THE UNIVERSITY OF CHICAGO

THE ECONOMICS OF PARENTING SKILL AND CHILD DEVELOPMENT

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
JUN HYUNG KIM

CHICAGO, ILLINOIS
JUNE 2019

To my parents.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my committee members, James J. Heckman (chair), Stéphane Bonhomme, and Alessandra Voena for advice, support, and the inspiration for economics research. I also would like to thank Kurt Hahlweg, Wolfgang Schulz, Nina Heinrichs at the University of Braunschweig and Tanja Zimmermann at the Hanover Medical School for their generous support. Their passion for psychology and the welfare of children made this paper possible. I am grateful of the support provided by the staffs and colleagues at the Center for the Economics of Human Development at the University of Chicago during the writing of this paper.

ABSTRACT

This paper develops and tests a model of parental punishment and child development that emphasizes the role of punishment as a communication channel from the parent to the child. The model mechanism is verified using an experimental data and tested using an observational data. In the model, the parent knows the long-run returns to the child's human capital. The child uses the parent's investment and punishment as noisy signals of the optimal level of effort. By using punishment consistently in response to the child's behavior, the child can be persuaded to behave optimally even in the absence of parental supervision, leading to long-run improvement in outcomes. Consistency of punishment depends on the parent's skill in managing parent-child interaction as the parent attempts to communicate to the child through punishment. The model implies that increasing the parent's skill in implementing punishment consistently can improve child development. Experimental evidence from Germany shows that improving the parent's skill can be achieved by education and training program, leading to better behavior of children as late as ten years after the program. Additional model implications are tested using nationally representative data from the US, showing that punishment can predict either better or worse child development and outcomes depending on how consistently it was implemented in response to child behavior.

CHAPTER 1

INTRODUCTION

Why do parents punish their children, and what effect does punishment have on child development? Think of parents reacting to problematic behavior of the child in an everyday context, for instance by yelling, implementing time-outs, grounding, or even using corporal punishment. Although such punishment behaviors are a part of everyday parenting in any household, their effects on child development are poorly understood. No paper in economics directly investigated the effect of punishment on child development, although this question has received attention in psychology.[1] Proposed answers include that the effects of punishment are good ([9], [49]), bad ([39], [64]), or unknown ([34], [63]). In order to improve our understanding of the role of parenting on child development and the consequences of endogenous parental responses to policy and market shocks, it is important to understand the true effect of parental punishment on child development.

Previous economics literature has focused on time, material, schooling and childcare as inputs to human capital development.[2] It is plausible to assume that these inputs will have monotone effects on child development. Opportunity costs of these inputs are easily imagined: spending more time with the child, or providing better environment to the child is costly to the parent.

What are not so straightforward, however, are decisions the parent faces when punishing the child in an everyday context.[3] Does it cost more for the parent to punish the child? Punishment may be more difficult and emotionally taxing for the parent when compared

1. Theoretical models have emphasized the interaction between the parent and the child ([55]), and the parenting style of the parents ([7], [24]). Empirical investigations have focused on the effects of parenting style ([8], [63]) and corporal punishment ([39], [34], [64]).

2. The literature on this topic is vast, including the classical work of [11]. More recent contributions include [25], on the material and time investment of the father and the mother; [18], on mandatory schooling; [37], on early childhood care; and [57], on neighborhood quality.

3. I consider punishment behavior such as time-out, grounding, privilege restriction or corporal punishment–behavior that can take place on a daily basis. A parent's choice to restrict support for a child's living expenses or college tuition could also be construed as punishment for misbehavior, but this paper does not account for such actions.

to not responding to the child's misbehavior. On the other hand, punishment may be an *easier* choice for the parent in some circumstances, perhaps because the parent may need to control her anger and frustration long enough to determine whether the child actually deserves punishment, or whether he can be motivated in some other way. Subsequent child development also depends on how the child responds to the punishment. In fact, for children showing problematic behavior, the primary effect of parental punishment is to make the problematic behavior *worse* ([55]). The parent's use of punishment, therefore, cannot be incorporated into economic analysis before the underlying choice problems of the parent and the child, and their connections to child development, are understood.

In this paper, I describe parent-child interaction as a noisy signaling process between a parent and a child in late childhood to early adolescence. The model affords a clear interpretation of parenting in a simple economic framework, clarifying the aspect of parental punishment that acts as an input in child human capital production. It generates predictions of parental and child behavior, demonstrating the varying effects of punishment on child development that are consistent with evidence from an experimental program in Germany and nationally representative data from the United States. And it points to a straightforward policy implications that can benefit child development.

The model interprets punishment as a noisy signal from the parent to the child.[4] Parental investment and child effort produce human capital in the child, generating utility for the child in the long run. The long-run return for the child's effort is *unknown* to the child but *known* to the parent.[5] The child learns about the long-run return of effort by observing parental investment and punishment. In particular, to the child, punishment signals how far he has deviated from the optimal behavior desired by the parent. As they parent tries to communicate this information to the child through punishment, the child may resist or avoid the

4. In this sense, the model closely resembles that of [50] by focusing on the parent's noisy communication to the child.

5. This setup is similar to that of [12], who analyzed a principal-agent model where the principal holds private information about the agent's productivity.

punishment entirely. Or, the child may comply with the parent's demands, but only to avoid immediate punishment, without understanding the intended message. The resistance of the child may escalate the parent-child interaction into conflict, leading the parent to punish *more harshly than intended* due to emotional over-reaction. On the other hand, anticipating this escalation, or faced with avoidance, the parent may give up on the punishment prematurely, leading the parent to punish *less than intended*.[6] Inconsistency in punishment refers to such over- and under-response with respect to the parent's intention. Parent's message to the child is obscured when the punishment signal is delivered inconsistently in this way. The ability of the parent to communicate to the child through consistent use of punishment is labeled *parenting skill* in this paper, and described as the precision of signal in the model.

As a direct application of the model, I evaluate a randomized intervention in Germany where the parents are educated and trained to consistently and effectively punish their children. The program curriculum emphasizes the behavioral techniques that allows the parent and the child to communicate more clearly in disciplinary situation, while avoiding emotional hostility that prevents effective communication. I show that parents who received training were less harsh and inconsistent in instances of punishment for at least 4 years after the intervention. Evaluation of child outcomes 10 years later revealed that children in early adolescence, between 13 and 16, showed significantly less problematic and aggressive behavior if their parents received parenting training when the children were between 2 and 6 years old. The fact that the nature of the punishment significantly changed in response to parent-specific education and training suggests that a relevant model of parent-child interaction should focus on parental characteristics that determine the effect of punishment. Both the nature of the curriculum and the pattern of parent-child interaction suggest that the consistency of punishment is determined by parenting skill as put forward in this paper. Finally, programs that target parents such as this can be implemented at a relatively lower cost and with significant improvement in child outcomes, while direct intervention on children

6. A low-skilled parent would often "lose control" or "give up" while trying to punish the child.

in the early childhood period require professionally trained personnel and often multiple years of exposure.[7]

In addition to the experimental data, I analyze the National Longitudinal Study of Youth 79 Child and Young Adult Supplement (CNLSY) to test the model's predictions. CNLSY is a nationally representative US data that contains measures of child behavior, human capital of parents and children, and rich parenting measures collected every two years until children reaches age 14. The data also include long-run child outcomes, such as wage earnings and college attendance before age 30. In this data I use measures of the parent's punishment use that are different from those used in experimental data that directly assesses the parent's inconsistency in punishment. I show that the effect of punishment on the child's developmental and life cycle outcomes can be positive or negative depending on parent characteristics that are correlated with the consistency in which the parent uses punishment in response to child misbehavior.

Finally, I review alternative models of parent-child interaction that are proposed in the literature. Some of these are based on the principal-agent framework, with a critical role played by parental incentive for the child's effort. I show that their implications are inconsistent with the empirical facts on parental punishment in the context of day-to-day parenting from the experimental data and CNLSY data that I present.

In Section 2 I review the related literature on parent-child interaction and punishment in economics and psychology. Section 2 provides a brief overview of the connection between emotional hostility in the household and the quality of parent-child communication in psychology literature, partly motivating the key mechanism in the baseline model. Section 3 presents the baseline model and generates testable predictions. Section 4 offers empirical analysis and discussion. Specifically, Section 1 applies the model to a randomized experiment, and Section 2 tests the model using CNLSY data. Section 3 discusses alternative

7. [30] is an example of a study that successfully combined both parenting intervention and direct child intervention.

models of parent-child interaction and Section 5 concludes. Throughout the paper, I adopt the convention of using ‘she’ for the parent and ‘he’ for the child without loss of generality.

CHAPTER 2

RELATED LITERATURE

1 Parent-Child Interaction

Reflecting the complex nature of parent-child interaction, the empirical literature reports significant heterogeneity in the relationship between this interaction and child development. Psychology literature shows that while combination of parental warmth and parental control of child behavior tends to predict the best child outcomes, the relationship between the parent's use of punishment and child outcomes is different at least in terms of socioeconomic status, ethnicity, and nationality ([63], [34], [49], [38]). Limited evidence based on the long-run evaluation of randomized interventions ([26], [48]) suggest that reducing the parent's harshness in punishment reduces the child's externalizing behavior at least 10 years out.[1] Harshness in these studies is measured as the parent's emotional over-reaction and inconsistency in punishment use (see Table D.1). A small empirical economics literature attending to parenting style also focuses on harshness measures, using data from Australia ([35]) and from Canada ([28]). These studies show that measures of the parent's harshness predict diminished cognitive and behavioral skills in adolescent children. These findings suggest it is not just the *use* of punishment, but the *quality* of punishment, that determines its effect on child outcomes.

Any useful economic model of parent-child interaction should be able to (1) endogenously determine the behavior of the parent and the child; (2) capture the heterogeneous effect of parenting on child development; and (3) generate policy-relevant implications. Many of the existing models are based on the principal-agent framework ([66], [16], [12], [4], [42], and [19]). As I explain in detail in Section 3, these models do not satisfactorily meet the three objectives outlined above when considering the parent's use of punishment as a part

1. A large literature on childhood parenting intervention exists, but for most of the interventions, follow-up attention does not extend beyond 6 months. See [51] for a review.

of everyday parenting. [36] and [27] take a different approach, focusing on historical and macroeconomic sources of variation in parent-child interactions over time and across societies. [50] models parental supervision as the child learns about the world. The model in this paper is an extension of [50] that accounts for the three objectives above.

This paper incorporates the parent's use of punishment as an additional input in the production of child human capital. Previous studies of human capital development considered multiple dimensions of child human capital ([22], [23], [3] and [41]), the parent's time and material investment ([43], [25]), child care ([65], [37]) and the quality of the neighborhood ([57]). Parent-child interaction, parenting style, and parental warmth/harshness, however,[2] have not been fully incorporated into the framework of human capital production, and their interactions with other inputs have not been examined. This model provides a theoretical foundation motivating empirical investigation of the effect of parent-child interaction in human capital development.

Finally, early childhood interventions have received much attention as a cost-effective and highly productive way of improving life cycle outcomes ([32]). The present study can improve our understanding of the role of the parent's response to changes in child behavior and skill, and can thus help improve interventions in households with children. In particular, a class of interventions specifically target parenting skill ([48]), the parent's monitoring ([14], [13], [15]) and the parent's knowledge ([21], [54]), showing that parent-specific interventions can be effective. The model in this paper proposes a channel in which parents respond to such interventions.

2 Parent-Child Communication and Emotional Hostility

This section briefly reviews the developmental psychology counterpart (see [55] and [31]) to the key mechanisms in the economic model presented in the following section. Psychologists

2. Operational meaning for parenting style and parent-child interaction are more distinguished in the field of psychology than they are in economics.

studying the children with aggressive behavior have noted that their parents tend to be both explosive and inconsistent in their punishment, and that the punishment often made the behavioral problems worse.

