For example, just as social defeat can cause depression-like behaviors in previously normal rodents, so too can social winning cause behavioral changes opposite to those of depression-like behavior in previously normal rodents (Zhou et al., 2017; Franklin et al., 2017).

4.2 Model

4.2.1 Utility Function

The parent gains utility from consumption c_i and suffers disutility from labor effort L_i and from providing parenting inputs ι_i, where i indexes the family dynasty. Affective capital e_i reduces both the level and marginal disutility of labor effort and parenting inputs. In addition, the parent altruistically gains utility from some function of her child's expected utility $\alpha_i \mathbb{E}\left[\tilde{V}\left(H_i', e_i', a_i\right)\right]$, where α_i represents the parent's altruism, $\mathbb{E}$ is the expectation operator, primes denote values for the offspring, $\tilde{V}\left(\centerdot\right)$ represents the child's indirect utility, H_i is educative human capital, and a_i are inherited assets. I thus specify the parent's utility according to the additively separable function

$$\ln c_i - \frac{\rho}{e_i} L_i - \frac{1}{e_i} \iota_i + \alpha_i \mathbb{E}\left[\tilde{V}\left(H_i', e_i', a_i'\right)\right]$$

where the parameter ρ adjusts the relative disutility of labor effort and parenting inputs to affective capital.

As written, affective capital reduces the disutility of labor effort L_i and parenting inputs ι_i. However, multiplying the utility function by affective capital e_i would rationalize identical choices. In this transformation of the model, affective capital increases the marginal utility of consumption and the marginal utility from investments in the child through the child's indirect utility function. In this economic model, there is no substantive distinction between depressed individuals finding tasks more unpleasant and finding the compensation for these tasks less rewarding. Fatigue is the mirror image of anhedonia.

$\tilde{V}\left(H_i, e_i, a_i\right)$ represents the child's indirect utility function under a simplifying assumption – that the child bears no children of her own (or, alternatively, that the child's altruism parameter were set to 0). I eliminate consideration of grandchildren purely for convenience. However, this assumption can be rationalized as multi-generational myopia – parents consider the intergenerational effects of their investments but do not consider the multigenerational effects. Alternatively, this assumption can be treated as a tractable approximation to multigenerational altruism. Given both the qualitative similarities between 'myopic' altruism and true altruism as well as the difficulty of using data to distinguish between different forms of altruism, I view this assumption as an innocuous one.

4.2.2 Budget Constraint

The parent spends from permanent income $\xi_i H_i L_i^{\phi} + a_i$. The parent generates income through a concave function of labor market effort L_i^{ϕ} ($\phi \in [0,1]$), which is multiplied by an exogenous shock ξ_i and educative human capital H_i. In addition, the parent may inherit assets a_i.

The parent spends permanent income on consumption c_i, affective investments I_i, educative inputs s_i ('schooling'), and bequests a_i' (which are discounted by the interest rate r). Altogether, the parent's budget constraint is given by

$$\xi_i H_i L_i^{\phi} + a_i = c_i + I_i + s_i + a_i' \frac{1}{1+r}$$

Thus, the parent's lagrangian is given by:

$$\begin{aligned}\ln c_i - \frac{\rho}{e_i} L_i - \frac{1}{e_i} \iota_i + \alpha_i \mathbb{E}\left[\tilde{V}\left(H_i', e_i', a_i'\right)\right] \\ + \lambda_i \left[\xi_i H_i L_i^{\phi} + a_i - c_i - I_i - s_i - a_i' \frac{1}{1+r}\right] \\ + \eta_i a_i',\end{aligned}$$

where λ_i is the Lagrange multiplier on the budget constraint and η_i is the Lagrange multiplier preventing negative bequests.

To summarize: the parent's choice variables are c_i, L_i, ι_i, $I_{i,t}$, $s_{i,t}$, and a_i'; the parent's stocks are H_i and e_i; and the parent's preference parameters are ρ and α_i.

4.2.3 Human Capital Production

I assume that human capital production is Cobb-Douglas:

$$e' = e^{\beta_{ee}} \iota^{\beta_{e\iota}} I^{\beta_{eI}} g_e' \varepsilon_e'$$

$$H' = H^{\beta_{HH}} e'^{\beta_{H,e,1}} s^{\beta_{H,I,1}} g_H' \varepsilon_H'$$

Parental choice variables determine two components of affective capital production: the parenting input component $\iota^{\beta_{e\iota}}$ and the affective goods component $I^{\beta_{eI}}$. Three components are not determined by parental choice variables: direct epigenetic bequests $e^{\beta_{ee}}$, genetics g_e', and shocks ε_e'. The parameter on direct epigenetic bequests β_{ee} can be thought of as the biologically determined intergenerational depreciation of epigenetic stock.

Parents' beliefs about the parameters in this production function may diverge from the actual parameters in the production function. In particular, parent i believes parameter $\beta_{k,k'}$

takes on the value $\beta_{k,k',i}$. Thus, variation in parental beliefs about the production function parameters can affect child outcomes by shifting what the parent believes to be optimal investment.

These production functions impose several strong assumptions. These assumptions permit a more tractable model, but the consequencefs of these assumptions should be kept in mind. First, Cobb-Douglas production features a constant elasticity of substitution. Combined with log utility in consumption, this production function implies that parents will not alter investments either to exploit or mitigate shocks to the child's human capital. Second, parents make investments in a single period. With Cobb-Douglas production, additional periods of investment do not add much content; parents would smooth goods investments across periods, devoting more resources towards more productive periods and with the discount rate biasing investments towards later periods. Thus, these first two assumptions implicitly ignore dynamic complementarity. The focus of the model is on intergenerational mobility rather than the timing of investments or the efficiency-equality tradeoffs of early-childhood interventions, so I view this sacrifice as acceptable. Third, these production functions do not permit self-investment. Children passively receive inputs, and the stocks do not change in adulthood. Fourth and relatedly, the child's educative human capital does not enter into his production of affective capital. This is a natural specialization given that individuals in the model do not self-invest, but it is unrealistic.

4.3 Equilibrium

4.3.1 Indirect utility function

I derive the indirect utility function $\tilde{V}\left(e',H',a'\right)$ for childless offspring by taking first-order conditions after setting the altruism parameter to zero. I take approximations around no bequests $a = 0$ so as to admit closed form solutions. Derivations are provided in Appendix I.

The solutions to the choice variables in terms of stocks and parameters are given by :

$$\begin{aligned} L^* &= \frac{e}{\rho}\phi - \frac{1}{\xi H}L^{1-\phi}a \\ &\approx \frac{e}{\rho}\phi \\ c^* &= \frac{e}{\rho}\phi\xi H\left(\frac{e}{\rho}\phi - \frac{1}{\xi H}L^{1-\phi}a\right)^{\phi-1} \\ &\approx H\xi\left(\frac{e}{\rho}\phi\right)^{\phi-1} \end{aligned}$$

and the indirect utility function is given by

$$\tilde{V}\left(H_i', e_i', a_i'\right) \approx \kappa_{\tilde{V}} + \phi\ln(e) + \ln H + \left(\frac{\rho}{\phi}\right)^{\phi}\frac{a}{\xi H e^{\phi}} + \ln\xi$$

where $\kappa_{\tilde{V}}$ is some constant as a function of parameters.

Thus, when the parent's credit constraints bind (so $a = 0$), the indirect utility function for the child is given by $\ln\left(He^{\phi}\right)$ plus a constant. The marginal utility of bequeathing assets when credit constraints bind depends inversely on the child's educative and affective capital. Intuitively, an inheritance is more valuable to children who are less productive or less willing to work.

4.3.2 Solutions to choice variables for the parent

Given the approximation to the child's indirect utility function, I solve for the parent's choice variables in terms of stocks and parameters under the assumption that the credit constraint binds ($a = 0$).

$$\begin{aligned}
L^* &= \phi\frac{e}{\rho}\left(1+\alpha\left[\beta_{eI}\left(1+\beta_{He}\right)+\beta_{Hs}\right]\right)-\frac{a}{H\xi}L^{1-\phi}\\
&\approx \phi\frac{e}{\rho}\left(1+\alpha\left[\beta_{eI}\left(1+\beta_{He}\right)+\beta_{Hs}\right]\right)\\
c^* &= \phi\frac{e}{\rho}H\xi\left(\phi\frac{e}{\rho}\left(1+\alpha\left[\beta_{eI}\left(1+\beta_{He}\right)+\beta_{Hs}\right]\right)-\frac{a}{H\xi}L^{1-\phi}\right)^{\phi-1}\\
&\approx H\xi\left(\phi\frac{e}{\rho}\right)^{\phi}\left(1+\alpha\left[\beta_{eI}\left(1+\beta_{He}\right)+\beta_{Hs}\right]\right)^{\phi-1}\\
I^* &= \alpha\beta_{eI}\left(1+\beta_{He}\right)c^*\\
s^* &= \alpha\beta_{Hs}c^*\\
\iota_i^* &= \alpha_i\beta_{e\iota,i}\left(1+\phi\beta_{He,i}\right)e_i
\end{aligned}$$

4.3.3 Discussion of equilibrium choices

Parental affective inputs are of special interest

$$\iota_i^* = \alpha_i\beta_{e\iota,i}\left(1+\phi\beta_{He,i}\right)e_i$$

Even though these solutions assume that the credit constraint binds, parental income does not determine affective inputs. This follows because: parental affective inputs lack a market price (in goods or the opportunity cost of time); and changes in other inputs to the production function cause offsetting income and substitution effects.

Yet because the parent's beliefs about the impact of affective inputs on the child's outcome may diverge from the actual impact, beliefs $\beta_{e\iota,i}$ and $\beta_{He,i}$ determine investment. The role of beliefs emphasizes that parenting behaviors without a market price nevertheless result from an economic decision weighing costs against perceived benefits. Parental affective capital determines the (psychic) costs of these investments. However, these investments result from choices and are not mere behaviors.

With measures of these beliefs and altruism, one could run the regression

$$\ln \iota_i^* = \ln(\alpha_i) + \ln \beta_{e\iota} \left(1 + \phi \beta_{He,i}\right) + \ln e_i$$

Because measures of affective capital, altruism, and beliefs lack any natural scale, such a regression would be useful only for indicating the relative explanatory power of these determinants as well as the for serving as a first stage in an instrumental variables regression. Such a regression specification depends heavily upon functional form assumptions, however – if the production function were not unit elasticity, then other omitted variables would also be generating parental affective inputs.

4.4 Intergenerational processes

4.4.1 Log-linear vector autoregressions in human capital

These solutions imply that human capital follows an intergenerational vector AR-1:

$$\begin{aligned}
\ln H' &= \kappa_H(\alpha_i) + (\beta_{He} + \phi \beta_{Hs}) \ln e' + (\beta_{HH} + (1+\phi)\beta_{Hs}) \ln H + \ln g'_H + \beta_{Hs} \ln \xi + \ln \varepsilon'_H \\
&= \kappa_H(\alpha_i) + b_{H1} \ln e' + b_{H2} \ln H + b_{H3} \ln g'_H + b_{H4} \ln \xi + \ln \varepsilon'
\end{aligned}$$

$$\begin{aligned}
\ln e' &= \kappa_e(\alpha_i, \beta_{e\iota,i}) + (\beta_{ee} + \beta_{eI}(1+\phi) + \beta_{e\iota}) \ln e + \beta_{eI} \ln H + \ln g'_e + \beta_{eI} \ln \xi + \ln \varepsilon_{e'} \\
&= \kappa_e(\alpha_i, \beta_{e\iota,i}) + b_{e1} \ln e + b_{e2} \ln H + b_{e3} \ln g'_e + b_{e4} \ln \xi + \ln \varepsilon_{e'}
\end{aligned}$$

where κ's are functions of parameters and the reduced form b coefficients are functions of structural parameters. Note that parental altruism and parental beliefs about parameters in the production function enter into the κ variables; to the extent that altruism and beliefs vary across families and covary with parental human capital, parental human capital is endogenous to parental altruism and parental beliefs.

A sufficient condition for this process to be mean-reverting is that the sum of all coefficients across both processes is less than one:

$$\beta_{He} + \beta_{HH} + (1 + 2\phi)\,\beta_{Hs} + \beta_{ee} + \beta_{eI}\,(1 + \phi) + \beta_{e\iota} + \beta_{eI} < 1$$

The reduced-form b coefficients describe the cumulative effect of changing parental human capital prior to conception on the child's adult human capital. For example, b_{e1} captures the effect of parental affective capital on the child's affective capital through direct epigenetic bequests β_{ee}, through increased expenditures following greater parental labor supply $\beta_{eI}\,(1 + \phi)$, and through parenting behaviors altered by parental affective capital $\beta_{e\iota}$. Increasing the parent's affective capital by one unit prior to conception would increase the child's affective capital by b_{e1}. But increasing the parent's affective capital by one unit at birth would increase the child's affective capital by $b_{e1} - \beta_{ee}$.

4.4.2 Discussion of intergenerational processes

Parameterizing this model appropriately allows it to be reconciled with results from several literatures. By setting β_{eI} and β_{Hs} close to zero, this model can be reconciled with studies finding small effects of parental income shocks on child outcomes, even in the event that credit constraints bind. Under such a restriction, parental educative human capital has no effect on the child's affective capital, and the effect of parental educative human capital operates through β_{HH}. By allowing $\beta_{e\iota}$ and β_{He} to be greater than zero, this model can be reconciled with adoption studies finding effects of adoptive parent characteristics on the adoptive child's depression and the adoptive child's educational attainment. By allowing for variance in g_H and g_e, this model can be reconciled with behavioral genetic studies finding large effects of genes on education and depression; and by allowing β_{ee} to be greater than zero, this model can be reconciled with the biological literature on epigenetic bequests. Thus, this model is sufficiently flexible so as to apportion causality to several mechanisms relevant to intergenerational mobility, allowing this model to be reconciled to the stylized facts of intergenerational mobility.

CONCLUSION

A biological literature indicates that stress-associated biological stocks both modify behavior in ways that may reduce income and propagate to subsequent generations – both directly through biological endowments and indirectly through parenting behaviors. This thesis synthesizes the biological literature and provides an economic theory of affective capital that captures key implications for intergenerational mobility.

The essential insight of the theory is that affective capital operates through the utility function, modifying the disutility of effortful actions. As a consequence, parental investment decisions may be inelastic to parental income, even when credit constraints bind. This does not imply that parental investment is an irrational and reflexive behavior but rather that parents consider psychic costs of investments when choosing to invest and that these psychic costs may vary across the population. The income-inelastic nature of affective capital renders it a prime candidate for understanding intergenerational mobility when credit constraints do not bind or are unimportant for parental investments. In addition, the model of affective capital is sufficiently flexible to apportion causality to genetics, epigenetic endowments, parenting behaviors, and stressors in the formation of affective capital. The relative importance of these inputs is important for the timing and nature of interventions. The inelastic nature of many of these inputs suggests that many generations may be required for low stocks to be remediated.

Descriptive statistics in the ALSPAC data are broadly consistent with the theory of affective capital. Depression is highly persistent both within and between generations, plays a large role in explaining young-adult child outcomes irrespective of controls and despite being measured with error, and predicts affectively-laden parenting behaviors.

Methylation data in the ALSPAC only weakly predict depression. These weak results can be interpreted in a few ways. A strong – and premature – interpretation would be that epigenetic data generally have little predictive content for depression (either epigenetic data at partum or adult epigenetic samples). However, ALSPAC samples, as large as they are, may be too small to extract a useful signal, especially when compounded by selection issues, measurement error in both methylation data and phenotypes, and the high dimensionality of methylation data. In addition, other epigenetic changes, like histone modifications, and epigenetic modifications in tissues other than whole blood might hold better predictive content.

Auxiliary data could improve the predictive content of the methylation data. For example, expression data linked with methylation data could allow the researcher to restrict attention to probes whose annotation is most predictive of DNA expression, thereby reducing the dimensionality of the methylation data. Likewise, expression data itself could be exploited for predicting depression.

Given the accumulated evidence, it is incumbent upon economists to better understand the role of depression in intergenerational mobility and labor markets more broadly.

REFERENCES

Orazio Attanasio, Sarah Cattan, Emla Fitzsimons, Costas Meghir, and Marta Rubio-Codina. Estimating the Production Function for Human Capital: Results from a Randomized Control Trial in Colombia. *IZA Discussion Papers*, 2016. URL `http://ideas.repec.org/p/iza/izadps/dp8856.html`.

Diana Baumrind. Current patterns of parental authority. *Developmental Psychology*, 4(1, Pt.2):1–103, 1971.

Jonathan P. Beauchamp, David Cesarini, Magnus Johannesson, Matthijs J. H. M. van der Loos, Philipp D. Koellinger Koellinger, Patrick J. F. Groenen, James H. Fowler, J. Niels Rosenquist, A. Roy Thurik, and Nicholas A. Christakis. Molecular Genetics and Economics. *The journal of Economic Perspectives*, 25(4):57–82, 2011. doi: 10.1038/nrg1470. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/pmc3306008/`.

Aaron T. Beck, Robert A. Steer, and Margery G. Carbin. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1):77–100, jan 1988. ISSN 02727358. doi: 10.1016/0272-7358(88)90050-5. URL `http://linkinghub.elsevier.com/retrieve/pii/0272735888900505`.

Gary S. Becker and Nigel Tomes. An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility on JSTOR, 1979. URL `http://www.jstor.org.proxy.uchicago.edu/stable/1833328?seq=1{\#}page{\_}scan{\_}tab{\_}contents`.

Gary S. Becker, Kevin M. Murphy, and JJrg L. Spenkuch. The Manipulation of Children's Preferences, Old Age Support, and Investment in Children's Human Capital. *SSRN Electronic Journal*, aug 2014. ISSN 1556-5068. doi: 10.2139/ssrn.2510866. URL `http://papers.ssrn.com/abstract=2510866`.

S. L. Berger, T. Kouzarides, R. Shiekhattar, and A. Shilatifard. An operational definition of epigenetics. *Genes & Development*, 23(7):781–783, apr 2009. ISSN 0890-9369. doi: 10.1101/gad.1787609. URL `http://genesdev.cshlp.org/content/23/7/781.short`.

Pietro Biroli. Gene and Economics Interaction in The Formation of Human Capital. The Case of Obesity. In *Health & Healthcare in America: From Economics to Policy*. Ashecon, jun 2014. URL `https://ashecon.confex.com/ashecon/2014/webprogram/Paper1171.html`.

Sandra E. Black, Paul J. Devereux, and Kjell G. Salvanes. Does Grief Transfer across Generations? Bereavements during Pregnancy and Child Outcomes. *American*

Economic Journal: Applied Economics, 8(1):193–223, jan 2016. ISSN 1945-7782. doi: 10.1257/app.20140262. URL `http://pubs.aeaweb.org/doi/10.1257/app.20140262`.

Lex Borghans, Angela Lee Duckworth, James J. Heckman, and Bas ter Weel. The Economics and Psychology of Personality Traits. *Journal of Human Resources*, 43 (4):972–1059, oct 2008. ISSN 0022-166X. doi: 10.3368/jhr.43.4.972. URL `http://jhr.uwpress.org/content/43/4/972.short`.

Andy Boyd, Jean Golding, John Macleod, Debbie A Lawlor, Abigail Fraser, John Henderson, Lynn Molloy, Andy Ness, Susan Ring, and George Davey Smith. Cohort Profile: The âChildren of the 90s'âthe index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology*, 42(1):111–127, feb 2013. ISSN 1464-3685. doi: 10.1093/ije/dys064. URL `http://www.ncbi.nlm.nih.gov/pubmed/22507743http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3600618https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dys064`.

Emma Bränn, Fotios Papadopoulos, Emma Fransson, Richard White, Åsa Edvinsson, Charlotte Hellgren, Masood Kamali-Moghaddam, Adrian Boström, Helgi B. Schiöth, Inger Sundström-Poromaa, and Alkistis Skalkidou. Inflammatory markers in late pregnancy in association with postpartum depression - A nested case-control study. *Psychoneuroendocrinology*, 79:146–159, 2017. ISSN 03064530. doi: 10.1016/j.psyneuen.2017.02.029. URL `http://www.sciencedirect.com/science/article/pii/S0306453016307879`.

G W Brown and P M Moran. Single mothers, poverty and depression. *Psychological medicine*, 27(1):21–33, jan 1997. ISSN 0033-2917. URL `http://www.ncbi.nlm.nih.gov/pubmed/9122302`.

Christopher R Browning, Kathleen A Cagney, and James Iveniuk. Neighborhood stressors and cardiovascular health: crime and C-reactive protein in Dallas, USA. *Social science & medicine (1982)*, 75(7):1271–9, oct 2012. ISSN 1873-5347. doi: 10.1016/j.socscimed.2012.03.027. URL `http://www.sciencedirect.com/science/article/pii/S0277953612002948`.

Pedro Carneiro and James J. Heckman. The Evidence on Credit Constraints in Post-Secondary Schooling*. *The Economic Journal*, 112(482):705–734, oct 2002. ISSN 0013-0133. doi: 10.1111/1468-0297.00075. URL `http://doi.wiley.com/10.1111/1468-0297.00075`.

Charlotte A.M. Cecil, Rebecca G. Smith, Esther Walton, Jonathan Mill, Eamon J. McCrory, and Essi Viding. Epigenetic signatures of childhood abuse and neglect: Implications for psychiatric vulnerability. *Journal of Psychiatric Research*, 83:184–

194, 2016. ISSN 00223956. doi: 10.1016/j.jpsychires.2016.09.010. URL `http://www.sciencedirect.com/science/article/pii/S0022395616303636`.

Frances A Champagne and Michael J Meaney. Stress during gestation alters postpartum maternal care and the development of the offspring in a rodent model. *Biological psychiatry*, 59(12):1227–35, jun 2006. ISSN 0006-3223. doi: 10.1016/j.biopsych.2005.10.016. URL `http://www.ncbi.nlm.nih.gov/pubmed/16457784`.

Yi-an Chen, Mathieu Lemire, Sanaa Choufani, Darci T Butcher, Daria Grafodatskaya, Brent W Zanke, Steven Gallinger, Thomas J Hudson, and Rosanna Weksberg. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics*, 8(2):203–9, feb 2013. ISSN 1559-2308. doi: 10.4161/epi.23470. URL `http://www.ncbi.nlm.nih.gov/pubmed/23314698http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3592906`.

Steven W Cole, Gabriella Conti, Jesusa M G Arevalo, Angela M Ruggiero, James J Heckman, and Stephen J Suomi. Transcriptional modulation of the developing immune system by early life social adversity. *Proceedings of the National Academy of Sciences of the United States of America*, 109(50):20578–83, dec 2012. ISSN 1091-6490. doi: 10.1073/pnas.1218253109. URL `http://www.pnas.org/content/109/50/20578.abstract`.

Rupert Conrad, Ingo Wegener, Katrin Imbierowicz, Reinhard Liedtke, and Franziska Geiser. Alexithymia, temperament and character as predictors of psychopathology in patients with major depression. *Psychiatry Research*, 165(1):137–144, 2009. ISSN 01651781. doi: 10.1016/j.psychres.2007.10.013. URL `http://www.sciencedirect.com/science/article/pii/S016517810700368X`.

Elisabeth Conradt, Katheleen Hawes, Dylan Guerin, David A. Armstrong, Carmen J. Marsit, Edward Tronick, and Barry M. Lester. The Contributions of Maternal Sensitivity and Maternal Depressive Symptoms to Epigenetic Processes and Neuroendocrine Functioning. *Child Development*, 87(1):73–85, jan 2016. ISSN 00093920. doi: 10.1111/cdev.12483. URL `http://doi.wiley.com/10.1111/cdev.12483`.

Jennifer E Cropley, Sally A Eaton, Alastair Aiken, Paul E Young, Eleni Giannoulatou, Joshua WK Ho, Michael E Buckland, Simon P Keam, Gyorgy Hutvagner, David T Humphreys, Katherine G Langley, Darren Henstridge, David Martin, Mark A Febbraio, and Catherine M Suter. Grand paternal inheritance of an acquired metabolic trait induced by ancestral obesity is associated with sperm RNA. Technical report, Cold Spring Harbor Laboratory, mar 2016. URL `http://biorxiv.org/lookup/doi/10.1101/042101`.

Pim Cuijpers, Filip Smit, Brenda W J H Penninx, Ron de Graaf, Margreet ten Have, and Aartjan T F Beekman. Economic costs of neuroticism: a population-based study. *Archives of general psychiatry*, 67(10):1086–93, oct 2010. ISSN 1538-3636. doi: 10.1001/archgenpsychiatry.2010.130. URL `http://www.ncbi.nlm.nih.gov/pubmed/20921124`.

Flavio Cunha and James Heckman. The Technology of Skill Formation. *American Economic Review*, 97(2):31–47, may 2007. ISSN 0002-8282. doi: 10.1257/aer.97.2.31. URL `https://www.aeaweb.org/articles?id=10.1257/aer.97.2.31`.

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

Jonathan de Quidt, Johannes Haushofer, Jonathan de Quidt, and Johannes Haushofer. Depression for Economists. 2016. URL `http://econpapers.repec.org/bookchap/nbrnberch/13831.htm`.

Brian G Dias and Kerry J Ressler. Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nature neuroscience*, 17(1):89–96, jan 2014. ISSN 1546-1726. doi: 10.1038/nn.3594. URL `http://dx.doi.org/10.1038/nn.3594`.

Brian G Dias, Stephanie A Maddox, Torsten Klengel, and Kerry J Ressler. Epigenetic mechanisms underlying learning and the inheritance of learned behaviors. *Trends in neurosciences*, 38(2):96–107, feb 2015. ISSN 1878-108X. doi: 10.1016/j.tins.2014.12.003. URL `http://www.sciencedirect.com/science/article/pii/S0166223614002203`.

Francis X. Diebold, Glenn D. Rudebusch, and S. Boragan Aruoba. The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics*, 131(1):309–338, 2006.

David M Dietz, Quincey Laplant, Emily L Watts, Georgia E Hodes, Scott J Russo, Jian Feng, Ronald S Oosting, Vincent Vialou, and Eric J Nestler. Paternal transmission of stress-induced pathologies. *Biological psychiatry*, 70 (5):408–14, sep 2011. ISSN 1873-2402. doi: 10.1016/j.biopsych.2011.05.005. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3217197{\&}tool=pmcentrez{\&}rendertype=abstract`.

Matthias Doepke and Fabrizio Zilibotti. Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission. jun 2014. URL `http://www.nber.org/papers/w20214`.

Júlia Matzenbacher dos Santos, Marcos Lazaro Moreli, Shikha Tewari, and Sandra Aparecida Benite-Ribeiro. The effect of exercise on skeletal muscle glucose uptake in type 2 diabetes: An epigenetic perspective. *Metabolism*, 64(12):1619–1628, 2015. ISSN 00260495. doi: 10.1016/j.metabol.2015.09.013.

Ruth Feldman, Magi Singer, and Orna Zagoory. Touch attenuates infants' physiological reactivity to stress. *Developmental Science*, 13(2):271–278, mar 2010. ISSN 1363755X. doi: 10.1111/j.1467-7687.2009.00890.x. URL `http://www.ncbi.nlm.nih.gov/pubmed/20136923`.

Gary Felsenfeld. A brief history of epigenetics. *Cold Spring Harbor perspectives in biology*, 6(1), jan 2014. ISSN 1943-0264. doi: 10.1101/cshperspect.a018200. URL `http://www.ncbi.nlm.nih.gov/pubmed/24384572http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3941222`.

Anne Fernald, Virginia A Marchman, and Adriana Weisleder. SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental science*, 16(2):234–48, mar 2013. ISSN 1467-7687. doi: 10.1111/desc.12019. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3582035{\&}tool=pmcentrez{\&}rendertype=abstract`.

Tamara B Franklin, Bianca A Silva, Zina Perova, Livia Marrone, Maria E Masferrer, Yang Zhan, Angie Kaplan, Louise Greetham, Violaine Verrechia, Andreas Halman, Sara Pagella, Alexei L Vyssotski, Anna Illarionova, Valery Grinevich, Tiago Branco, and Cornelius T Gross. Prefrontal cortical control of a brainstem social behavior circuit. *Nature Neuroscience*, 2017. doi: 10.1038/nn.4470. URL `http://vyssotski.ch/Publications/NatNeurosci2017.pdf`.

A. Fraser, C. Macdonald-Wallis, K. Tilling, A. Boyd, J. Golding, G. Davey Smith, J. Henderson, J. Macleod, L. Molloy, A. Ness, S. Ring, S. M. Nelson, and D. A. Lawlor. Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology*, 42(1):97–110, feb 2013. ISSN 0300-5771. doi: 10.1093/ije/dys066. URL `https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dys066`.

Paul Gilbert. Evolution and depression: issues and implications. *Psychological Medicine*, 36(03):287, oct 2005. ISSN 0033-2917. doi: 10.1017/S0033291705006112. URL `http://www.journals.cambridge.org/abstract{\_}S0033291705006112`.

Camilla A M Glad, Johanna C Andersson-Assarsson, Peter Berglund, Ragnhildur Bergthorsdottir, Oskar Ragnarsson, and Gudmundur Johannsson. Reduced DNA methylation and psychopathology following endogenous hypercortisolism - a genome-wide study. *Scientific reports*, 7:44445, mar 2017. ISSN 2045-2322. doi: 10.

1038/srep44445. URL http://www.ncbi.nlm.nih.gov/pubmed/28300138http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5353706.

José María Gómez, Miguel Verdú, Adela González-Megías, and Marcos Méndez. The phylogenetic roots of human lethal violence. *Nature*, sep 2016. ISSN 0028-0836. doi: 10.1038/nature19758. URL http://www.nature.com/doifinder/10.1038/nature19758.

Richard L. Gorsuch. *Factor Analysis*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 2nd edition, 1983. ISBN 978-0898592023.

David W. Haley and Kathy Stansbury. Infant Stress and Parent Responsiveness: Regulation of Physiology and Behavior during Still-Face and Reunion. *Child Development*, 74(5):1534–1546, 2003. ISSN 0009-3920. doi: 10.1111/1467-8624.00621. URL www.jstor.org/stable/3696192http://www.jstor.org/stable/pdfplus/10.2307/3696192.pdf?acceptTC=true.

Margaret D. Hanson and Edith Chen. Daily stress, cortisol, and sleep: The moderating role of childhood psychosocial environments. *Health Psychology*, 29(4):394–402, 2010. ISSN 1930-7810. doi: 10.1037/a0019879. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/a0019879.

James J Heckman and Stefano Mosso. The Economics of Human Development and Social Mobility. *National Bureau of Economic Research Working Paper Series*, No. 19925, 2014. URL http://www.nber.org/papers/w19925.

James J. Heckman, Jora Stixrud, and Sergio Urzua. The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics*, 24(3):411–482, 2006. URL http://ideas.repec.org/a/ucp/jlabec/v24y2006i3p411-482.html.

Robin Holliday. Epigenetics: An overview. *Developmental Genetics*, 15(6):453–457, 1994. ISSN 0192-253X. doi: 10.1002/dvg.1020150602. URL http://doi.wiley.com/10.1002/dvg.1020150602.

Lotte C. Houtepen, Christiaan H. Vinkers, Tania Carrillo-Roa, Marieke Hiemstra, Pol A. van Lier, Wim Meeus, Susan Branje, Christine M. Heim, Charles B. Nemeroff, Jonathan Mill, Leonard C. Schalkwyk, Menno P. Creyghton, René S. Kahn, Marian Joëls, Elisabeth B. Binder, and Marco P. M. Boks. Genome-wide DNA methylation levels and altered cortisol stress reactivity following childhood trauma in humans. *Nature Communications*, 7:10967, mar 2016. ISSN 2041-1723. doi: 10.1038/ncomms10967. URL http://www.nature.com/doifinder/10.1038/ncomms10967.

Brian M Iacoviello, Lauren B Alloy, Lyn Y Abramson, and Jimmy Y Choi. The early course of depression: a longitudinal investigation of prodromal symptoms and their relation to the symptomatic course of depressive episodes. *Journal of abnormal psychology*, 119(3):459–67, aug 2010. ISSN 1939-1846. doi: 10.1037/a0020114. URL `http://www.ncbi.nlm.nih.gov/pubmed/20677835http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2916193`.

Oliver P. John and Sanjay Srivastava. *The Big Five Trait taxonomy: History, measurement, and theoretical perspectives.* Guilford Press, 1999. ISBN 1-57230-483-9 (Hardcover).

Randy W. Kamphaus, Martha D. Petoskey, and Anna Walters Morgan. *A history of intelligence test interpretation.* Guilford Press, 1997. ISBN 1-57230-147-3 (Hardcover).

Tim Kautz and Wladimir Zanoni. Measuring and Fostering Non-Cognitive Skills in Adolescence: Evidence from Chicago Public Schools and the OneGoal Program. *Working Paper*, 2014. URL `http://home.uchicago.edu/{~}tkautz/OneGoal{\_}TEXT.pdf`.

Michael P. Keane and Kenneth I. Wolpin. The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment. *International Economic Review*, 42 (4):1051–1103, nov 2001. ISSN 0020-6598. doi: 10.1111/1468-2354.00146. URL `http://doi.wiley.com/10.1111/1468-2354.00146`.

Eric I Knudsen, James J Heckman, Judy L Cameron, and Jack P Shonkoff. Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences of the United States of America*, 103 (27):10155–62, jul 2006. ISSN 0027-8424. doi: 10.1073/pnas.0600888103. URL `http://www.pnas.org/content/103/27/10155.abstract`.

Stephen A Krawetz, Adele Kruger, Claudia Lalancette, Rebecca Tagett, Ester Anton, Sorin Draghici, and Michael P Diamond. A survey of small RNAs in human sperm. *Human reproduction (Oxford, England)*, 26(12):3401–12, dec 2011. ISSN 1460-2350. doi: 10.1093/humrep/der329. URL `http://www.ncbi.nlm.nih.gov/pubmed/21989093http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3212879`.

Julia C Lemos, Matthew J Wanat, Jeffrey S Smith, Beverly A S Reyes, Nick G Hollon, Elisabeth J Van Bockstaele, Charles Chavkin, and Paul E M Phillips. Severe stress switches CRF action in the nucleus accumbens from appetitive to aversive. *Nature*, 490(7420):402–6, oct 2012. ISSN 1476-4687. doi: 10.1038/nature11436. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3475726{\&}tool=pmcentrez{\&}rendertype=abstract`.

Debra Lerner, David A Adler, Hong Chang, Leueen Lapitsky, Maggie Y Hood, Carla Perissinotto, John Reed, Thomas J McLaughlin, Ernst R Berndt, and William H Rogers. Unemployment, job retention, and productivity loss among employees with depression. *Psychiatric services (Washington, D.C.)*, 55(12):1371–8, dec 2004. ISSN 1075-2730. doi: 10.1176/appi.ps.55.12.1371. URL `http://ps.psychiatryonline.org/doi/abs/10.1176/appi.ps.55.12.1371`.

M.Christine Lovejoy, Patricia A Graczyk, Elizabeth O'Hare, and George Neuman. Maternal depression and parenting behavior. *Clinical Psychology Review*, 20(5):561–592, aug 2000. ISSN 02727358. doi: 10.1016/S0272-7358(98)00100-7. URL `http://www.sciencedirect.com/science/article/pii/S0272735898001007`.

Katherine R. Luking, David Pagliaccio, Joan L. Luby, and Deanna M. Barch. Depression Risk Predicts Blunted Neural Responses to Gains and Enhanced Responses to Losses in Healthy Children. *Journal of the American Academy of Child & Adolescent Psychiatry*, 55(4):328–337, 2016. ISSN 08908567. doi: 10.1016/j.jaac.2016.01.007.

Lundi Ly, Donovan Chan, and Jacquetta M. Trasler. Developmental windows of susceptibility for epigenetic inheritance through the male germline. *Seminars in Cell & Developmental Biology*, 43:96–105, 2015. ISSN 10849521. doi: 10.1016/j.semcdb.2015.07.006.

S Maccari, H J Krugers, S Morley-Fletcher, M Szyf, and P J Brunton. The consequences of early-life adversity: neurobiological, behavioural and epigenetic adaptations. *Journal of neuroendocrinology*, 26(10):707–23, oct 2014. ISSN 1365-2826. doi: 10.1111/jne.12175. URL `http://www.ncbi.nlm.nih.gov/pubmed/25039443`.

Matteo Manera and Angelo Marzullo. Modelling the load curve of aggregate electricity consumption using principal components. *Environmental Modelling & Software*, 20(11):1389–1400, 2005. ISSN 13648152. doi: 10.1016/j.envsoft.2004.09.019.

Asaf Marco, Tatiana Kisliouk, Tzlil Tabachnik, Noam Meiri, and Aron Weller. Overweight and CpG methylation of the Pomc promoter in offspring of high-fat-diet-fed dams are not "reprogrammed" by regular chow diet in rats. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 28(9):4148–57, sep 2014. ISSN 1530-6860. doi: 10.1096/fj.14-255620. URL `http://www.fasebj.org/content/28/9/4148.short`.

Emma L Marczylo, Akwasi A Amoako, Justin C Konje, Timothy W Gant, and Timothy H Marczylo. Smoking induces differential miRNA expression in human spermatozoa: a potential transgenerational epigenetic concern? *Epigenetics*, 7 (5):432–9, may 2012. ISSN 1559-2308. doi: 10.4161/epi.19794. URL `http://www.ncbi.nlm.nih.gov/pubmed/22441141`.

Lisa A. Martin, Harold W. Neighbors, Derek M. Griffith, Sampson N, Walters EE, and Kessler RC. The Experience of Symptoms of Depression in Men vs Women. *JAMA Psychiatry*, 70(10):1100, oct 2013. ISSN 2168-622X. doi: 10.1001/jamapsychiatry.2013.1985. URL `http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/jamapsychiatry.2013.1985`.

Grant S. McCall and Nancy Shields. Examining the evidence from small-scale societies and early prehistory and implications for modern theories of aggression and violence. *Aggression and Violent Behavior*, 13(1):1–9, 2008. ISSN 13591789. doi: 10.1016/j.avb.2007.04.001.

Courtney A Miller and J David Sweatt. Covalent modification of DNA regulates memory formation. *Neuron*, 53(6):857–69, mar 2007. ISSN 0896-6273. doi: 10.1016/j.neuron.2007.02.022. URL `http://www.ncbi.nlm.nih.gov/pubmed/17359920`.

A. K. Murashov, E. S. Pak, M. Koury, A. Ajmera, M. Jeyakumar, M. Parker, O. Williams, J. Ding, D. Walters, and P. D. Neufer. Paternal long-term exercise programs offspring for low energy expenditure and increased risk for obesity in mice. *The FASEB Journal*, 30(2):775–784, feb 2016. ISSN 0892-6638. doi: 10.1096/fj.15-274274. URL `http://www.fasebj.org/cgi/doi/10.1096/fj.15-274274`.

T Narita, T Sato, S Hirano, M Gota, K Sakado, and T Uehara. Parental child-rearing behavior as measured by the Parental Bonding Instrument in a Japanese population: factor structure and relationship to a lifetime history of depression. *Journal of affective disorders*, 57(1-3):229–34, jan 2000. ISSN 0165-0327. URL `http://www.ncbi.nlm.nih.gov/pubmed/10708836`.

Z Nemoda, R Massart, M Suderman, M Hallett, T Li, M Coote, N Cody, Z S Sun, C N Soares, G Turecki, M Steiner, and M Szyf. Maternal depression is associated with DNA methylation changes in cord blood T lymphocytes and adult hippocampi. *Translational Psychiatry*, 5(4):e545, apr 2015. ISSN 2158-3188. doi: 10.1038/tp.2015.32. URL `http://www.ncbi.nlm.nih.gov/pubmed/25849984http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4462598http://www.nature.com/doifinder/10.1038/tp.2015.32`.

Tim F Oberlander, Joanne Weinberg, Michael Papsdorf, Ruth Grunau, Shaila Misri, and Angela M Devlin. Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1) and infant cortisol stress responses. *Epigenetics*, 3(2):97–106, jan 2008. ISSN 1559-2308. URL `http://www.ncbi.nlm.nih.gov/pubmed/18536531`.

G Parker. The Parental Bonding Instrument: psychometric properties reviewed. *Psychiatric developments*, 7(4):317–35, jan 1989. ISSN 0262-9283. URL `http:`

//www.ncbi.nlm.nih.gov/pubmed/2487899.

W J Peyrot, S H Lee, Y Milaneschi, A Abdellaoui, E M Byrne, T Esko, E J C de Geus, G Hemani, J J Hottenga, S Kloiber, D F Levinson, S Lucae, N G Martin, S E Medland, A Metspalu, L Milani, M M Noethen, J B Potash, M Rietschel, C A Rietveld, S Ripke, J Shi, G Willemsen, Z Zhu, D I Boomsma, N R Wray, and B W J H Penninx. The association between lower educational attainment and depression owing to shared genetic effects? Results in ~25,000 subjects. *Molecular psychiatry*, 20(6):735–43, jun 2015. ISSN 1476-5578. doi: 10.1038/mp.2015.50. URL http://dx.doi.org/10.1038/mp.2015.50.

Anna C Phillips, Douglas Carroll, and Geoff Der. Negative life events and symptoms of depression and anxiety: stress causation and/or stress generation. *Anxiety, stress, and coping*, 28(4):357–71, 2015. ISSN 1477-2205. doi: 10.1080/10615806.2015.1005078. URL http://www.ncbi.nlm.nih.gov/pubmed/25572915http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4772121.

John P Polheber and Robert L Matchock. The presence of a dog attenuates cortisol and heart rate in the Trier Social Stress Test compared to human friends. *Journal of behavioral medicine*, 37(5):860–7, oct 2014. ISSN 1573-3521. doi: 10.1007/s10865-013-9546-1. URL http://www.ncbi.nlm.nih.gov/pubmed/24170391.

B Reed, N Fang, B Mayer-Blackwell, S Chen, V Yuferov, Y Zhou, and M J Kreek. Chromatin alterations in response to forced swimming underlie increased prodynorphin transcription. *Neuroscience*, 220: 109–18, sep 2012. ISSN 1873-7544. doi: 10.1016/j.neuroscience.2012.06.006. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3412925{\&}tool=pmcentrez{\&}rendertype=abstract.

A. B. Rodgers, C. P. Morgan, S. L. Bronson, S. Revello, and T. L. Bale. Paternal Stress Exposure Alters Sperm MicroRNA Content and Reprograms Offspring HPA Stress Axis Regulation. *Journal of Neuroscience*, 33(21):9003–9012, may 2013. ISSN 0270-6474. doi: 10.1523/JNEUROSCI.0914-13.2013. URL http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.0914-13.2013.

Ali B. Rodgers, Christopher P. Morgan, N. Adrian Leu, and Tracy L. Bale. Transgenerational epigenetic programming via sperm microRNA recapitulates effects of paternal stress. *PNAS*, 112(44):13699–13704, oct 2015. doi: 10.1073/pnas.1508347112. URL http://www.pnas.org/content/early/2015/10/14/1508347112.abstract?sid=01e653a0-6478-4d6e-bad0-a2334bc38298.

Tania L Roth and J David Sweatt. Annual Research Review: Epigenetic mechanisms and environmental shaping of the brain during sensitive periods of development. *Journal of child psychology and psychiatry, and allied disciplines*, 52(4):398–

408, apr 2011. ISSN 1469-7610. doi: 10.1111/j.1469-7610.2010.02282.x. URL `http://doi.wiley.com/10.1111/j.1469-7610.2010.02282.x`.

John Salamone, Merce Correa, Susana Mingote, Suzanne Weber, and Andrew Farrar. Nucleus Accumbens Dopamine and the Forebrain Circuitry Involved in Behavioral Activation and Effort-Related Decision Making: Implications for Understanding Anergia and Psychomotor Slowing in Depression. *Current Psychiatry Reviews*, 2(2):267–280, may 2006. ISSN 15734005. doi: 10.2174/157340006776875914. URL `http://www.ingentaconnect.com/content/ben/cpsr/2006/00000002/00000002/art00009`.

Lindsey Sherdell, Christian E Waugh, and Ian H Gotlib. Anticipatory pleasure predicts motivation for reward in major depression. *Journal of abnormal psychology*, 121(1):51–60, feb 2012. ISSN 1939-1846. doi: 10.1037/a0024945. URL `http://www.ncbi.nlm.nih.gov/pubmed/21842963http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3335300`.

Michael J. Sheriff, Charles J. Krebs, and Rudy Boonstra. The ghosts of predators past: population cycles and the role of maternal programming under fluctuating predation risk. *Ecology*, 91(10):2983–2994, oct 2010. ISSN 0012-9658. doi: 10.1890/09-1108.1. URL `http://doi.wiley.com/10.1890/09-1108.1`.

A K Short, K A Fennell, V M Perreau, A Fox, M K O'Bryan, J H Kim, T W Bredy, T Y Pang, and A J Hannan. Elevated paternal glucocorticoid exposure alters the small noncoding RNA profile in sperm and modifies anxiety and depressive phenotypes in the offspring. *Translational Psychiatry*, 6(6):e837, jun 2016. ISSN 2158-3188. doi: 10.1038/tp.2016.109. URL `http://www.nature.com/doifinder/10.1038/tp.2016.109`.

Gary Solon. Theoretical models of inequality transmission across multiple generations. *Research in Social Stratification and Mobility*, 35:13–18, 2014. ISSN 02765624. doi: 10.1016/j.rssm.2013.09.005.

Christopher Taber and Stephen V. Cameron. No Title. *Journal of Political Economy*, 112(1):132–182, 2004.

R A Thompson and C A Nelson. Developmental science and the media. Early brain development. *The American psychologist*, 56(1):5–15, jan 2001. ISSN 0003-066X. URL `http://www.ncbi.nlm.nih.gov/pubmed/11242988`.

Conrad Hal Waddington. The Epigenotype. *International Journal of Epidemiology*, 41(1):10–13, 1942.

Ian C G Weaver, Nadia Cervoni, Frances A Champagne, Ana C D'Alessio, Shakti Sharma, Jonathan R Seckl, Sergiy Dymov, Moshe Szyf, and Michael J Meaney. Epigenetic programming by maternal behavior. *Nature neuroscience*, 7(8):847–54, aug 2004. ISSN 1097-6256. doi: 10.1038/nn1276. URL `http://www.ncbi.nlm.nih.gov/pubmed/15220929`.

Simone R Witzmann, Jonathan D Turner, Sophie B Mériaux, Onno C Meijer, and Claude P Muller. Epigenetic regulation of the glucocorticoid receptor promoter 1(7) in adult rats. *Epigenetics*, 7(11):1290–301, nov 2012. ISSN 1559-2308. doi: 10.4161/epi.22363. URL `http://www.ncbi.nlm.nih.gov/pubmed/23023726http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3499330`.

Eric S Wohleb, Nicole D Powell, Jonathan P Godbout, and John F Sheridan. Stress-induced recruitment of bone marrow-derived monocytes to the brain promotes anxiety-like behavior. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 33(34):13820–33, aug 2013. ISSN 1529-2401. doi: 10.1523/JNEUROSCI.1671-13.2013. URL `http://www.jneurosci.org/content/33/34/13820.short`.

Qian Zhao, Jing Hou, Bo Chen, Xue Shao, Ruiming Zhu, Qian Bu, Hui Gu, Yan Li, Baolai Zhang, Changman Du, Dengqi Fu, Jueying Kong, Li Luo, Hailei Long, Hongyu Li, Yi Deng, Yinglan Zhao, and Xiaobo Cen. Prenatal cocaine exposure impairs cognitive function of progeny via insulin growth factor II epigenetic regulation. *Neurobiology of Disease*, 82:54–65, 2015. ISSN 09699961. doi: 10.1016/j.nbd.2015.05.014.

Tingting Zhou, Hong Zhu, Zhengxiao Fan, Fei Wang, Yang Chen, Hexing Liang, Zhongfei Yang, Lu Zhang, Longnian Lin, Yang Zhan, Zheng Wang, and Hailan Hu. History of winning remodels thalamo-PFC circuit to reinforce social dominance. *Science*, 357(6347), 2017. URL `http://science.sciencemag.org.proxy.uchicago.edu/content/357/6347/162.full`.

Frederick J. Zimmerman and Wayne Katon. Socioeconomic status, depression disparities, and financial strain: what lies behind the income-depression relationship? *Health Economics*, 14(12):1197–1215, dec 2005. ISSN 1057-9230. doi: 10.1002/hec.1011. URL `http://www.ncbi.nlm.nih.gov/pubmed/15945040`.

APPENDIX A

A - GLOSSARY

Affective Related to moods or feelings. For example, both major depressive disorder and bipolar disorder are affective disorders.

Allostasis Long-lasting changes to an organism's biology. That is, homeostasis refers to how the body keeps certain biological variables, such as blood sugar or hormone levels, in a particular range. Allostasis refers to how this particular range might shift. For example, high body-fat composition may can cause type-II diabetes, in which blood sugar levels remain persistently elevated above the normal range.

Anhedonia Reduced anticipation of reward ("anticipatory anhedonia") or experience of reward ("consumatory anhedonia"). Anticipatory anhedonia is a core symptom of depression; however, consumatory anhedonia is not (by contrast, consumatory anhedonia is common in schizophrenia). Thus, a person with depression struggles to imagine enjoyment from ordinarily rewarding activity, but tends to experience as much enjoyment as a non-depressed person from the same activities. In individuals with depression, variation in anticipated enjoyment predicts effort on tasks to attain enjoyable things, but variation in experienced enjoyment does not.

Annotation data Meta data for methylation data. Specifically, data from the Infinium 450k describes the extent of methylation on regions of DNA measured by each *probe*. Annotation data describe the region of DNA measured by the probe.

Anticipatory Anhedonia Reduced ability to imagine future enjoyment from ordinarily rewarding activities.

CpG site A cytosine-guanine sequence of nucleotides in DNA. Typically, methylation of cytosine can occur only when the cytosine is followed by a guanine. Hence, a CpG site is a location on DNA that is amenable to methylation. It can be thought of as a switch that may be toggled 'on' or 'off'.

Daughter cell When a cell divides into two copies of the original, the copy cells are called 'daughter cells'.

Depression-like behavior Murine models of depression function as analogies of human depression. A good murine-model of depression generates generates symtoms in rodents that resemble those in humans–that is, depression-like behavior. Subjective symptoms of depression, such as guilt, sadness, and hopelessness cannot be measured in rodents. However, more objective symptoms can be measured. For example, depressed humans exhibit anticipatory anhedona – an inability to intuit how one will enjoy ordinarily rewarding activities. Humans exhibiting this symptom do not pursue to the same extent things that they would ordinarily would enjoy. So, an analogous symptom in murine models would be the "sucrose preference test". Specifically, rodents ordinarily prefer water laced with sugar. Rodents exhibiting depression-like behavior exhibit an attenuated preferenced for sugar-water. Similarly, depressed humans give up on difficult tasks earlier than non-depressed humans. So, an analogous symptom in murine models would be "learned helplessness". Specifically, rodents exhibiting depression-like behavior will more quickly give up trying to escape from a pool of water or from being suspended upside-down by the tail.

Germline Cells that determine the genetics of the next generation. That is, genetic changes to *somatic* cells cannot be inherity, but genetic changes to cells in the germline can be inherited.

Monocytes White blood cells. Three types of cells normally exist in blood: erythrocytes (red blood cells), thrombocytes (platelets), and monocytes (white blood cells). Neither erythrocytes nor thrombocytes contain nuclear DNA, but monocytes do; hence, DNA extracted from blood samples comes from monocytes. Many different types of monocytes exist (including natural-killer cells, helper t-cells, and plasma b-cells), and their relative frequencies may differ over time and across individuals. Because cellular differentiation (changing of generalist cells into tissue-specific cells) entails epigenetic differentiation (different cell types express different genes, the same genes to different extents, and the same genes in different ways, which epigenetic modifications help to accomplish), methylation may vary across different types of monocytes. Consequently, in principle, methylation may vary across samples due to differences in monocyte composition rather than due to differences in methylation with identical cell types.

Methylation Any epigenetic modification of DNA. Methylation consists of the addition of a methyl group (carbon and three hydrogens) to the back of a cytosine nucleotide, typically at a CpG site. DNA methylation may serve several distinct roles. Among these, methylation at a CpG site in the promoter of a gene tends to prevent transcription machinery from reading that gene. As a consequence, less of the gene encodes is made, which shifts the metabolism of the cell.

Cystosine methylation is extremely chemically stable, so that cells remove methylation not by removing the methylation group but by hydroxylating the methyl group (adding another chemical group), which DNA repair machinery reads as DNA damage. The DNA repair machinery excises the hydroxylated cytosine and replaces it with an unmethylated cytosine. This extreme chemical stability is part of the reason why epigenetic modifications may persist long after what causes them.

When cells divide, a copy of DNA is made so that each cell has a copy of the DNA. The cellular machinery that makes copies of the DNA reads both methylated cytosine and regular cytosine as regular cytosine. Thus, for both daughter cell to have a methylated cytosine, different cellular machinery must copy the methylation to the DNA copy. This appears to occur in many cases, in which case the methylation is 'heritable' to daughter cells (in this context, 'heritability' refers to epigenetic modifications being transferred to daughter cells but not necessary to offspring).

Nucleotide DNA is sequence of nucleotides strung together. Only four nucleotides may compose this DNA string: guanine, adenine, thiamine, and cytosine. These nucleotides constitute the alphabet of the genome. For protein encoding regions of DNA, a three nucleotide sequence composes a 'word' specifying an amino acid. Thus, a protein encoding gene can be thought of as a sentence describing how to construct a protein machine.

Psychosocial stress Stress related to social status and social relationships. For example, conflict with others or defeat in such a conflict. Social-defeat stress is one experimental method of apply psychosocial stress in murine models.

RNA RNA is analogous to DNA: both consist of sequences of four possible nucleotides. DNA is copied from DNA and provides the blueprints for an organism; in contrast, RNA is copied from DNA, functioning as intermediate instructions. RNA serves many specific functions. Two derserve mention. Messenger-RNA (mRNA) is a copy of the protein-encoding parts of a gene; cellular machinery reads the instructions provided by mRNA to construct proteins. In contrast, microRNA do not code for protein but can interact with mRNA to modify its translation.

Social-defeat stress A type of psychosocial stress in which an animal is subordinated by a conspecific. For example, one experimental model of social defeat stress in rats involves introducing a smaller male rat into the cage of a larger, dominant rat. The larger rat defends his territory by attacking the smaller rat; the smaller rat attempts to escape but cannot because the exits to the cage are sealed off. The larger rat attacks the smaller rat and the smaller rat attempts to ward off future attacks through submissive behaviors. This stressor initially causes anxiety-like behavior in the smaller rat. Repeating this stressor several times robustly causes depression-like behavior in the smaller rat.

Transcription Making an mRNA copy of a gene. The cell constructs proteins from these mRNA copies. Epigenetic changes that reduce the rate of transcription reduce the availability of mRNA copies and so reduce the quantity of protein produced.

APPENDIX B

B - AN EPIGENETICS PRIMER, CONTINUED

Although the terminology is evolving, “epigenetics” generally refers to alterations to the DNA and associated structures that persistently change how the DNA is expressed without altering the heritable DNA code itself (see Berger et al., 2009, for a discussion of terminology). These alterations can persist within generations long after whatever caused them. And some of these alterations may propagate across generations, possibly through both direct and indirect mechanisms. Many of these alterations can be measured, which allows the researcher to proxy for the biological changes caused by past environment.

Such a high-level description of epigenetics will leave some readers uncomfortable, so I review terminology and describe some epigenetic mechanisms, beginning with an overview of DNA transcription.

DNA consists of sequences of four possible *nucleotides* (smallish molecules which form the alphabet of the genome). A gene is a stretch of DNA preceded by a promoter and start *codon* and followed by a stop codon. The promoter is a sort of volume-control knob that facilitates the reading of DNA when appropriately bound by other cell machinery. A codon is a sort of word composed of three nucleotides. Some of these words specify actions (start and stop codons), other words specify ingredients in a protein. RNA polymerase is a protein that attaches to the DNA near the promoter and makes RNA copies of the gene. The cell edits the RNA, producing mRNA that codes for a particular protein. When more mRNA is produced, more of the associated protein is produced, which can shift the metabolic equilibrium within the cell and within the body generally. *Methylation* – perhaps the most studied epigenetic mechanism – refers to the addition of a methyl group to the back of a cytosine nucleotide at a *CpG site*, a two-nucleotide stretch of DNA which is amenable to methylation. While technically this changes the nucleotide from cytosine to methyl-cytosine, the transcription machinery reads methyl-cytosine as cytosine – so the protein encoding sequence remains intact. However, methylation in the promoter often impedes transcription machinery from attaching to the promoter to start transcription. In turn, this impeded transcription reduces mRNA concentrations, and ultimately reduces concentration of the associated protein. Changes in the concentration of particular proteins can shift homeostatic equilibria, causing far-reaching effects.

The effects of methylation on other stretches of DNA can cause other effects, but these are less well understood. Even so, methylation is perhaps the best understood and most easily measured epigenetic mechanism. Because my data includes methylation measures but no other epigenetic measures, I do not discuss these other epigenetic mechanisms any further. [1]

Epigenetics serve three broad functions. First, epigenetic mechanisms contribute to cellular differentiation. Skins cells differ from liver cells in large part because of which genes are transcribed and which genes are not.

Second, epigenetics mediate many durable changes of an organism's biology in response to its environment. As a particularly relevant example, chronic and severe stressors increase methylation of promoters to mineralocorticoid receptors in the hypothalamus (among other things), which decreases the expression of minteralocorticoid receptors in the

---

1. By way of analogy, DNA is a book, genes are sentences, start and stop codons are punctuation marks delimiting sentences, exons are clauses within a sentence, and introns are parentheticals. RNA polymerase is a copy machine; it finds sentences by searching for promoters. Whenever it finds a sentence, it makes a copy. Post-transcriptional editing removes parentheticals and pastes clauses together in different combinations. When these edited copies drift towards a ribosome, the ribosome makes a protein in accordance with the instructions in the sentence. Methylation of cytosines in the promoter makes it more difficult for the copier to find the promoter. Consequently, methylation in the promoter reduces the number of copies that can float to the ribosomes and so reduces the quantity of that protein created by the ribosome. Likewise, histone acetylation opens the book at the associated page, allowing the copier to make copies of sentences on those pages.

hypothalamus (Witzmann et al., 2012)[2]. Because these receptors mediate negative feedback on stress hormone levels, the increased methylation increases average stress hormone levels, even long after the stressors have vanished. This long-run shift exemplifies *allostasis* – essentially, long-lasting changes in homeostatic equilibria. And the *allostatic load* refers to the cumulative burden of environmental shocks leading to allostasis. So epigenetics may be thought of as one allostatic mechanism and psychosocial stress as one allostatic load.

This biological terminology resembles market dynamics with which economists are familiar. By way of analogy, an upward shift in the oil supply curve would be akin to an allostatic load, and the long-run downward shift in the demand curve following the adoption of more fuel efficient technology would be akin to allostasis. So too, temporary variations in the environment can yield persistent shifts in biological equilibria.

Third, epigenetics may transmit environmental adaptations or maladaptations across generations. Evidence in animal models strongly supports the existence of such transmission. Evidence also supports the phenotypic relevance of such transmission, but this evidence is more limited. Transmission may be mediated through the gametes, through epigenetic information in semen, through in-utero environment, or through postnatal environment, but limited evidence indicates the relative importance of these transmission channels. As is generally true, animal models provide stronger causal evidence than human studies.

This epigenetic transmission may be relevant for a wide variety of behaviors and health outcomes in addition to those associated with psychosocial stress. For example, maternal prenatal exposure to cocaine causes epigenetic changes which mediate cognitive deficits in offspring; and maternal prenatal exposure to a high-fat diet causes epigenetic changes

---

[2] As other examples, epigenetic alterations contribute to long-term memory formation (Miller and Sweatt (2007)), and epigenetic alterations may mediate many of the effects of obesity on causing insulin resistance (dos Santos et al., 2015). Broadly speaking, epigenetic changes within tissue types function as a sort of metabolic memory.

which mediate increased obesity in offspring (Zhao et al., 2015; Marco et al., 2014)[3]. Mounting evidence indicates that environmental shocks even prior to conception can determine epigenetic endowments. For example, Dias and Ressler (2014) exposed female mice to odor fear conditioning prior to conception, pairing an electric foot shock to a specific smell. The first and second-generation pups, even those raised apart, were averse to this particular smell, and not others, as though they too had been exposed to the fear conditioning. In offspring, the promoter to the associated smell receptor was hypomethylated, implying increased expression of this receptor, and connectivity was enhanced between neurons processing this smell and and the fear center of the brain (see Dias et al. (2015) for evidence on underlying epigenetic mechanisms). This result provides insight into the stunning specificity and durability of epigenetic endowments formed prior to conception.

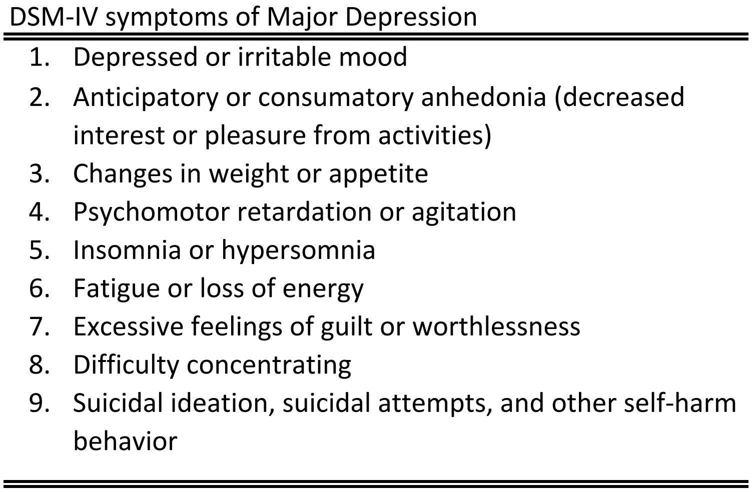
Likewise, evidence supports the transfer epigenetic endowments through the male germline or semen. For example, experimental variation in paternal mouse exercise prior to conception affects the risk of obesity in offspring (Murashov et al., 2016); these changes in offspring methylation may be mediated through changes in small RNA concentrations in semen (Cropley et al., 2016). As a more relevant example, changes in stress hormones of paternal mice change the small RNA profile in semen, which mediates changes in offspring depression-like behavior (Short et al., 2016). Mechanisms similar to those found in mice appear to exist in people as well. For example, human semen also contains a large variety of small RNAs, including microRNA (Krawetz et al., 2011), and the profile of small RNAs associate with facets of the father's environment (e.g. Marczylo et al. (2012)). See Ly et al. (2015) for a review.

---

[3] Prenatal exposure of mice to cocaine causes methylation changes that mediate impaired cognition in adult offspring (Zhao et al., 2015). Prenatal exposure of mice to a high fat diet changes methylation on a gene in pups that regulates food intake, ultimately causing obesity in the pups (which a normal diet after weening fails to reverse) (Marco et al., 2014).