Development psychology describes a cycle of negative reinforcement between a parent and a child. The parent's use of punishment is typically initiated as an attempt to change the child's behavior. If the child immediately complies and the parent approves, the child learns the desirable behavior and no conflict occurs. However, the child may choose to ignore or resist the parent's attempt to change his behavior. This situation can arise in any household, depending in part on the child's pro-social skills.[3] When a child resists, some parents can maintain emotional stability and enforce firm but mild restrictions on the child so he clearly recognizes the desired behavior. Other parents, however, may succumb to emotional over-reaction, punishing too harshly for a given misbehavior. Here, the child may resist further, thereby escalating the tension and leading the parent to over-punish. Or, the parent may choose to avoid further confrontation by giving up on trying to modify the child's behavior, thus under-punishing the child.

Such parent-child interaction characterized by emotional hostility and confrontation is detrimental to the development of pro-social skills. First, facing the parent's hostility, the child may not invest in learning the behavior preferred by the parent and may not internalize it as the new norm. Instead the child may try to avoid the parent's anger by complying with the request–but only to placate the parent's anger, or by putting up resistance until she relents. Also, the child may attribute the cause of the punishment to the parent's emotional state rather than his own misbehavior. In this way, explosive punishment tend to be also inconsistent, and inconsistent punishment tend to be explosive. Child behavior deteriorates in response to inconsistent and explosive punishments.

The baseline model presented in the next section describes the parent-child interaction

3. Pro-social skill is a child's ability to understand the motivations of others and comply with external demands, and is cultivated throughout childhood.

in the context of parental punishment described above for the entire childhood.

CHAPTER 3

SIGNALING MODEL OF PARENT-CHILD INTERACTION

The model presented in this paper describes parent-child interactions from late childhood to early adolescence. This time period is chosen because, at this age, the child has some perception of the future state of the world, but the parent's role in shaping the child's view is still significant. There are two periods in the model. In the first period, the child effort and parental investment produces the child's human capital. In the second period, the child makes an effort choice once more. Long-run outcome of the child is realized at the end of second period as a function of human capital and the child's second-period effort. The child in the first period knows that the second period exists, but does not know the marginal productivity of effort in producing long-run outcome. Therefore, when the child chooses effort in the first period (before parent's choices), the child completely relies on his belief about his own long-run productivity. The altruistic parent, on the other hand, knows this value (denoted R in the model), and wishes to inform the child of this value so that the child's life cycle utility is maximized. Punishment is one particular channel in which this information is communicated to the child. Specifically, the parent attempts to punish the child by how far his first-period effort has deviated from the optimal level determined by R. Avoidance or resistance of punishment by the child may lead to parent-child conflict. Parents differ in their ability to manage parent-child interaction during punishment, so that punishment is implemented consistently and the children accurately learn the message sent by the parent. This ability, called parenting skill in the model, determines the technology of child human capital production in the household, and can be fostered by policy instrument evaluated in Section 1.

Table 3.1 summarizes the notations used in the model to provide a convenient reference throughout this section.

Table 3.1: Model Notations

Child's $t=2$ preference: $R(\theta_2+\beta_2 a_2)-\frac{\tau}{2}a_2^2$	
R	returns to human capital in $t=2$
θ_2	child's human capital in $t=2$
β_τ	child's effort productivity in $t=2$
τ	child's $t=2$ effort cost parameter
a_2	child's $t=2$ effort
Child's $t=1$ preference: $-(M-a_1)^2-\frac{v}{2}a_1^2$	
M	realized parental message
a_1	child's $t=1$ effort
v	child's $t=1$ effort cost parameter
Parent's $t=2$ preference: $\alpha_2\times$(child's $t=2$ preference)	
α_2	parent's altruism parameter for the child's $t=2$ utility
Parent's $t=1$ preference: $-\frac{c}{2}\varphi^2+\alpha_1\times$(child's $t=1$ preference)	
α_1	parent's altruism parameter for the child's $t=1$ utility
c	parent's investment cost parameter
φ	parent's investment choice
Human capital production: $\theta_2=\beta_\theta\theta_1+\beta_a a_1+\beta_I I$	
β_θ	marginal productivity of baseline human capital
β_a	marginal productivity of the child's $t=1$ effort
β_I	marginal productivity of parental investment
I	realized parental investment
Child's posterior belief: $\rho_2=\frac{1}{s_2}\left(\rho_1 s_1+s_M\lambda_M^2\tilde{M}+s_I\lambda_I^2\tilde{I}\right)$	
s_M	precision of signal M
λ_M	$\delta\beta_a/v$ (δ: discount rate), related to the payoff of a_2
s_I	precision of investment I
λ_I	$\delta\alpha_2\beta_I/c$, related to the payoff of I
ρ_1, s_1	child's prior beliefs
s_2	$s_1+s_M\lambda_M^2+s_I\lambda_I^2$, posterior belief precision
$\tilde{M}$	M/λ_M, the signal 'observed by the child'
$\tilde{I}$	I/λ_I, the signal 'observed by the child'

1 Timing, Technology and Preferences

In this model, there is one child and one parent, with one period of childhood and one period of adulthood. The timing of the model is

1. Child chooses action a_1 based on prior belief. $(t = 1)$

2. Parent responds with investment I and message M. $(t = 1)$

3. Child updates belief after observing I and M. Child's human capital develops, taking I and M as inputs. $(t = 1)$

4. Child chooses action a_2, after observing second period human capital. $(t = 2)$

5. Child generates life cycle outcome. $(t = 2)$

The child's action in the second period is costly and can be interpreted as post-secondary education or labor supply decision. Child outcome can be interpreted as the outcomes generated as a function of human capital accumulated up to post-secondary education, such as years of schooling or earnings.

The child's human capital at time t is denoted θ_t and follows linear technology:

$$\theta_2 = \beta_\theta \theta_1 + \beta_I I + \beta_a a_1 \tag{3.1}$$

where θ_1 is the baseline human capital, I is parental investment and a_1 is child's own action. Once baseline human capital, parental investment and child effort are determined, human capital growth is deterministic.

The nature of parent-child interaction is such that there is no communication channel between the parent and the child other than the realization of I and M. The child receives disciplinary message from the parent, produced as

$$M \sim \mathcal{N}\left(\mu, s_M^{-1}\right) \tag{3.2}$$

Producing M imposes disutility on the child through the punishment function, which is

$$Punish\,(M, a_1) = (M - a_1)^2\,. \tag{3.3}$$

The only cost of punishment from the parent's perspective is that she is altruistic towards the child. Since the parent generates M after a_1 is determined, punishment does not affect a_1. It only affects a_2 by signaling mechanism. Assuming quadratic function on the punishment function puts a natural upper limit on how much utility the parent can generate for the child, even if the parent completely disregards the child's long-run outcome. When the child's choice a_1 deviates from the optimal level the parent desires, the parent can noisily signal how far the child has deviated, at the altruistic cost of causing pain on the child. Quadratic punishment function unfortunately implies that the parent disapproves of too much effort as much as too little effort. While some parents may actually send a negative signal when the child puts in too much effort,[1] the analysis below will focus on the case where the parent tries to signal the child to put in more effort.

The precision of the parent's message s_M is called "parenting skill" in this model and it is heterogeneous in the population. If parenting skill is large, the realized punishment is close to what the parent originally intended, and the child receives messages close to what the parent originally intended. If parenting skill is low, both the message and the punishment may deviate significantly from what the parent originally intended. The parent's inability to control her message as well as the severity of the punishment captures the pattern of parent-child interaction described in Section 2 for the entire childhood. The implications of heterogeneous s_M are the focus of this paper.

1. Parents who believe they are heavily discriminated against in the labor market may tell their children that investment in human capital accumulation is wasted. While it would be interesting to analyze such cases, I do not discuss it further in this paper.

Investment input is produced in childhood as

$$I \sim \mathcal{N}\left(\varphi,\ s_I^{-1}\right) \tag{3.4}$$

where the parent chooses φ. Producing I will incur utility cost for the parent, through the household budget constraint. Precision parameter s_I is heterogeneous in the population and exogenous to the model. I includes purchasing toys and other educational materials for the child, and spending time actively engaging with the child. The stochastic component captures the fact that actual accumulation of human capital is different for each dollar and for time spent by the parent. Books are different in quality and content, and parents are different in their ability to hold their children's attention.

The preference of the child is

$$R\left(\theta_2 + \beta_{\mathcal{T}} a_2\right) - \frac{\tau}{2} a_2^2,\ t = 2 \tag{3.5}$$

$$-\frac{v}{2} a_1^2 - \left(M - a_1\right)^2,\ t = 1 \tag{3.6}$$

$\beta_{\mathcal{T}}$ is the marginal productivity of the child's effort in $t = 2$. Thus, the child's payoff is $R\left(\theta_2 + \beta_{\mathcal{T}} a_2\right)$, which is realized after the child has made effort choice towards human capital at the beginning of $t = 2$. R determines the productivity in which human capital and the child's effort in $t = 2$ produces child outcome. Higher R would justify more costly investment in human capital. τ is a parameter for convex effort cost, ensuring optimal internal effort exists as a function of R.

In the first period, the child does not derive direct utility from θ_1. The child is however non-myopic and pays convex effort cost $-\frac{v}{2} a_1^2$ towards human capital production. In addition, the child receives punishment from the parent in the form of $\left(M - a_1\right)^2$, signaling what his optimal behavior should have been from the parent's perspective. The child has nonzero discount rate $\delta \in (0, 1)$, and takes account of future utility benefit when deciding how much effort to choose in $t = 1$.

Preference of the parent is

$$\alpha_2 \left(R\left(\theta_2 + \beta_\tau a_2\right) - \frac{\tau}{2} a_2^2 \right),\ t = 2 \tag{3.7}$$

$$-\frac{c}{2}\varphi^2 + \alpha_1 \left(-\frac{v}{2} a_1^2 - \left(M - a_1\right)^2 \right),\ t = 1 \tag{3.8}$$

where α_1 and α_2 are the altruism parameters of the parent in each time period, both of which are positive. $-\frac{c}{2}\varphi^2$ is convex cost of investment, with cost parameter c. Utility cost of investment implicitly accounts for material and the time cost of investing in the child.

2 Perfect-Information Model

The main implications of the model are drawn from an imperfect information setting, where the child does not know R but forms a belief on its true value. To initiate the analysis, the perfect information model is presented here.

The child's choice in the second period is given by the objective function

$$\max_{a_2} R\left(\theta_2 + \beta_\tau a_2\right) - \frac{\tau}{2} a_2^2$$

where human capital θ_2 is given at the beginning of the period.

In the first period, the child's choice a_1 is given by

$$\begin{aligned} \max_{a_1} & -\frac{v}{2} a_1^2 - E\left[\left(M - a_1\right)^2\right] \\ & + \delta E\left[R\left(\theta_2 + \beta_\tau a_2\right) - \frac{\tau}{2} a_2^2\right] \end{aligned}$$

The expectations are taken with respect to the realization of investment and discipline inputs.

The parent's problem is

$$\max_{\mu,\varphi} -\frac{c}{2}\varphi^2 + \alpha_1 E\left[-\frac{v}{2}a_1^2 - (M-a_1)^2\right]$$
$$+\alpha_2 \delta E\left[R\left(\theta_2 + \beta_\tau a_2\right) - \frac{\tau}{2}a_2^2\right]$$

whereδ is the discount rate. It is not essential whether the discount rate of the parent and the child are the same. The equilibrium strategies in the perfect information case are summarized in the following lemma.

Lemma 1. *In the perfect information model, optimal effort of the child in each period is increasing in R:*

$$a_2^\dagger = \frac{\beta_\tau}{\tau}R$$
$$a_1^\dagger = \frac{\delta\beta_a}{v}R \equiv \lambda_M R$$

Optimal parental investment is increasing in R, while optimal punishment is minimized:

$$\varphi^\dagger = \frac{\alpha_2\delta\beta_I}{c}R \equiv \lambda_I R$$
$$\mu^\dagger = a_1^\dagger$$

The child's optimal solution in the second period is increasing in returns to skill R and marginal productivity of effort β_τ, while decreasing in the cost parameter τ. Since the parent is altruistic and the value of R is public knowledge, the child always chooses optimal effort choice in period 1, denoted $a_1^\dagger$. The child's first period effort is increasing in R, β_a and discount rate δ, while decreasing in cost parameter v. Since the child is already making the best possible choice, the parent will seek to minimize punishment by equating the mean of signal to be the same as the child's effort choice. Finally, due to linear human capital production function and linear second period utility, investment does not interact with the child's second period effort. Investment choice is increasing in R, β_I, δ, and second period altruism α_2, while decreasing in cost parameter c.