APPENDIX C

C - ALSPAC SURVEY DATA DESCRIPTION



DSM-IV symptoms of Major Depression

1. Depressed or irritable mood
2. Anticipatory or consumatory anhedonia (decreased interest or pleasure from activities)
3. Changes in weight or appetite
4. Psychomotor retardation or agitation
5. Insomnia or hypersomnia
6. Fatigue or loss of energy
7. Excessive feelings of guilt or worthlessness
8. Difficulty concentrating
9. Suicidal ideation, suicidal attempts, and other self-harm behavior

Table C.1: List of DSM-IV depression symptoms

A. Variable Description

i. Depression inventories I interpret depression as a manifestation of low affective capital, so measures of depression may be useful for tracing the intergenerational process of affective capital.

Depression symptoms include fatigue, anhedonia (difficulty feeling or anticipating enjoyment), difficulty concentrating, sad or irritable mood, psychomotor retardation (slowed movement) or agitation, and others, which depression inventories assess (See Table C.1 for the list DSM-IV depression symptoms). Inventories, either completed by the subject or by an observer (such as a clinician or parent), score the subject based on the number and severity of symptoms. Often, guidelines for scoring these inventories indicate that major depression is likely when the score exceeds some threshold. While this threshold would seem to indicate a discrete nature to major depressive episodes, prodromal and residual symptoms surrounding depressive episodes are virtually identical but smaller in magnitude to those in a depressive episode (Iacoviello et al., 2010). In my interpretation, depression inventories measure the left tail of a continuous trait; thresholds merely indicate when clinical treatment may be warranted. In this vein, I primarily use scores calculated from the inventories rather than using indicators when inventory scores exceed the designated threshold. I chiefly use factor scores from a single factor model, as the factor scores reduce measurement error relative to a simple index [1].

Psychologists rigorously test the *validity* of depression inventories to ensure that these inventories do indeed measure depression and do so with limited error; even so, considerable measurement error may exist in these inventories. While some disagreement exists on appropriate standards for validity, valid instruments must feature good psychometric properties: individual items should strongly correlate with each other (sometimes assessed with Cronbach's alpha), the depression inventory scores from the same inventory administered twice in a short time interval should strongly correlate (*test-retest validity*), and different depression inventories administered contemporaneously should strongly correlate. In practice, the test-retest correlations of most depression inventories range between 0.7 to 0.9. Correlations between different inventories typically run lower – sometimes as low as 0.6 when administered to a non-clinical population (Beck et al., 1988). This latter finding suggests that around half the variance in depression inventory scores administered to a non-clinical population is measurement error[2]. Measurement error may be greater if

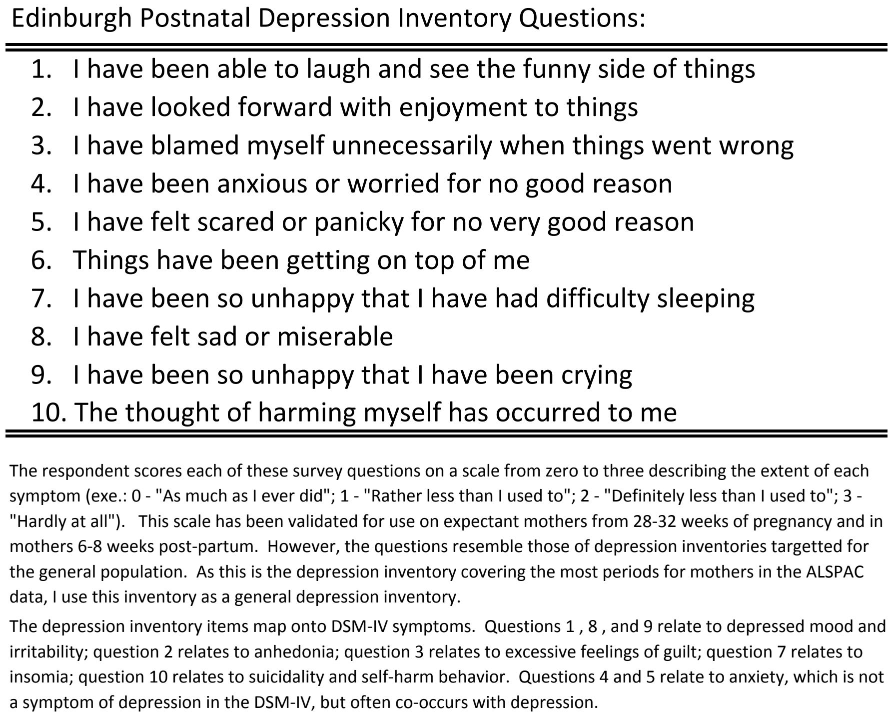
---

[1] A single factor model more heavily weights those items that more strongly correlate with the other items. Putatively, each item in the depression inventory is measuring the same latent trait – depression. Thus, those items that more strongly correlate with each other must also more strongly correlate with the latent trait being measured. Increasing the weight of these items increases the signal-to-noise ratio when compared with the uniform weights of a simple index.

[2] As an example, let θ be depression, $\dot{\theta}_1$ be one depression inventory score, $\dot{\theta}_2$ be another, ε_1 be measurment error for inventory 1 and ε_2 be measurement error for inventory 2. Normalize the random variables to be unit variance. Also assume that ε_1 and ε_2 are independent. Express the depression inventory in terms of the data generating process $\dot{\theta}_j = \rho_j \theta + \left(1 - \rho_j^2\right)^{1/2} \varepsilon_j$, where ρ is the correlation between depression inventory j $\dot{\theta}_j$ and actual depression θ. Then the correlation between depression inventory 1 and 2 equals $\rho_1 \rho_2$. Supposing $\rho_1 = \rho_2 = \rho x$, this implies that the variance in depression inventory 1 or 2 explained by actual depression equals the correlation between depression inventories. Of course, measurement error may be greater if ε_1 is positively correlated with ε_2.

some variable other than depression systematically generates correlation between survey items assessing depression generally. For example, alexithymia (difficulty describing or identify emotions) is elevated in depression (see Conrad et al., 2009, for example). Thus, alexithymia in depressed patients could bias scores downwards on items inquiring about negative emotions. Similarly, machismo might lead a subject to systematically answer questions as though the subject experiences no symptoms even when the subject does. This sort of measurement error would be correlated across different inventories, so high correlations between inventories cannot exclude the possibility of large measurement error. In addition, these sources of measurement error may compromise the application of depression inventories on males in particular. For example, Martin et al. (2013) explore evidence of sex-bias in depression inventories, comparing the estimated prevalence of depression in males using traditional depression inventories and using depression inventories specialized to assess depression in males. Martin et. al. find that traditional inventories indicate around 5 percent prevalence of depression in males, but the male-specialized depression inventories indicate 26 percent prevalence of depression in males. These male-specialized inventories are not available in the ALSPAC.

The ALSPAC data administers several different depression inventories to participant mothers, fathers, and children at multiple timepoints. The ALSPAC administers the Edinburgh Post-Natal Depression Survey (EPDS) with greatest frequency, so I primarily rely on this inventory for my analysis (see Table C.2 for a list of EPDS inventory questions). The participant scores inventory items by the extent to which the subject feels he experiences the listed symptoms. While the EPDS targets post-natal depression rather than depression generally, the questions are substantively identical to those of other more general depression inventories, such as the Hamilton depression inventory (for comparison, see Table C.3 for a list of Hamilton inventory questions). For this reason, I treat the EPDS as a general depression inventory. The ALSPAC also utilizes the Crown-Crisp experiential index



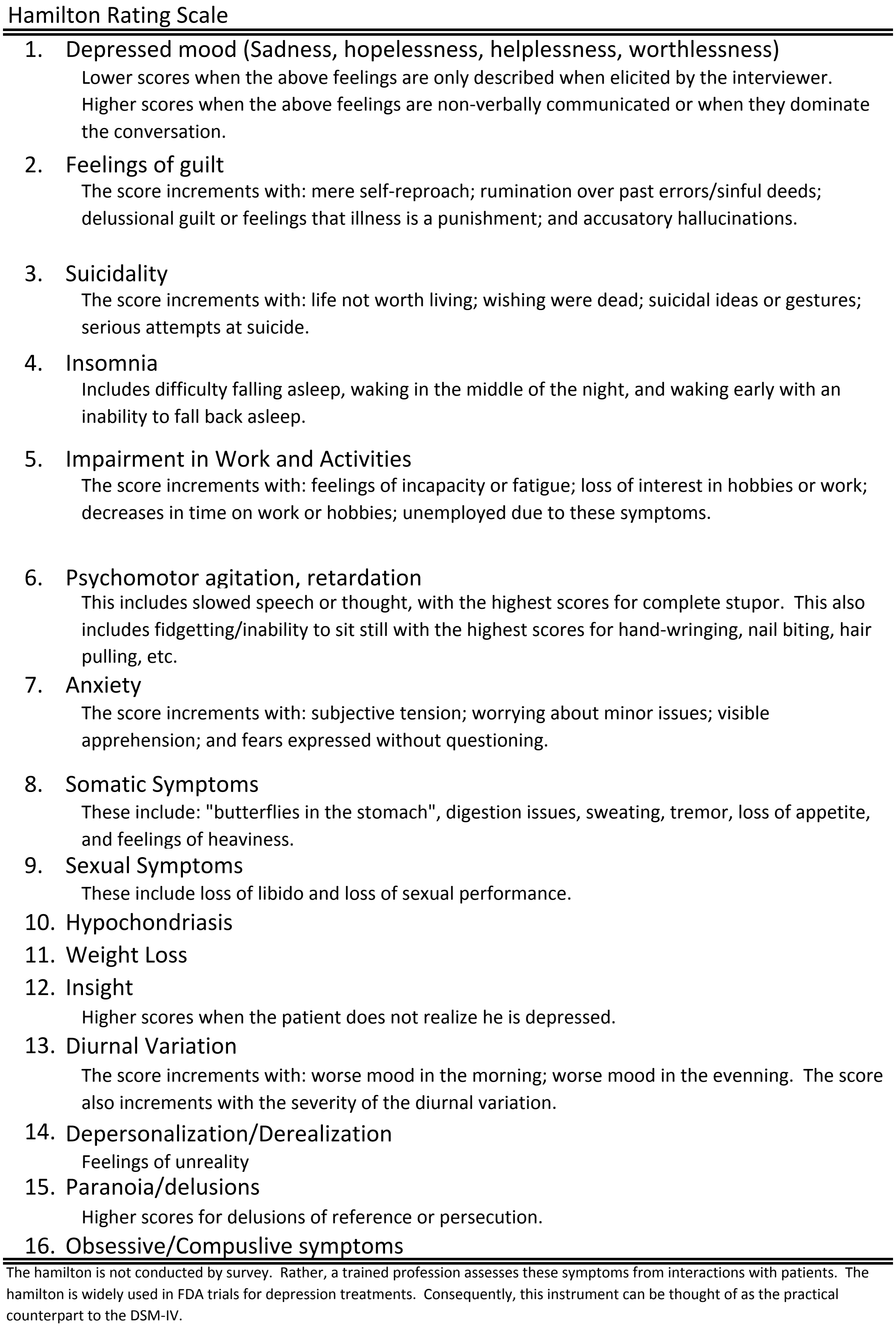
Edinburgh Postnatal Depression Inventory Questions:

1. I have been able to laugh and see the funny side of things
2. I have looked forward with enjoyment to things
3. I have blamed myself unnecessarily when things went wrong
4. I have been anxious or worried for no good reason
5. I have felt scared or panicky for no very good reason
6. Things have been getting on top of me
7. I have been so unhappy that I have had difficulty sleeping
8. I have felt sad or miserable
9. I have been so unhappy that I have been crying
10. The thought of harming myself has occurred to me

The respondent scores each of these survey questions on a scale from zero to three describing the extent of each symptom (exe.: 0 - "As much as I ever did"; 1 - "Rather less than I used to"; 2 - "Definitely less than I used to"; 3 - "Hardly at all"). This scale has been validated for use on expectant mothers from 28-32 weeks of pregnancy and in mothers 6-8 weeks post-partum. However, the questions resemble those of depression inventories targetted for the general population. As this is the depression inventory covering the most periods for mothers in the ALSPAC data, I use this inventory as a general depression inventory.

The depression inventory items map onto DSM-IV symptoms. Questions 1 , 8 , and 9 relate to depressed mood and irritability; question 2 relates to anhedonia; question 3 relates to excessive feelings of guilt; question 7 relates to insomia; question 10 relates to suicidality and self-harm behavior. Questions 4 and 5 relate to anxiety, which is not a symptom of depression in the DSM-IV, but often co-occurs with depression.

Table C.2: List of EPDS depression inventory items



Hamilton Rating Scale

1. Depressed mood (Sadness, hopelessness, helplessness, worthlessness)
   Lower scores when the above feelings are only described when elicited by the interviewer. Higher scores when the above feelings are non-verbally communicated or when they dominate the conversation.
2. Feelings of guilt
   The score increments with: mere self-reproach; rumination over past errors/sinful deeds; delussional guilt or feelings that illness is a punishment; and accusatory hallucinations.
3. Suicidality
   The score increments with: life not worth living; wishing were dead; suicidal ideas or gestures; serious attempts at suicide.
4. Insomnia
   Includes difficulty falling asleep, waking in the middle of the night, and waking early with an inability to fall back asleep.
5. Impairment in Work and Activities
   The score increments with: feelings of incapacity or fatigue; loss of interest in hobbies or work; decreases in time on work or hobbies; unemployed due to these symptoms.
6. Psychomotor agitation, retardation
   This includes slowed speech or thought, with the highest scores for complete stupor. This also includes fidgetting/inability to sit still with the highest scores for hand-wringing, nail biting, hair pulling, etc.
7. Anxiety
   The score increments with: subjective tension; worrying about minor issues; visible apprehension; and fears expressed without questioning.
8. Somatic Symptoms
   These include: "butterflies in the stomach", digestion issues, sweating, tremor, loss of appetite, and feelings of heaviness.
9. Sexual Symptoms
   These include loss of libido and loss of sexual performance.
10. Hypochondriasis
11. Weight Loss
12. Insight
    Higher scores when the patient does not realize he is depressed.
13. Diurnal Variation
    The score increments with: worse mood in the morning; worse mood in the evenning. The score also increments with the severity of the diurnal variation.
14. Depersonalization/Derealization
    Feelings of unreality
15. Paranoia/delusions
    Higher scores for delusions of reference or persecution.
16. Obsessive/Compuslive symptoms

The hamilton is not conducted by survey. Rather, a trained profession assesses these symptoms from interactions with patients. The hamilton is widely used in FDA trials for depression treatments. Consequently, this instrument can be thought of as the practical counterpart to the DSM-IV.

Table C.3: List of Hamilton depression inventory items

Short Mood and Feelings Questionaire items:

1. I felt miserable or unhappy
2. I didn't enjoy anything at all
3. I felt so tired I just sat around and did nothing
4. I was very restless
5. I felt I was no good anymore
6. I cried a lot
7. I found it hard to think properly or concentrate
8. I hated myself
9. I was a bad person
10. I felt lonely
11. I thought nobody loved me
12. I thought I could never be as good as the other kids
13. I did everything wrong

The respondent scores each of these survey questions on a scale from zero to two describing the extent of each symptom: 0 - "Not True"; 1 - "Sometimes"; 2 - "True".

Table C.4: List of SMFQ inventory items

to measure maternal and paternal depression, but I rely less on this inventory because it is applied less frequently.

ALSPAC assesses the child's depression with the Moods and Feelings inventory (MF) and the Short Mood and Feelings Questionnaire inventory (SMFQ). For earlier timepoints, the mother assesses the child using these instruments, and for later timepoints the child assesses himself using these instruments (see Table C.4 for a list of SMFQ inventory items).

ii. Educational variables While US data often codes educational attainment in years of education, the ALSPAC codes educational attainment in terms of certifications, as is typical of UK data. Parental education is indicated by one of five mutually exclusive categories (in order): certificate of secondary education, vocational degree, O-levels, A-levels, and post-secondary degree. Parents with a certificate of secondary education, vocational degree, or O-levels qualification all feature similar outcomes, whereas parents attaining A-levels or a post-secondary degrees face better outcomes. For this reason, I often summarize educational attainment with a dummy variable indicating A-level qualifications or a post-secondary degree.

For the children, the ALSPAC includes a greater variety of certifications. This partly reflects changes in the educational options available to younger UK cohorts and partly reflects a more thorough assessment of child educational outcomes. Nevertheless, as with older cohorts, A-levels certification is the single certification most predictive of better outcomes.

Because children are about 20 years old in the latest available survey, few children who will complete advanced degrees have done so. Pursuit of advanced degrees can be proxied by educational enrollment of students at timepoint 20.

iii. Income, proxies for income, employment, occupation The ALSPAC includes coarse measures of household income for the mother at timepoints 8 and 19 and personal income for the child at timepoint 20. These variables lump income into buckets; twenty percent of mothers report incomes in the highest bucket (greater than 48 thousand pounds per year).

In principle, consumption data in the ALSPAC could be used to infer permanent income by assuming constant income expansion paths; however, poor data quality in the consumption variables prevents this application. Consumption variables include rent/mortgage payments at timepoints 4, 7, 8, 11 from surveys of the father as well as food and other expenditures. These measures lump individuals into expenditure buckets. A large share of respondents reported expenditures placing them in the highest bucket. For example, about half of the participants fell into the highest consumption bucket for food expenditure. Consequently, little variance in consumption variables can be used to predict household income. Forecasts of household income using these consumption variables only weakly correlate with observed household income.

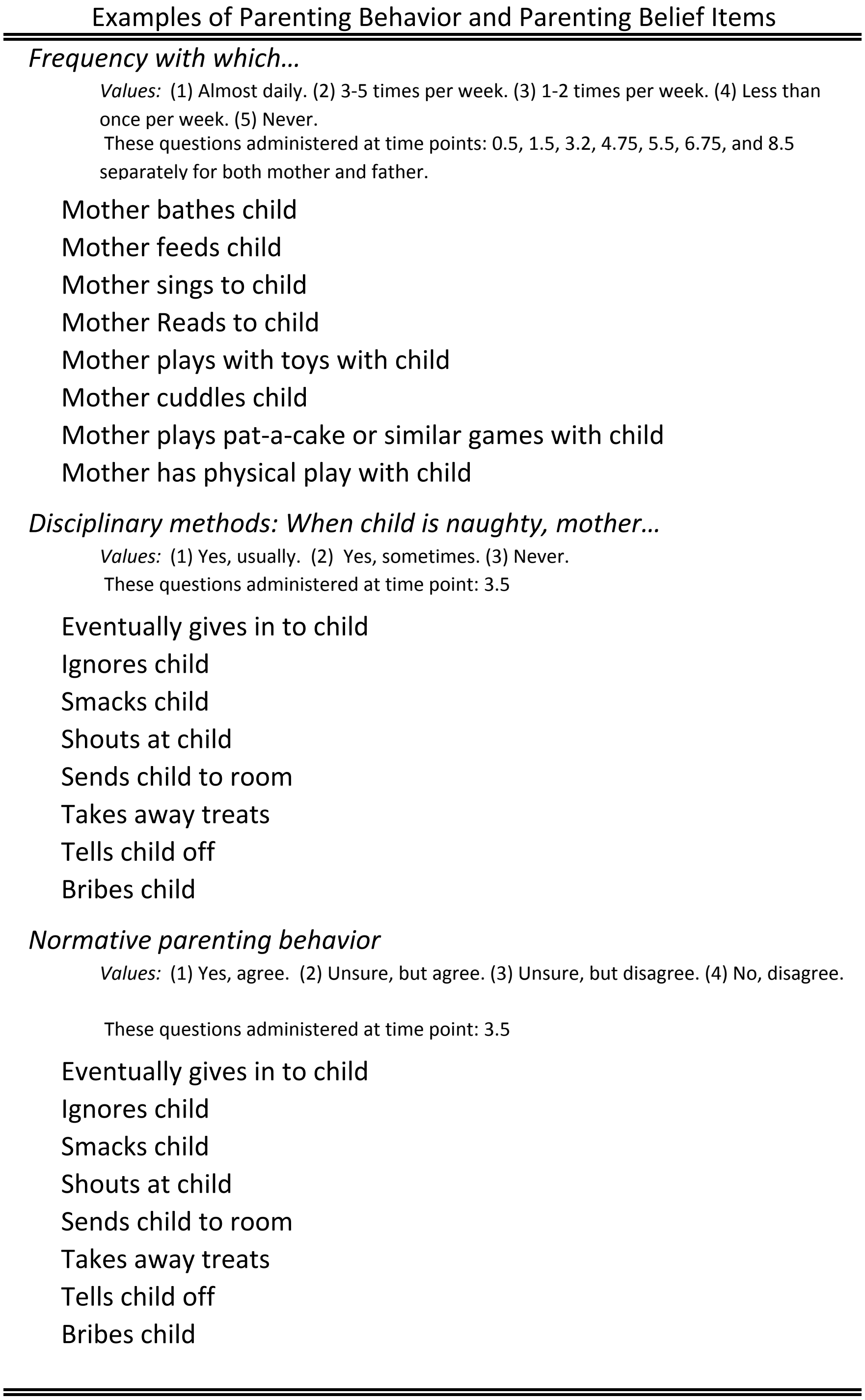
The ALSPAC includes measurements of maternal and paternal unemployment at time-points -0.25, 0, 0.25, 0.75, 1.75, 9, 11, and 19. The measurements include questions about recent job loss, e.g. "My partner lost his job since my child was 1 year old", as well as explicit questions about employment status and receipt of unemployment benefits.

The ALSPAC also includes a measure of "social class", or broad occupational categories, for the parents, and this variable is assessed near the mother's due date (time-point 0). Social class categories include "managers and senior officials", "professional occupations", "administrative and secretarial occupations", "associate professional and technical professions", "skilled trade occupations", "personal service occupations", "sales and customer service operations", "process, plant, and machine operatives", and "elementary occupations."

iv. Parenting behaviors The ALSPAC includes a rich set of instruments assessing parenting behaviors over multiple timepoints (see Table C.5 for examples of measured parenting behaviors). Assessed behaviors include time spent in various activities with children, such as physical play, cuddling, reading, and managing hygiene. These time-spent behaviors were measured at timepoints 0.5, 1.5, 3.2, 4.75, 5.5, 6.75, and 8.5.

Assessed parenting behaviors also include disciplinary methods, such as the use of "time-outs" and corporal punishment as well as the manner in which the mother responds to child tantrums. Different questions were administered in different periods, including timepoints 2, 3.5, 6.5, and 16.5.

v. Parenting beliefs The ALSPAC also includes questions regarding how parents think they *ought* to parent. These question include what parents believe to be appropriate punishments (e.g. "Sometimes a good smack is the best way to discipline a child"), appropriate nurturing interactions (e.g. "Cuddling a baby is very important", "Talking to a baby is im-



Examples of Parenting Behavior and Parenting Belief Items

Frequency with which…

Values: (1) Almost daily. (2) 3-5 times per week. (3) 1-2 times per week. (4) Less than once per week. (5) Never.

These questions administered at time points: 0.5, 1.5, 3.2, 4.75, 5.5, 6.75, and 8.5 separately for both mother and father.

Mother bathes child
Mother feeds child
Mother sings to child
Mother Reads to child
Mother plays with toys with child
Mother cuddles child
Mother plays pat-a-cake or similar games with child
Mother has physical play with child

Disciplinary methods: When child is naughty, mother…

Values: (1) Yes, usually. (2) Yes, sometimes. (3) Never.

These questions administered at time point: 3.5

Eventually gives in to child
Ignores child
Smacks child
Shouts at child
Sends child to room
Takes away treats
Tells child off
Bribes child

Normative parenting behavior

Values: (1) Yes, agree. (2) Unsure, but agree. (3) Unsure, but disagree. (4) No, disagree.

These questions administered at time point: 3.5

Eventually gives in to child
Ignores child
Smacks child
Shouts at child
Sends child to room
Takes away treats
Tells child off
Bribes child

Table C.5: Subsample of parenting behaviors measured.

portant”), and appropriate effort (e.g. “A baby should fit into parent’s routine”). Different sets of questions are administered at timepoints 0, 1, 1.75, and 2.75.

vi. Stressors The ALSPAC includes a large set of variables measuring stressors – some plausibly exogenous, others clearly endogenous.

Plausibly exogenous stressors include bereavement and muggings/burglaries (conditional on expectations of muggings and burglaries). Bereavement variables include: *loss of a friend or relative*, measured at timepoints -0.3, 0.6, 3, 9, and 19; *loss of a child*, measured at timepoints -0.3, 0.6, 3, 9, and 19; and *loss of partner*, measured at timepoints -0.3, 0.6, 3, and 19.

Variables describing muggings and house/car burglary events are recorded at timepoints -0.3, 0.2, 0.7, 1.75, 2.75, 9, and 11. Variables describing expectations of burglary and mugging risk include the respondents assessment of the severity of muggings and burglaries in the neighborhood, with scores ranging over: 1 - “A serious problem”; 2 - “A minor problem”; and 3 - “Not a problem”. These expectation variables were recorded at timepoints -0.5, 5, 7, and 19.

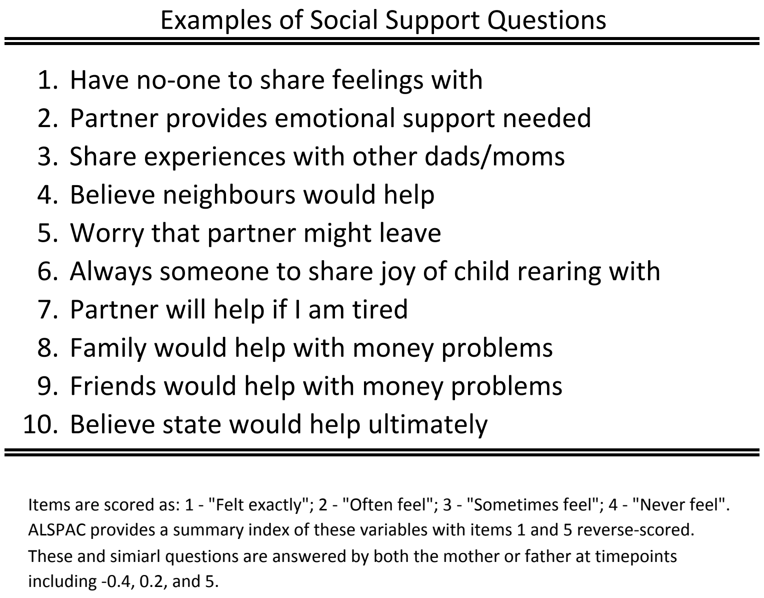
Variables related to endogenous stressors include a social stress index, measures of marital quality, measures of neighborhood characteristics, and measures of social support. The social stress index (“weighted life events score”) includes items such as “convicted of a crime since previous survey”, “became homeless since pervious survey”, “argued with family or friends since previous survey”, “separated from partner since previous survey”, and “had problems at work since previous survey.” The index gives greater weight to events when the respondent reports that the event “affected me a great deal”. A full list of items composing the index is available in Table C.6 . This social stress index was administered at timepoints including -0.25, 0.25, 0.75, 1.75, 9, 11, and 19.

Items Composing the Life Events Index
Partner died
One child died
Friend or relative died
One child ill
Partner was ill
Friend or relative ill
Admitted to hospital
In trouble with law
Divorced
Partner not wanted your child
Have been ill
Partner lost job
Partner had problems at work
Partner had problems at work
Lost job
Partner went away
Partner in trouble with law
You and partner separated
Income reduced
Argued with Partner
Argued with family/friends
Moved house
Partner physically hurt you
Became homeless
Major financial problems
Have married
Partner physical-cruel to child
Convicted
Started a new job
Partner emotionally cruel to you
Partner emotionally cruel to child

The weighted life-events score calculates the sum of individual item scores, where events that did not occur receive a score of zero and events that did occur receive a score in proportion to how much the respondent claimed the event affected the respondent.

ALSPAC administered this life-events survey at multiple timepoints to both the mother and father, including timepoints -0.3, .0.3, 0.8, 1.8, 9, 11, and 19.

Table C.6: List of “life events” stressors



Examples of Social Support Questions

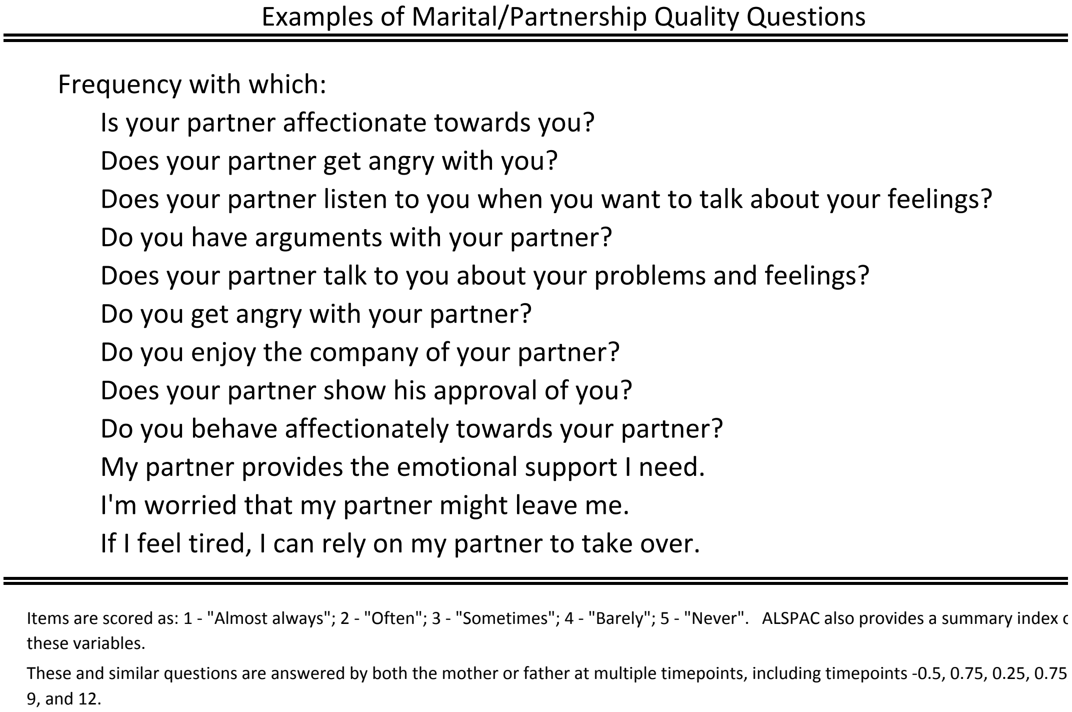
1. Have no-one to share feelings with
2. Partner provides emotional support needed
3. Share experiences with other dads/moms
4. Believe neighbours would help
5. Worry that partner might leave
6. Always someone to share joy of child rearing with
7. Partner will help if I am tired
8. Family would help with money problems
9. Friends would help with money problems
10. Believe state would help ultimately

Items are scored as: 1 - "Felt exactly"; 2 - "Often feel"; 3 - "Sometimes feel"; 4 - "Never feel". ALSPAC provides a summary index of these variables with items 1 and 5 reverse-scored. These and simiarl questions are answered by both the mother or father at timepoints including -0.4, 0.2, and 5.

Table C.7: List of “life events” stressors

Measures of social support assess the extent to which the respondent believes family, friends, and neighbors would provide assistance when needed. Question prompts include “I have no one to share my feelings with” and “friends would help if I had money problems” (see Table C.7).