3 Imperfect-Information Model

The imperfect information model is the baseline model of this paper. The child starts the period with exogenous prior belief about the value of R. The child's prior belief is described as $R \sim \mathcal{N}\left(\rho_1, s_1^{-1}\right)$. Additionally, the child knows the values of the parent's precision parameters s_M and s_I in equations (3.4) and (3.2). After observing signals I and M, the child's belief is updated to posterior $\mathcal{N}\left(\rho_2, s_2^{-1}\right)$. The child starts the second period with this posterior belief. In the expectation notation below, the child's belief at time t is denoted as subscript Ch, t.

Characterization of Optimal Policies and Belief Updating

I start with the child's $t = 2$ choice, move on to parental investment and punishment, and then look at the child's $t = 1$ choice. At the beginning of the second period, the child's objective function is

$$\max_{a_2} E_{Ch,2}\left[R\left(\theta_2 + \beta_\tau a_2\right) - \frac{\tau}{2}a_2^2\right]$$

The value of θ_2 is already observed, so the only uncertainty in $t = 2$ is in the value of R. First-order condition is

$$E_{Ch,2}\left[R\right]\beta_\tau = \tau a_2$$

The child's optimal choice in $t = 2$ is

$$a_2^* = \frac{\beta_\tau}{\tau}\rho_2$$

where $\rho_2 = E_{Ch,2}\left[R\right]$. The child's second period effort is increasing in the child's beliefs regarding the value of R.

Plugging in the optimal choice of the child, his $t = 2$ utility is

$$R\theta_2 + \frac{\beta_\tau^2 \left(R\rho_2 - \rho_2^2\right)}{2\tau}$$

which is concave and has a unique optimum when $\rho_2 = R$. Thus, the parent has an incentive to send as accurate a signal as possible.

The parent chooses parenting after the child has chosen first-period effort a_1, taking account of the child's updating of his beliefs. The parent's objective function is

$$\max_{\mu,\varphi} - \frac{c}{2}\varphi^2 + \alpha_1 E\left(-\frac{v}{2}a_1^2 - (M - a_1)^2\right) \tag{3.9}$$

$$+ \alpha_2 \delta E\left(R\left(\theta_2 + \beta_\tau a_2\right) - \frac{\tau}{2}a_2^2\right) \tag{3.10}$$

where the expectation is taken over the realization of the inputs M and I. Since a_2 depends on ρ_2, whose value depends on the realization of M and I, expectation is taken over the realization of a_2 and a_2^2 as well.

I express μ as $\mu = wa_1^\dagger + (1 - w)\, a_1$, a linear combination of two values $a_1^\dagger$ and a_1. The parent chooses μ by choosing w, any real number. Then, the parent's choice of $w = 0$ is equivalent to $\mu = a_1$, minimizing the punishment for the child. Choice of $w = 1$ is equivalent to $\mu = a_1^\dagger$, meaning that the parent is sending an unbiased signal to the child.

I make the simplifying assumption that the child is naïve in interpreting the parent's signal, so that the child believes the parent is always producing unbiased signals with respect to message M and investment I. In other words, the child always assumes the parent chooses

$w = 1$ and $\varphi = \lambda_I R$, which are perfect information solutions.[2][3]

Under these restrictions on the child's beliefs, his posterior belief updating follows

$$\rho_2 = \frac{s_1 \rho_1 + s_M \lambda_M^2 \tilde{M} + s_I \lambda_I^2 \tilde{I}}{s_1 + s_M \lambda_M^2 + s_I \lambda_I^2} \tag{3.11}$$

$$s_2 = s_1 + s_M \lambda_M^2 + s_I \lambda_I^2 \tag{3.12}$$

where

$$\tilde{I} \equiv \frac{I}{\lambda_I} \sim \mathcal{N}\left(\frac{\varphi}{\lambda_I},\ \lambda_I^{-2} s_I^{-1}\right) \tag{3.13}$$

and

$$\tilde{M} \equiv \frac{M}{\lambda_M} \sim \mathcal{N}\left(\frac{\mu}{\lambda_M},\ \lambda_M^{-2} s_M^{-1}\right) \tag{3.14}$$

Since the child believes $w = 1$ and $\varphi^* = \lambda_I R$, he believes $\tilde{I} \sim \mathcal{N}\left(R,\ \lambda_I^{-2} s_I^{-1}\right)$ and $\tilde{M} \sim \mathcal{N}\left(R,\ \lambda_M^{-2} s_M^{-1}\right)$.

Equation (3.11) describes how the child's belief is actually updated. The parent's ex-

---

2. This assumption states that the child is not fully sophisticated with respect to the parent's strategy, but instead formulates a heuristic guess. By deviating from the classical game theory setting, the model gains in the tractability of expression for equilibrium strategies and comparative statics. One could argue this assumption more closely captures how the parent's strategies are perceived by the child who has never raised a child. The restrictive assumption nonetheless remains a weakness of the model. The difficulty of solving a model with classical assumption arises from the fact that agents' preferences in childhood are different from those they express in adulthood. Ongoing work seeks to build a tractable model that can improve on this aspect.

3. Consider the model where the child believes the parent is choosing w to be some value $\tilde{w}$. Then the child would interpret the parent's signal M as

$$\frac{M - (1 - \tilde{w})\, a_1}{\tilde{w} \lambda_M}$$

which has distribution $\mathcal{N}\left(R,\ s_M^{-1} \tilde{w}^{-2} \lambda_M^{-2}\right)$ from the child's perspective, when in fact it is distributed $\mathcal{N}\left(\frac{\mu - (1-\tilde{w}) a_1}{\tilde{w} \lambda_M},\ s_M^{-1} \tilde{w}^{-2} \lambda_M^{-2}\right)$. It can be shown that whenever the parent knows the child's belief of $w = \tilde{w}$, the parent wants to deviate w to a different value. The baseline model analyzes a special case with $\tilde{w} = 1$. If $\tilde{w}$ were higher, shifting mean might be harder (because of the higher denominator) but the child has higher posterior belief precision, or higher confidence. If $\tilde{w}$ were lower, shifting mean might be easier but the child's posterior belief precision is lower. In other words, the child is more easily swayed but also remains more skeptical.

pected value of the child's $t = 2$ belief is

$$\rho_2^e \equiv E\left[\rho_2\right] = \frac{s_1\rho_1 + s_M\lambda_M^2\frac{\mu}{\lambda_M} + s_I\lambda_I^2\frac{\varphi}{\lambda_I}}{s_1 + s_M\lambda_M^2 + s_I\lambda_I^2}$$

The parent can choose how the child's belief will be updated in expectation by choosing φ and μ.

From the child's perspective, his expected belief at $t = 2$ from $t = 1$ perspective is

$$E_{\mathcal{H}_1}\left[\rho_2\right] = E_{Ch,1}\left[\frac{s_1\rho_1 + s_M\lambda_M^2 R + s_I\lambda_I^2 R}{s_1 + s_M\lambda_M^2 + s_I\lambda_I^2}\right] = \rho_1$$

In other words, the assumption of naïvete is consistent with the assumption that the child thinks his own belief will remain the same in expectation.

With these restrictions, it is actually possible to obtain the child's optimal choice first, characterized by the objective function

$$\begin{aligned}\max_{a_1} -\frac{v}{2}a_1^2 - E_{Ch,1}\left[(M - a_1)^2\right]\\ + \delta E_{Ch,1}\left[R\left(\theta_2 + \beta_\tau a_2\right) - \frac{\tau}{2}a_2^2\right]\end{aligned}$$

The child's optimal solution mirrors the perfect information solution:[4]

$$a_1^* = \frac{\delta\beta_a\rho_1}{v} = \lambda_M\rho_1$$

The parent's optimal choice of φ and μ, given the child's choice of a_1, is characterized by the objective function

$$\max_{\varphi,\mu} -\frac{c}{2}\varphi^2 + \alpha_1 E\left[-\frac{v}{2}a_1^2 - (M - a_1)^2\right] \tag{3.15}$$

$$+\,\delta\alpha_2 E\left[R\left(\theta_2 + \beta_\tau a_2\right) - \frac{\tau}{2}a_2^2\right] \tag{3.16}$$

4. This is true for any value of $\tilde{w}$ the child assumes the parent has chosen.

Solving the parent's objective function (3.15) produces optimal choices of the parent w^* and φ^*.

Theorem 1. *In the imperfect information model, optimal choice of the child is increasing in ρ_1 and ρ_2, respectively, and it equals perfect-information choice when $\rho_t = R$. Optimal investment φ is increasing in R and $(R - \rho_1)$. Optimal punishment is increasing in w^* and $(R - \rho_1)^2$, where w^* is determined by equation (3.17).*

Proof. See Appendix for proof. □

Optimal choice of w is described as

$$w^* = \frac{s_1 \Psi + s_M \lambda_M^2 \Psi}{\frac{\alpha_1}{\alpha_2} + s_M \lambda_M^2 \Psi} \tag{3.17}$$

for constants $\lambda_M \equiv \frac{\delta \beta_a}{v}$ and $\Psi \equiv \delta \frac{1}{2} \frac{\beta_\tau^2}{\tau} \frac{s_M}{s_2^2} \frac{1}{\mathcal{B}+1}$. Optimal investment depends on $(R - \rho_1)$ because of the effect of investment on the child's belief in addition to its effect on the child's human capital. Optimal punishment depends on the parent's choice of w^*, which determines how severe the punishment is in response to underlying child misbehavior, which is proportional to $(R - \rho_1)^2$. Since the child believes M has mean $a_1^\dagger$, punishment is a signal to the child that he has deviated from the optimal behavior. Choosing bigger w^* signals to the child that the optimal behavior $a_1^\dagger$ has higher value.

This paper is primarily interested in the comparative statics with respect to s_M, in part because the parenting component of early childhood intervention programs can be interpreted as improving parent's ability to communicate with their children (see [51]). One such program is reviewed in depth in Section 1.[5] The next theorem presents the main comparative statics result:

Theorem 2. *The parent's optimal punishment severity w^* converges to a value between 0 and 1 as $s_M \to \infty$. This value is higher as the child's $t = 2$ effort becomes more productive.*

5. Improving school quality and neighborhood environment can be interpreted as increasing s_I.

Parent's investment choice φ^* *converges to perfect information choice as* $s_I \to \infty$.

In other words,

$$\lim_{s_M \to \infty} w^* = w^{lim} \in (0, 1)$$
$$\lim_{s_I \to \infty} \varphi^* = \varphi^\dagger$$

The value of w^{lim} depends on the parameters that determine the value of second-period effort, such as $\beta_\mathcal{T}$. The key trade off in punishment choice is between the child's pain in the first period and the improvement in the child's outcome in the second period, as summarized in Table 3.2. As s_M increases, w^* tends to increase as well. For very high value of $\beta_\mathcal{T}$, w can even be greater than 1 because the marginal benefit of sending a stronger signal far outweighs the cost of punishment. As the signal precision increases further, w^* eventually converges to a value less than 1. If $\beta_\mathcal{T}$ is very small, there is little value in the child's adult period effort and the parent chooses not to send much of a signal at all, no matter the parenting skill. w^{lim} is always between 0 and 1 for $\beta_\mathcal{T} > 0$ and a positive cost of punishment. This comparative static is described in Figure 3.1.