Measures of marital/partnership quality assess perceived affection, anger, and support between partners. Examples include the frequency with which “my partner expresses affection to me”, “my partner discusses his feelings”, and “I express anger towards my partner”. Such questions were assessed at timepoints -0.5, 0.75, 0.25, 0.75, 4, 8, 9, and 12 (see Table C.8).



Examples of Marital/Partnership Quality Questions

Frequency with which:
- Is your partner affectionate towards you?
- Does your partner get angry with you?
- Does your partner listen to you when you want to talk about your feelings?
- Do you have arguments with your partner?
- Does your partner talk to you about your problems and feelings?
- Do you get angry with your partner?
- Do you enjoy the company of your partner?
- Does your partner show his approval of you?
- Do you behave affectionately towards your partner?
- My partner provides the emotional support I need.
- I'm worried that my partner might leave me.
- If I feel tired, I can rely on my partner to take over.

Items are scored as: 1 - "Almost always"; 2 - "Often"; 3 - "Sometimes"; 4 - "Barely"; 5 - "Never". ALSPAC also provides a summary index c these variables.

These and similar questions are answered by both the mother or father at multiple timepoints, including timepoints -0.5, 0.75, 0.25, 0.75 9, and 12.

Table C.8: List of marital quality questions

B. Survey and item non-response

While the ALSPAC features minimal survey non-response rates, (see C.9), the ALSPAC also features high item non-response rates; these non-responders compose a selected subsample (see Tables C.10 and C.11).

Of particular note, depression is highly predictive of non-response. For example, mothers who complete only one or two of eight EPDS depression inventories score about seven-tenths of a standard deviation higher on their completed depression inventories than mothers who complete all eight of the of the EPDS depression inventories. And 8.6 percent of mothers did not complete any of the eight depression inventories. In general, non-response corresponds with higher depression, worse educational attainment, and lower household income.

This selected non-response presents three issues. First, those purported to be most adversely affected by low affective capital are also those least measured. Thus, item non-response eliminates much of the most useful variation, reducing power and potentially introducing bias. These missing values can be imputed only by making strong ignorability assumptions.

Second, depression inventory non-response will tend to introduce non-classical measurement error into depression inventory scores. That is, suppose that depression inventory scores in a non-selected sample yield classical measurement error about true depression levels. And suppose the sample selects against those with high depression. Then those scoring high in depression in the selected sample will tend to feature lower true depression levels than their depression inventory scores indicate. This non-classical measurement error will tend to bias downwards any estimated associations between depression and worse outcomes.

Survey	Time	Took Survey	Means in subsamples		
			surv b Depres. Factor	surv a Educ >= A_Levels	surv n HH Income
a		13719	-0.010	0.358	20049
			(0.0091)	(0.0043)	(83.63)
b		13364	0.000	0.357	20028
			(0.0091)	(0.0043)	(83.76)
c		12578	-0.023	0.353	20075
			(0.0093)	(0.0043)	(83.78)
d		12609	-0.024	0.361	20071
			(0.0093)	(0.0044)	(84.01)
e		11848	-0.040	0.369	20101
			(0.0095)	(0.0045)	(84.26)
f		11345	-0.051	0.373	20117
			(0.0096)	(0.0046)	(84.38)
g		10430	-0.063	0.385	20137
			(0.0099)	(0.0048)	(85.21)
h		9754	-0.079	0.393	20132
			(0.0102)	(0.0050)	(86.02)
j		9618	-0.075	0.391	20091
			(0.0103)	(0.0051)	(85.93)
k		14741	0.000	0.353	19997
			(0.0091)	(0.0043)	(83.40)
l		14736	0.000	0.353	19997
			(0.0091)	(0.0043)	(83.40)
m		14744	0.000	0.353	19997
			(0.0091)	(0.0043)	(83.37)
n		15099	0.000	0.353	19999
			(0.0091)	(0.0043)	(82.06)
p		15180	0.000	0.353	19998
			(0.0091)	(0.0043)	(82.07)
q		15208	0.000	0.353	19998
			(0.0091)	(0.0043)	(82.07)
r		15202	0.000	0.353	20000
			(0.0091)	(0.0043)	(82.08)
s		15188	0.000	0.353	19999
			(0.0091)	(0.0043)	(82.10)
t		14732	0.000	0.353	19997
			(0.0091)	(0.0043)	(83.40)

Standard errors are in parenthesis. The first two columns of data indicate the number of observations who (1) took the survey (2) repsonded to at least 50% of the questions in the survey. The third column of data indicates the average share of questions that study participants did not answer. The final three column are means of variables among those who took the survey indicated by the row. The final row in

Table C.9: Survey non-response

Evidence of non-random item non-response

	dependent variables from bivariate regressions						
RHS variables	mean maternal depression	mean paternal depression	mean child depression	mother attained A_levels	father attained A_levels	child attained A_levels	household income, t=19
maternal dep. inv. missing	**0.077**	**0.045**	**0.036**	**-0.056**	**-0.045**	**-0.031**	**-951.5**
count	14128	11267	9672	12509	10793	4022	3499
range: 0-8	(0.0044)	(0.0053)	(0.0058)	(0.0022)	(0.0027)	(0.0041)	(337.7)
	3.3E-68	8.1E-18	7.2E-10	4.5E-143	3.1E-60	6.0E-14	4.9E-03
paternal dep. inv. missing	**0.086**	**0.090**	**0.033**	**-0.052**	**-0.057**	**-0.044**	**-2000.7**
count	14128	11267	9672	12509	10793	4022	3499
range: 0-5	(0.0050)	(0.0075)	(0.0060)	(0.0025)	(0.0029)	(0.0044)	(169.5)
	1.5E-66	2.5E-33	2.9E-08	4.4E-92	6.2E-84	5.1E-23	1.5E-31
child dep. inv. missing	**0.068**	**0.032**	**0.041**	**-0.066**	**-0.048**	**-0.081**	**-1403.4**
count	14128	11267	9672	12509	10793	4022	3499
range: 0-6	(0.0045)	(0.0051)	(0.0072)	(0.0023)	(0.0025)	(0.0050)	(209.6)
	5.1E-51	2.5E-10	2.0E-08	7.4E-184	1.5E-77	4.5E-56	2.5E-11
in methylation subsample	**-0.129**	-0.029	**-0.088**	**0.164**	**0.101**	**0.080**	**2052.9**
indicator variable	14128	11267	9672	12509	10793	4022	3499
	(0.0289)	(0.0329)	(0.0292)	(0.0166)	(0.0170)	(0.0198)	(730.2)
	8.4E-06	3.7E-01	2.5E-03	5.3E-23	3.5E-09	5.9E-05	5.0E-03
missing household income	**0.197**	0.049	**0.110**	**-0.224**	**-0.133**	**-0.096**	**.**
var.: t=19	14128	11267	9672	12509	10793	4022	.
indicator variable	(0.0176)	(0.0199)	(0.0198)	(0.0098)	(0.0104)	(0.0147)	.
	6.7E-29	1.3E-02	3.0E-08	1.5E-114	1.5E-37	7.7E-11	.
missing child educational	**0.221**	**0.111**	**0.103**	**-0.185**	**-0.151**	**.**	**-2481.9**
certification	14128	11267	9672	12509	10793	.	3499
indicator variable	(0.0173)	(0.0194)	(0.0198)	(0.0095)	(0.0102)	.	(559.0)
	3.0E-37	1.1E-08	2.4E-07	1.3E-82	1.1E-49	.	9.3E-06

Number of observations beneath coefficient estimates. Standard errors in parenthesis. P-values from two-sided t-test in scientific notation. Bolded coefficient estimates are significant at the 99% significance level.

The RHS variables describe non-response. The first three of these variables provide a count of missing depression inventories over the longitudinal study for the mother, father, and child respectively. The fourth RHS variable indicates whether the mother-child pair were included in the methylation subsample. The fifth RHS variable indicates if the child's educational certifications are missing.

The dependent variables cover key variables we worry might be related to nonresponse. The first three variables are standardarized averages of depression inventories across the longitudinal design for the surveys in which individuals responded to the depression inventories. Thus, the coefficient of mean maternal depression on the maternal depression inventory missing count describes the relationship between maternal depression inventory scores *when the mother did take the inventory* and the number of depression inventories the mother failed to take. The next three variables describe maternal, paternal, and child educaitonal attainment, using a-levels completion as a convenient summary variable.

Table C.10: Selective item non-response

RHS variables	LHS variable: Number of depression inventories that the respondent did not complete											
	Mother				Father				Child			
Mean Dep. Inv.	**0.370**			**0.093**	**0.160**			**0.190**	**0.086**			0.010
	(0.0182)			(0.0150)	(0.0125)			(0.0238)	(0.0147)			(0.0287)
	1.1E-90			6.9E-10	1.9E-37			1.8E-15	6.7E-09			7.2E-01
A-Levels		**-0.799**		-0.047		**-0.587**		**-0.298**		**-0.780**		**-0.531**
		(0.0330)		(0.0263)		(0.0304)		(0.0449)		(0.0475)		(0.0509)
		2.0E-126		7.2E-02		1.7E-81		4.0E-11		9.9E-59		8.4E-25
ln HH income t=19			**-0.086**	-0.011			**-0.505**	**-0.186**			**-0.254**	-0.056
			(0.0238)	(0.0228)			(0.0439)	(0.0390)			(0.0372)	(0.0398)
			2.8E-04	6.3E-01			5.0E-30	1.8E-06			9.4E-12	1.6E-01
Obs.	14128	12509	3499	3450	11267	10793	3499	3013	9672	4022	3499	2021

Standard errors in parenthesis. P-values in scientific notation. Bolded coefficient estimates are significant at the 99% significance level.

This table shows the wultivariate regression results of number of depression inventories that the respondent did not complete on covariates. The mean depression inventory is the standardized average of depression inventory factors for those depression inventories that the respondent completed. A-levels is a dummy indicating that the respondent completed A-levels certification. ln HH income is the log of household income (for the mother) at time point 19. Mean depression scores and the A-levels dummy are for the mother, father, or child depending on whether the outcome variable is for the mother, father, or child, respectively.

Table C.11: Selective non-response to depression inventories

These first two issues may be especially problematic for uncovering associations between methylation and depression. Because ALSPAC analyzed methylation data from a random sample only among those mother-child pairs who participated in all five clinic visits, the methylation data is restricted to those with perfect clinic attendance rates. Uncovering associations between depression and methylation in this subsample requires using limited variation in biased depression inventory scores among those least depressed.

Third, the patterns of non-response are themselves consistent with the model of affective capital. The model of affective capital predicts that those with less affective capital experience greater disutility from engaging in effortful tasks, such as survey completion. Consequently, low affective capital itself causes item nonresponse. Consistent with this hypothesis, parental depression inventory scores are significantly associated with item nonresponse even when conditioning on parental education, but parental education is significantly associated with item non-response only when not conditioning on parental depression inventory scores.

Autocorrelations in mother depression inventory factor scores

time	survey	b	c	e	f	g	h	r	t
-0.3	b	1.000							
		(0.000)							
0	c	0.644	1.000						
		(0.007)	(0.000)						
0.2	e	0.510	0.560	1.000					
		(0.008)	(0.008)	(0.000)					
0.7	f	0.502	0.545	0.597	1.000				
		(0.009)	(0.008)	(0.008)	(0.000)				
1.8	g	0.471	0.508	0.543	0.601	1.000			
		(0.009)	(0.009)	(0.008)	(0.008)	(0.000)			
2.8	h	0.458	0.483	0.502	0.562	0.608	1.000		
		(0.009)	(0.009)	(0.009)	(0.009)	(0.008)	(0.000)		
11.2	r	0.376	0.383	0.394	0.407	0.457	0.474	1.000	
		(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.000)	
19	t	0.356	0.379	0.386	0.393	0.430	0.426	0.475	1.000
		(0.015)	(0.015)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.000)

Auto-correlations are in depression factors estimated from the Edinburgh Postnatal Depression Inventory. Standard errors are in parenthesis. "time" refers to the approximate number of years since the study child's birth. Negative numbers indicate the approximate number of years before the study child's due date.

Table C.12: Maternal Depression Autocorrelations

C. Inter- and intra-generational correlation structure of depression inventories.

i.a Intragenerational depression autocorrelation structure The biological literature suggests that long-lived stocks predispose individuals to depression. Such long-lived stocks are essential for affective capital to be playing a role in intergenerational mobility. To investigate the persistence of depression, I estimated the autocorrelations of maternal depression inventory factors based on the Edinburgh Post-Natal Depression Inventory administered across multiple surveys (see Table C.12 and Figure C.1). The reader should keep in mind that the test-retest correlations for most depression inventories are in the range of 0.75 to 0.9; and correlations between different depression inventories in non-clinical settings can

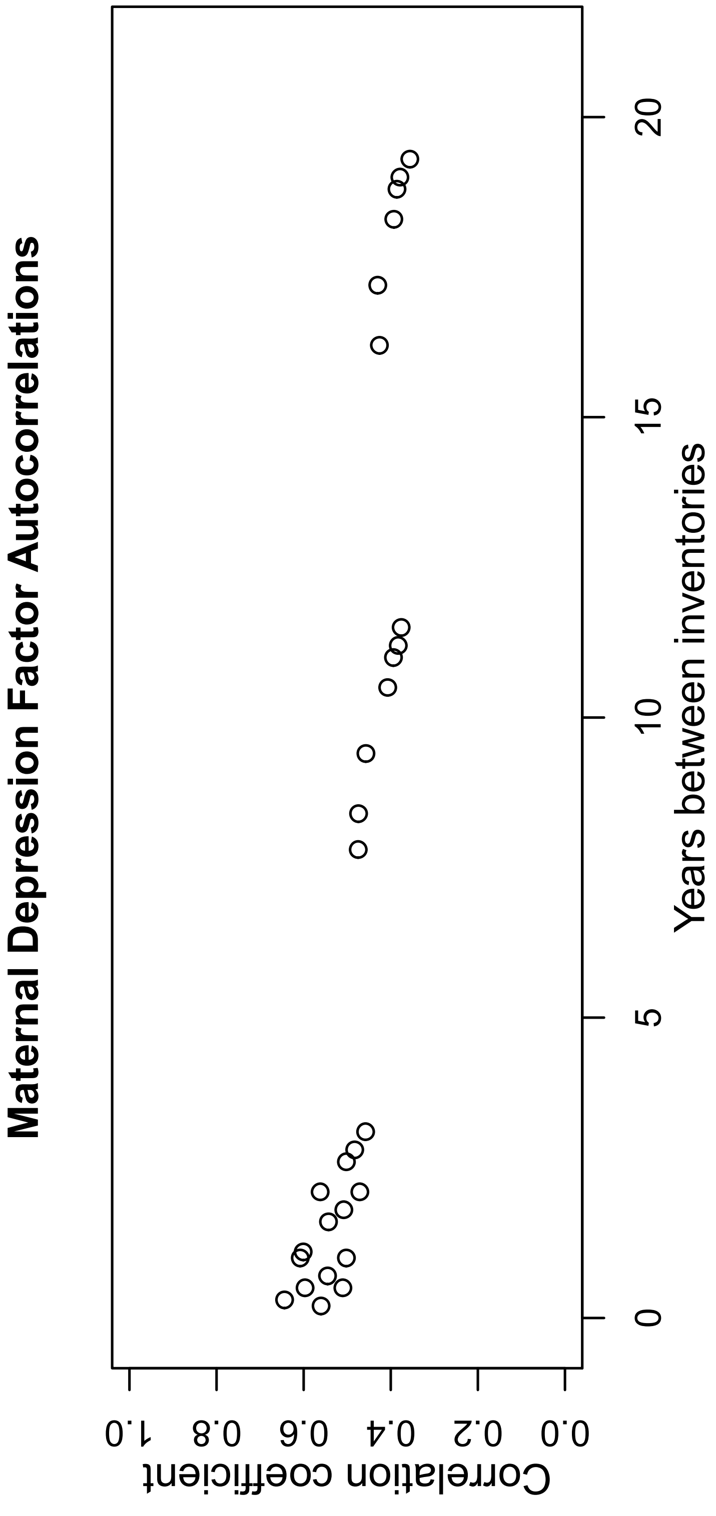


Figure C.1. Maternal Depression Autocorrelations plotted against time between measurements.

run as low as 0.6, indicating substantial measurement error in depression inventories (for example, see Beck et al., 1988). This measurement error in the tests will drive down both short-run and long-run autocorrelations.

The observed correlation structure is inconsistent with an AR-1 process. An homogeneous AR-1 process fit to the short-run autocorrelation structure would imply much smaller correlations at the longest intervals than those we observe in the data. Specifically, the correlation between the depression inventory factors over intervals less than one year range from about 0.5 to about 0.64. These correlations drop only modestly to around 0.47 when considering intervals of eight years and 0.42 when considering intervals of sixteen years. In my interpretation, correlations substantially below one over the shortest time-intervals indicate transient measurement error, while correlations asymptoting to a positive number over the longer intervals indicates variation in depression caused by person-specific constants.

Consonantly, these autocorrelations are consistent with large, transient measurement error and person-specific constants in an AR-1 process. Imposing this structure implies that transient measurement error explains nearly half of the variance in the depression inventory factors, while person-specific constants explain most of the remaining variance. In particular, suppose depression inventories follow a mean-reverting AR-1 process with person-specific constants and a period specific measurement error term. That is, depression $d_{i,t}$ for mother i at timepoint t is given by

$$d_{i,t+1} = \bar{d}_i + \left(d_{i,t} - \bar{d}_i\right)\delta + \varepsilon_{i,t+1}$$

where δ is an auto-regressive parameter with $\delta \in [0,1)$ and $\varepsilon_{i,t}$ are i.i.d. shocks, so $d_{i,t}$ reverts to $\bar{d}_i$. Depression inventory factor scores are given by

$$\tilde{d}_{i,t} = d_{i,t} + e_{i,t}$$

where $e_{i,t}$ is i.i.d. measurement error.

These equations imply the moment conditions

$$\sigma^2_{\bar{d}_i} = \sigma^2_{\varepsilon} / \left(1 - \delta^2\right)$$

$$\sigma^2_{d} = \sigma^2_{\varepsilon} / \left(1 - \delta^2\right) + \sigma^2_{\bar{d}_i}$$

$$\sigma^2_{\tilde{d}} = \sigma^2_{d} + \sigma^2_{e}$$

$$\mathbb{C}\left[d_{i,t}, d_{i,t+\tau}\right] = \delta^{\tau}\sigma^2_{d} + \left(1 - \delta^{\tau}\right)\sigma^2_{\bar{d}_i}$$

$$\lim_{\tau\to\infty} \mathbb{C}\left[d_{i,t}, d_{i,t+\tau}\right] = \sigma^2_{\bar{d}_i}$$

which feature a natural graphical interpretation. In the scatterplot of autocorrelations against elapsed time, the curve fitting the scatterplot will asymptote at $\sigma^2_{\bar{d}_i}$ and will intersect the y-axis at $1 - \sigma^2_{e}$. Thus, the length between 1 and the y-intercept indicates the share of variance due to measurement error, the height of the asymptote indicates the share of variance due to the permanent component $\bar{y}_i$, and the region inbetween indicates the share of variance due to the transitory component $\sigma^2_{\varepsilon} / \left(1 - \delta^2\right)$. Graphical inspection indicates that around 35% of the variation in depression inventories is measurement error while an additional 40% of the variation is due the permanent component, which leavese the transitory component explaining the smallest portion of the variance.

Altogether, these autocorrelations are consistent with a long-lived biological stock predisposing individuals to depression as is necessary for affective capital to be playing a meaningful role in intergenerational mobility.

i.b Intergenerational depression autocorrelation structure The intergenerational transfer of stocks predisposing individuals to depression – whether the mechanism is through genetics, direct epigenetic endowments, or parenting behaviors – is essential for affective capital to be playing a role in intergenerational mobility. To investigate the extent to which depression persists across generations, I estimate correlations between mother and child depression over the survey periods (See Figure C.2s).

Most saliently, the child depression factors after age 9 are all fairly strongly correlated with the maternal depression factors from all periods. The relative timing of the depression factors for the mother and child does not appear to relate to the magnitude of the correlation coefficients. This should not be surprising given the persistence of maternal depression and the noise generating time-series variation. In addition, this result is consistent with depression inventory scores in the child being generated by a highly persistent state variable which remains stable after early childhood, such as one due to genetic endowments, epigenetic endowments, or early childhood parenting inputs. Altogether, after accounting for measurement error, the intergenerational correlation in depression is probably around 0.5 or higher[3].

Two other features of these correlations deserve mention. First, the child's depression factors at ages 10 and 12 are more strongly correlated with maternal depression inventories than the child's depression factors at ages 14 and 18. This may partly be due to the survey instruments used -- the depression factors for children at ages 10 and 12 were constructed from survey responses provided by the mother about the child, while the later depression

---

[3] Suppose that materenal depression inventory scores are given by $\dot{\theta} = \rho\theta + \left(1-\rho^2\right)^{1/2}\varepsilon$ and child depression inventory scores are given by $\dot{\theta}' = \rho'\theta' + \left(1-\rho'^2\right)^{1/2}\varepsilon'$, where random variables $\dot{\theta}$, θ, and ε are standardized, ρ is the correlation between the depression inventory scores θ and actual depression θ, and ε is measurement error. Then the correlation between maternal and child depression is given by

$$Corr\left(\theta,\theta'\right) = Corr\left(\dot{\theta},\dot{\theta}'\right)/\rho\rho'$$

Suppose$\rho = 0.6$ as an upper bound as argued before, and suppose that $\rho = \rho'$, which is also a reasonable upper bound on the quality of depression measurements of children. Taking an observed intergenerational correlation $Corr\left(\dot{\theta},\dot{\theta}'\right)$ of 0.3, this implies that $Corr\left(\theta,\theta'\right) = 0.5$. This is a back-of-the-envelope lower bound. The intergenerational correlation in depression is probably higher.

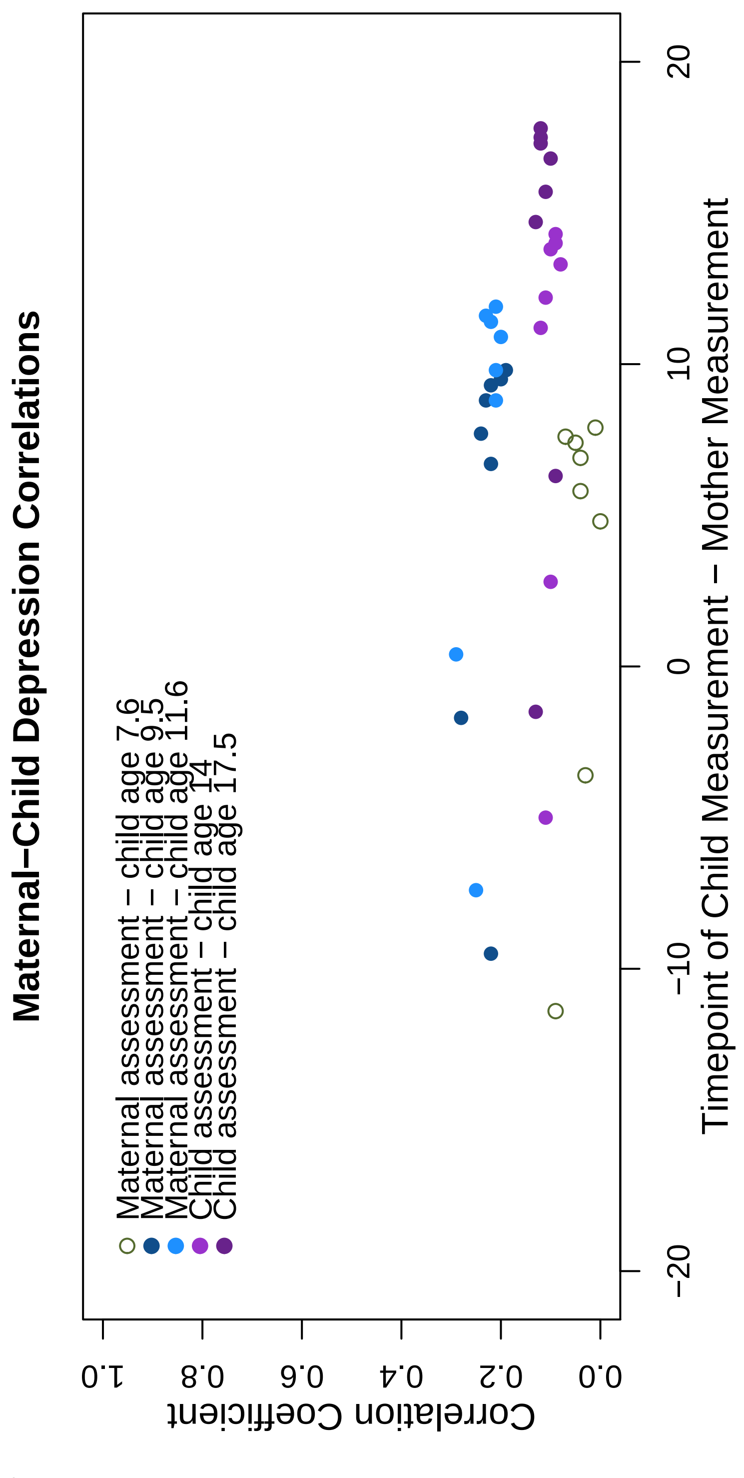


Figure C.2. Mother-child depression correlations over the survey duration factors were constructed from surveys conducted by the child. This difference in utilized survey instruments might generate higher correlations in one of two ways. One, the mother might project her own depression symptoms when assessing her child, which would bias upwards the estimated correlation. However, this source of bias should be seen with the timepoint 8 child depression inventories as well. So, the dearth of significantly positive correlations between maternal depression and child timepoint 8 depression would seem to rule out this source of bias. Two, the mother might do a better job assessing the child's mood than the child does at assessing his own, which would imply there is greater measurement error in the depression factors constructed from the adolescents' survey responses.

Second, the depression factor for the age 8 child is significantly associated with the mother's depression factor for only two of the eight periods in which the mother was assessed, and all estimated correlations are small. This may be because the instrument used to assess depression in this period lacks validity or it could be because children of this age do not reliably manifest depressive symptoms at this age even if these children have already accrued the relevant biological stock. Either way, the available depression inventories for children under the age of 9 are not of much use.

APPENDIX D

D - ANALYSIS OF SURVEY DATA, CONTINUED

Stressors and parental depression

a. Bereavement shocks : For bereavement shocks to be useful as instruments for identifying the effects of parental affective capital on child outcomes or for identifying the methylation changes caused by stress and associated with depression, these bereavement shocks need to predict large and persistent changes in parental depression. In addition, bereavement shocks should not predict depression in years prior to bereavement.

In the ALSPAC sample, bereavement shocks associate with only small, transient, and often insignificant increases in depression scores (see Tables D.1 , D.2 , and D.3). Loss of a friend or relative reliably predicts increased depression scores in that period; however, the magnitudes are small, and only rarely does bereavement significantly associate with increased depression in subsequent surveys. Likewise, only in a few surveys does the mother's loss of her partner significantly associate with her depression in that survey, and mother's loss of her partner often significantly associates with lower depression in subsequent periods, contrary to expectations. Loss of a child significantly associates with much larger increases in depression scores for the period in which the mother lost her child, but these coefficients are estimated with less precision due to infrequency with which children die. Depression in later periods does not robustly associate with loss of a child in earlier periods.

The small and transient effects of bereavement on depression can be interpreted in two ways. First, these small effects may indicate that psychosocial stress in adults generally does not cause lasting changes to affective capital. For example, genetics are fixed and the epigenetic stock may be established early in life. Second – and my preferred interpretation – these small effects may indicate that bereavement in particular does not belong to the class

Death of a friend or relative and mother's depression

Bereavement		Edinburgh Post-Natal Depression Factor				
	survey	b	f	h	r	t
	time	-0.3	0.6	2.7	11.2	19.0
survey:	time:					
b	-0.3	**-0.155**	**-0.109**	**-0.072**	**-0.083**	-0.068
		(0.028)	(0.031)	(0.032)	(0.038)	(0.047)
f	0.6	-0.046	**-0.072**	-0.006	**-0.087**	-0.067
		(0.026)	(0.025)	(0.028)	(0.033)	(0.044)
h	2.7	-0.019	-0.042	**-0.054**	-0.022	-0.056
		(0.023)	(0.023)	(0.023)	(0.027)	(0.036)
p	9.2	-0.021	**-0.046**	-0.036	**-0.091**	**-0.113**
		(0.023)	(0.023)	(0.024)	(0.025)	(0.033)
t	19.0	-0.043	-0.037	-0.055	-0.063	**-0.093**
		(0.031)	(0.032)	(0.032)	(0.034)	(0.033)

The rows indicate the survey in which bereavement was assessed. 'Time' indicates the number of years before the mother's due date or years after the birth of the child. The columns indicate the respondent's depression inventory factor, using the Edinburgh Post-Natal Depression inventory. Parameter esimates come from bivariate regressions of the depression inventory factor on a bereavement indicator. Standard errors are in parenthesis beneath the estimates. Bolded estimates are significant at the 95% confidence level using two-tailed t-test (not corrected for multiple hypothesis testing).

Table D.1: Maternal depression and death of a friend or relative

Death of a child and mother's depression

Bereavement		Edinburgh Post-Natal Depression Factor				
	survey	b	f	h	r	t
	time	-0.3	0.6	2.7	11.2	19.0
survey	time					
b	-0.3	-0.539	-0.038	-0.648	-0.119	**-1.060**
		(0.358)	(0.255)	(0.373)	(0.355)	(0.472)
f	0.6	-0.655	-0.644	0.339	**0.560**	**0.518**
		(0.404)	(0.421)	(0.365)	(0.230)	(0.016)
h	2.7	-0.062	**0.239**	**-0.637**	0.045	**0.371**
		(0.172)	(0.111)	(0.227)	(0.157)	(0.131)
p	9.2	**-0.697**	-0.498	-0.103	**-0.949**	-0.040
		(0.333)	(0.397)	(0.349)	(0.396)	(0.264)
t	19.0	0.149	-0.033	0.199	**0.564**	**-1.369**
		(0.321)	(0.335)	(0.276)	(0.211)	(0.394)

The rows indicate the survey in which bereavement was assessed. 'Time' indicates the number of years before the mother's due date or years after the birth of the child. The columns indicate the respondent's depression inventory factor, using the Edinburgh Post-Natal Depression inventory. Parameter esimates come from bivariate regressions of the depression inventory factor on a bereavement indicator. Standard errors are in parenthesis beneath the estimates. Bolded estimates are significant at the 95% confidence level using two-tailed t-test (not corrected for multiple hypothesis testing).

Table D.2: Maternal depression and death of a child

Death of a partner and mother's depression

Bereavement		Edinburgh Post-Natal Depression Factor				
	survey	b	f	h	r	t
	time	-0.3	0.6	2.7	11.2	19.0
survey	time					
b	-0.3	-0.231	-0.313	0.034	-0.028	-0.168
		(0.240)	(0.264)	(0.315)	(0.336)	(0.449)
f	0.6	0.518	-0.476	0.214	0.318	**0.632**
		(0.272)	(0.457)	(0.274)	(0.523)	(0.224)
p	9.2	-0.450	0.039	-0.256	0.084	-0.453
		(0.317)	(0.216)	(0.277)	(0.197)	(0.292)
t	19.0	-0.340	0.050	-0.019	**-0.711**	**-1.316**
		(0.321)	(0.231)	(0.275)	(0.331)	(0.203)

The rows indicate the survey in which bereavement was assessed. 'Time' indicates the number of years before the mother's due date or years after the birth of the child. The columns indicate the respondent's depression inventory factor, using the Edinburgh Post-Natal Depression inventory. Parameter esimates come from bivariate regressions of the depression inventory factor on a bereavement indicator. Standard errors are in parenthesis beneath the estimates. Bolded estimates are significant at the 95% confidence level using two-tailed t-test (not corrected for multiple hypothesis testing).

Table D.3: Maternal depression and death of partner of stressors that cause large lasting changes to affective capital. Rather, social stressors that indicate to an individual his position within a social hierarchy may be more relevant. The biological literature on social defeat stress in animal models provides suggestive support for such a proposition. In experiments utilizing social defeat, a male rat is forced into the territory of a stronger, socially dominant male rat; the stronger rat attacks the weaker rat, which reliably causes depression-like behaviors in the smaller rat. This social defeat experiment resembles fights seen among rats in wild mischiefs, where rats establish their position in a hierarchy through fighting. Essentially, those stressors that cause depression are very often indicative of social status. Such a class of stressors is inherently endogenous.

b. Burglaries and Muggings: Neither burglaries nor mugging events robustly associate with maternal or paternal depression. In contrast, concerns about the prevalence of burglaries and muggings in the mother's neighborhood do strongly and significantly associate with maternal depression.

Parental depression and parental socioeconomic outcomes

In the ALSPAC data, adolescent depression predicts adverse young adult outcomes. To what extent do these adverse young adult outcomes anticipate adverse middle-aged outcomes?

Parental depression predicts a variety of worse socioeconomic outcomes for parents: lower household income, lower SES occupation, lower educational attainment, and higher unemployment (see Tables D.4 , D.5 , and D.6). For example, a one standard deviation decrease in depression predicts about half the change in household income associated with an A-levels qualification – despite potentially severe error in the measurement of paternal depression. Conditioning on other covariates does little to reduce the predictive content of parental depression. Parental depression factors also significantly associate with parental

Regressions of unemployment variables on long-run depression averages

	No. of surveys with job loss of:				Reported job loss in any survey of:			
	Father		Mother		Father		Mother	
RHS vars	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Father Dep. Mean	0.076	0.080			0.047	0.042		
	0.012	*0.011*			*0.012*	*0.009*		
	(0.007)	(0.009)			(0.004)	(0.005)		
	{2.3E-26}	{4.6E-19}			{8.1E-28}	{5.1E-17}		
Mother Dep. Mean			0.027	0.028			0.024	0.024
			0.004	*0.004*			*0.004*	*0.004*
			(0.004)	(0.005)			(0.003)	(0.004)
			{4.1E-13}	{3.1E-09}			{1.3E-13}	{2.4E-09}
A_levels - father		-0.070				-0.052		
		0.002				*0.004*		
		(0.015)				(0.009)		
		{3.7E-06}				{1.4E-08}		
A_levels - mother				-0.027				-0.026
				0.001				*0.001*
				(0.009)				(0.007)
				{2.0E-03}				{3.6E-04}
Included occupational dummies?		Yes		Yes		Yes		Yes
R2	0.012	0.015	0.004	0.005	0.012	0.013	0.004	0.006
Obs.	11,242	8,753	13,653	10,102	11,242	8,753	13,653	10,102

Coefficients from regression of unemployment outcome on depression and covariates. Partial explained variance in italix (coefficient squared time variance of corresponding RHS variable divided by variance of LHS variable). Heteroskedastic robust standard errors in parenthesis. P-values in curly braces from a two-sided t-test.

Depression variables are the average of available EPDS depression inventory factors for each respondent over all available periods.

Table D.4: Job loss and depression

Regressions of Household income on parental depression, education, and occupation

		LHS: Household income, period t=19		
RHS	(1)	(2)	(3)	(4)
Mother Depression	-3,349***			-2,311***
	(357)			(354)
Father Depression	-1,156***			-1,276***
	(323)			(331)
Mother A-levels		8,433***	4,846***	5,072***
		(571)	(689)	(709)
Father A-levels		5,898***	2,553***	2,382***
		(571)	(622)	(638)
Father: Occupational Category				
Manager/Senior Official			8,340***	35,003***
			(1,131)	(2,402)
Professional			2,847***	29,969***
			(621.6)	(2,276)
Administrative and Secretarial			3,255***	29,961***
			(1,087)	(2,392)
Skilled Trade Occupation			-1,709**	25,264***
			(868.7)	(2,241)
Personal Service Occupation			-2,329**	25,048***
			(1,160)	(2,379)
Process, Plant, and Machine Operatives			-3,420	23,485***
			(2,267)	(2,947)
Process, Plant, and Machine Operatives			-3,546	24,225***
			(4,421)	(4,327)
Mother: Occupational Category				
Manager/Senior Official			35,999***	8,429***
			(1,195)	(2,477)
Professional			30,451***	2,869
			(825.9)	(2,305)
Administrative and Secretarial			28,143***	1,136
			(686.5)	(2,251)
Skilled Trade Occupation			26,547***	-7.248
			(1,384)	(2,569)
Personal Service Occupation			25,050***	-2,676
			(1,223)	(2,442)
Process, Plant, and Machine Operatives			27,096***	
			(2,290)	
Observations	3,186	3,243	2,870	2,681
R-squared	0.04	0.131	0.199	0.218

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table D.5: Parental depression is predictive of parental household income

Bivariate regressions of depression on SES variables

	LHS: Depression factor mean	
RHS: SES variable	Mother	Father
Means by Occupational Category		
Manager/Senior Official	-0.201***	-0.046
	(0.027)	(0.041)
Professional	-0.0951***	-0.0638**
	(0.015)	(0.018)
Administrative and Secretarial	-0.124***	-0.0558**
	(0.027)	(0.016)
Skilled Trade Occupation	-0.0280*	0.053
	(0.016)	(0.037)
Personal Service Occupation	0.028	0.001
	(0.028)	(0.033)
Process, Plant, and Machine Operatives	0.045	-0.012
	(0.052)	(0.069)
Armed Forces	0.511***	0.011
	(0.174)	(0.666)
Attained A-levels	-0.141***	-0.0334*
	(0.018)	(0.020)
Houshold income in 10k lbs.	-0.101***	-0.0645**
t=19	(0.009)	(0.010)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.6: Means of depression by parental occupation, educational attainment, and household income.

unemployment, including after conditioning on parental occupation and parental educational attainment. This association does not appear to be driven by unemployment causing depression, as depression scores prior to unemployment and current depression scores correlate about equally strongly with unemployment (see Table D.7).

Bivariate Regressions of Mother Job Loss on Depression

Mother Dep. Inv. Fac. Score	Job Loss Just Prior to This Survey						
timepoints	-0.25	0.25	0.75	1.75	9	11	19
-0.25	**0.025**	**0.010**	0.001	**0.006**	0.005	0.003	0.003
	0.009	0.003	0.000	0.001	0.001	0.000	0.000
	(0.003)	(0.002)	(0.002)	(0.003)	(0.004)	(0.003)	(0.003)
	{0.000}	{0.000}	{0.605}	{0.017}	{0.191}	{0.304}	{0.325}
0	**0.024**	**0.004**	-0.001	**0.004**	**0.006**	0.001	0.004
	0.009	0.000	0.000	0.001	0.001	0.000	0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.004)	(0.004)	(0.003)
	{0.000}	{0.066}	{0.712}	{0.090}	{0.089}	{0.716}	{0.223}
0.25	**0.023**	**0.005**	0.000	**0.006**	0.004	0.003	0.004
	0.008	0.001	0.000	0.001	0.000	0.000	0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
	{0.000}	{0.049}	{0.849}	{0.014}	{0.294}	{0.319}	{0.270}
0.75	**0.011**	0.002	0.000	**0.008**	0.006	0.004	0.003
	0.002	0.000	0.000	0.003	0.001	0.000	0.000
	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)
	{0.000}	{0.357}	{0.860}	{0.001}	{0.102}	{0.314}	{0.409}
1.75	**0.017**	**0.005**	0.002	**0.004**	0.004	0.000	**0.008**
	0.005	0.001	0.000	0.001	0.001	0.000	0.002
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
	{0.000}	{0.056}	{0.467}	{0.052}	{0.202}	{0.883}	{0.013}
2.75	**0.018**	0.003	0.003	**0.004**	0.001	0.002	**0.006**
	0.005	0.000	0.000	0.001	0.000	0.000	0.001
	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
	{0.000}	{0.181}	{0.168}	{0.093}	{0.696}	{0.604}	{0.071}
11	**0.013**	0.000	0.000	**0.004**	**0.010**	0.002	0.004
	0.004	0.000	0.000	0.001	0.003	0.000	0.000
	(0.003)	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)	(0.003)
	{0.000}	{0.948}	{0.873}	{0.078}	{0.012}	{0.433}	{0.214}
19	**0.010**	0.001	0.002	0.001	0.007	0.005	**0.007**
	0.003	0.000	0.000	0.000	0.001	0.001	0.002
	(0.004)	(0.003)	(0.003)	(0.003)	(0.005)	(0.004)	(0.003)
	{0.009}	{0.777}	{0.637}	{0.636}	{0.168}	{0.279}	{0.027}

Coefficents from bivariate regressions of job loss indicator on maternal depression inventory factor. Underlined coefficients are from corresponding timepoints. Explained variance under coefficient estimates. Standard errors in parenthesis. P-values in curly braces. Bold coefficient estimates are significant at the 90% confidence interval, two-sided t-test. Heteroskedstic robust standard errors. Depression inventory factors calculated using EPDS surveys from respective timepoints.

Table D.7: Maternal depression and job loss

APPENDIX E

E - EXTENDED REVIEW OF RELATED ECONOMIC RESEARCH

My research builds on the skill formation literature, which explores the multifaceted nature of human capital, the determinants of its formation, and its ramifications for mobility and inequality.

While seminal economic work on intergenerational mobility treated human capital as univariate and, essentially, identical to IQ (e.g. Becker and Tomes (1979)), such a unidimensional view of human capital fails to adequately explain many labor market outcomes Heckman et al. (2006); thus, the skill formation literature has expanded human capital to include personality traits, health, genetics and preferences. In their survey of the skill formation literature, Heckman and Mosso Heckman and Mosso (2014) encapsulate this expansive interpretation of human capital in the notion of "skills", where skills include all personal characteristics that provide the capacity to act. Much work has focused on personality traits, like conscientiousness, openness, neuroticism, and grit Borghans et al. (2008), sometimes termed "non-cognitive skills" Cunha and Heckman (2008). Non-cognitive skills are manifold. In some papers, non-cognitive skills refer to conscientiousness, a personality trait describing organization and industriousness. In other papers, non-cognitive skills (including socio-emotional skills) refer to mental health – e.g. in depression in mothers or behavioral problems in children Attanasio et al. (2016). A variety of such non-cognitive skills are found to be plastic yet durable. Their plasticity renders them susceptible to modification; their durability allows the modifications to endure.

Skills also include preferences, so preference formation can be viewed as part of the skill formation literature. For example, Becker et. al. Becker et al. (2014) develop a model in which parents nurture feelings of filial obligation and guilt in their children so as to induce these children to support parents in old age. With preferences so altered, parents can rely on children to earn a return on the investments parents place in their children, thus increasing parental investments. In the same vein, Doepke and Zilibotti (2014) consider how parental preferences interact with the socioeconomic environment to dictate a choice between parenting archetypes described in Baumrind Baumrind (1971). In this economic framework, authoritative parents shape the choices of children by altering the preferences of children while authoritarian parents shape the choices of children by imposing punitive constraints. In both cases, the preferences of parents and children play a key role in dictating economically important investments.

Likewise, skills include health, genetic, and epigenetic traits to the extent that health, genetic, and epigenetic traits shape the capacity to act. In this vein, Biroli Biroli (2014) considers a genetic variant that alters the preference for calories. Consequently, exogenous variation in preferences drives variation health investment and ultimately inequality in health outcomes. Also, Cole et. al. Cole et al. (2012) probe the effects of depriving rhesus macaque of maternal rearing on the transcriptome (which is related to epigenetics). The stark changes in transcription indicate that early how early-life privations may drive later-life variation in health and behavioral outcomes.

This skill formation literature pays special attention to the production of these skills, most directly by estimating technologies of skill formation. In what has become the standard dynamic model, the skills of parents, skills of children, and investments of parents enter into a dynamically nested CES production function; the skills of one period feed into the production of skills in subsequent periods Heckman and Mosso (2014). Discussion of this technology develops two key ideas: sensitive periods and dynamic complementarity. Sensitive periods refer to ages during which investments are most productive at producing skills . Considerable evidence indicates that investments are most productive in the youngest children Knudsen et al. (2006); Thompson and Nelson (2001); interventions targeting adolescents or adults tend to perform poorly, while interventions targeting children often yield substantial and durable effects Heckman and Mosso (2014). On the other hand, dynamic complementarity refers to how early investments increase the productivity of later investments. Generally, this literature finds that different skills tend to be more substitutionary early in childhood, but highly complementary in later childhood and highly complementary across time. Both ideas guide interpretation of the evidence as well as policy recommendations.

The dynamics of skill formation shed insight into a parallel literature on credit constraints. Limited evidence supports the proposition that credit constraints impede investment in adolescents. For example Carneiro and Heckman (2002) finds low returns to college education for marginal students, Taber and Cameron (2004) finds that the mincer return to college exceeds the instrumented return to college, and Keane and Wolpin (2001) estimate a structural model in which borrowing constraints are tight yet do not much impact educational choices, all of which are difficult to reconcile with credit constraints lowering educational attainment. However, a dynamic framework permits parents to be credit constrained when children are young but not when children are older. Indeed, even with limited access to capital markets, younger parents may choose to build a stock of precautionary savings, sacrificing investment in younger children . In this light, bequests do not imply the absence of credit constraints. Rather, credit constraints may operate insidiously, preventing efficient investment in children and thereby reducing optimal investment in adolescents.

The dynamic technology of skill formation motivates considering a larger variety of human capital inputs just as the literature has considered a larger variety of human capital. Indeed, because non-cognitive skills feedback into production of future non-cognitive skills, the non-cognitive skills of children in one period must be inputs into the production of non-cognitive skills in the subsequent period. However, the skill formation literature considers a larger variety of inputs than those directly implied by a dynamic model, including: the vocabulary to which children are exposed (need to find other citation, Fernald et al. (2013)); parental altruism and preference formation; parental time, nutrition Attanasio et al. (2016), and others.

Explicitly or implicitly, the skill formation literature develops the implications of dynamic complementarity and sensitive periods for economic mobility and, ultimately, inequality. For example, both sensitive periods and dynamic complementarity in the technology of skill formation may indicate that investments promoting skill formation are most productive when focused on the young, particularly when credit constraints bind for younger parents. Likewise, sensitive periods and dynamic complementarity indicate that investing relatively more in the least able children can reduce inequality and improve efficiency even if the child's skills in any one period complement investments Cunha and Heckman (2007). Furthermore, the early substitutability of cognitive and non-cognitive skills combined with the relative malleability of non-cognitive skills indicates that interventions designed to reduced inequality in capacities ought to target non-cognitive skills in the young. These ideas justify early intervention on grounds of both efficiency and fairness. Indeed, this explains why the literature has given so much space to evaluating early childhood interventions.

My project contributes to the skill formation literature by developing the scope of human capital and by probing the technology of its formation.

My project explores a neglected dimension of human capital, both by developing the concept of affective capital as as form of habit capital within an intergenerational framework as well as by utilizing novel measurements of this human capital. In a habit formation model, engaging in the habit forming activity reduces utility levels today but increases the habit stock tomorrow; this larger habit stock increases utility levels whiles reducing in magnitude the marginal disutility of the habit forming activity. In my framework, affective capital follows the same pattern but within an intergenerational context. That is, investing in affective capital reduces the utility of the current generation but increases the stock of affective capital for the next generation; this larger affective capital stock increases utility levels while reducing in magnitude the disutility of investing in the third generation[1]. This perspective parallels the skill formation literature's interest in preference formation. Affective capital also provides a theoretical justification for linking investments and outcomes with certain non-cognitive skills. For example, Attanasio et al. (2016) utilize parental depression inventory scores and child behavioral problems as measures of non-cognitive skills, which is somewhat unusual given that non-cognitive skills more often to personality traits like conscientiousness. My approach helps to make sense of such a choice.

I measure affective capital using epigenetic data. Epigenetics are well suited to proxy for certain aspect of human capital because they are durable yet plastic. When epigenetic changes do not alter behavior outright, they can serve as proxies for other biological changes causing behavior. However, epigenetics have not been much utilized in economic research, perhaps because data containing both epigenetics and rich measures of economic outcomes have only recently become available. I develop methodology that allows such newly available epigenetic data to proxy for aspects of human capital.

My project also probes the technology of skill formation. For example, using epigenetics allows me to directly measure human capital endowments at partum – an otherwise challenging feat – and distinguishing these human capital endowment from early life investments may play a key role in understanding the technology of skill formation. As developed in Cunha and Heckman (2007), the complementarity between skills intra- and inter-temporally informs optimal timing of interventions. Interventions targeting a skill in adults will fail to achieve a good return if production of that skill becomes inelastic in adulthood, suggesting that interventions should instead target children, where increments in skill

---

[1] Surprisingly, this similarity is more than superficial. Chronic stress, especially in childhood, cause changes to the brain and to behavior that strongly resemble the effects of drug abuse [FIND CITATIONS].

can compound while the technology is most productive. However, other skills may remain elastic in adulthood. If these other skills can be endowed to children or if these other skills enhance parental investment in children, then interventions targeting adults may indeed earn a good return for the subsequent generation – even if remediating such skills does not much improve the labor market outcomes of the current generation. Affective capital may be such a skill. Indeed, an incipient early-childhood intervention based in Charlottesville, Virginia, treats the depressive symptoms of parents via cognitive-behavioral therapy in a bid to improve parental inputs directed at children.

My project also provides insight into the appropriate measures of investment. Of note, a rich psychology literature has probed the importance of maternal attachment to the child Narita et al. (2000); Parker (1989); Roth and Sweatt (2011). Maternal attachment does not requires goods. It does require time. But more importantly, it requires that the mother uses the available time to interact with the child in particular ways. A specification of the technology of skill formation might neglect the manner of these interactions, instead estimating reduced-form parameters on parental skills. For example, maternal depression associates with less maternal attachment, so maternal affective capital could be treated as though it were itself an investment in the child's affective capital. However, doing so ignores that depressed mothers can choose to express affective engagement through additional effort Lovejoy et al. (2000). Ignoring this choice would introduce potentially non-classical measurement error into a crucial variable. Some particular goods may also be important for forming affective capital. For example, pet ownership may protect children from stressors by blunting the hormonal stress response Polheber and Matchock (2014). Other goods may be less efficacious. More generally, time and goods as broad categories may be too course to measure the inputs most important for skill formation. Finer measures require theoretical justification, and affective capital builds on such justifications.

APPENDIX F

F - IDENTIFICATION OF LATENT VARIABLES USING EPIGENETIC DATA

The Latent Variables Simultaneous Equations Model

The LVSEM treats (demeaned) methylation data as a linear function of unobserved mean-zero latent variables

$$\underbrace{M_i}_{m\times 1} = \underbrace{\Gamma}_{m\times k}\underbrace{\theta_i}_{k\times 1} + \underbrace{\varepsilon_i}_{m\times 1} \quad \text{LVSEM eqn. 1 (measurement system)}$$

where M_i is the column vector of methylation beta values for person i, θ_i is a column vector of latent k variables for person i, Γ is a parameter matrix mapping latent variables onto methylation data, and ε_i is a vector of residuals orthogonal to the latent variables. The coefficient matrix can be partitioned into column vector corresponding to each latent variable $\Gamma = \begin{pmatrix}\Gamma_1 & \Gamma_2 & \dots & \Gamma_K\end{pmatrix}$. I often refer to Γ_j as the "methylation pattern" associated with latent variable j. Without loss of generality, elements in the residual vector are orthogonal to each other (correlated residuals can be decomposed into an uncorrelated component and another latent variable). LVSEM eqn. 1 is equivalent to a measurement system in a linear factor model, so I refer to this system as the measurement system.

In turn, the latent variables may be functions of latent variables and observed variables,

$$\theta_i = \psi\theta_i + bx_i + \phi_i \quad \text{LVSEM eqn. 2 (latent variable DGP)}$$

and other observed variables may be functions of these latent variables (and observed variables)

$$Y = \alpha\theta_i + \beta w_i + \varepsilon_i \quad \text{LVSEM eqn. 3 (outcome equation)}$$

where α and β are coefficients, Y_i is an outcome variable, w_i is a vector of observed covariates, and ε_i is the residual.

One latent variable, $\theta_{1,i}$, caused in part by psychosocial stress and causing depression, is of particular interest. While in principal, different stressors may alter different epigenetic stocks, and different epigenetic (and non-epigenetic) stocks may cause different depression symptoms, I abstract away from these, reifying affective capital as a single latent variable $\theta_{1,i}$.

This perspective sheds light on several questions: how might depression-associated biological stock $\theta_{1,i}$ be forecasted using methylation data and a noisy estimate of $\theta_{1,i}$; how might $\theta_{1,i}$ be disentangled from correlated latent variables $\theta_{-1,i}$ so as to admit a causal model; and how might methylation data be used to refine interpretation of depression inventory scores?

Forecasting from the perspective of the LVSEM

Calculate a leave-one-out forecast of depression-related epigenetic stock $\hat{\theta}_{1,i}$ for each person i with the formula

$$\hat{\theta}_{1,i} = \dot{\theta}_{1,-i} M'_{-i} M_i$$

and rescale the forecasts to be mean zero and unit variance across the N observations, where $\dot{\theta}_{1,-i}$ is the $1 \times (n-1)$ row vector of standardized depression inventory scores of all persons other than person i and M_{-i} is the $m \times (n-1)$ matrix of methylation column vectors for each person other than i.

This formula is equivalent to running m bivariate regressions of depression inventories on each methylation probe, forecasting depression for each regression, constructing a weighted average of these forecasts where the weights are proportional to the variance of the associated methylation probe, and standardizing the weighted average. These weights can be rationalized as those that would arise if the bivariate regressions were estimated with ridge regression and the ridge regression penalty were arbitrarily large.

The LVSEM motivates this forecasting method. Forecasting depression-related epigenetic stock using methylation data does not require disentangling the depression-related stock from other latent variables in the measurement system. Define Γ_1^f as the parameter vector mapping $\theta_{1,i}$ onto M_i when $\theta_{1,i}$is observed and the other confounding latent variables are not observed. That is, decompose $\theta_{-1,i}$ into a component correlated with $\theta_{1,i}$ and a residual:

$$\theta_{-1,i} = \rho\theta_{1,i} + \theta_{-1,i}^r$$

then the measurement system can be rewritten as

$$\begin{aligned} M_i &= (\Gamma_1 + \Gamma_{-1}\rho)\,\theta_{1,i} + \left(\Gamma_{-1}\theta_{-1,i}^r + \varepsilon_i\right) \\ &= \Gamma_1^f\theta_{1,i} + \varepsilon^f \end{aligned}$$

where $(\Gamma_1 + \Gamma_{-1}\rho) \equiv \Gamma_1^f$ and $\left(\Gamma_{-1}\theta_{-1,i}^r + \varepsilon_i\right) \equiv \varepsilon_i^f$. By construction,$\varepsilon_i^f$ is orthogonal to $\theta_{1,i}$.

Preforming regression algebra on this equation yields the forecast estimator. Construct the row vector $\theta_1 = \begin{pmatrix} \theta_{1,1} & \theta_{1,2} & \dots & \theta_{1,N} \end{pmatrix}$. Define

$$\begin{aligned} \hat{\Gamma}_1^f \equiv& M\theta_1' \left(\theta_1\theta_1'\right)^{-1} \\ =& \Gamma_1^f + \varepsilon^f\theta_1' \left(\theta_1\theta_1'\right)^{-1} \end{aligned}$$

note that $\theta_1\theta_1'$ is a scalar, so $\hat{\Gamma}_1^f = \kappa_1 M\theta_1'$ for some scalar κ_1.

Define

$$\begin{aligned}\hat{\theta}_i &\equiv \left(\hat{\Gamma}_1^{f\prime}\hat{\Gamma}_1^{f}\right)^{-1}\hat{\Gamma}_1^{f\prime}M_i\\ &= \left(\hat{\Gamma}_1^{f\prime}\hat{\Gamma}_1^{f}\right)^{-1}\hat{\Gamma}_1^{f\prime}\Gamma_1^{f}\theta_i + \left(\hat{\Gamma}_1^{f\prime}\hat{\Gamma}_1^{f}\right)^{-1}\hat{\Gamma}_1^{f\prime}\varepsilon_i^{f}\end{aligned}$$

Note that $\left(\hat{\Gamma}_1^{f\prime}\hat{\Gamma}_1^{f}\right)^{-1}$ is a scalar. So $\hat{\theta}_i = \kappa_2\hat{\Gamma}_1^{f\prime}M_i$ for some scalar κ_2. Combining these equations yields $\hat{\theta}_i = \kappa_1\kappa_2\theta_1 M' M_i$. Because $\hat{\theta}_i$ is renormed to be mean zero and unit variance, the scalars can be omitted from the equation, thus yielding the forecast equation.

A noisy measure of $\theta_{1,i}$ is sufficient to identify Γ_1^f with attenuation bias. Consequently, noise in the proxy biases forecasts towards zero. Such bias is irrelevant because the forecasts are standardized; asymptotically, noise in the proxy will not affect the rank ordering of forecasts and will not affect correlations between the forecast and the thing forecasted.

The leave-one-out forecast for person i omits person i from the estimation of $\kappa_1\hat{\Gamma}_1^f$, so $\kappa_1\hat{\Gamma}_{1,-i}^f = M_{-i}, \theta_{1,-i}$. Consequently, any correlation between the forecast $\hat{\theta}_i$ and observed depression inventory scores should not be due to overfitting.

The forecast can be interpreted in three ways. First, it can be directly interpreted in terms of the mapping implied by the LVSEM. Second, it can be interpreted as a ridge-regression forecast using a large number of bivariate regressions and arbitrarily large ridge penalty. Third, the forecast can be interpreted in terms of weighting depression similarity of i and others in terms of epigenetic similarity. That is, the term $M'_{-i}M_i$ describes the epigenetic similarity of person i with each other person. Thus the forecast$\dot{\theta}_{-i}\left(M'_{-i}M_i\right)$ predicts person i's depression from the depression levels of epigenetically similar persons.

One practical feature of this forecast is its computation time. For each person i, computation time when grouping matrix multiplication appropriately is $O\left(N^2K\right)$ where N is the number of observations and K is the number of measurements. Because N is small and K is large in my sample, a computation time linear in the number of measurements is desirable.

The forecast must be interpreted warily. Any correlation between depression inventory scores and the leave-one-out forecast of depression indicates that methylation data in monocytes have predictive content. It is tempting to interpret such a forecast as a forecast of the epigenetic stock caused by stressors and causing depression. Such an interpretation is not warranted. $\hat{\Gamma}_1^f$ confounds Γ_1 with $\rho\Gamma_{-1}$. So, for example, smoking and obesity both correlate with depression and both involve changes in methylation patterns. Consequently, Γ_1 could be a zero vector while $\rho\Gamma_{-1}$ is not. In this case, a forecast of depression using methylation data exploits the predictive content of these correlated epigenetic stocks. It is especially tempting to overinterpret the forecast of maternal depression using the child's cord blood. A high correlation indicates substantial information in the cordblood but does not necessarily imply epigenetic endowments of affective capital.

Identifying affective capital requires disentangling θ_1 from the other latent variables $\theta_{-1,i}$.

Identifying θ_1 in the LVSEM

In general, identifying θ_1 with methylation data requires identifying the full coefficient matrix Γ mapping all latent variables onto methylation data. This is a serious threat to identification as the number of latent variables generating methylation data is potentially quite large, and identifying the methylation patterns associated with all latent variables may be unrealistic. Under what assumptions can one identify $\theta_{1,i}$ without identifying the methylation patterns Γ_{-1} associated with all other latent variables θ_{-1}?

In the special case that$\Gamma'_{-1}\Gamma_1 = 0$ (either in the full methylation data or in a subsample of methylation data), then identifying Γ_1 is sufficient to identify θ_1 (plus independent forecast error, which approaches zero under some mild assumptions as the number of measurements grows). Essentially, if and only if the methylation patterns corresponding with other epigenetic stocks are sufficiently dissimilar to the methylation pattern for θ_1 can one ignore these other stocks when estimating θ_1; if other stocks feature similar methylation patterns to θ_1, then a change in one of these other stocks would be mistaken for a change in θ_1. Failure to identify the methylation patterns associated with these other stocks prevents disentangling changes in these other stocks from changes in θ_1. More precisely, an estimate of θ_1 using only Γ_1 yields

$$\begin{aligned}\hat{\theta}_{1,i} =& \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1 M_i \\ =& \theta_{1,i} + \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_{-1}\theta_{-1,i} + \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\varepsilon_i\end{aligned}$$

In expectation, $\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\varepsilon_i = 0$. However, $\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_{-1}\theta_{-1,i}$ introduces correlation between $\hat{\theta}_{1,i}$ and $\theta_{-1,i}$ except in the event that Γ_1 and Γ_{-1} are orthogonal. This is a different problem than that of the latent variables being correlated; the latent variables could all be orthogonal and the identification problem would remain.

This problem of identification threatens the use of estimated epigenetic stocks in other analyses. In particular, suppose some outcome Y_i (e.g. income) in LVSEM eqn. 3 is a linear function of epigenetic stocks

$$Y_i = \theta_{1,i}\beta_1 + \theta_{-1,i}\beta_{-1} + \varepsilon_i$$

Furthermore, as a 'best case' scenario, suppose that the latent variables are all mutually orthogonal, so that $\theta_{-1,i}$ can be omitted from a regression of Y_i on $\theta_{1,i}$ without biasing the estimate of $\theta_{1,i}$. Because estimates of $\theta_{1,i}$ are confounded with $\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_{-1}\theta_{-1,i}$, the estimate of $\theta_{1,i}$ will be correlated with $\theta_{-1,i}$ even though $\theta_{1,i}$ is not, causing omitted variable bias.

Utilizing exclusion restrictions

Exclusion restrictions on observed variables x_i that generate the latent variables can be used to identify a subset of vectors in the coefficient matrix Γ without requiring identification of all parameters in the coefficient matrix. A sufficient assumption to identify Γ_1 is that at least one observed variable $x_{1,i}$, conditional on other observed variables $x_{2,i}$, affects only $\theta_{1,i}$ without affecting other latent variables itself. Formally, partition x_i from LVSEM eqn. 2 into $x_i = \begin{pmatrix} x_{1,i} & x_{2,i} \end{pmatrix}$ where $x_{1,i}$ is a scalar and $x_{2,i}$ is a scalar or row vector and partition b into $b = \begin{pmatrix} b_1 & b_2 \end{pmatrix}$ where b_1 is a column vector and b_2 is a column vector or a matrix; likewise, partition ψ into column vectors $\psi = \begin{pmatrix} \psi_1 & \psi_2 & \cdots & \psi_k \end{pmatrix}$ which map each latent variable onto the others. $x_{1,i}$ plays a role similar to that of an instrument in instrumental variables regression.