Discussion of Model Implications

I discuss three key predictions of the model. Throughout this section, I maintain the assumption that the child's adult choice is relevant ($\beta_\mathcal{T}$ not too close to 0), and that the parent wants to motivate the child's effort, or

$$\rho_1 < R. \tag{3.18}$$

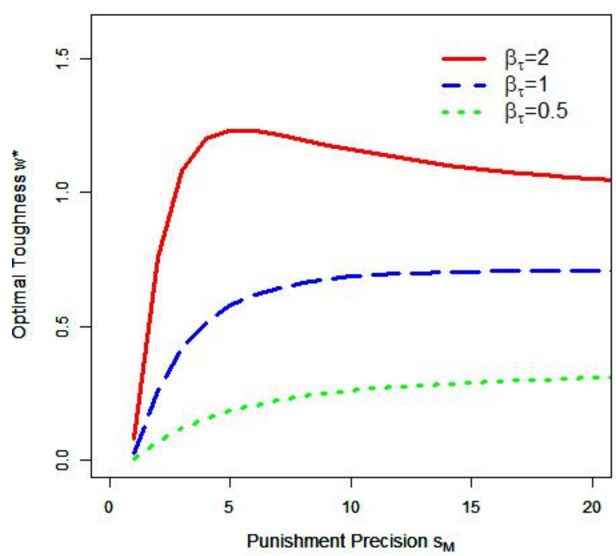
The first prediction looks at the determinant of parental punishment.

Proposition 1. *Child misbehavior is more likely to be punished as* s_M *increases.*

Table 3.2: Parent's Signaling Technology

	Punishment	Investment
Parent's intended choice	μ	φ
Realized signal	$M \sim \mathcal{N}\left(\mu,\ s_M^{-1}\right)$	$I \sim \mathcal{N}\left(\varphi,\ s_I^{-1}\right)$
Child's 'naïve' view	$E\left[M\right] = \lambda_M R \neq \mu$	$E\left[I\right] = \lambda_I R \neq \varphi$
$t = 1$ cost	$-\alpha_1 \left(M - a_1\right)^2$	$-\frac{c\varphi^2}{2}$
$t = 2$ benefit	$a_2^* = \frac{\beta_\tau}{\tau} \rho_2$	$\theta_2,\ a_2^*$

Figure 3.1: Comparative statics of w^* with respect to s_M



Expression for the expected value of punishment is

$$E\left[(M-a_1)^2\right] = s_M^{-1} + w^2\left(a_1^\dagger - a_1\right)^2 \tag{3.19}$$

s_M^{-1} is the component of average punishment that is unrelated to the child's behavior. This component decreases as s_M increases, decreasing punishment on average. $w^2\left(a_1^\dagger - a_1\right)^2$ is the component of the punishment that is determined in response to child misbehavior. As seen in Theorem 2 and Figure 3.1, increase in s_M tends to increase w. In other words, parents with higher parenting skill punish less on average, but are more likely to punish when the children misbehave.

The second prediction pertains to the heterogeneous effect of parental punishment on the child's outcome.

Proposition 2. *The child's expected outcome increases in punishment as s_M increases.*

Building on the first proposition, the second proposition generates relevant implications for human capital development policy and literature. The child's long-run outcome is generated in the model as

$$R\left(\theta_2 + \beta_\tau a_2^*\right)$$

where the child's effort is

$$\begin{aligned} a_2^* &= \frac{\beta_\tau}{\tau}\rho_2 \\ \rho_2 &= \frac{1}{s_2}\left(s_1\rho_1 + s_M\lambda_M^2\frac{M}{\lambda_M} + s_I\lambda_I^2\frac{I}{\lambda_I}\right) \end{aligned}$$

The parent also has altruistic concern over the child's convex cost of effort, so that the parent wishes to signal accurate value of R. As before, continue to assume $\rho_1 < R$. If the parent chooses $w = 0$ in $t = 1$, so that no punishment signal is sent, then $E\left[\frac{M}{\lambda_M}\right] = \rho_1$ and child outcome realization is sub-optimal. Note from equation (3.19) that even if the parent chooses $w = 0$, the child still receives some punishment purely by mistake. If the parent chooses $w^* > 0$, punishment is an informative signal so that ρ_2 moves towards R in expectation. The signal becomes stronger as s_M increases, since w^* increases in s_M and the weight on punishment in the child's posterior belief ρ_2 increases. The effect of punishment on child outcome increases in s_M.

The next proposition uses measures of the subjective expectation of the parent and the child to test for the existence of a signaling mechanism.

Proposition 3. *The child's subjective expectation of his own outcome increases in punishment as s_M increases.*

When asked during childhood what his own future outcome would be like, the child already has observed at least some of the realization of parental inputs, but not the actual value of R. Therefore, parental investment and punishment should revise his own expectation of the $t = 2$ outcome the same way they affect his actual outcome. Specifically,

$$\begin{aligned} &E_{Ch,2}\left[R\theta_2\right] + \beta_T E_{Ch,2}\left[Ra_2\right] - \frac{\tau}{2} E_{Ch,2}\left[a_2^2\right] \\ =&\rho_2\theta_2 + \rho_2\beta_T a_2 - \frac{\tau}{2} a_2^2 \end{aligned}$$

This revision should not take place for the parent's expectation of child outcome, since the parent knows the value of R and the mean of her own parental investment and punishment choices.

One interesting implication of the model is that improving the parent's monitoring of the child's behavior has the same effect as increasing parenting skill s_M. With classical measurement error on the child's $t = 1$ effort choice a_1, signal in punishment use becomes noisier. Allowing the parent to more precisely respond to the child's behavior by improving monitoring technology can allow her to both reduce unnecessary punishment, as [4] has shown, and allow her to use punishment only when necessary. This is what [13] has found, after a randomized field experiment that sent the parents of high school students text messages regarding the academic participation of their children at school. One of the findings of that study was that parents who received the information treatment were *more* likely to use moderate forms of punishment, such as restriction on privileges.

The model also suggests a reason for heterogeneity in the effects of parenting style and parental punishment on child outcomes. For instance, classical studies by [7] and others working in the field of developmental psychology determined that combining parental warmth towards the child with control over the child's behavior through the use of punishment (authoritative parenting style) was superior to using only control (authoritarian parenting style) or only showing warmth (permissive parenting style). However, subsequent literature has shown that the effect is heterogeneous by ethnicity, nationality and the parent's socioeconomic status. For example, [29] found that the authoritarian parenting style had a negative effect on adolescent academic achievement for White and Asian children but not for Black or Hispanic children in the US. One way to interpret this heterogeneity is to notice that cultural norms on appropriate parental punishment might be different across different ethnic groups. Excluding parental abuse, which does obvious damage to the child's physical development and mental health, my model suggests that the punishment affects the child through the child's belief in himself, which then determines the child's actions in both the short run and the long run. It is thus important to consider how the children might interpret different forms of parental investment and punishment behavior, which would depend on the particular cultural norms of the child's community. A parental behavior that might feel too harsh or inconsistent for one child may not be be interpreted the same for another, in a different community.

The model ultimately argues that well-implemented punishment of child misbehavior can actually increase the child's confidence about his future potential. This claim is consistent with the empirical findings in the parenting style literature, which show that some form of parental control is typically superior to no control at all ([8], [56]), especially if the parental control is focused on clear and consistent communication to the child. Intuitively, if the parent thinks the child can succeed in the future by putting in the effort, then she will be more inclined to spend time and provide materials for him. Also, the parent would try harder to correct misbehavior, since the marginal benefit of correcting the child's behavior is significant. Suppose, however, the parent has low parenting skill, so that her attempts at correcting the child's behavior often results in parent-child conflict and mutual hostility. In that case, punishment results in disutility for both the parent and the child, while the child does not learn the optimal behavior in the eyes of the parent. There, the parent finds little benefit in trying to modify the child's behavior. Parental punishment nonetheless occurs and could even occur quite a lot, since some conflict is unavoidable in the household and the could escalate through the reinforcement cycle described in Section 2. The right policy to improve child outcomes through parents would then be to improve the quality of parental punishment by helping the parents use punishment as a precise signal while avoiding mutual hostility, as opposed to trying to curtail parental punishment in the first place.

CHAPTER 4

EMPIRICAL ANALYSIS

1 Model Application: Randomized Experiment

I apply the model in this paper to analyze and interpret a randomized intervention whose specific goal was to educate and train parents in behavioral techniques that would allow them to use punishment as a clear communication to their children while avoiding mutual hostility. Thus, the policy implication of the model has a direct empirical counterpart that I will show to be effective. The curriculum of the program is based on the Level 4 Triple P Positive Parenting Program.[1] Triple P provides relatively short education and training sessions to the parents of preschool children. The curriculum trains parents how to use quiet-time and time-out techniques, and how to interact with their children before, during, and after punishment. Parents are taught to use clear and concise messages and to avoid responding to resistance from their children based on trivial issues. "An active skills training process is incorporated into Triple P" that teaches the parents to select their own goals for their children, implement effective discipline strategies, and self-evaluate the success with their chosen goals ([61]).

The sample that I focus on here consists of parallel implementations of the Triple P program in the German city of Braunschweig between 2001 and 2002, and involving 477 households in 33 preschools at baseline. The program consisted of 4 weeks of training and another 4 weeks of phone feedback. Randomization was done at the level of preschools their children attended. Participants were the parents who volunteered into the study after learning about the goals of the program, but prior to randomization. Thus, the sample does not represent the population of young mothers in Germany, although the study still identify

1. Level 1 is media-based information campaign and the intensity of program intervention increases with each level. The highest level available is level 5, targeting households with known risk factors. See [60] for an overview and [61] for a discussion of theoretical foundations. Triple P program was also used as a part of the treatment in [30].

the effect of intervention for the parents who would voluntarily participate in the program if it were made available.[2] The training sessions were attended by the parents without their children.

The first study consisted of 280 households in 18 preschools, where 186 households in 11 preschools were randomized into the intervention group and then offered training, while those in the control group did not receive training. Households were not targeted based on any child risk factors, household variables related to socioeconomic status, or neighborhood quality. The second study consisted of 197 households in 15 preschools at baseline, all of which received the training. The second study targeted preschools in disadvantaged neighborhood. Households in the second study were randomized into one of four groups based on whether or not they received monetary participation incentive, and whether they were in group-training setting or individual-training setting. Other than the 'individual' treatment and the 'unpaid' treatment in the second study, all program participants received training in a small group setting with a monetary participation incentive, following the same training curriculum. Baseline descriptive statistics is presented in Table 4.1. Because of the study design, there were significant differences between the control group and the combined intervention group in terms of neighborhood index, individual sessions and participation incentive recipiency. I control for these and other variables in the table in the empirical analysis.

A previous study in [48] analyzed the first study, exploiting the randomized training status to evaluate the program. They found that the intervention program reduced externalizing behavior and improved subjective mental well-being of children 10 years after the intervention, noting as well that the observed change in parenting mediated the effect of the intervention. [48] suffers from a problem with limited statistical power, since the design is based on cluster randomization with only 17 clusters.[3] The second study was previously

2. This research design also ensures that the intervention affected parents through skills training, and not simply through knowledge provision, i.e. telling the parents to avoid harsh parenting. The parents who volunteered to the program after hearing about the program description most likely understood that harsh parenting could be harmful to their children.

3. One cluster is dropped in that study since it had only one participant, and was not feasible for statistical

Table 4.1: Program Baseline Sample Characteristics

	Control	Intervention1	Intervention2	p-value
Problematic Behavior	-0.216	-0.055	0.157	0.21
Poverty Status	0.074	0.050	0.637	0.59
Mother graduated HS	0.553	0.541	0.228	0.87
Single Mother	0.106	0.049	0.192	0.17
Low Neighborhood	0.234	0.242	1.000	0.97
Middle Neighborhood	0.426	0.296	0.000	0.64
High Neighborhood	0.340	0.462	0.000	0.66
Individual Session	0.000	0.000	0.574	0
No Participation Incentive	0.000	0.000	0.320	0
Sample Size	94	186	197	
Number of Clusters	7	11	15	

Note: Intervention 1 refers to the first study with both the intervention and the control group. Intervention 2 refers to the second study which randomized on participation incentive and individual setting. Problematic behavior is a baseline measure of Child Behavior Checklist externalizing behavior subscale at baseline, normalized to be mean 0 and standard deviation 1 in the sample. A household is defined to as in poverty status if it makes less than 2,000 Deutsche Marks in monthly household income, where the exchange rate is approximately 1 DM $\approx$ 0.54 USD in 2001-2002. High school graduation equaled 1 if respondent's last degree attained was upper secondary or attended college. Low, middle and high neighborhood quality scale is constructed using objective kindergarten social structure index (OKS). The scale is based on the rate of unemployment, number of families on welfare, number of immigrants and quality of housing in the particular neighborhood ([6]). Fifth column is the p-value from the two-sided t-test against the null of zero in regressing the baseline variable on the training recipiency status, conditional on the indicator for belong in the second study.

analyzed in [44] who found that while paid-group showed improvement child outcomes, the unpaid-group showed even better child outcomes. The group with individual training setting also showed better child outcomes than those in the group training setting.

The study described in this paper alleviates the statistical power concern by combining the two studies, resulting in 33 clusters. The randomization design for the whole study can be interpreted as conditional randomization, where those in the low-quality neighborhood are more likely to be randomized into the training program. Target parents were those who could speak German and who had a child between ages 2.6 and 6. Parenting behavior and child outcomes were followed for the first 4 years after the intervention, and then again at 10 years after the intervention. Many of the measures between the two studies exactly overlap, and I only use the measures that overlap in the analysis.

An important advantage of the randomized intervention as an application of the model is that the curriculum can be interpreted as an investment in parenting skill that increases s_M. Measures of the parent's dysfunctional parenting, shown in Table D.1, reflect her over-reactive and inconsistent approach to punishment ([5], [53]). Reduction in the scale of dysfunctional parenting is interpreted as an increase in s_M. A key outcome measure is the child's externalizing behavior using Child Behavior Checklist (CBCL, [2] and [1]), which is based on the mother's self-reporting. Externalizing behavior consists of the child's delinquent and aggressive behavior toward others, including his own parents. The externalizing behavior measure also accounts for the child disobeying his parent, which corresponds to $\left(a_1^{\dagger} - a_1^{*}\right)^2$ in the model.

Baseline analysis is based on the linear model

$$Y_{i_k} = \alpha + \beta Z_k + \gamma X_{i_k} + \epsilon_{i_k} \tag{4.1}$$

for individual i in preschool cluster k. Z_{i_k} is an indicator that equals 1 if cluster k is in

inference which relied on wild cluster bootstrap tests.

the intervention group, and 0 if it is in the control group. X_{i_k} includes the child's baseline externalizing behavior, and indicators for single mother status, mother's high school graduation, neighborhood quality, participation incentive recipiency and individual training session. Error terms are clustered at the preschool level. I use sum scores standardized to be mean zero and standard deviation 1 for both the dysfunctional parenting of the parent and the externalizing behavior of the child. I repeat the analysis using alternative ways of constructing the measure and find that the effects are not sensitive to alternative specifications. For statistical inference, I use a one-sided t-test, assuming that the intervention effect always has a non-negative effect on parenting behavior and child outcomes.[4] I use mother's self-reported parenting measures. Measures reported by father are available but the rate of response is too low to be useful in this study (only 296 observations are available for father's parenting). If, for example, fathers play the role of discipline while mothers play the role of caregiver in many households, analysis solely based on the maternal report of parenting suffers from serious measurement error. In the study sample, the correlation between mother's harsh parenting score and father's harsh parenting score is 0.46. In addition, parenting research suggests that maternal and paternal levels of involvement on the child are in general strongly positively correlated ([58]). These findings do not reject but lower the possibility that parenting behaviors are very different between mother and father in the sample.

Table 4.2 presents the effect of receiving parent training on the measure of the parent's dysfunctional parenting. The measure of parenting is constructed as a summary measure for the parenting that the parent implemented for the four years after the intervention, and approximately covers ages 7-10. As the table indicates, the reduction in dysfunctional parenting is large and significant, at more than 0.5 standard deviation units. Table 4.4 presents the effect of parent training on the child's externalizing behavior, measured 10

4. There is insufficient study of the effect of parent training on the adolescent outcomes of children. In fact, this is one of very few studies that presents such a result. Literature surveys in [60] and [51] do not report negative effects of parent training on child outcomes. If there is evidence that parent training can actually harm child outcomes, then the use of a one-sided t-test is inappropriate.

years after the intervention based on the child's self-reporting. These are measured when the child is between 13 and 16. The effect is to reduce the child's problematic behavior by 0.2 standard deviation units. Considering that the intervention program lasted only 8 weeks, and only interacted with the parents, it is remarkable that significant effects on the child's behavior were found 10 years later. These results can now be interpreted in the context of human capital development: parenting skill during early childhood is an input toward production of the child's behavioral skills in early adolescence.

Table 4.2: Program Effect on Harsh Punishment

	(1)	(2)
Parenting Training	-0.563***	-0.543***
	(0.123)	(0.070)
Observations	463	462
R^2	0.050	0.202
Adjusted R^2	0.048	0.188
Control	X	O

Note: Response of parenting measures. Error terms clustered at preschool level. Controls include: baseline child behavior, single mother status, mother's education, neighborhood quality, indicator for participation incentive, indicator for group setting.

One concern over the study design is that the measure of dysfunctional parenting relies on the parent's self-reporting. While these measures are unlikely to suffer from recall bias, as the parents are asked to report their current parenting behavior, they may still suffer self-report bias, where the stigma associated with harsh parenting or child misbehavior causes parents to under-report dysfunctional parenting or child misbehavior. While it is impossible to directly consider the extent to which self-report bias is a problem at this point, the results from Table 4.3, which estimated equation (4.1) on measures collected immediately after the intervention program, suggest it does not pose a serious problem. The magnitude of the training on the parent's dysfunctional parenting is similar to the size of the effects found in 4.4, at -0.594 standard deviation units reduction. But more importantly, there is no effect on the child's externalizing behavior measured right after the intervention. If self-report bias is a major concern, it is difficult to explain why the parent self-reported measures show effect for parenting behavior but not for child misbehavior.

Table 4.3: Program Pre-Post Effects

	(1) Harsh Parenting	(2) Problematic Behavior
Parent Training	-0.594***	0.052
	(0.052)	(0.613)
Observations	466	466
R^2	0.223	0.622
Adjusted R^2	0.210	0.615
Control	O	O

Note: Standard errors clustered at preschool level. Outcomes measured immediately after the program.

Table 4.4: Program Effect on Child's Problematic Behavior, 10yrs

	(1)	(2)
Parenting Training	-0.217^{+}	-0.257*
	(0.129)	(0.140)
Observations	358	353
R^2	0.008	0.040
Adjusted R^2	0.005	0.018
Control	X	O

Note: Error terms clustered at preschool level. Controls include: baseline child behavior, single mother status, mother's education, neighborhood quality, indicator for participation incentive, indicator for group setting. Outcome is child externalizing behavior at 10 yr follow up.

Attrition is another concern in a study with long follow-up period. While the attrition rate for the first study is only 11%, the attrition rate for the second study was much greater at 43%. The analysis of the attrition pattern suggests poverty and single mother status were predictive of attrition status. I estimate a logit model for the probability of remaining in the sample, and use inverse probability weighting (IPW) to repeat the analysis in the equation (4.1). The estimates in Table 4.5 show that, while the precision of the estimates are slightly lower in the IPW regression, intervention effect estimates are similar to the main analysis in Table 4.4 and remain significantly different from zero at 10% significance level.

Table 4.5: Intervention Effect with Inverse Probability Weighting

	(1) Harsh parenting	(2) Harsh parenting	(3) Problem behavior	(4) Problem behavior
Parent Training	-0.577***	-0.542***	-0.216+	-0.238*
	(0.146)	(0.073)	(0.131)	(0.138)
Observations	449	449	342	342
R^2	0.045	0.227	0.008	0.054
Adjusted R^2	0.043	0.213	0.005	0.032
Control	X	O	X	O

Note: Error terms clustered at preschool level. Controls include: baseline child externalizing behavior, single mother status, mother's high school graduation status, neighborhood quality index, participation incentive recipiency, and indicator for individual training setting. Weights are the inverse of the predicted probability of child outcome being measured at 10 year follow up. This probability is estimated by logit model with predictors including intervention group indicator, poverty status indicator, indicator for being in the second study, baseline child externalizing behavior, mother's high school graduation status, single mother status and neighborhood quality index.

Measures of parenting based on harshness and inconsistency in punishment ([5]) are used widely in labor panel data, in addition to developmental psychology studies on parenting training. [35] and [28] are examples that used these measures to investigate the association of these measures with adolescent child outcomes, finding that parental harshness during childhood predicted problematic behavior in adolescents. My model makes explicit the underlying behavioral assumptions kept implicit in these papers, and the experimental evidence aligns with the conclusions in these studies. Parental harshness and inconsistency in punishment are interpreted as measures of parenting skill, determining the child human capital production technology in the household.

As a second application of the model, I test model predictions using CNLSY data that is based in the US and with different measures of parental punishment.

2 Model Application: CNLSY

I use the National Longitudinal Study of Youth 1979 (NLSY79) and its Child and Young Adult Supplement (CNLSY) to test the model's predictions. NLSY79 is a nationally representative sample from the US, covering 12,686 individuals who were between 14 and 22 years of age at the time of the first survey in 1979. They were surveyed annually until 1994 and then on a biennial basis thereafter. The latest survey available at the time of this paper's writing is 2014 version. CNLSY79 is a separate survey of all the children born from NLSY79 women, collected biennially since 1986. Of interest to this study are the assessments of academic achievement, maternal reports of child behavior, measures of home environment, parenting, and the child's long run education and wage outcomes measured until age 30.

The data also contains measures of academic achievement in the form of the Peabody Individual Achievement Test (PIAT), measured every two years between ages 5 and 14. The data contain Mathematics, Reading Recognition and Reading Comprehension assessments. I use the three measures as noisy measures of cognitive skill to construct standardized scores for the cognitive skill. They correspond to the child's human capital at each age. The child's behavior is measured by the Behavior Problems Index (BPI), which is based on maternal reports of child behavior collected every two years for children between ages 5 and 14. Typical measures include the parent's ratings on whether the child argues too much, is disobedient at home, is disobedient at school, and is stubborn, sullen, or irritable. Maternal reports of child behavior are suitable for this study, since, in this model, parental investment and punishment are based on the parent's evaluation of the child's behavior. The BPI measure, which includes the parent's perception of the child's disobedience, is assumed to correspond to the deviation of childhood behavior from the parent's optimal behavior. BPI is based on six subscales: Headstrong, Anxious, Antisocial, Dependent, Hyperactive and Peer conflict. I use the PIAT tests as noisy measures of the child's human capital and BPI subscales as noisy measures of the child's behavior, and estimate standardized scores using factor analysis.

The data contain rich measures of parenting behavior. I use cognitive stimulation subscore from the Home Observation for Measurement of the Environment (HOME) Short-Form Inventory, which is conducted every two years for children from birth to 14. The HOME score combines both the mother's self-report and the interviewer's observations, and it is widely used as a measure of home environment characteristics that can promote the child's cognitive development ([17]). While the exact component of the measure differs by the age of the child, it typically accounts for the number of books at home, how often the parent speaks or reads to the child, and how often she takes the child to museums or other cognitively stimulating settings. I interpret them as measures of parental investment I_t, since they require nontrivial contributions of household resources.