In principle, $\theta_{1,i}$ may affect other latent variables, so that multiple entries in ψ_1 may be non-zero. This could be through direct biological mechanisms or through behavioral mechanisms. For example, chronic changes in stress hormone levels in tandem with depression can increase cardiovascular disease risk and cardiovascular disease may involve its own set of associated methylation patterns. In this case, the entry in ψ_1 mapping θ_1 onto the epigenetic latent variable(s) associated with heart disease would be non-zero. Likewise, changes in diet or substance use caused by depression may in turn alter the methylation patterns associated with diet or substance-use. Depending on the application, these channels can be ignored. Assuming $\psi_1 = \vec{0}'$ implicitly includes epigenetic stocks downstream of θ_1 as part of θ_1. Under this interpretation, Γ_1 would represent the cumulative changes in methylation through immediate and downstream mechanisms (both directly biological and behavioral mechanisms). Imposing $\psi_1 = \vec{0}'$ represents a strong structural assumption, which may limit external validity of any findings. For example, if depression increases smoking behavior, then assuminxg $\psi_1 = \vec{0}'$ includes in Γ_1 those methylation changes caused by smoking in proportion to the amount that depression causes smoking. In a different population where smoking is quite rare even among the depressed, assuming $\psi_1 = \vec{0}'$ would identify a different parameter vector for Γ_1. Forecasts trained in one population would be less accurate when applied to another.

When $x_{1,i}$ is excluded from the equations generating other latent variables $\theta_{-1,i} \perp x_{1,i} \mid x_{2,i}$ and $x_{2,i}$ is not, regressing methylation data on $x_{1,i}$ and $x_{2,i}$ identifies Γ_1, but does not identify the methylation patterns for other stocks associated with $x_{2,i}$ unless additional exclusion restrictions for these stocks are utilized. For expository purposes, suppose that $x_{1,i}$ and $x_{2,i}$ both cause $\theta_{1,i}$, only $x_{2,i}$ causes $\theta_{2,i}$, and no observed variables cause only $\theta_{2,i}$. Then one cannot identify Γ_2 using these observed variables, but one can identify the component of Γ_2 residual to Γ_1. That is, the methylation pattern Γ_2 can be decomposed into a linear combination of Γ_1 and a residual vector $\dot{\Gamma}_2$: $\Gamma_2 = \delta\Gamma_1 + \dot{\Gamma}_2$. This residual vector $\dot{\Gamma}_2$ can be identified.

In the event that Γ_1 is identified but only the residual vector of Γ_2 is identified, one can nevertheless use epigenetic stocks estimated using Γ_1 and the residual vector to Γ_2 to identify parameters on θ_1 and θ_2 in the outcome equation (LVSEM eqn. 3). Although the estimate of θ_1 is confounded with variation in θ_2 because Γ_2 is not identified and $\Gamma_1'\Gamma_2$ may be non-zero, regressing an outcome variable on these estimates of θ_1 and θ_2 nevertheless recovers the same coefficients that would be recovered if estimates of θ_1 and θ_2 were not confounded. The intuition follows from residual regression.

Estimating θ_1 with observed variables satisfying such an exclusion restriction places heavy demands on the data. Because any variable that could plausibly satisfy the exclusion restriction (e.g. bereavement) is unlikely to explain much of the variation of θ_1 in observational data, sampling variation in Γ_1 will tend to be larger than if a proxy for θ_1 which explains much of the variation in θ_1 (e.g. depression inventory scores) were used to estimate $\Gamma_1 + \rho\Gamma_{-1}$.

I consider using bereavement shocks as x_1 variables satisfying the aforementioned exclusion restriction; controlling for age and other covariates, bereavement is assumed to alter the epigenetic stocks that psychosocial stress generally alters without altering other epigenetic stocks. Unfortunately, more plausibly exogenous categories of bereavement (loss of a pre-adolescent child) feature too little variance to be of use. Less plausibly exogenous categories of bereavement (death of a friend or relative) are highly significant predictors of increased depression in the contemporaneous survey, but do not predict long-run changes in depression rates and do not explain much of the variation in depression rates, indicating that these shocks are either underpowered for my sample or are not in the set of shocks that cause lasting epigenetic changes. Stressors that explain a large share of variance in depression (relationship quality, social support, neighborhood quality) do not plausibly satisfy the exclusion restriction.

Affective Capital

In this section I develop a latent variable model for reifying affective capital as an epigenetic stock. This latent variable model identifies affective capital by utilizing observed variables that are assumed to cause affective capital.

Utilizing the methylation data requires some form of data reduction. For each of the roughly six-thousand observations in the ALSPAC with methylation data, the methylation data contains hundreds of thousands of continuous variables[1] each representing the extent of methylation at a site on monocyte DNA. Using this data requires either preselecting

---

[1] The beta-value at a methylation site describes the ratio of methylation signal intensity to total signal intensity from the methylation probe; thus, the beta-value lies between zero and one. Logit transformation of the beta-value yields the M-value; thus, the M-value lies between negative infinity and positive infinity, with positive values indicating that more CpG islands are methylated than not.

a small subset of theoretically relevant methylation sites or extracting summary variables reflecting patterns variation within the methylation data as a whole. My approach assumes that a linear function of small set of latent variables generates the methylation data, and one of these latent variables is affective capital. Thus, my approach resembles a linear factor model, at least superficially.

For this reason, description of my approach would first benefit from an overview of linear factor models and an explanation why these linear factor models may not be well-suited to the task at hand. In a linear factor model, latent variables ('factors', 'latent factors', or 'common factors') co-occurring in linear equations that determine the observed methylation variables ('measurements') generate the covariance between methylation variables ('measurements'). Factor analysis recovers the coefficients on these latent variables ('loadings') by interpreting the covariance of the methylation variables ('measurements'), and in this sense resembles principal components analysis (indeed, sometimes exploratory factor analysis begins with principal components analysis; other times exploratory factor analysis utilizes maximum likelihood estimation or a method of moments). However, the measurement covariance matrix can be decomposed in many ways, so identification of the loadings and the factors requires additional assumptions. Some methods of factor analysis assume that the factors are uncorrelated, which implies that factors occurring in mutually exclusive sets of measurements do not contribute to the covariance between those measurements. However, even with this assumption, the 'rotation' of the factors is unidentified. That is, the factors can be thought of as dimensions of the data, and a change of basis (rotating the axes) will not alter the explanatory power of the factors; since parameters cannot be identified if changes in the parameters do not change the distribution of the data, the rotation is not identified. A variety of rules of thumb suggest how to rotate the data for different purposes (for a review of factor analysis, see Gorsuch, 1983).

Factor analysis effectively serves two purposes: data reduction for data description and the construction of reliable indices. While the seemingly arbitrary assumptions on orthogonality of the factors and factor rotation would seem to preclude use of estimated factors for recovering structural parameters on these factors, the factors can nevertheless be used for data description. For example, survey responses may covary when the survey questions pertain to similar subject matter; factor analysis (or alternatively principal components analysis) can summarize these responses as though all the questions were assessing some smaller dimension of characteristics. Personality psychology makes use of this approach when constructing the ’big five’ personality factors (John and Srivastava, 1999). Researchers rotate the factors so as to facilitate interpretation of the personality factors. Under the preferred rotation, questions pertaining to organization and industriousness strongly ’load’ (to load on a factor is to have non-zero coefficients on that factor) have on one factor (termed ’conscientiousness’) but not on other factors, questions pertaining to the tendency to experience negative emotions load on a different factor (termed ’neuroticism’) but not other factors, and so on. Because the factors load on subsets of questions that can be easily interpreted, the factors themselves may be easily interpreted. These factors may then be used in other empirical work, tracing the relationships between personality characteristics and life outcomes. For instance, Cuijpers et al. (2010) find that the neuroticism factor predicts higher healthcare expenditures and worse employment outcomes. Examples of other descriptive applications of factor analysis include principal components analysis of the yield curve (exe. Diebold et al., 2006) and factor analysis of electrical load curve (Manera and Marzullo, 2005).

Factor analysis can also recover reliable indices. When researchers design a set of survey questions to assess precisely one trait but do not expect that responses to any single question will strongly predict the trait, the researchers may assume that the survey responses load on a single factor and extract this factor using factor analysis (or sometimes just a simple index). For instance, IQ tests can extract the ’g’ factor (Kamphaus et al., 1997), and depression inventories can extract a depression factor. In this context, the factors may be used in causal structural models because questions were designed to assess a theoretically motivated factor.

Even when questions load on multiple factors, theory can allow the factors to be utilized in structural work. A theoretically motivated dedicated measure exemplifies such an approach. Specifically, theory may justify that a set of questions loads on two independent factors, but a subset of questions (the ’dedicated measures’) loads on only one of these factors. This dedicated measure assumption pins down the rotation of the matrix because this subset of questions may covary with other questions only through the factor upon which this subset loads. Covariance between questions outside of this subset determines the loadings on the other factor. Some economists have made use of this approach. For example, Kautz and Zanoni (2014) assumes that high school test scores load on only the cognitive ability factor but high school GPA and other academic outcomes load on both cognitive ability and conscientiousness. Other theoretically motivated interpretation of factors includes the use of principal components as controls in genome-wide association studies (Beauchamp et al., 2011). That is, the law of independent assortment combined with non-assortative mating implies that alleles on different genetic loci should be uncorrelated within a given ethnic or geographic population. However, genetic drift can cause shifts in allelic frequencies between populations in what is called ’genetic stratification’. Thus, correlation in alleles can be interpreted as arising from ethnic or geographic background. Then the principal components capture this ethnic background and may be used to control for ethnic background when regressing outcomes on alleles (thus precluding so-called ’chopstick effects’ from the estimated effects of the allele).

While factor analysis might appear useful for recovering an affective capital latent variable, several problems riddle this approach. These problem stem from a currently limited understanding of the epigenome and how the epigenome affects phenotype. Although a rich biology literature indicates that stressors affect methylation broadly, and that stressors often alter methylation at specific sites, the current research does not preclude that other environmental features do not also alter methylation at these sites. Consequently, one cannot justify treating methylation at such sites as dedicated measures of affective capital. As an alternative, one might be tempted use non-methylation measures as dedicated measures of affective capital, and use the covariance between these non-methylation measures (such as depression inventories) and methylation to identify the loadings on affective capital in the methylation data. Unfortunately, such an approach would require that the affective capital factor does not covary with other epigenetic factors. If, for example, those with low affective capital also exercise less or smoke more, and exercise and smoking affect methylation, then such a factor analytic approach will attribute the changes in methylation due to exercise and smoking to the affective capital factor. While this confoundedness might not be too problematic for more descriptive work, this confoundedness will prevent using affective capital in a structural model, where, for example, the econometrican may want to utilize an exogenous stressor that shifts affective capital without changing smoking behavior or exercise.

My latent variable approach does not place assumptions on the covariance structure of the factors, does not require assumption on the rotation of loadings, and does not require dedicated measures; instead, my approach utilizes exclusion restrictions on some observed variables purported to cause the latent factors (in combination with some weak assumptions regarding the coefficients on the latent variables, to be formally elucidated). These exclusion restrictions allow identification of factor loadings from a regression of measurements on the causal variables. Thus, my approach offers a useful alternative to more standard factor analytic approaches when: theory indicates that certain observed variables exogenously cause a latent variable after conditioning on covariates; theory cannot justify excluding all but one factor from some measurements; subsequent analysis requires requires distinguishing a factor of interest from other factors that correlate with it.

Concretely, the ALSPAC data includes variables describing whether individuals have been mugged and whether their homes have been burglarized (as well as other quasi-exogenous stressors, such as bereavement). Muggings and burglaries cause substantial stress, so that rates of depression spike after falling victim to either (Phillips et al., 2015; Browning et al., 2012). Thus, it is reasonable to assume that these events cause changes in methylation typically caused by stressors. However, the risk of being mugged or burglarized will be be a function of location of residence, which may be correlated with other things that cause changes in methylation (for example, obesity may cause changes in methylation patterns; obesity is correlated with socioeconomic status; and socioeconomic status may correlate with living in high crime areas). Consequently, changes in methylation associated with adverse life events cannot be interpreted as the effects of those adverse life events. But the ALSPAC data affords controls for confounding variables. Specifically, the ALSPAC data includes variables describing individual's expectations about the likelihood of being mugged or having their homes burglarized. Under a rational expectations assumption, being mugged or having one's house burglarized is independent of other things that affect methylation after conditioning on the expectation of these events occurring.

It bears emphasizing that my empirical strategy makes use of burglary and other conditionally exogenous stressors not because these stressors are a primary, or even secondary, mechanism retarding intergenerational mobility, but rather because burglary victimization, conditional on individual's rational expectations over burglary victimization, allows identification of a latent variable whose transmission may in fact substantially retard intergenerational mobility. With the latent variable's methylation coefficients identified, the extent of the factor's intergenerational transmission can be measured as well as the mechanisms governing its transmission.

To preview my formal discussion: first, I set up a linear measurement system of latent latent variables and two linear processes where observed variables generate the latent variables; second, I describe the conditions under which one can identify the latent variable coefficients from the measurement system and observed variables; third, I show how violation of a central assumption prevents unbiased estimates of the latent variable but, nevertheless, may not imperil subsequent structural analysis using the misidentified latent variable. Fourth, I discuss how the number of measures and observations affect error on latent variable estimates. Fifth, I discuss conditions under which one can account for measurement error in latent variable estimates in finite samples.

Formally, I treat methylation levels at K different sites $M_i \equiv \begin{pmatrix} M_{1,i} & M_{2,i} & \dots & M_{K,i} \end{pmatrix}'$ as measures of latent variables $\theta_i \equiv \begin{pmatrix} \theta_{AK,i} & \theta_{2,i} & \theta_{3,i} & \dots & \theta_{J,i} \end{pmatrix}'$ in a linear latent variable model with coefficients $\underset{K\times J}{\Gamma} \equiv \begin{pmatrix} \Gamma_1 & \Gamma_2 & \Gamma_3 & \dots & \Gamma_J \end{pmatrix}$

$$M_i = \Gamma\theta_i + \varepsilon_i$$

where $\varepsilon_i = \begin{pmatrix} \varepsilon_i^1 & \varepsilon_i^2 & \dots & \varepsilon_i^K \end{pmatrix}'$ are the uniquenesses (that is, $\varepsilon^k \perp \varepsilon^{k'}$ if $k \neq k'$). The observed univariate variables x_i and w_i generate two of the latent variables $\theta_{AK,i}$ and θ_2, where x_i would be stressful life events such as burglaries and w_i would be expectations of those burglaries occurring (my approach does not require that x_i and w_i be univariate, but this assumption speeds exposition):

$$\begin{aligned} \theta_{AK,i} &= x_i\beta_{AK,x} + w_i\beta_{AK,w} + \varepsilon_{AK,i} \\ \theta_{2,i} &= w_i\beta_{2,w} + \varepsilon_{2,i} \end{aligned}$$

I normalize the coefficients $\beta_{AK,x} = 1$ and $\beta_{2,w} = 1$, which defines the units of $\theta_{AK,i}$ and $\theta_{2,i}$ and is permissible under the assumption that $\beta_{AK,x} \neq 0$, $\beta_{2,w} \neq 0$. I also assume $\Gamma_1 \neq 0$

and $\Gamma_2 \neq 0$. Crucially, I assume that the equations generating all latent variables other than θ^{AK} exclude x_i: $\mathbb{E}\left[\varepsilon_{2,i} \mid x_i\right] = 0$, $\mathbb{C}\left(\theta_{j,i}, x_i\right) = 0 \; \forall j \neq AK, 2$. However, identification of the latent variable coefficients does not require that $\mathbb{E}\left[\varepsilon_{AK,i} \mid x_i\right] = 0$ as I can simply treat the component of $\varepsilon_{AK,i}$ correlated with x_i as part of x_i, which simply affects interpretation of the units of $\theta_{AK,i}$. I also assume that $\mathbb{C}\left(\theta_{j,i}, w_i\right) = 0 \; \forall j \neq AK, 2$, which is to say that conditioning variables w_i are uncorrelated with the "residual latent variables", although I suspect that this assumption can be relaxed (or reinterpreted) under certain conditions without harming identification of structural parameters on Γ_1 or $\theta_{AK,i}$.

Substituting the equations generating the latent variables into the measurement system yields the reduced form equation

$$\begin{aligned} M_i &= \Gamma_1 x_i + \left(\Gamma_2 + \Gamma_1 \beta_{AK,w}\right) w_i + \begin{pmatrix} \Gamma_3 & \Gamma_4 & \ldots & \Gamma_J \end{pmatrix} \begin{pmatrix} \theta_{3,i} & \theta_{4,i} & \ldots & \theta_{J,i} \end{pmatrix}' \\ &\quad + \varepsilon_i + \varepsilon_{AK,i}\Gamma_1 + \varepsilon_{2,i}\Gamma_2 \\ &= \Gamma_1 x_i + \tilde{\Gamma}_2 w_i + \tilde{\varepsilon}_i \end{aligned}$$

Note that the residual from the reduced form regression $\tilde{\varepsilon}_i$ contains the latent variables $\theta_{3,i}$ through $\theta_{J,i}$ (which I term the "residual latent variables"), uniquenesses ε_i, and the residuals from the equations generating $\theta_{AK,i}$ and $\theta_{2,i}$. Together, I term the sum of the residual latent variables and the uniquenesses ε_i the "measurement residual". I include the residual latent variables in this model to aid discussion of measurement error in forecasts of the θ_{AK} and θ_2 latent variables.

Claim 0: The structural parameter vector Γ_1 and the reduced form parameter vector $\tilde{\Gamma}_2$ are identified.

Proof: Under the above assumptions, $\mathbb{E}\left[\tilde{\varepsilon}_i \mid x_i, w_i\right] = 0$ (with methylation levels demeaned), thus OLS assumptions are satisfied (as the sample size grows large; the number of measurements do not affect identification of the coefficients for any given measurement).

Claim 1: The structural latent variable coefficient vector Γ_2 can be recovered if Γ_1 is orthogonal to Γ_2 in the sense that $\Gamma_1'\Gamma_2 = 0$; the structural latent variable coefficient vector Γ_2 cannot be identified without restrictions on the parameter space.

Proof: To prove the first half of the claim, note that

$$\begin{aligned} & \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1' M_i \\ = \quad & x_i + \beta_{AK,w} w_i + \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2 w_i + \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\tilde{\varepsilon}_i \quad \text{[eqn. 1]} \end{aligned}$$

and $\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2 w_i = 0$ if $\Gamma_1'\Gamma_2 = 0$. Thus, β_{AK} is identified. With β_{AK} identified, Γ_2 is identified from the equation $\tilde{\Gamma}_2 - \Gamma_1\beta_{AK,w} = \Gamma_2$. It is trivial to prove the second half of this claim when there is only one w variable: the first measurement contains three parameters and two independent variables, and each subsequent measurement adds two additional parameters and two additional moment restrictions. Thus, identifying all of the parameters requires some restriction on the parameter space. It can be shown that when there are multiple w variables, Γ_2 still cannot be identified if $\Gamma_1 \not\perp \Gamma_2$, but instead $\Gamma_2 - s\Gamma_1$ is identified for some scalar s. Note that $\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2$ is a scalar, so that the regression estimate of $\beta_{AK,w}$ converges to$\overset{\circ}{\beta}_{AK,w} \equiv \beta_{AK,w} + \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2$, and the estimate of the coefficient vector Γ_2 converges to $\overset{\circ}{\Gamma}_2 \equiv \Gamma_2 - \Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2$.

Intuitively, any resemblance between Γ_1 and $\tilde{\Gamma}_2$ could be caused by either a resemblance between the coefficient vectors Γ_1 and Γ_2, w_i having an effect on $\theta_{AK,i}$, or any combination of the two. Assuming that $\Gamma_1 \perp \Gamma_2$ essentially imposes that Γ_2 does not resemble Γ_1. When this assumption is false, the above estimation residualizes the weights of Γ_2.

For future reference, recall the definition

$$\overset{\circ}{\Gamma}_2 \equiv \Gamma_2 - \Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2$$

Claim 2: If Γ_2 is estimated under the assumption that $\Gamma_1 \perp \Gamma_2$, and this assumption is false, then forecasts of both $\theta_{AK,i}$ and $\theta_{2,i}$ will be inconsistent.

Proof: The proof is in secion C.

Claim 3: Define $\theta_{AK-2,i}$ as the vector $\begin{pmatrix} \theta_{AK,i} & \theta_{2,i} \end{pmatrix}'$; likewise, define $\overset{\circ}{\theta}_{AK-2,i}$ as the forecast of $\theta_{AK-2,i}$ by using Γ_1 and $\overset{\circ}{\Gamma}_2$. Furthermore, suppose that the forecast error contributed by the residual latent variables $\theta_3, \ldots, \theta_J$ and the uniquenesses ε_i is small. Then regressing a variable on $\overset{\circ}{\theta}_{AK-2,i}$ will identify the same coefficients as those from regressing on $\theta_{AK-2,i}$.

Proof: I have proven this by verifying that $\left(\overset{\circ}{\theta}{}'_{AK-2,i}\overset{\circ}{\theta}_{AK-2,i}\right)^{-1}\overset{\circ}{\theta}{}'_{AK-2,i}\theta_{AK-2,i}$ equals the identity matrix. The proof consists of long and tedious matrix algebra, which are included in section C. Because $\left(\overset{\circ}{\theta}{}'_{AK-2,i}\overset{\circ}{\theta}_{AK-2,i}\right)^{-1}\overset{\circ}{\theta}{}'_{AK-2,i}\theta_{AK-2,i}$ equals the identity matrix, premultiplying a regression term $\theta_{AK-2,i}\beta$ by the pseudo-inverse $\left(\overset{\circ}{\theta}{}'_{AK-2,i}\overset{\circ}{\theta}_{AK-2,i}\right)^{-1}\overset{\circ}{\theta}{}'_{AK-2,i}$ will yield the regression coefficient β. Combined with the fact that $\overset{\circ}{\theta}_{AK-2,i}$ is orthogonal to the error term whenever $\theta_{AK-2,i}$, this implies that regression of an outcome on $\overset{\circ}{\theta}_{AK-2,i}$ will asymptotically yield the same result as the regression of the outcome on $\theta_{AK-2,i}$.

Intuitively, this result can be explained in light of residual regression. Specifically, regressions on the inconsistent latent variable estimate nevertheless yields consistent regression parameters because $\overset{\circ}{\Gamma}_2$ is the component of Γ_2 residual to Γ_2; and regression on a forecast calculated from 'regression' using Γ_1 and $\overset{\circ}{\Gamma}_2$ residualizes the loading coefficients anyway.

In light of claim 3, the somewhat strong assumption that $\Gamma_1 \perp \Gamma_2$ *may not be too problematic for using the estimated factors in subsequent analysis.*

Claim 4: Error in the latent variable estimates due to the uniqueness and residual latent variables will grow arbitrarily small as the number of measures grows if (1) the variance of the latent variable coefficients for additional measurements does not approach zero or infinity, and the expected value of latent variable coefficients for additional measurements do not approach infinity in absolute value, and (2) if either (A) the latent variable coefficients for the residual latent variables are orthogonal to the latent variable coefficients Γ_1 and Γ_2 or (B) the contribution of the residual latent variables to the covariance between measures grows arbitrarily small.

Condition [1] stipulates that the additional measures added remain relevant but not too relevant, which allows a law of large numbers to be applied, averaging out the forecasts of a factor from many different measurements. Condition [A] stipulates that changes in the measures caused by changes in residual latent variables do not affect forecasts of θ_{AK} or θ_2. Condition [B] essentially stipulates that the residual factors account for a trivial portion of the covariance of measures.

Proof sketch: Proof sketch is in section C.

Claim 5: The variance-covariance structure of the latent variables θ_{AK} and θ_2 is identified if the latent variable coefficients Γ_1 and Γ_2 are known and are orthogonal to the latent variable coefficients for the residual factors; if the variance-covariance structure of θ_{AK} and θ_2 is identified and the latent variable coefficients Γ_1 and Γ_2 are orthogonal to the latent variable coefficients for the residual latent variables, then the variance of the measurement error on θ_{AK} and θ_2 can be estimated (such as when the number of measurements is small).

This claim allows the econometrician to adjust for measurement error on the latent variables when using these latent variables in subsequent analysis.

Proof: Proof is in the appendix.

To summarize, my approach allows identification of a latent variable of interest (as well as its coefficients) in a linear measurement system when the econometrician observes variables purported to cause only the latent variable of interest after conditioning on control variables. This approach requires: a linearity assumption (which can be somewhat relaxed), an exclusion restriction (on the variables purported to cause the latent variable of interest), orthogonality of coefficients on latent variables (violations of which may not impede use of the latent variable in subsequent structural work), and that the number of measures is large (under some regularity conditions). The advantage of this latent variable approach is that it allows researchers to extract meaningful latent variables from epigenetic data for subsequent structural work without requiring the typical assumptions utilized in factor analysis which are unlikely to apply with methylation data. The key assumption required for this approach is that the econometrician observes variables that he can assume cause only the latent variable of interest (after controlling for other observed variables).

C) Proofs and proof sketches for latent variable approach

Claim 2: If Γ_2 is estimated under the assumption that $\Gamma_1 \perp \Gamma_2$, and this assumption is false, then forecasts of both $\theta_{AK,i}$ and $\theta_{2,i}$ will be inconsistent.

Proof: The proof is a tedious computation but it should be unsurprising.

Throughout the following derivations and proofs I make heavy use of the fact that Γ_1 and Γ_2 are column vectors, so $\Gamma_i'\Gamma_j$ for $i, j \in \{1,2\}$ is a scalar; thus transposing or commuting this scalar is a permissible manipulation. To save space, define $\Omega \equiv \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'$.

Call $\overset{\circ}{\theta}' \equiv \left(\overset{\circ}{\Gamma}{}'\overset{\circ}{\Gamma}\right)^{-1}\overset{\circ}{\Gamma}{}' M$, the forecast of θ using Γ_1 and $\overset{\circ}{\Gamma}_2$ when Γ_1 and Γ_2 are not orthogonal. This equals

$$\begin{aligned}
[1]:\ \overset{\circ}{\theta} &= \left(\begin{pmatrix}\Gamma_1' \\ \Gamma_2' - \Omega\Gamma_2\Gamma_1'\end{pmatrix}\begin{pmatrix}\Gamma_1 & \Gamma_2 - \Gamma_1\Gamma_2'\Omega'\end{pmatrix}\right)^{-1}\begin{pmatrix}\Gamma_1' \\ \Gamma_2' - \Omega\Gamma_2\Gamma_1'\end{pmatrix}M \\
[2]: &= \begin{pmatrix}\Gamma_1'\Gamma_1 & \Gamma_2'\Gamma_1 - \Omega\Gamma_2\Gamma_1'\Gamma_1 \\ \Gamma_2'\Gamma_1 - \Omega\Gamma_2\Gamma_1'\Gamma_1 & \Gamma_2'\Gamma_2 - 2\Omega\Gamma_2\Gamma_1'\Gamma_2 + \Omega\Gamma_2\Gamma_1'\Gamma_1\Gamma_2'\Omega'\end{pmatrix}^{-1}\begin{pmatrix}\Gamma_1' \\ \Gamma_2' - \Omega\Gamma_2\Gamma_1'\end{pmatrix}M \\
[3]: &= \begin{pmatrix}\Gamma_1'\Gamma_1 & \Gamma_2'\Gamma_1 - \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\Gamma_1 \\ \Gamma_2'\Gamma_1 - \Gamma_1'\Gamma_2 & \Gamma_2'\Gamma_2 - 2\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\Gamma_2 + \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1}\end{pmatrix}^{-1}\begin{pmatrix}\Gamma_1' \\ \Gamma_2' - \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\end{pmatrix}M \\
[4]: &= \begin{pmatrix}\Gamma_1'\Gamma_1 & \Gamma_2'\Gamma_1 - \Gamma_1'\Gamma_2 \\ \Gamma_2'\Gamma_1 - \Gamma_1'\Gamma_2 & \Gamma_2'\Gamma_2 - 4\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2 + \Gamma_1'\Gamma_2\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1}\end{pmatrix}^{-1}\begin{pmatrix}\Gamma_1' \\ \Gamma_2' - \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\end{pmatrix}M \\
[5]: &= \begin{pmatrix}\Gamma_1'\Gamma_1 & 0 \\ 0 & \Gamma_2'\Gamma_2 - 4\Gamma_1'\Gamma_2\left(\Gamma_1'\Gamma_1\right)^{-1} + 2\Gamma_1'\Gamma_2\left(\Gamma_1'\Gamma_1\right)^{-1}\end{pmatrix}^{-1}\begin{pmatrix}\Gamma_1' \\ \Gamma_2' - \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\end{pmatrix}M \\
[6]: &= \frac{1}{\left(\Gamma_2'\Gamma_2\Gamma_1'\Gamma_1 - 2\Gamma_1'\Gamma_2\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_1\right)}\begin{pmatrix}\Gamma_2'\Gamma_2 - 2\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2 & 0 \\ 0 & \Gamma_1'\Gamma_1\end{pmatrix}\begin{pmatrix}\Gamma_1' \\ \Gamma_2' - \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\end{pmatrix}M \\
[7]: &= \frac{1}{\left(\Gamma_2'\Gamma_2\Gamma_1'\Gamma_1 - 2\Gamma_1'\Gamma_2\right)}\begin{pmatrix}\Gamma_2'\Gamma_2 - 2\Gamma_1'\Gamma_2\left(\Gamma_1'\Gamma_1\right)^{-1} & 0 \\ 0 & \Gamma_1'\Gamma_1\end{pmatrix}\begin{pmatrix}\Gamma_1' \\ \Gamma_2' - \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\end{pmatrix}M \\
[8]: &= \frac{1}{\left(\Gamma_2'\Gamma_2\Gamma_1'\Gamma_1 - 2\Gamma_1'\Gamma_2\right)}\begin{pmatrix}\Gamma_2'\Gamma_2\Gamma_1' - 2\Gamma_1'\Gamma_2\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1' \\ \Gamma_1'\Gamma_1\Gamma_2' - \Gamma_1'\Gamma_2\Gamma_1'\end{pmatrix}M
\end{aligned}$$

where step 1 comes from the regression formula, step 2 multiplies vectors inside the inverse, step 3 substitute in Ω, step 4 simplifies by utilizing the fact that $\Gamma_i'\Gamma_j$ is a scalar, step 5 cancels more terms, step 6 computes the inverse of a 2x2 matrix, step 7 simplifies further, and step 8 multiplies the inverse matrix times $\overset{\circ}{\Gamma}'$. This derivation shows that the unbiasedness of the forecast of θ from $\overset{\circ}{\Gamma}$ is ensured only when Γ_1 and Γ_2 are orthogonal (in the statistical sense). When they are orthogonal, then the forecast is

$$\begin{aligned}
&\begin{pmatrix}\left(\Gamma_2'\Gamma_2\Gamma_1'\Gamma_1 - 2\Gamma_1'\Gamma_2\right)^{-1}\left(\Gamma_2'\Gamma_2\Gamma_1' - 2\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\right) \\ \left(\Gamma_2'\Gamma_2\Gamma_1'\Gamma_1 - 2\Gamma_1'\Gamma_2\right)^{-1}\left(\Gamma_1'\Gamma_1\Gamma_2' - \Gamma_1'\Gamma_2\Gamma_1'\right)\end{pmatrix}M \\
= \quad &\begin{pmatrix}\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1' \\ \left(\Gamma_2'\Gamma_2\right)^{-1}\Gamma_2'\end{pmatrix}M
\end{aligned}$$

as one would expect.

Claim 3: Define $\theta_{AK-2,i}$ as the vector $\begin{pmatrix} \theta_{AK,i} & \theta_{2,i} \end{pmatrix}'$; likewise, define $\overset{\circ}{\theta}_{AK-2,i}$ as the forecast of $\theta_{AK-2,i}$ by using Γ_1 and $\overset{\circ}{\Gamma}_2$. Furthermore, suppose that the forecast error contributed by the residual factors $\theta_{3,} \ldots, \theta_J$ and the uniquenesses ε_i is small. Then regressing a variable on $\overset{\circ}{\theta}_{AK-2,i}$ will identify the same coefficients as those from regressing on $\theta_{AK-2,i}$.