In addition to parental investment, the data contain measures of parental punishment. Specifically, I use measures based on questions such as "About how many times, if any, have you had to spank [Child First Name] in the past week?" This question is repeated for four other types of punishment: grounding, taking away television or other privileges, taking away allowance, and sending the child to his room. The responses are in frequencies. Parents are asked about spanking for children ages 0 through 14 and are asked about all 5 types of punishment for children ages 5 through 14. Asking for the parent's activity in the past week reduces concern for recall bias. Self-report bias may be a concern if the parents choose to under-report punishment use out of concern for social stigma. However, based on the survey of corporal punishment in Section C in the Appendix (which may be the most stigmatizing form of punishment), both the support for corporal punishment use and the self-report of corporal punishment use by parents remain significant across socioeconomic backgrounds. Based on the nationally representative survey in the US, even in 2011, 12% of the parents within the 90th-percentile of household income endorsed the use of physical punishment, and 10% of them reported having used it in the past week ([59]). Other studies surveyed in the Appendix suggest parents may under-report punishment use in the household not because they find it stigmatizing, but because it happens so frequently and mundanely that the parents often forget having punished their children. Figures 4.1 through 4.10 and Table 4.6 describes mother-reported punishment use for children, across ages and household income. Table 4.6 shows that income better predicts parental investment than does parental punishment. Figures 4.9 and 4.8 show that among the parents with household income in the 10th sample decile, over 60% with children ages 7-8 reported having used some type of punishment. The percentage declines as children age but remains almost 40% for those 13 to 14 years old. Figure 4.6 shows that only about 15% of the parents report having never used punishment between children ages 5 through 14. On the other hand, almost 30% of the parents always responded that they used some type of punishment whenever they are asked. While it is impossible to rule out self response bias due to social stigma, it seems that the concern is not significant, at least in American society.

Table 4.6: Descriptive Model of Parenting

	(1) punishment	(2) investment
Log Net HH income	-0.016***	0.037***
	(0.002)	(0.006)
Father presence at home	-0.004	0.172***
	(0.009)	(0.018)
AFQT	-0.041***	0.305***
	(0.006)	(0.014)
firstborn	0.060***	0.290***
	(0.007)	(0.013)
Child is female	-0.069***	0.121***
	(0.007)	(0.015)
BLACK	0.013	0.069
	(0.014)	(0.038)
NON-BLACK, NON-HISPANIC	-0.002	0.187***
	(0.014)	(0.035)
Constant	0.774***	-0.763***
	(0.029)	(0.067)
Observations	8687	8605
R2	0.041	0.270

Note: Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. Punishment model is based on linear probability model.

Figure 4.1: Punishment Use at Age 6

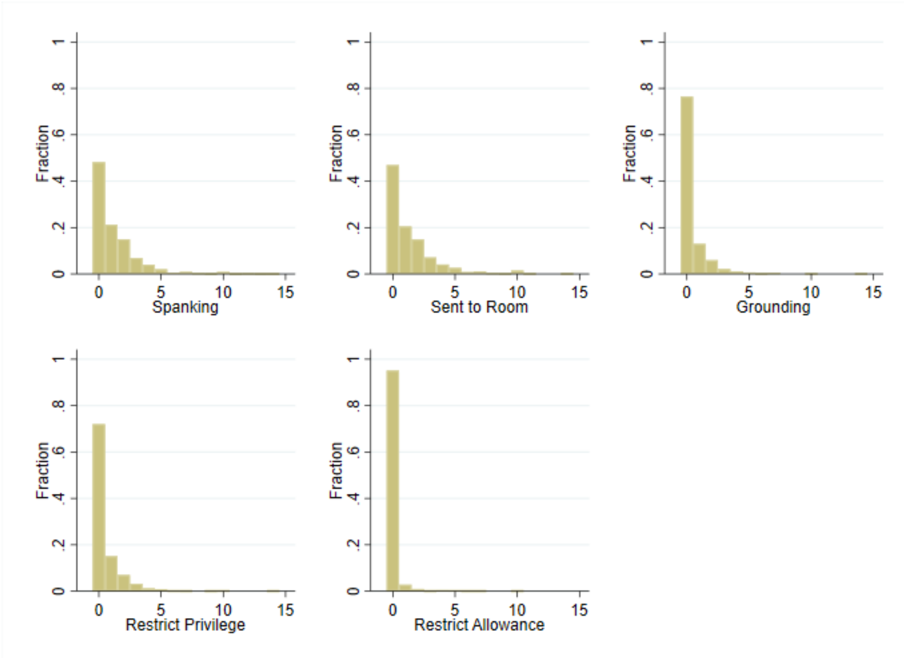


Figure 4.2: Punishment Use at Age 8

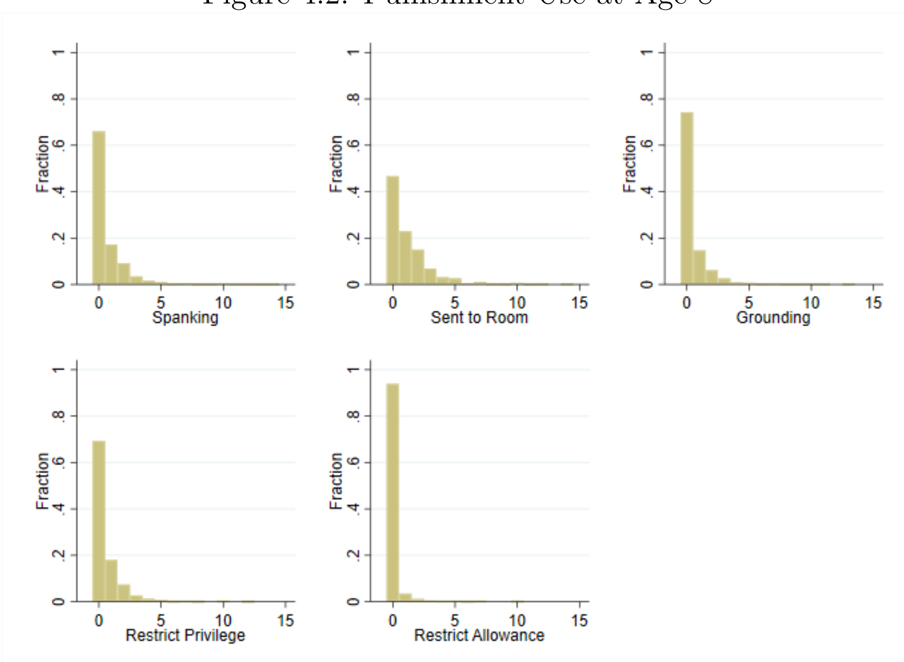


Figure 4.3: Punishment Use at Age 10

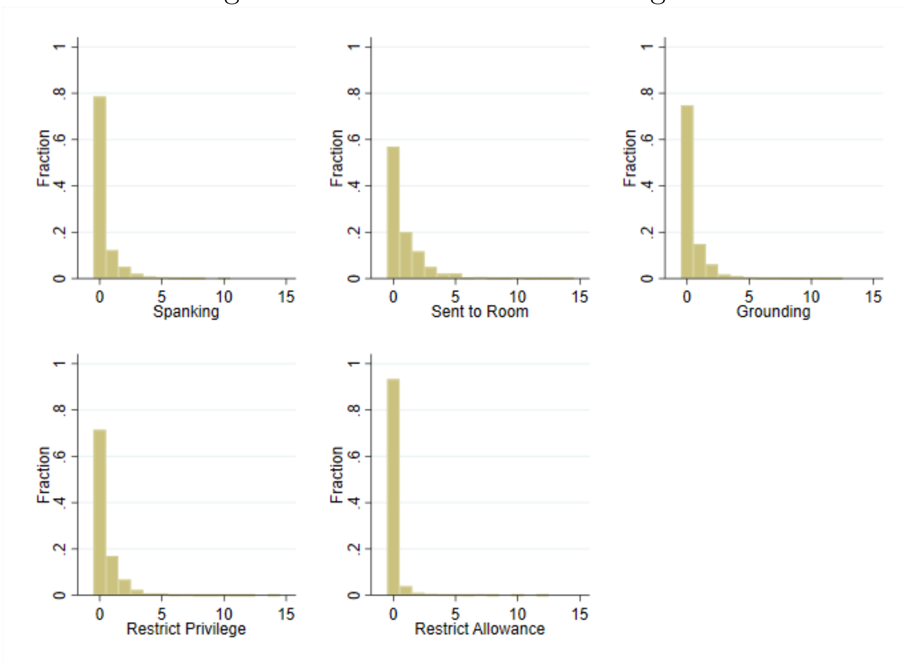


Figure 4.4: Punishment Use at Age 12

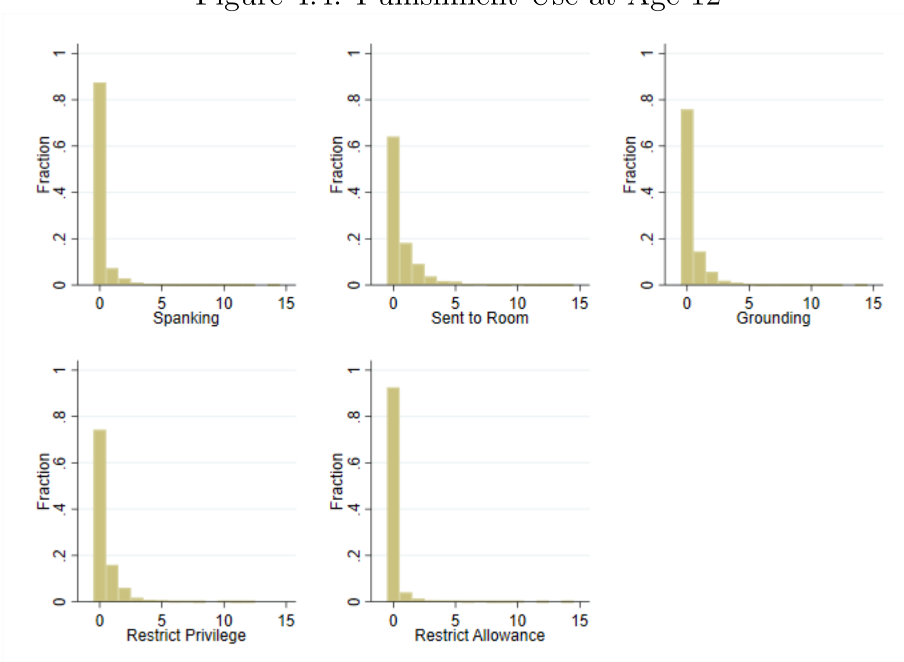


Figure 4.5: Punishment Use at Age 14

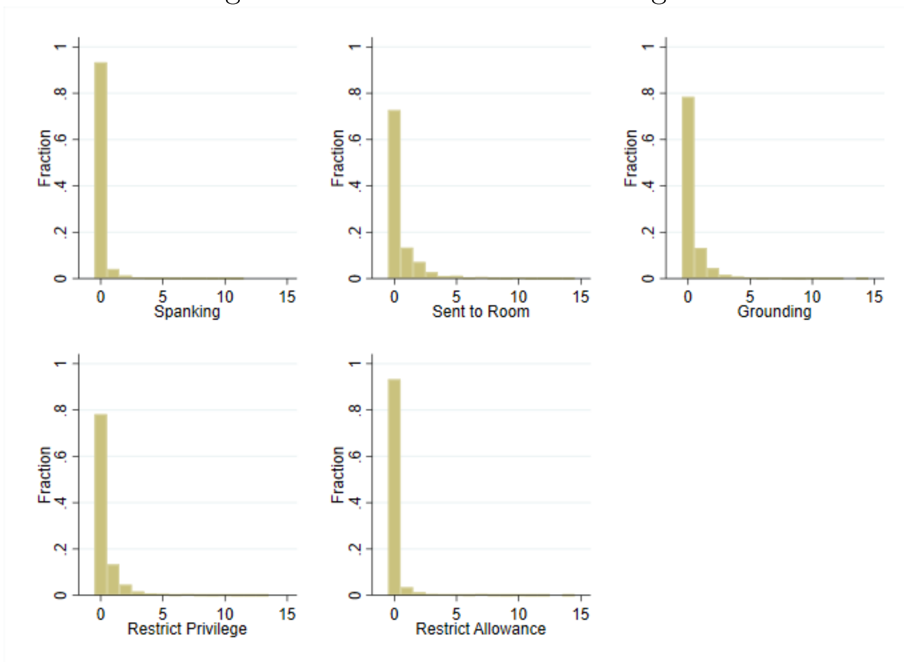


Figure 4.6: Average Punishment Use for Each Child

Figure 4.7: Average Punishment Frequency for Each Child Across 5 Punishment Types