Proof: I have proven this by verifying that $\left(\overset{\circ}{\theta}{}'_{AK-2,i}\overset{\circ}{\theta}_{AK-2,i}\right)^{-1}\overset{\circ}{\theta}{}'_{AK-2,i}\theta_{AK-2,i}$ equals the identity matrix, but the proof consists of long and tedious matrix algebra. Intuitively, this result can be explained in light of residual regression (residual regression refers to the fact that a regression on variables *h* and *z* will yield the same estimates as a regression on *h* and the residual to the regression of *z* on *h*; this is why regression captures the ceteris paribus relationships between dependent and independent variables). Specifically, regressions on the inconsistent factor estimates nevertheless yields consistent regression parameters because $\overset{\circ}{\Gamma}_2$ is the component of Γ_2 residual to Γ_1; and regression on a forecast calculated from 'regression' using Γ_1 and $\overset{\circ}{\Gamma}_2$ residualizes the loading coefficients anyway.

First, some auxilary computations demonstrate that $\overset{\circ}{\Gamma}\left(\overset{\circ}{\Gamma}{}'\overset{\circ}{\Gamma}\right)^{-1}\overset{\circ}{\Gamma}{}' = \Gamma\left(\Gamma'\Gamma\right)^{-1}\Gamma'$. Compute components of $\overset{\circ}{\Gamma}\left(\overset{\circ}{\Gamma}{}'\overset{\circ}{\Gamma}\right)^{-1}\overset{\circ}{\Gamma}{}'$:

$$
\begin{aligned}
&\left(\overset{\circ}{\Gamma}{}'\overset{\circ}{\Gamma}\right)\\
= &\begin{pmatrix} \Gamma_1'\Gamma_1 & \Gamma_1'\Gamma_2 - \Gamma_1'\Gamma_1\Gamma_2'\Omega' \\ \Gamma_1'\Gamma_2 - \Gamma_1'\Gamma_1\Gamma_2'\Omega' & \Gamma_2'\Gamma_2 - 2\Gamma_2'\Gamma_1\Gamma_2'\Omega' + \Omega\Gamma_2\Gamma_1'\Gamma_1\Gamma_2'\Omega' \end{pmatrix}\\
= &\begin{pmatrix} \Gamma_1'\Gamma_1 & \Gamma_1'\Gamma_2 - \Gamma_1'\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} \\ \Gamma_1'\Gamma_2 - \Gamma_1'\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} & \Gamma_2'\Gamma_2 - 2\Gamma_2'\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} + \left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} \end{pmatrix}\\
= &\begin{pmatrix} \Gamma_1'\Gamma_1 & \Gamma_1'\Gamma_2 - \Gamma_2'\Gamma_1 \\ \Gamma_1'\Gamma_2 - \Gamma_2'\Gamma_1 & \Gamma_2'\Gamma_2 - 4\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} + 2\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} \end{pmatrix}\\
= &\begin{pmatrix} \Gamma_1'\Gamma_1 & 0 \\ 0 & \Gamma_2'\Gamma_2 - 2\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} \end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&\left(\mathring{\Gamma}'\mathring{\Gamma}\right)^{-1}\\
={}& \frac{1}{\Gamma_1'\Gamma_1\left(\Gamma_2'\Gamma_2-2\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1}\right)}\begin{pmatrix}\Gamma_2'\Gamma_2-2\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} & 0\\ 0 & \Gamma_1'\Gamma_1\end{pmatrix}\\
&= \frac{1}{\Gamma_1'\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_2'\Gamma_1}\begin{pmatrix}\Gamma_2'\Gamma_2-2\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} & 0\\ 0 & \Gamma_1'\Gamma_1\end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&\mathring{\Gamma}\left(\mathring{\Gamma}'\mathring{\Gamma}\right)^{-1}\\
={}& \frac{1}{\Gamma_1'\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_2'\Gamma_1}\begin{pmatrix}\Gamma_1 & \Gamma_2-\Gamma_1\Gamma_2'\Omega'\end{pmatrix}\begin{pmatrix}\Gamma_2'\Gamma_2-2\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} & 0\\ 0 & \Gamma_1'\Gamma_1\end{pmatrix}\\
={}& \frac{1}{\Gamma_1'\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_2'\Gamma_1}\begin{pmatrix}\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} & \Gamma_2\Gamma_1'\Gamma_1-\Gamma_1\Gamma_2'\Omega'\Gamma_1'\Gamma_1\end{pmatrix}\\
={}& \frac{1}{\Gamma_1'\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_2'\Gamma_1}\begin{pmatrix}\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} & \Gamma_2\Gamma_1'\Gamma_1-\Gamma_1\Gamma_2'\Gamma_1\end{pmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&\mathring{\Gamma}\left(\mathring{\Gamma}'\mathring{\Gamma}\right)^{-1}\mathring{\Gamma}'\\
={}& \kappa\begin{pmatrix}\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1} & \Gamma_2\Gamma_1'\Gamma_1-\Gamma_1\Gamma_2'\Gamma_1\end{pmatrix}\begin{pmatrix}\Gamma_1'\\ \Gamma_2'-\Omega\Gamma_2\Gamma_1'\end{pmatrix}\\
={}& \kappa(\left[\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1}\right]\Gamma_1'\\
&+\left[\Gamma_2\Gamma_1'\Gamma_1-\Gamma_1\Gamma_2'\Gamma_1\right]\left[\Gamma_2'-\Omega\Gamma_2\Gamma_1'\right])\\
={}& \kappa\Gamma_1\Gamma_2'\Gamma_2\Gamma_1'-2\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'+\left[\Gamma_2\Gamma_1'\Gamma_1\Gamma_2'-\Gamma_1\Gamma_2'\Gamma_1\Gamma_2'\right]\\
&+\left[-\Gamma_2\Gamma_1'\Gamma_1\Omega\Gamma_2\Gamma_1'+\Gamma_1\Gamma_2'\Gamma_1\Omega\Gamma_2\Gamma_1'\right]\\
={}& \kappa(\Gamma_1\Gamma_2'\Gamma_2\Gamma_1'-2\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'+\left[\Gamma_2\Gamma_1'\Gamma_1\Gamma_2'-\Gamma_1\Gamma_2'\Gamma_1\Gamma_2'\right]\\
&+\left[-\Gamma_2\Gamma_1'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'+\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\right])\\
={}& \kappa(\Gamma_1\Gamma_2'\Gamma_2\Gamma_1'-2\Gamma_1\Gamma_2'\Gamma_1\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'+\Gamma_2\Gamma_1'\Gamma_1\Gamma_2'-\Gamma_1\Gamma_2'\Gamma_1\Gamma_2'\\
&+-\Gamma_2\Gamma_1'\Gamma_2\Gamma_1'+\Gamma_1\Gamma_1'\Gamma_2\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1')\\
={}& \kappa(\Gamma_1\Gamma_2'\Gamma_2\Gamma_1'-2\Gamma_1\Gamma_1'\left[\Gamma_1'\Gamma_2\left(\Gamma_1'\Gamma_1\right)^{-1}\right]+\Gamma_2\Gamma_2'\left[\Gamma_1'\Gamma_1\right]-\Gamma_1\Gamma_2'\left[\Gamma_2'\Gamma_1\right]\\
&+-\Gamma_2\Gamma_1'\left[\Gamma_1'\Gamma_2\right]+2\Gamma_1\Gamma_1'\left[\Gamma_1'\Gamma_2\left(\Gamma_1'\Gamma_1\right)^{-1}\right])\\
={}& \kappa\left(\Gamma_1\Gamma_1'\left[\Gamma_2'\Gamma_2\right]+\Gamma_2\Gamma_2'\left[\Gamma_1'\Gamma_1\right]-\Gamma_1\Gamma_2'\left[\Gamma_2'\Gamma_1\right]+-\Gamma_2\Gamma_1'\left[\Gamma_2'\Gamma_1\right]\right)\\
={}& \kappa\left(\Gamma_1\Gamma_2'\Gamma_2\Gamma_1'+\Gamma_2\Gamma_1'\Gamma_1\Gamma_2'-\Gamma_1\Gamma_2'\Gamma_1\Gamma_2'+-\Gamma_2\Gamma_2'\Gamma_1\Gamma_1'\right)
\end{aligned}
$$

where $\kappa=\frac{1}{\Gamma_1'\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_2'\Gamma_1}$. So $\mathring{\Gamma}\left(\mathring{\Gamma}'\mathring{\Gamma}\right)^{-1}\mathring{\Gamma}'=\frac{1}{\Gamma_1'\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_2'\Gamma_1}\left(\Gamma_1\Gamma_2'\Gamma_2\Gamma_1'+\Gamma_2\Gamma_1'\Gamma_1\Gamma_2'-\Gamma_1\Gamma_1'\Gamma_2\Gamma_2'+-\Gamma_2\Gamma_1'\Gamma_2\Gamma_1'\right)$
Now compute $\Gamma\left(\Gamma'\Gamma\right)^{-1}\Gamma'$:

$$
\begin{aligned}
& \Gamma\left(\Gamma'\Gamma\right)^{-1}\Gamma' \\
= \quad & \kappa\begin{pmatrix}\Gamma_1 & \Gamma_2\end{pmatrix}\begin{pmatrix}\Gamma_2'\Gamma_2 & -\Gamma_1'\Gamma_2 \\ -\Gamma_1'\Gamma_2 & \Gamma_1'\Gamma_1\end{pmatrix}\begin{pmatrix}\Gamma_1' \\ \Gamma_2'\end{pmatrix} \\
= \quad & \kappa\begin{pmatrix}\Gamma_1\Gamma_2'\Gamma_2-\Gamma_2\Gamma_1'\Gamma_2 & -\Gamma_1\Gamma_1'\Gamma_2+\Gamma_2\Gamma_1'\Gamma_1\end{pmatrix}\begin{pmatrix}\Gamma_1' \\ \Gamma_2'\end{pmatrix} \\
= \quad & \kappa\left(\left[\Gamma_1\Gamma_2'\Gamma_2-\Gamma_2\Gamma_1'\Gamma_2\right]\Gamma_1'+\left[-\Gamma_1\Gamma_1'\Gamma_2+\Gamma_2\Gamma_1'\Gamma_1\right]\Gamma_2'\right) \\
= \quad & \kappa\left(\Gamma_1\Gamma_2'\Gamma_2\Gamma_1'-\Gamma_2\Gamma_1'\Gamma_2\Gamma_1'-\Gamma_1\Gamma_1'\Gamma_2\Gamma_2'+\Gamma_2\Gamma_1'\Gamma_1\Gamma_2'\right)
\end{aligned}
$$

This shows that $\overset{\circ}{\Gamma}\left(\overset{\circ}{\Gamma}'\overset{\circ}{\Gamma}\right)^{-1}\overset{\circ}{\Gamma}' = \Gamma\left(\Gamma'\Gamma\right)^{-1}\Gamma'$.

This fact allows us to solve for $\overset{\circ}{B}$ in terms of B:

$$
\begin{aligned}
& & & \operatorname{plim}\overset{\circ}{B} \\
[1] & & = \quad & \left(\overset{\circ}{\theta}'\overset{\circ}{\theta}\right)^{-1}\overset{\circ}{\theta}'\theta B \\
[2] & & = \quad & \left(\left(\overset{\circ}{\Gamma}'\overset{\circ}{\Gamma}\right)^{-1}\overset{\circ}{\Gamma}'MM'\overset{\circ}{\Gamma}\left(\overset{\circ}{\Gamma}'\overset{\circ}{\Gamma}\right)^{-1}\right)^{-1}\left(\overset{\circ}{\Gamma}'\overset{\circ}{\Gamma}\right)^{-1}\overset{\circ}{\Gamma}'MM'\Gamma\left(\Gamma'\Gamma\right)^{-1}B \\
[3] & & = \quad & \left(\left(\overset{\circ}{\Gamma}'\overset{\circ}{\Gamma}\right)^{-1}\overset{\circ}{\Gamma}'MM'\overset{\circ}{\Gamma}\left(\overset{\circ}{\Gamma}'\overset{\circ}{\Gamma}\right)^{-1}\right)^{-1}\left(\overset{\circ}{\Gamma}'\overset{\circ}{\Gamma}\right)^{-1}\overset{\circ}{\Gamma}'MM'\overset{\circ}{\Gamma}\left(\overset{\circ}{\Gamma}'\overset{\circ}{\Gamma}\right)^{-1}\overset{\circ}{\Gamma}'\Gamma\left(\Gamma'\Gamma\right)^{-1}B \\
[4] & & = \quad & \overset{\circ}{\Gamma}'\Gamma\left(\Gamma'\Gamma\right)^{-1}B \\
[5] & & = \quad & \begin{pmatrix}\Gamma_1' \\ \Gamma_2'-\Omega\Gamma_2\Gamma_1'\end{pmatrix}\begin{pmatrix}\Gamma_1 & \Gamma_2\end{pmatrix}\begin{pmatrix}\Gamma_1'\Gamma_1 & \Gamma_1'\Gamma_2 \\ \Gamma_1'\Gamma_2 & \Gamma_2'\Gamma_2\end{pmatrix}^{-1}B \\
[6] & & = \quad & \begin{pmatrix}\Gamma_1'\Gamma_1 & \Gamma_1'\Gamma_2 \\ \Gamma_2'\Gamma_1-\Omega\Gamma_2\Gamma_1'\Gamma_1 & \Gamma_2'\Gamma_2-\Omega\Gamma_2\Gamma_1'\Gamma_2\end{pmatrix}\begin{pmatrix}\Gamma_1'\Gamma_1 & \Gamma_1'\Gamma_2 \\ \Gamma_1'\Gamma_2 & \Gamma_2'\Gamma_2\end{pmatrix}^{-1}B \\
[7] & & = \quad & \kappa\begin{pmatrix}\Gamma_1'\Gamma_1 & \Gamma_1'\Gamma_2 \\ \Gamma_2'\Gamma_1-\Omega\Gamma_2\Gamma_1'\Gamma_1 & \Gamma_2'\Gamma_2-\Omega\Gamma_2\Gamma_1'\Gamma_2\end{pmatrix}\begin{pmatrix}\Gamma_2'\Gamma_2 & -\Gamma_1'\Gamma_2 \\ -\Gamma_1'\Gamma_2 & \Gamma_1'\Gamma_1\end{pmatrix}B \\
[8] & & = \quad & \kappa\begin{pmatrix}\Gamma_1'\Gamma_1\Gamma_2'\Gamma_2-\Gamma_1'\Gamma_2\Gamma_1'\Gamma_2 & -\Gamma_1'\Gamma_1\Gamma_1'\Gamma_2+\Gamma_1'\Gamma_2\Gamma_1'\Gamma_1 \\ \Gamma_2'\Gamma_1\Gamma_2'\Gamma_2-\Omega\Gamma_2\Gamma_1'\Gamma_1\Gamma_2'\Gamma_2-\Gamma_2'\Gamma_1\Gamma_1'\Gamma_2+\Omega\Gamma_2\Gamma_1'\Gamma_1\Gamma_1'\Gamma_2 & -\Gamma_2'\Gamma_1\Gamma_1'\Gamma_2+\Omega\Gamma_2\Gamma_1'\Gamma_1\Gamma_1'\Gamma_2+\Gamma_2'\Gamma_2\Gamma_1'\Gamma_1-\Omega\Gamma_2\Gamma_1'\Gamma_2\Gamma_1'\Gamma_1\end{pmatrix}B \\
[9] & & = \quad & \kappa\begin{pmatrix}\Gamma_1'\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_1'\Gamma_2 & 0 \\ \Gamma_2'\Gamma_1\Gamma_2'\Gamma_2-\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_2'\Gamma_1+\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\Gamma_1\Gamma_1'\Gamma_2 & \chi\end{pmatrix}B \\
[10] & & = \quad & \kappa\begin{pmatrix}\Gamma_1'\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_1'\Gamma_2 & 0 \\ 0 & \Gamma_2'\Gamma_2\Gamma_1'\Gamma_1-2\Gamma_2'\Gamma_1\end{pmatrix}B \\
[11] & & = \quad & B
\end{aligned}
$$

where $\kappa \equiv \frac{1}{\Gamma_1'\Gamma_1\Gamma_2'\Gamma_2-2\Gamma_1'\Gamma_2}$ and $\chi \equiv -2\Gamma_2'\Gamma_1+\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\Gamma_1\Gamma_1'\Gamma_2+\Gamma_2'\Gamma_2\Gamma_1'\Gamma_1-\left(\Gamma_1'\Gamma_1\right)^{-1}\Gamma_1'\Gamma_2\Gamma_1'\Gamma_2\Gamma_1'\Gamma_1$

Step [1] uses OLS, where $\overset{\circ}{\theta}\perp\varepsilon$. Step [2] substitutes in for $\overset{\circ}{\theta}$. Step [3] uses the fact that $\Gamma\left(\Gamma'\Gamma\right)^{-1} = \Gamma\left(\Gamma'\Gamma\right)^{-1}\Gamma'\Gamma\left(\Gamma'\Gamma\right)^{-1} = \overset{\circ}{\Gamma}\left(\overset{\circ}{\Gamma}'\overset{\circ}{\Gamma}\right)^{-1}\overset{\circ}{\Gamma}'\Gamma\left(\Gamma'\Gamma\right)^{-1}$. Step [4] cancels $\left(\overset{\circ}{\Gamma}'\overset{\circ}{\Gamma}\right)^{-1}\overset{\circ}{\Gamma}'MM'\overset{\circ}{\Gamma}\left(\overset{\circ}{\Gamma}'\overset{\circ}{\Gamma}\right)^{-1}$ with its inverse. Step [5] substitutes in the definition of $\overset{\circ}{\Gamma}$and Γ. Step [6] multiplies vectors. Step [7] calculates the inverse of a 2x2 matrix. Step [8] multiplies the two matrices. Step [9] cancels terms and substitutes in for the definition of Ω. Step [10] cancels terms. Step [11] divides the scalars inside the matrix by the scalar multiplying the matrix, yielding the identity matri

Claim 4: Error in the factor estimates due to the uniqueness and residual factors will grow arbitrarily small as the number of measures grows if (1) the variance of the loadings for additional measurements does not approach zero or infinity, and the expected value of loadings for additional measurements do not approach infinity or negative infinity, and (2) if either (A) the loadings for the residual factors are orthogonal to the loadings Γ_1 and Γ_2 or (B) the contribution of the residual factors to the covariance between measures grows arbitrarily small.

Condition [1] stipulates that the additional measures added remain relevant but not too relevant, which allows a law of large numbers to be applied. Condition [A] stipulates that changes in the measures caused by changes in residual factors do not affect forecasts of θ_{AK} or θ_2. Condition [B] essentially stipulates that the residual factors account for a trivial portion of the covariance of measures.

Proof sketch: The measurement system residual causes error in forecasts of the factors θ_{AK} and θ_2. The mean of the uniquenesses across measurements will tend towards zero as the number of measurements grows large by the law of large numbers. If the (absolute value) of weights on these measurements do not shrink to zero or approach infinity in expectation as the number of measurements grows arbitrarily large (e.g., all weights are drawn from a normal distribution), then the measurement error on the factors caused by the uniquenesses will shrink to zero. However, the mean of the residual factors is unaffected by the number of measurements. If the weights for these residual factors are correlated with Γ_1 and Γ_2, then the residual factors generate a correlation in measurement error for θ_{AK} and θ_2 across measurements, and this does not vanish asymptotically. But if instead the factor loadings for the residual factors are orthogonal to the factor loadings for $\theta_{AK,i}$ and $\theta_{2,i}$, then this error correlated across measurements does not affect forecasts of θ_{AK} and θ_2.

Claim 5: The variance-covariance structure of the factors θ_{AK} and θ_2 is identified if the loadings Γ_1 and Γ_2 are known and are orthogonal to the loadings for the residual factors; if the variance-covariance structure of θ_{AK} and θ_2 is identified and the loadings Γ_1 and Γ_2 are orthogonal to the loadings for the residual factors, then the variance of the measurement error on θ_{AK} and θ_2 can be estimated (such as when the number of measurements is finite).

Proof: Define $\gamma_{j,k}$ as the factor loading for factor j of measurement k. That is, $\gamma_{1,k}$ would be the row k parameter in Γ_1.

The covariance between measures is a function of the factor loadings Γ_1 and Γ_2, the variance and covariance terms of θ_{AK} and θ_2, and the covariance of the measurement residuals caused by residual factors:

$$\begin{aligned}\mathbb{C}\left[M_{i,k_1}, M_{i,k_2}\right] &= \gamma_{1,k_1}\gamma_{1,k_2}\mathbb{V}\left[\theta_{AK}\right] + \gamma_{2,k_1}\gamma_{2,k_2}\mathbb{V}\left[\theta_2\right] \\ &\quad + \left(\gamma_{1,k_1}\gamma_{2,k_2} + \gamma_{1,k_2}\gamma_{2,k_1}\right)\mathbb{C}\left[\theta_{AK}, \theta_2\right] + f_{k_1,k_2}\end{aligned}$$

with the γ's already identified, a regression of the measure covariances on the above function of the γ's will identify the variance and covariance terms of θ_{AK} and θ_2 if the OLS assumptions are satisfied – if f_{k_1,k_2} are orthogonal to the γ's. Given the structure of the problem, the residual factors generate the residual covariance terms f_{k_1,k_2}. If the factor loadings for the residual factors are orthogonal to the factor loadings Γ_1 and Γ_2, then f_{k_1,k_2} is orthogonal to the γ's.

D) Alternative factor analytic methods for identifying affective capital with epigenetic markers

I have considered two alternative methods for identifying the latent factor θ_{AK}. These alternative methods are more standard but depend upon implausible assumptions. The first alternative method uses dedicated measures while the second alternative method uses the covariance between methylation measures and depression inventory scores, which are affected by the same chronic stressors that affect the methylation levels.

1) Dedicated Measures

Experimental murine evidence finds that chronic stress changes methylation levels at particular sites. In humans, methylation levels at analogous sites associate with a past experience of severe stressors. Let Ω be the set of sites at which stressors are known to alter methylation. Let M_i^k be the level of methylation at site k for person i.

Assume that methylation at sites in Ωload on only affective capital (that is, the Ω sites are dedicated measures).

$$M_i^k = \theta_i^{AK}\gamma_k^{AK} + e_i^k \; \forall k \in \Omega$$

where γ_k^{AK} is the loading on affective capital for measure k. This assumption implies that the error terms e_i^k are uniquenesses $e_i^k \perp e_i^{k'}$, $k \neq k'$. Assuming θ_i^k has unit variance, three measures are sufficient to identify γ_k up to sign and to identify the variance of the uniquenesses.

While this approach identifies the loadings for affective capital under assumption the dedicated measures assumption, the assumption is implausible. The sites in Ω are associated with genes involved in many pathways, and the methylation at these sites may be affected by aspects of the environment unrelated to chronic stress and unimportant for the effects of chronic stsress, generating correlation in the uniquenesses and thus error in the factor forecasts (and potentially non-classical error).

One might be tempted to include another factor $\theta_i^{\neg AK}$allowing for measures in Ω to be correlated with measures not in Ω:

$$M_i^k = \begin{cases} \theta_i^{AK}\gamma_k^{AK} + \theta_i^{\neg AK}\gamma_k^{\neg AK} + e_i^k & \forall k \in \Omega \\ \theta_i^{\neg AK}\gamma_k^{\neg AK} + e_i^k & \forall k \notin \Omega \end{cases}$$

This approach relaxes the assumption that sites in $\Omega$are dedicated measures of θ^{AK}, but does so at the expense of making the stronger and less plausible assumption that the other measures are dedicated measures of the alternative factor $\theta^{\neg AK}$. That is, the set of sites Ωmust include all sites whose methylation is affected by stressors, and only one other factor generates correlation in methylation between sites in Ω and sites outside of Ω.

2) Covariance between measures of affect and epigenetics

Because the stressors affecting epigenetics also cause depression, it is reasonable to assume that depression inventory items correlate with the methylation changes caused by stressors.

Assume a linear factor model where only the affective capital latent variable θ^{AK} has non-zero loadings in both the methylation measures and depression inventory scores. One additional factor may generate correlation within depression inventory items and several different factors may generate correlation within methylation measures without impairing identification. Factors are orthogonal to each other and of unit variance.

Formally, the measurement system can be expressed as

$$\begin{aligned} M_i^k &= \gamma_k^{AK}\theta^{AK} + \theta^{RC}\Gamma_k^{RC} + \varepsilon^M \\ D_i^l &= \gamma_l^D\theta^{AK} + \gamma_l^{CE}\theta^{CE} + \varepsilon^D \end{aligned}$$

where γ_k^{AK} is the scalar loading for affective capital on methylation measure k, θ^{AK} is the affective capital latent variable, θ_k^{RC} is a 'residual correlation' vector of factors, Γ_k^{RC} is a vector of loadings, ε^M and ε^D are uniquenesses, M_i^k is methylation measure k, D_i^l is depression inventory item l, and θ^{CE} is a 'correlated error' factor. While θ^{CE} may be literally interpreted as a measurement error that is correlated across depression inventory items, it captures any component of depression inventory scores unrelated to epigenetics.

Because θ_k^{RC} is uncorrelated with θ^{AK} and θ^{CE}, the factor structure is similar to a triangular loading pattern, which is identified up to sign if there are enough measures with different loadings. Allowing additional factors to generate correlations among the methylation measures does not thwart identification. To prove this, note that factor loadings for the depression inventory items are identified up to a rotation matrix without using the methylation measures. Thus, the factor estimates will be a rotation of the true factors plus some measurement error. In turn, the estimated coefficients from a regression of methylation measures on these factor estimates (with the correction for measurement error) will be a rotation of the coefficients from a regression of methylation measures on the true factors. Because the methylation measures are assumed to load on only one of the depression inventory factors, the correct rotation matrix will yield regression coefficients of zero for one of the factors. This rotation matrix comes from the set non-trivial of 2x2 rotation matrices and is characterized by one parameter; a unique value of this rotation parameter will yield regression coefficients of zero for one of the factors. Thus, the rotation is identified.

Although this model is identified, it depends on the assumption that $\theta^{RC} \perp \theta^{AK}$, which is not likely to be true. Specifically, depression correlates with socioeconomic status, diet, exercise, and other things that may directly or indirectly affect epigenetics. As a consequence of this assumption being violated, the correct loadings for the methylation measures will not be identified. This becomes problematic when θ^{AK} is used as an outcome in subsequent analysis. That is, an instrument may shift parental inputs into the child's affective capital. If these inputs do not also shift θ^{RC} in proportion to the population covariance between θ^{RC} and θ^{AK}, the instrumental variable regression will yield incorrect estimates of the effect of these parental inputs on the child's affective capital, and the structural model will understate the role of parenting inputs in the population distribution of affective capital.

APPENDIX G

G - REVIEW OF HUMAN STUDIES ASSOCIATING METHYLATION TO DEPRESSION AND RELATED PHENOTYPES

Research exploring associations between human outcomes and methylation falls into two broad camps: Methylome-Wide Association Studies (MWAS) and candidate methylation studies. In an MWAS, researchers perform univariate regressions of some outcome or exposure on a large number of methylation probes, reporting significant probes after controlling for the family-wise error rate (most often using the Benjamini-Hoxburg correction). Candidate studies focus research on sites that prior biological research indicates may be important for the relevant outcome. Below I report relevant results from both strands of literature.

Recent MWAS studies have recovered methylome-wide significant probes for a variety of stress and depression associated phenotypes. Several prominent examples, which I revisit in the empirical section, include:

- **Depression–** Nemoda et al. (2015) predict history of maternal depression using the infinium 450k on 44 neonatal cord blood samples, 62 post-mortem adult offspring hippocampal slices, and 38 antenatal maternal blood samples (which are acutely small sample sizes). They find 145 methylome-wide significant probes in the cord blood, 2520 in the hippocampal slices, and none in the maternal blood.

- **Cortisol abnormalities–** Depression often features alterations in the diurnal variation of cortisol as well as increases in average cortisol levels, so methylation changes that associate with cortisol abnormalities are also likely to associate with depression.

Indeed, many mouse models find that psychosocial stress causes aberrant methylation near the NR3C1 mineralocorticoid (cortisol) receptor gene in a variety of tissues.

  - Glad et al. (2017) predict free urinary cortisol and CRH stimulated cortisol response in remitted Cushing's disease patients (N=48 women) and controls (N=16 women) using infinium 450k methylation data. Several of their epigenome-wide significant probes are near or on the NR3C1 mineralocorticoid receptor gene.
  - Houtepen et al. (2016) perform a methylome-wide association study on stress reactivity in sample of 85 adults. Several probes were methylome-wide significant, including one for the KITLG gene, which had previously been implicated in the stress response. Childhood trauma predicted stress reactivity and predicted methylation of the KITLG gene, but adult trauma did not.

- **Inflammation** – Depression associates with greater systemic inflammation. Bränn et al. (2017) predict post-partum depression using a large set of inflammatory biomarkers in prepartum mothers. For genes corresponding to the Bonferroni significant inflammatory biomarkers, the authors investigate methylation on 29 associated probes. Only two are significant at the 95 percent level, but none after a Bonferroni correction.

- **Abuse** – Cecil et al. (2016) perform a methylome-wide association study in a sample of 124 high-risk inner-city youth on three dimensions of self-reported abuse: sexual abuse, physical abuse, and physical neglect. They find a large set of methylome-wide significant probes for each category of abuse. There is no overlap in methylome-wide significant probes for each category of abuse.

Several facets of these studies deserve mention. First, all of these studies suffer from small sample sizes, and they are underpowered to reliably detect weaker associations. Second, no single probe is methylome-wide significant in more than one of the above studies. This might be surprising in light of the similar exposures/phenotypes considered. These first two facets suggest that depression is likely to feature modest associations with a large number of probes, but strong associations with no probes. Third, stronger relationships exist between methylation and depression when brain tissue is sampled instead of blood. This suggests that methylation data from whole blood performs worse at predicting depression-associated biological stock than methylation data from other tissue samples.

Candidate probe studies focus attention on probes near genes that have been previously implicated in the relevant phenotype: for stress and depression, probes near the NR3C1 mineralocorticoid receptor gene (which is involved in stress hormone signaling) have garnered the most attention. For example, Oberlander et al. (2008) investigate the relationship between maternal depression, infant methylation at a site near the NR3C1 gene in blood samples, and infant stress reactivity, finding that increased methylation at the site associated with both higher maternal depression and increased infant stress reactivity. See also Conradt et al. (2016).

I revisit these studies in my data analysis with two aims. First, I verify the extent to which methylation for the above probes are indeed associated with depression. This verification can be interpreted as testing the joint hypothesis that the above probes are not associated with depression and that the ALSPAC sample is of sufficient quality to uncover associations between depression and methylation data. Second, I seek to utilize results of these studies as a prior, allowing me to restrict attention to probes that are putatively less noisy. This can permit a more accurate forecast of depression using methylation data.

APPENDIX H

H - EPIGENETIC ANALYSIS, CONTINUED

I compute bivariate regressions of all remaining methylation probes for different measures of depression over different timepoints. For the mother, I utilized depression inventory factors in close proximity to the period in which blood was drawn as well as an average of non-missing depression inventories over all periods. The average of depression scores may have two advantages over the time-point specific depression measure. First, the average will better proxy for long-lived stocks than the one-time depression scores, and so may do a better job predicting persistent epigenetic modifications. Second, such an average imputes missing values, thus increasing power.

I apply a t-test for on the null hypothesis that the probe M-value is not associated with the depression measure. For each sample and depression measure set, I store the p-values for each test.

7.C.i Second-stage regression: annotation data are informative

For the p-values collected on bivariate regression of time-point 15 mother methylation on the time-point 12 maternal depression factor, I regress the stored p-values on annotation dummies (see Table H.1). Negative coefficients from this second-stage regression indicate that methylation for correspondingly annotated probes are more likely to be significantly associated with maternal depression (in sample) after conditioning on other annotation dummies.

Annotation data fall into four categories: *reference gene group*, *differentially methylation region (DMR)*, *enhancer*, and *regulatory feature group*. Within each category, dummies are mutually exclusive. Within each category, dummies may sum to zero, indicating that the probe has not been classified as belonging to any of the subcategories. *Reference*

Regression of P-values on Probe Annotation Dummies

p-values derived from bivariate regressions of methylation beta values on depression inventory scores

Category	Annotation Dummy	Est.	Std. Err.	t-stat	p-value
	(Intercept)	0.41	0.004	112.0	0.000
Refgene Group	TSS200	-0.05	0.007	-7.2	0.000
Refgene Group	TSS1500	0.03	0.008	3.3	0.001
Refgene Group	`5'UTR`	-0.03	0.012	-2.6	0.010
Refgene Group	Body	0.01	0.006	2.2	0.026
Refgene Group	`1stExon`	-0.02	0.014	-1.7	0.098
Refgene Group	`3'UTR`	0.06	0.020	3.1	0.002
DMR	DMR	-0.03	0.009	-3.2	0.002
DMR	RDMR	0.06	0.016	3.4	0.001
DMR	CDMR	0.03	0.021	1.5	0.126
Enhancer	Enhancer	0.04	0.006	6.4	0.000
Regulatory Feature Group	Promoter_Associated_Cell_type_specific	-0.16	0.015	-10.6	0.000
Regulatory Feature Group	Unclassified_Cell_type_specific	-0.04	0.008	-5.4	0.000
Regulatory Feature Group	Promoter_Associated	-0.23	0.005	-45.8	0.000
Regulatory Feature Group	NonGene_Associated	-0.23	0.030	-7.8	0.000
Regulatory Feature Group	Unclassified	-0.13	0.008	-16.9	0.000
Regulatory Feature Group	Gene_Associated	0.01	0.049	0.3	0.773
Regulatory Feature Group	Gene_Associated_Cell_type_specific	0.06	0.038	1.6	0.109
Regulatory Feature Group	NonGene_Associated_Cell_type_specific	0.09	0.146	0.6	0.545

These results are a second stage regression. In the first stage, the maternal Edinburgh Post-Natal Depression Survey factor from survey r (time-point 11) was separately regressed on the beta values for each of the approximately 500,000 methylation probes assayed at time point 14. The p-values from these regressions were assembled into a table matched with probe indicators. The p-values were not adjusted for multiple hypothesis testing. In the second stage, the p-values were regressed on a set of dummy variables indicating different functional categories of the DNA regions measured by each probe.

Data on functional categories came from an annotation table downloaded from the company which developed the methylation assay. These dummies fall into several different categories and were determined with different methodolgies. The Refgene Group category came from analyzing the DNA sequence, and determining if the region of DNA measured by the probe fell into certain standard sequences. For example, the transcription start site is a particular sequence of DNA nucleotides that always indicate the begining of transcription of a gene. "TSS200" indicats that the DNA measured by the methylation probe was within 200 base pairs (distance measured by nucleotide count) ahead of the transcription start site. Similarly, the "3'UTR" is a segment at the beginning of a gene which is transcribed into mRNA but is not translated into protein. The DMR category refers to probes at which DNA methylation has been experimentally determined to vary, usual when holding genetics constant by comparing twins. For example, the CMR dummy indicates that this genes is differentially methylated when comparing individuals with or without certain cancers. Unsurprisingly, this particular dummy is not significant. The Enhancer dummy refers to a region of DNA that is known to increase expression of DNA when bound by certain factors. [PROVIDE MORE DESCRIPTION]

In short, these regression results indicate that what epigenetically distinguishes depressed individuals from non-depressed individuals (for whatever reason -- sampling variation or some causal relationship) is more likely to manifest in some probe categories than others. In this sense, these results indicate which probe categories are likely to contain more signal on underlying epigenetic stocks. However, these results do not indicate that changes in methylation at particular probe categories are indicative of depression. For example, suppose that methylation patterns across probes are generated by a smaller dimensional set of latent epigenetic factors. Mean values of these factors may vary across depressed and non-depressed individuals due to: sampling variation, epigenetic stocks causing depression, epigenetic stocks caused by depression, or epigenetic stocks mutually caused by what causes depression. The fact that some probe categories are more likely to have low p-values in the first stage regression indicates that these probe categories are more likely to load on various epigenetic stocks than other categories are. In light of this interpretation, restricting subsequent analysis to the regulatory feature groups **Promoter Associated Cell Type Specific, Promoter Associated, NonGene Associated,** and **Unclassified,** and the RefGene group **TSS200** probes should improve the signal-to-noise ratio.

Table H.1: Classifying predictive content of methylation variables in terms of annotation dummies *gene group* dummies indicate the location of the DNA measured by the probe with respect to a neighboring gene. So, for example, the measured DNA might include a region of DNA that codes for protein ('Body') or might be in a non-coding region immediately upstream of the gene ('TSS200'). *Differential methylation region* dummies indicate that prior research has found that methylation on this region of DNA varies across subpopulations. For example, 'CDMR' indicates that other researchers have found aberrant methylation at this probe in those with cancer. *Enhancer* dummies indicate that the region covered by the probe increases expression of a gene when bound by appropriate regulatory proteins. *Regulatory feature group* dummies indicate roles in transcription for the associated DNA.

Of note, this second-stage regression indicates that probes annotated as 'promoter associated' feature much lower p-values in the first-stage regression. This is unsurprising, as much of the biological research on methylation has explored how changes in methylation on promoters alters transcription of genes in response to environment. About 36,000 probes are thus annotated (for comparison, there are about 20,000 genes in the human genome, so about one and a half probes are indicated as 'promoter associated' for each gene).

These results indicate which probes are more likely to contain information relevant for distinguishing two subpopulations.

However, these results do not imply that altered methylation on these probes is associated with depression. In particular, sampling variation can generate correlation between depression and epigenetic stocks even if these epigenetic stocks are not correlated with depression in the population. In turn, annotation data may strongly predict the magnitude of coefficients for correlated stocks in the measurement system's methylation pattern. Even if the in-sample non-zero correlation between depression and these unobserved latent epigenetic stocks were not significantly different from zero, these correlated stocks can generate extremely significant coefficients on annotation dummies in the second-stage regression. Thus, the second stage regression does not provide much information about the relationship between methylation data and depression per se. Rather, this second stage regression indicates which probes are more likely to hold predictive content generally.

7.C.ii Evidence of depression-associated methylation in different probe sets

In this section, I search for methylome-wide significant probes from the first-stage regression. In addition, I restrict analysis to probe sets which prior research indicates may be of particular importance. These include: “Nemoda” probes, which were found in children to be methylome wide significant in predicting maternal depression; “Cecil” probes, which were found in adults to be methylome-wide significant in predicting adult self-reports of childhood abuse; “Glad” probes, which were found to be methylome-wide significant in predicting aberrant cortisol levels or cortisol responses; “Houtepan” probes, which were found to be associated with aberrant stress response; “Savitz” probes, which point to genes that are aberrantly expressed in the monocytes of those with depression; and the collection of all these probes together (note that, strictly speaking, the only replication I perform is for Nemoda probes using cord-blood samples and maternal depression measures). In addition, I include the set of probes near the NR3C1 gene, as prior biological research implicates methylation near this gene in the etiology of depression-like behavior in mice.

Table 7.B.i in the supplementary appndix displays results (See QQ-plots in the supplementary appendix for visualization of the distribution of p-values with respect to the null of a uniform distribution). For each sample/depression measure pair, I report five statistics: the number of Bonferroni significant probes at a false-discovery rate of 10%, the number of Benjamini-Hoxburg significant probes at a false-discovery rate of 10%, the mean p-value of probes, the p-value of the one-sided t-test of the null hypothesis that the mean first-stage p-value is greater 0.5, and the p-value of the Kolgomorov-Smirnov test that the first-stage p-values are distributed uniformly. Bonferroni multiple hypothesis testing is overly conser-

vative. Benjamini-Hoxburg multiple-hypothesis testing is less conservative, but assumes that the p-values across tests are uncorrelated, which is not likely to be a valid assumption. The mean p-value from the first-stage regression gives some indication of the number of and extent to which probes are predictive of the depression measure; if methylation at none of the probes were associated with depression in population, then we would expect a mean of first-stage p-values near 0.5.

For only a few of the sample/depression pairs are there any methylome-wide significant probes. This is true for both Bonferroni multiple-hypothesis testing and the less conservative Benjamini-Hoxburg multiple hypothesis testing. And this is true whether considering the full set of probes or considering a subset of probes that prior research indicates are more likely to be relevant. There are some notable exceptions to this rule. First, mother's timepoint 15 sample using her timepoint 12 depression factor yields 3771 methylome-wide significant probes when using the Benjamini-Hoxburg correction. Second, the child's cord blood sample using his timepoint 16 depression factor yields 4646 methylome-wide significant probes. And the child's timepoint 15 sample using his timepoint 16 depression factor yields 11035 methylome-wide significant probes. While it would be tempting to interpret these results as indicative of a large depression-related epigenetic signal, it is important to recognize that these results are not robust to how depression is measured.

The relative dearth of methylome-wide significant probes is puzzling when contrasted with the often low mean of p-values estimated in the first-stage regression. For example, using p-values from the first-stage regression of mother's time-point 15 methylation data on time-point 12 depression inventory, the average p-value is 0.14. About 40% of these probes are significant under single-hypothesis testing at a 90% significance level. However, none of these probes survive multiple-hypothesis testing at a 90% false-discovery rate.

It is also surprising that so few of the probe sets motivated by prior research yield significant results after multiple hypothesis testing. Strictly speaking, only one of these probe sets for only one sample/depression measure pair can be strictly interpreted as a replication – the Nemoda probes using maternal depression scores and child cord blood (none of the probes in this probe set are significant after multiple hypothesis testing). However, the stressors and phenotypes assessed by Cecil (abuse), Glad (cortisol reactivity and levels), and Houtepen (stress response) all relate to depression. Methylation associated with these stressors and phenotypes should also be associated with depression. Similarly, in only two of the sample/depression pairs are there any NR3C1 probes that are significant after multiple hypothesis testing.

APPENDIX I

I - DERIVATION OF THE ECONOMIC MODEL

Utility Function

The parent gains utility from consumption c_i and suffers disutility from labor effort L_i and from providing parenting inputs ι_i, where i indexes the family dynasty. Affective capital e_i reduces both the level and marginal disutility of labor effort and parenting inputs. In addition, the parent altruistically gains utility from some function of her child's expected utility $\alpha_i \mathbb{E}\left[\tilde{V}\left(H_i', e_i', a_i\right)\right]$, where α_i represents the parent's altruism, $\mathbb{E}$ is the expectation operator, primes denote values for the offspring, $\tilde{V}(\centerdot)$ represents the child's indirect utility, H_i is educative human capital, and a_i are inherited assets. I thus specify the parent's utility according to the additively separable function

$$\ln c_i - \frac{\rho}{e_i} L_i - \frac{1}{e_i} \sum_t \iota_{i,t} + \alpha_i \mathbb{E}\left[\tilde{V}\left(H_i', e_i', a_i'\right)\right]$$

the parameter ρ adjusts the relative response of labor effort and parenting inputs to affective capital.

$\tilde{V}\left(H_i, e_i, a_i\right)$ represents the child's indirect utility function under a simplifying assumption – that the child bears no children of her own (or, alternatively, that the child's altruism parameter were set to 0). I eliminate consideration of grandchildren purely for convenience. However, this assumption can be rationalized as multi-generational myopia – parents consider the intergenerational effects of their investments but do not consider the multigenerational effects. Alternatively, this assumption can be treated as a tractable approximation to multigenerational altruism. Given both the qualitative similarities between 'myopic' altruism and true altruism as well as the difficulty of using data to distinguish between different forms of altruism, I view this assumption as an innocuous one.

Budget Constraint

The parent spends from permanent income $\xi_i H_i L_i^{\phi} + a_i$. The parent generates income through a concave function of labor market effort L_i^{ϕ} ($\phi \in [0,1]$), which is multiplied by an exogenous shock ξ_i and educative human capital H_i. In addition, the parent may inherit assets a_i.

The parent spends permanent income on consumption c_i, affective investments $I_{i,t}$, educative inputs $s_{i,t}$ ('schooling'), and bequests a_i', which are discounted according to an appropriate rate r. Altogether, the parent's budget constraint is given by

$$\xi_i H_i L_i^{\phi} + a_i = c_i + \sum_t I_{i,t} + \sum_t s_{i,t} + a_i' \frac{1}{1+r}$$

Thus, the parent's lagrangian is given by:

$$\begin{aligned} & \ln c_i - \frac{\rho}{e_i} L_i - \frac{1}{e_i} \iota_i + \alpha_i \mathbb{E}\left[\tilde{V}\left(H_i', e_i', a_i'\right)\right] \\ +\lambda_i & \left[\xi_i H_i L_i^{\phi} + a_i - c_i - \sum_t I_{i,t} - \sum_t s_{i,t} - a_i' \frac{1}{1+r}\right] \\ & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + \eta_i a_i', \end{aligned}$$

where λ_i is the Lagrange multiplier on the budget constraint and η_i is the Lagrange multiplier preventing non-negative bequests.

To summarize: the parent's choice variables are c_i, L_i, ι_i, $I_{i,t}$, $s_{i,t}$, and a_i'; the parent's stocks are H_i and e_i; and the parent's preference parameters are ρ and α_i.

Human Capital Production

Because the ALSPAC data includes methylation data for the parents and child when the child is born, for the child when the child is age 7, and for the parents and child when the child is age 14, I estimate the production of the child's affective capital with parental investments occurring in two periods of childhood (for the theoretical section I consider only one period of investment). Subscripts denote age $t \in \{0,1,2,3,5\}$, where 0 refers to when the child was born, 1 refers to when the child was age 7, 2 refers to when the child is age 14, and 3 refers to early adulthood (when parents have children), and 5 refers to later adulthood (when adult's the child is 14). I skip period 4 to reflect that the ALSPAC data does not measure parental epigenetics when the child is age 7. Inputs are indexed to the child's age. Although I will consider

Parents endow children with affective capital epigenetics at birth according to a linear function of parental affective capital at partum $e'_0 = \alpha_0 + \delta_{e,e,0} e_3 + \varepsilon'_e$. The parameter $\delta_{e,e,0}$ can be thought of as the biologically determined intergenerational depreciation of affective capital mediated through unobserved prenatal envrionment and seminal small RNA content. α_0 is a constant and ε'_e is an exogenous shock.

A CES production function produces the child's affective capital in periods 1 and 2 using the child's prior affective capital e'_t, parental affective inputs ι_t (e.g., cuddling), parental goods inputs I_t, a summary of the child's genes relevant for producing affective capital g_e, and shocks ε_{t+1}.

$$e'_{t+1} = \left(\beta_{e,e,t+1} e_t'^{\sigma_e} + \beta_{e,\iota,t+1} \iota_t^{\sigma_e} + I_t^{\sigma_e} \beta_{e,I,t+1} + g_{e,t}'^{\sigma_e}\right)^{\frac{1}{\sigma_e}} \varepsilon'_{e,t+1}$$

$$\varepsilon'_{t+1} = \nu_{e,t+1} \phi_{e,t+1}$$

$$H'_{t+1} = \left(\beta_{H,H,t+1} H_t'^{\sigma_H} + \beta_{H,e,t+1} e_t'^{\sigma_H} + s_t^{\sigma_H} \beta_{H,s,t+1} + g_{H,t}'^{\sigma_H}\right)^{\frac{1}{\sigma_H}} \varepsilon'_{H,t+1}$$

$$\varepsilon'_{H,t+1} = \nu_{H,t+1} \omega_{H,t+1}$$

$$\varepsilon'_{e,t+1} = \nu_{e,t+1} \omega_{e,t+1}$$

$$H'_0 = 1$$

The shocks ε can be decomposed into an exogenous component ν and an endogenous component ξ, which will be useful for control function estimation.

Parents' beliefs about the parameters in this production function may diverge from the actual parameters in the production function. In particular, parent i believes parameter $\beta_{k,k,t}$ takes on the value $\beta_{k,k,t,i}$. Thus, variation in parental beliefs about the production function parameters can affect child outcomes by shifting what the parent believes to be optimal investment.

The child's educative human capital is assumed not to enter into his production of affective capital prior to age 14. Indeed, prior to age 14, the child passively receives all inputs. This assumption is partly due to limited measurements on child cognitive ability. I also assume that parental education, parental cognitive ability, and parental affective capital do not enter into the production function directly; rather, these affect the choice of inputs ι_t and I_t, and these also associate with the endowments e'_0 and g'_e. However, to the extent that available measures do not account for their inputs, non-inputs like parental education and cognitive ability may proxy for inputs.

Equilibrium

For this section, I specialize the model so as to admit closed-form solutions. Then I discuss the equilibrium decisions of parents, focusing on the case where credit constraints bind.

I assume a unit elasticity of substitution (a.k.a Cobb-Douglas) for the production of educative and affective human capital. I also assume a single period of investment (with Cobb-Douglas production, allowing for multiple investment periods complicates the model without providing additional insight); however, I include the child's affective capital as an input to his educative human capital. The specialized model is given by:

$$
\begin{aligned}
\ln c_i - \frac{\rho}{e_i} L_i - \frac{1}{e_i}\iota_i + \alpha_i \mathbb{E}\left[\tilde{V}\left(e', H', a'\right)\right] \\
+\lambda\left[\xi_{it} H_i L_i^{\phi} + a_i - c_i - I_i - s_i - a_i' \frac{1}{1+r}\right] \\
+\eta a_i',
\end{aligned}
$$

$$e' = e^{\beta_{ee}} \iota^{\beta_{e\iota}} I^{\beta_{eI}} g_{e,}' \varepsilon_{e,}'$$

$$H' = H^{\beta_{HH}} e'^{\beta_{H,e,1}} s^{\beta_{H,I,1}} g_H' \varepsilon_H'$$

For the remainder of this section I suppress the familial dynasty index i.

Indirect utility function

I derive the indirect utility function $\tilde{V}\left(e', H', a'\right)$ for childless offspring by taking first-order conditions after setting the altruism parameter to zero. I take approximations around null bequests $a = 0$ so as to admit closed form solutions. Then the first order conditions are given by

$$
\begin{aligned}
[c]: &\quad \frac{1}{\lambda} = c \\
[L]: &\quad \frac{\rho}{e} = \phi \lambda \xi H L^{\phi-1} \\
[\lambda]: &\quad \xi_{it} H_i L_i^{\phi} + a_i = c_i
\end{aligned}
$$

Combine the first-order conditions for consumption and labor effort, yielding

$$c = \frac{e}{\rho} \phi \xi H L^{\phi-1}$$

Solve for labor effort by substituting into the budget constraint, yielding

$$\begin{aligned}\frac{\rho}{e}\left(\xi H L^{\phi}+a\right) &= \phi\xi H L^{\phi-1}\\ L^{*} &= \frac{e}{\rho}\phi-\frac{1}{\xi H}L^{1-\phi}a\\ &\approx \frac{e}{\rho}\phi\end{aligned}$$

thus solving for labor effort (when a is close to zero). Note that labor effort is decreasing in a ($\frac{dL^{*}}{da}=\frac{-\beta L^{1-\phi}}{1+(1-\phi)L^{-\phi}a}$) and for this reason consumption is also increasing in a.

Substitute into the prior equation, yielding

$$\begin{aligned}c^{*} &= \frac{e}{\rho}\phi\xi H\left(\frac{e}{\rho}\phi-\frac{1}{\xi H}L^{1-\phi}a\right)^{\phi-1}\\ &\approx H\xi\left(\frac{e}{\rho}\phi\right)^{\phi}\end{aligned}$$

thus solving for consumption (when a is close to zero).

Derive the indirect utility function by substituting these solution into the utility function and approximating around $a=0$

$$\begin{aligned}&\ln\left(\tfrac{e}{\rho}\phi\xi H\left(\tfrac{e}{\rho}\phi-\tfrac{1}{\xi H}L^{1-\phi}a\right)^{\phi-1}\right)-\tfrac{\rho}{e}\left(\tfrac{e}{\rho}\phi-\tfrac{1}{\xi H}L^{1-\phi}a\right)\\ =\;&\ln\tfrac{\phi}{\rho}+\ln e+\ln H+\ln\xi+(\phi-1)\ln\left(\tfrac{e}{\rho}\phi-\tfrac{1}{\xi H}L^{1-\phi}a\right)+\tfrac{\rho}{e}\tfrac{1}{\xi H}L^{1-\phi}a-\phi\\ \approx\;&\ln\tfrac{\phi}{\rho}-\phi+\ln e+\ln H+(\phi-1)\ln\left(\tfrac{e}{\rho}\phi-\tfrac{1}{\xi H}\left(\tfrac{e}{\rho}\phi\right)^{1-\phi}a\right)+\tfrac{\rho}{e}\tfrac{1}{\xi H}\left(\tfrac{e}{\rho}\phi\right)^{1-\phi}a+\ln\xi\\ =\;&\ln\tfrac{\phi}{\rho}-\phi+\ln e+\ln H+(\phi-1)\ln\left(\tfrac{e}{\rho}\phi-\tfrac{1}{\xi H}\tfrac{e}{\rho}\phi\left(\tfrac{\rho}{e\phi}\right)^{\phi}a\right)+\tfrac{\rho}{e}\tfrac{e}{\rho}\phi\tfrac{1}{\xi H}\left(\tfrac{\rho}{e\phi}\right)^{\phi}a+\ln\xi\\ =\;&\ln\tfrac{\phi}{\rho}-\phi+(\phi-1)\ln\left(\tfrac{\phi}{\rho}\right)+(\phi-1)\ln(e)+\ln e+\ln H+(\phi-1)\ln\left(1-\left(\tfrac{\rho}{\phi}\right)^{\phi}\tfrac{a}{\xi He^{\phi}}\right)+\phi\left(\tfrac{\rho}{\phi}\right)^{\phi}\tfrac{a}{\xi He^{\phi}}+\ln\xi\\ =\;&\phi\left(\ln\left(\tfrac{\phi}{\rho}\right)-1\right)+\phi\ln(e)+\ln H+(\phi-1)\ln\left(1-\left(\tfrac{\rho}{\phi}\right)^{\phi}\tfrac{a}{\xi He^{\phi}}\right)+\phi\left(\tfrac{\rho}{\phi}\right)^{\phi}\tfrac{a}{\xi He^{\phi}}+\ln\xi\\ \approx\;&\phi\left(\ln\left(\tfrac{\phi}{\rho}\right)-1\right)+\phi\ln(e)+\ln H+(1-\phi)\left(\tfrac{\rho}{\phi}\right)^{\phi}\tfrac{a}{\xi He^{\phi}}+\phi\left(\tfrac{\rho}{\phi}\right)^{\phi}\tfrac{a}{\xi He^{\phi}}+\ln\xi\\ =\;&\kappa_{\tilde{V}}+\phi\ln(e)+\ln H+\left(\tfrac{\rho}{\phi}\right)^{\phi}\tfrac{a}{\xi He^{\phi}}+\ln\xi\end{aligned}$$

where $\kappa_{\tilde{V}}$ is a function of parameters. The first approximation ignores income effects on labor effort caused by inherited assets, evaluating labor supply with no assets inherited. The second approximation takes a first-order approximation around $a = 0$, which allows us to evaluate the marginal utility of bequests at $a = 0$

Thus, when the parent's credit constraints bind (so $a = 0$), the indirect utility function for the child is given by $\ln\left(He^{\phi}\right)$. The marginal utility of bequeathing assets when credit constraints bind depends inversely on the child's educative and affective capital. Intuitively, an inheritance is more valuable to children who are less productive or less willing to work.

First order conditions

Substitute the child's indirect utility function into the parent's utility function. Then the first-order conditions are given by:

$$
\begin{aligned}
[c]: &\quad c^{*} = \frac{1}{\lambda} \\
[L]: &\quad \frac{\rho}{e} = \lambda^{*}\xi H L^{\phi-1} \\
[\iota]: &\quad \iota^{*} = \alpha\beta_{\mathbf{e}\iota}\left(\mathbf{1} + \phi\beta_{\mathbf{He}}\right)\mathbf{e} \\
[I]: &\quad I^{*} = \alpha\beta_{eI}\frac{1}{\lambda} = \alpha\beta_{eI}\left(1+\beta_{He}\right)c \\
[s]: &\quad s^{*} = \alpha\beta_{Hs}\frac{1}{\lambda} = \alpha\beta_{Hs}c \\
\left[a'\right]: &\quad \lambda\frac{1}{1+r} + \eta = \left(\frac{\rho}{\phi}\right)^{\phi}\mathbb{E}\left[\frac{1}{\xi' H' e'^{\phi}}\right] \\
[\eta]: &\quad \eta a' = 0 \\
[\lambda]: &\quad \xi_{it}H_{i}L_{i}^{\phi} + a_{i} - c_{i} - I_{i} - s_{i} - a_{i}'\frac{1}{1+r}
\end{aligned}
$$

Constraint on non-negativity of bequests binds ($a' = 0$)

Assume that constraint on non-negative bequests binds. Then $\eta > 0$ and $a' = 0$

Combine the first-order conditions for consumption and labor, yielding

$$c = \xi \phi \frac{e}{\rho} H L^{\phi - 1}$$

Substitute variables into the budget constraint, approximating with $a = 0$, yielding

$$\xi H L^{\phi} + a = \xi \phi \frac{e}{\rho} H L^{\phi - 1} \left(1 + \alpha \left[\beta_{eI} \left(1 + \beta_{He}\right) + \beta_{HS}\right]\right)$$

$$\begin{aligned} L^* &= \phi \frac{e}{\rho} \left(1 + \alpha \left[\beta_{eI} \left(1 + \beta_{He}\right) + \beta_{Hs}\right]\right) - \frac{a}{H\xi} L^{1-\phi} \\ &\approx \phi \frac{e}{\rho} \left(1 + \alpha \left[\beta_{eI} \left(1 + \beta_{He}\right) + \beta_{Hs}\right]\right) \end{aligned}$$

Solve for the remaining choice variables approximating with $a = 0$, yielding

$$\begin{aligned} c^* &= \phi \frac{e}{\rho} H \xi \left(\phi \frac{e}{\rho} \left(1 + \alpha \left[\beta_{eI} \left(1 + \beta_{He}\right) + \beta_{Hs}\right]\right) - \frac{a}{H\xi} L^{1-\phi} \right)^{\phi - 1} \\ &\approx H \xi \left(\phi \frac{e}{\rho} \right)^{\phi} \left(1 + \alpha \left[\beta_{eI} \left(1 + \beta_{He}\right) + \beta_{Hs}\right]\right)^{\phi - 1} \end{aligned}$$

$$I^* = \alpha \beta_{eI} \left(1 + \beta_{He}\right) c^*$$

$$s^* = \alpha \beta_{Hs} c^*$$

Thus all choice variables are expressed in terms of parameters and stocks.

Discussion of equilibrium choices

Of special interest, parental affective inputs ι is a linear function of parental affective capital:

$$\iota_i^* = \alpha_i \beta_{e\iota,i} \left(1 + \phi \beta_{He,i}\right) e_i$$

Parental income does not determine these inputs because they lack a market price (in goods or time). Yet because the parent's beliefs about the impact of affective inputs on the child's outcome may diverge from the actual impact, beliefs $\beta_{e\iota,i}$ and $\beta_{He,i}$ determine investment. These determinants emphasizes that parenting behaviors without a market price nevertheless result from an economic decision weighing costs against perceived benefits.

With measures of these beliefs and altruism, one could run the regression

$$\ln \iota_i^* = \ln(\alpha_i) + \ln \beta_{e\iota} \left(1 + \phi \beta_{He,i}\right) + \ln e_i$$

Because measures of affective capital, altruism, and beliefs lack any natural scale, such a regression would be useful only for indicating the relative explanatory power of these determinants as well as the for serving as a first stage in an instrumental variables regression.

Intergenerational processes

These solutions imply that human capital follows an intergenerational vector AR-1:

$$\begin{aligned} \ln H' &= \kappa_H(\alpha_i) + (\beta_{He} + \phi \beta_{Hs}) \ln e' + (\beta_{HH} + (1+\phi)\beta_{Hs}) \ln H + \ln g'_H + \beta_{Hs} \ln \xi + \ln \varepsilon'_H \\ &= \kappa_H(\alpha_i) + b_{H1} \ln e' + b_{H2} \ln H + b_{H3} \ln g'_H + b_{H4} \ln \xi + \ln \varepsilon' \end{aligned}$$

$$\begin{aligned}\ln e' &= \kappa_e\left(\alpha_i, \beta_{e\iota,i}\right) + \left(\beta_{ee} + \beta_{eI}\left(1+\phi\right) + \beta_{e\iota}\right)\ln e + \beta_{eI}\ln H + \ln g'_e + \beta_{eI}\ln\xi + \ln\varepsilon_{e'} \\ &= \kappa_e\left(\alpha_i, \beta_{e\iota,i}\right) + b_{e1}\ln e + b_{e2}\ln H + b_{e3}\ln g'_e + b_{e4}\ln\xi + \ln\varepsilon_{e'}\end{aligned}$$

where κ's are functions of parameters and the reduced form b coefficients are functions of structural parameters. Note that parental altruism and parental beliefs about parameters in the production function enter into the κ variables; to the extent that altruism and beliefs vary across families and covary with parental human capital, parental human capital is endogenous.

A sufficient condition for this process to be mean-reverting is that the sum of all coefficients across both processes is less than one:

$$\beta_{He} + \beta_{HH} + \left(1+2\phi\right)\beta_{Hs} + \beta_{ee} + \beta_{eI}\left(1+\phi\right) + \beta_{e\iota} + \beta_{eI} < 1$$

The reduced-form b coefficients describe the cumulative effect of changing parental human capital prior to conception on the child's adult human capital. For example, b_{e1} captures the effect of parental affective capital on the child's affective capital through direct epigenetic bequests β_{ee}, through increased expenditures allowed by greater parental labor supply $\beta_{eI}\left(1+\phi\right)$, and through parenting behaviors altered by parental affective capital $\beta_{e\iota}$. Increasing the parent's affective capital by one unit prior to conception would increase the child's affective capital by b_{e1}. But increasing the parent's affective capital by one unit at birth would increase the child's affective capital by $b_{e1} - \beta_{ee}$.

By setting β_{eI} and β_{Hs} close to zero, this model can be reconciled with studies finding small effects of parental income shocks on child outcomes. Under such a restriction, parental educative human capital has no effect on the child's affective capital, and the effect of parental educative human capital operates through β_{HH}; by allowing $\beta_{e\iota}$ and β_{He} to be greater than zero, this model can be reconciled with adoption studies finding effects of adoptive parent characteristics on the adoptive child's depression and the adoptive child's educational attainment; by allowing for variance in g_H and g_e, this model can be reconciled with behavioral genetic studies finding large effects of genes on education and depression; and by allowing β_{ee} to be greater than zero, this model can be reconciled with the biological literature on epigenetic bequests. Thus, this model can be made consistent with the stylized factos of intergenerational mobility and provides a framework for apportioning several different causes of intergenerational mobility.