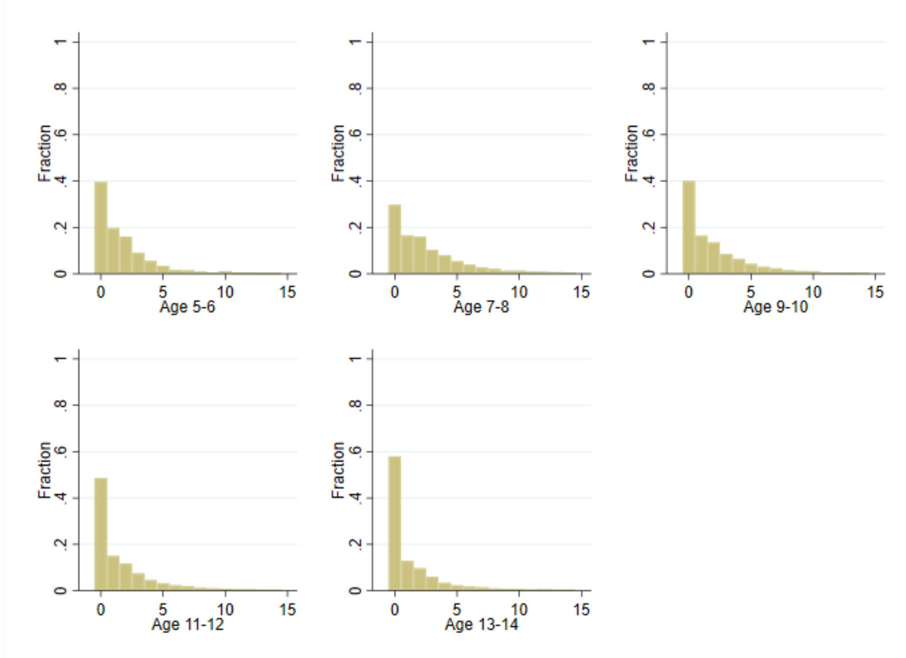


Figure 4.8: Average Punishment Frequency for Each Child Across 5 Punishment Types, Bottom Income Decile

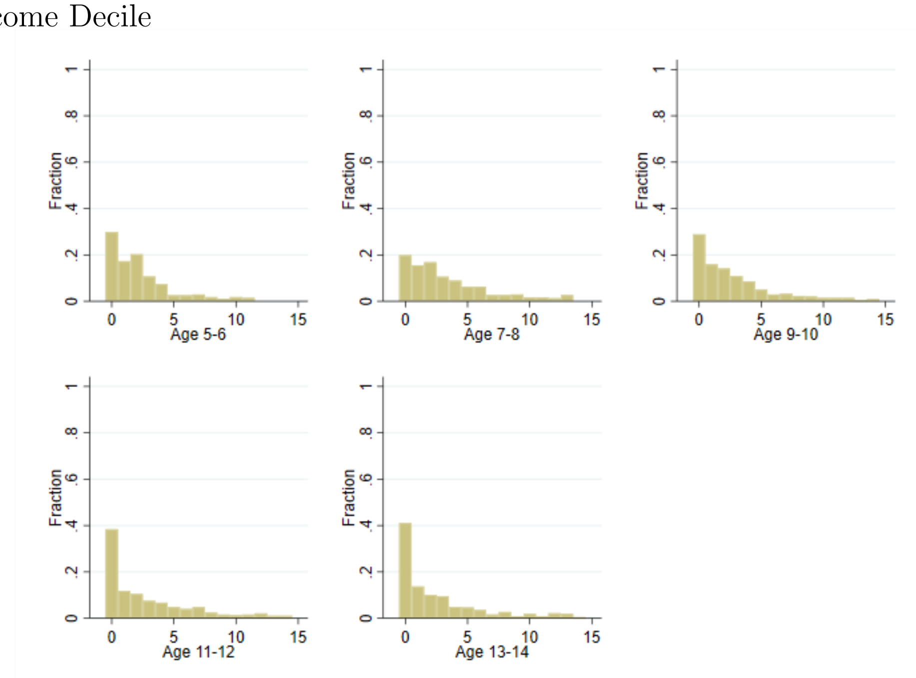


Figure 4.9: Average Punishment Frequency for Each Child Across 5 Punishment Types, Top Income Decile

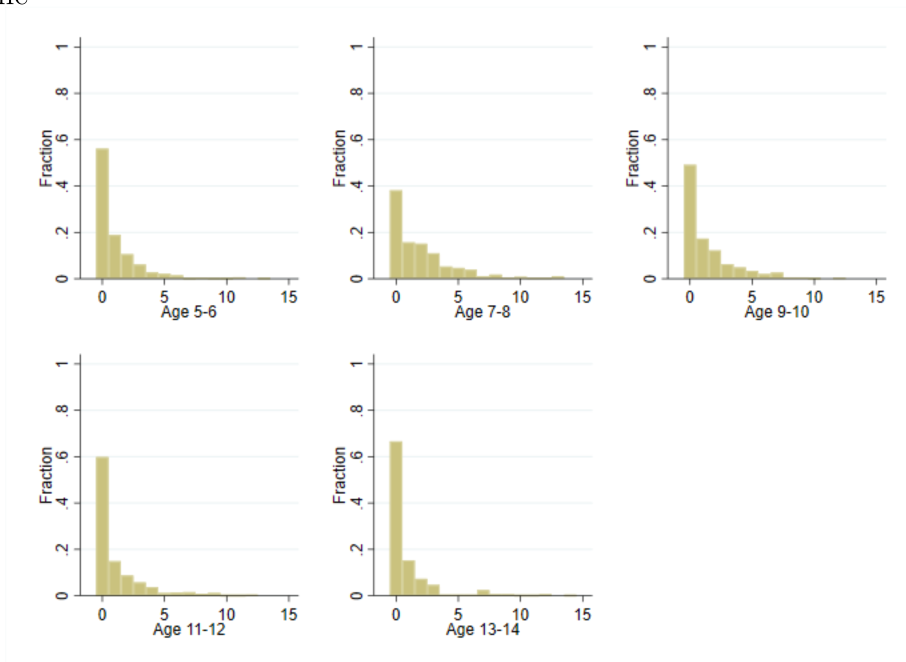
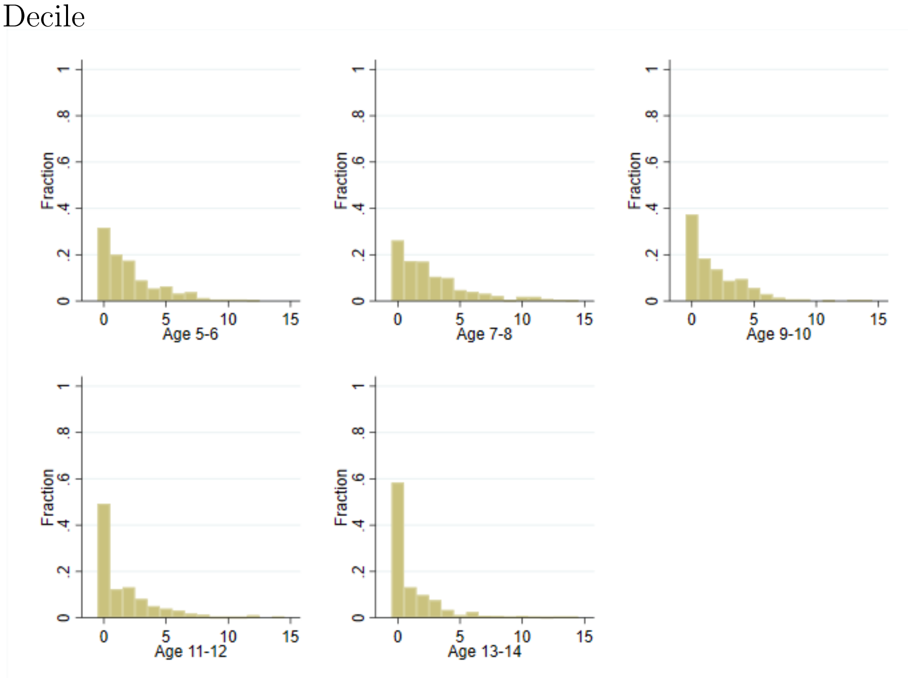


Figure 4.10: Average Punishment Frequency for Each Child Across 5 Punishment Types, 5th Income Decile



In my analysis, I primarily use the binary indicator for whether any type of punishment took place in the past week from the point of the interview, in part because the primary variation is in 0-1 margin. For a binary punishment measure $\mathbb{M}_{i,t}$ for person i at age t, I assume the measure is generated as

$$\mathbb{M}_{i,j,t} = \mathbb{I}\left[\left(M_{i,t} - a_{i,t}\right)^2 - \epsilon_{i,t} > 0\right] \tag{4.2}$$

for mean zero i.i.d error term $\epsilon_{i,t}$. $\epsilon_{i,t}$ captures the noise from the random choice of a particular week from the two-year interaction period, and other measurement noise.

CNLSY data lack direct measures of parenting skill that are present in Section 1. Instead, I use measures of the parent's human capital as measures of parenting skill. While imperfect, the parent's human capital and parenting skill are related for two reasons. First, parenting skill is directly produced by human capital. Parents with more human capital would be more likely to read or learn about good child-rearing practices appropriate for child development, and have relatives or friends with greater parenting skill. Second, the parenting skill of the first generation would be transmitted to the second generation. Parents would primarily imitate the parenting behavior they received from their own parents.[5] Also, high-skilled parenting in the first generation would lead to higher human capital in the second generation (per model implication), leading to higher parenting skill in the second generation. CNLSY data contain the Air Force Qualification Test (AFQT), which was taken by most of the parents in the sample. The test was administered in 1980, when the parents were adolescents.

Model predictions from Section 3 are summarized below:

1. Child misbehavior is more likely to be punished as s_M increases.
2. Child outcome improves in punishment as s_M increases.
3. A child's expectation of his own outcome improves in punishment as s_M increases.

5. Animal studies using rats and observation studies on humans show that mothering behavior is transmitted across generations, even when the mother and the child are not biologically related. See [52].

I test each of these predictions in detail below.

Prediction 1: Child misbehavior is more likely to be punished as s_M increases.

Equation (3.19) shows that, for parents with high skill, punishment is low on average, while the coefficient on child misbehavior is higher. I estimate the following equation for parent-child pair i and age t:

$$\mathbb{M}_{i,t} = \beta_0 + \beta_1 AFQT_i + (\beta_2 + \beta_3 AFQT_i) \times BPI_{i,t} + X_{i,t}\gamma + \epsilon_{i,t} \tag{4.3}$$

$BPI_{i,t}$ corresponds to the deviation of child's behavior from the parent's desired optimal behavior, which is $\left(a_1^{\dagger} - a_1^{*}\right)^2$. $AFQT_i$ equals s_M in this model so $\beta_2+\beta_3 AFQT_i$ corresponds to $(w^*)^2$. Then Prediction 1 is equivalent to $\beta_3 > 0$. I estimate a series of linear probability models, using different measures of punishment in the data. $X_{i,t}$ includes BPI, log household net income,[6] father's presence at home, indicators for race, child's gender, child's birth year, parent's birth year and birth order. The error term is clustered at the household level.[7] I estimate the equation using a punishment indicator and cognitive stimulation scale. The first two columns of Table 4.7 show the estimates from the linear probability model on the punishment indicator, which equals 1 if any one of the five punishments are used on the child. Both the fixed effect and random effect specification show that the coefficient estimate for the interaction of BPI and AFQT is positive and significant. Combined with the positive and significant coefficient for BPI itself, the estimates show that punishment is more likely when the child misbehaves, even more so when the parenting skill is high. This interaction is not observed for the investment model in the third and fourth column. More investment is chosen when the child has greater cognitive skill and shows better behavior, implying that parental investment may be a complement to these skills. I repeat this analysis in Table 4.8

6. Net income is deflated to 2,000 value. I applied inverse hyperbolic sine transformation rather than log to accommodate zero values.

7. Precision of key coefficient estimates do not change meaningfully when I cluster at the state level.

using the years of a parent's education when the child in question is born, and in Table 4.9 using a probit model. The findings are qualitatively similar to the ones using AFQT.

Table 4.7: Parenting

	(1) Pun.	(2) Pun.	(3) Inv.	(4) Inv.
BPIxAFQT	0.018***	0.014*	-0.001	-0.009
	(0.004)	(0.007)	(0.009)	(0.011)
BPI	0.143***	0.111***	-0.104***	-0.036***
	(0.004)	(0.007)	(0.008)	(0.011)
AFQT	-0.004	0.000	0.244***	0.000
	(0.006)	(.)	(0.015)	(.)
PIAT	-0.014**	-0.004	0.113***	0.043***
	(0.005)	(0.008)	(0.009)	(0.011)
Inv.	0.002	0.005		
	(0.004)	(0.006)		
Observations	22686	22686	23906	23906
Model	RE	FE	RE	FE
R2	0.182	0.133	0.334	0.206

Note: Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. Pun. is any punishment between child ages 8 and 14. Inv. Investment is estimated factor for cognitive stimulation scale for ages 8 through 14. Additional controls include log household net income, indicators for father's presence at home, race, child's birth year, parent's birth year, child's gender and child's birth order. AFQT is parent's AFQT score. BPI is scaled so that higher value indicates more noncompliance.

Table 4.10 repeats the analysis for each of the punishment measures in the data using a random effects model, while Table 4.11 uses a fixed effects model. Interestingly, the coefficient estimate for the interaction term is positive only for send-to-room and privilege restrictions, while being negative and significant for spanking, grounding and allowance restrictions. Also, coefficient estimates for AFQT are insignificant but positive for send-to-room and privilege restrictions, while being negative and significant for the other measures. Coefficients for BPI are greater for send-to-room and privilege restrictions than for other punishment methods. It seems that the overall interaction is driven by the parent's choice of using the send-to-room versus other types of punishment.

Table 4.8: Parenting, with Parent's Education Level

	(1) Pun.	(2) Pun.	(3) Inv.	(4) Inv.
BPIxYrsEduc	0.004*	0.002	0.004	0.002
	(0.002)	(0.003)	(0.004)	(0.005)
BPI	0.084***	0.078*	-0.149**	-0.061
	(0.022)	(0.037)	(0.048)	(0.061)
YrsEduc	0.001	0.000	0.110***	0.000
	(0.003)	(.)	(0.006)	(.)
PIAT	-0.012*	-0.005	0.126***	0.046***
	(0.005)	(0.008)	(0.009)	(0.011)
Inv.	-0.000	0.005		
	(0.005)	(0.006)		
Observations	20456	20456	21232	21232
Model	RE	FE	RE	FE
R2	0.181	0.123	0.366	0.224

Note: Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. Pun. is any punishment between child ages 8 and 14. Inv. Investment is estimated factor for cognitive stimulation scale for ages 8 through 14. Additional controls include log household net income, indicators for father's presence at home, race, child's birth year, parent's birth year, child's gender and child's birth order. BPI is scaled so that higher value indicates more noncompliance. Punishment is interacted with the years of the parent's education at the time of teh child's birth.

Table 4.9: Parenting, with Probit Model

	(1) Pun.	(2) Spanking	(3) SentToRm	(4) Privilege	(5) Grounding	(6) Allowance
main						
BPIxAFQT	0.038**	0.007	0.048***	0.017	-0.009	-0.068**
	(0.014)	(0.015)	(0.014)	(0.014)	(0.015)	(0.023)
BPI	0.427***	0.324***	0.355***	0.339***	0.346***	0.241***
	(0.014)	(0.014)	(0.015)	(0.014)	(0.015)	(0.021)
AFQT	-0.006	-0.106***	0.031	0.013	-0.149***	-0.190***
	(0.018)	(0.021)	(0.020)	(0.020)	(0.022)	(0.034)
PIAT	-0.042**	-0.100***	-0.040**	-0.023	-0.058***	-0.056**
	(0.014)	(0.016)	(0.015)	(0.016)	(0.017)	(0.022)
Inv.	0.007	-0.068***	0.054***	0.089***	0.023	0.023
	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)	(0.021)
Observations	22665	22593	19891	19904	19924	19746
R2	0.091	0.119	0.060	0.055	0.109	0.117

Note: Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. Pun. is any punishment between child ages 8 and 14. Additional controls include log household net income, indicators for father's presence at home, race, child's birth year, parent's birth year, child's gender and child's birth order. AFQT is parent's AFQT score. BPI is scaled so that higher value indicates more noncompliance.

Table 4.10: Punishment Measures, RE

	(1)	(2)	(3)	(4)	(5)
	SentToRm	Spanking	Privilege	Grounding	Allowance
BPIxAFQT	0.015**	-0.018***	0.002	-0.023***	-0.024***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.003)
BPI	0.126***	0.092***	0.105***	0.098***	0.034***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)
AFQT	0.013	-0.030***	0.003	-0.041***	-0.023***
	(0.007)	(0.006)	(0.006)	(0.006)	(0.003)
PIAT	-0.016**	-0.030***	-0.008	-0.015***	-0.007*
	(0.005)	(0.004)	(0.005)	(0.004)	(0.003)
Inv.	0.017***	-0.019***	0.026***	0.006	0.003
	(0.005)	(0.004)	(0.004)	(0.004)	(0.003)
Observations	19892	22599	19905	19925	19779
Model	RE	RE	RE	RE	RE
R2	0.123	0.214	0.095	0.171	0.106

Note: Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. (1): sent-to-room; (2): spanking; (3): privilege restriction; (4): grounding; (5): allowance restriction. These are indicators for any use between child ages 8 and 14. Inv. Investment is estimated factor for cognitive stimulation scale for ages 8 through 14. Additional controls include log household net income, indicators for father's presence at home, race, child's birth year, parent's birth year, child's gender and child's birth order. AFQT is parent's AFQT score. BPI is scaled so that higher value indicates more noncompliance.

Table 4.11: Punishment Measures, FE

	(1) SentToRm	(2) Spanking	(3) Privilege	(4) Grounding	(5) Allowance
BPIxAFQT	0.002	-0.017**	0.005	-0.017*	-0.010
	(0.008)	(0.007)	(0.008)	(0.007)	(0.006)
BPI	0.096***	0.064***	0.077***	0.077***	0.028***
	(0.008)	(0.007)	(0.008)	(0.007)	(0.005)
AFQT	0.000	0.000	0.000	0.000	0.000
	(.)	(.)	(.)	(.)	(.)
PIAT	-0.015	-0.029***	-0.003	-0.002	0.001
	(0.009)	(0.008)	(0.008)	(0.007)	(0.005)
Inv.	0.012	-0.009	0.017**	0.002	0.003
	(0.007)	(0.006)	(0.006)	(0.006)	(0.004)
Observations	19892	22599	19905	19925	19779
Model	FE	FE	FE	FE	FE
R2	0.087	0.088	0.067	0.107	0.050

Note: Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. (1): sent-to-room; (2): spanking; (3): privilege restriction; (4): grounding; (5): allowance restriction. These are indicators for any use between child ages 8 and 14. Inv. Investment is estimated factor for cognitive stimulation scale for ages 8 through 14. Additional controls include log household net income, indicators for father's presence at home, race, child's birth year, parent's birth year, child's gender and child's birth order. AFQT is parent's AFQT score. BPI is scaled so that higher value indicates more noncompliance.

The analysis here assumes that observed punishment is primarily initiated by the parent's perception of the child's misbehavior. Data do not contain measures for the reason of punishment, but do contain measures of how the parent would respond to the child's tantrum. Tables 4.12 and 4.13 estimate equation (4.3) without the interaction term. Between the parents who responded they would use the send-to-room and parents who responded they would use spanking, the primary difference is that the constant term estimate is much higher for the punishment method that the parent responded to use. Coefficient on BPI is similar across the models. Thus, it seems that the parent indeed uses these punishments in response to child misbehavior, but BPI does not capture all aspects of it.

Table 4.12: Send-to-Room Use

	(1)	(2)	(3)
PIAT	-0.013**	-0.013	0.002
	(0.004)	(0.008)	(0.009)
BPI	-0.121***	-0.131***	-0.110***
	(0.004)	(0.007)	(0.008)
Constant	0.473***	0.662***	0.507***
	(0.025)	(0.044)	(0.053)
Observations	8149	3676	3120
R2	0.088	0.086	0.053

Model (1) is for the entire sample. Model (2) is for the parents who uses send-to-room in response to tantrum. Model (3) is for the parents who uses spanking in response to tantrum. Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. SRM: mother responds to child tantrum with . SPK: mother responds to tantrum with .

Prediction 2: Child outcome improves in punishment as s_M increases

To test the effect of parental punishment on child outcomes, I first examine the child's college attendance at age 22 and log wage earnings at age 26. I estimate

$$y_{i,j} = \beta_0 + \beta_1 AFQT_i + (\beta_2 + \beta_3 AFQT_i) \times Punish_i + X_i\gamma + \epsilon_{i,j} \tag{4.4}$$

Table 4.13: Spanking Use

	(1)	(2)	(3)
PIAT	-0.055***	-0.039***	-0.014
	(0.003)	(0.007)	(0.008)
BPI	-0.107***	-0.106***	-0.135***
	(0.003)	(0.006)	(0.007)
Constant	0.394***	0.373***	0.552***
	(0.020)	(0.037)	(0.051)
Observations	8444	3679	3140
R2	0.157	0.098	0.102

Model (1) is for the entire sample. Model (2) is for the parents who uses send-to-room in response to tantrum. Model (3) is for the parents who uses spanking in response to tantrum. Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. SRM: mother responds to child tantrum with . SPK: mother responds to tantrum with .

where X includes punishment variables, indicators for race, child gender, and child's birth year, parent's birth year, log household income, father's presence at home, parental investment and its interaction with AFQT. Error term is clustered at the household level. Household income, father's presence at home, and parental investment measures are averaged within each child for ages 8 through 14. Punishment equals 1 if any punishment takes place between ages 8 and 14. $\beta_2 + \beta_3 AFQT_i$ corresponds to the effect of punishment on the child's outcome through ρ_2, which determines the child's effort in $t = 2$. We therefore expect $\beta_3 > 0$.

Table 4.14 shows that, consistent with the model's prediction, we see that the coefficient for punishment-AFQT is positive, large, and significant for all the models considered. Punishment does not affect college attendance at age 22 for the parents with a high AFQT score, but the children are less likely to attend college if they are punished by parents with a low AFQT score. Punishment predicts a higher wage at age 26 if the parent has a high AFQT score, while predicting lower wage when the parent has a low score.

Tables 4.15, 4.16, and 4.17 repeat the analysis for subsamples based on high and low levels of household income, child sex and race subsamples. Hispanic families stand out from

Table 4.14: Child Outcomes

	(1) lnWage	(2) lnWage	(3) lnWage	(4) College	(5) College	(6) College
Punishment	-0.392	-0.307	-0.334	-0.182***	-0.189***	-0.211***
	(0.279)	(0.301)	(0.341)	(0.042)	(0.047)	(0.047)
Pun.xAFQT	0.789**	0.746*	0.652*	0.158***	0.172***	0.153***
	(0.283)	(0.310)	(0.329)	(0.037)	(0.040)	(0.045)
Investment	0.512**	0.527**	0.408*	0.132***	0.124***	0.136***
	(0.156)	(0.170)	(0.169)	(0.020)	(0.022)	(0.024)
Inv.xAFQT	-0.024	-0.080	-0.048	0.030*	0.031	0.041*
	(0.147)	(0.159)	(0.174)	(0.014)	(0.017)	(0.018)
AFQT	-0.619	-0.620	-0.513	-0.075*	-0.095*	-0.098*
	(0.321)	(0.344)	(0.348)	(0.036)	(0.040)	(0.042)
Observations	1299	1166	1166	807	734	734
R^2	0.091	0.129	0.177	0.349	0.380	0.434
Adjusted R^2	0.073	0.076	0.062	0.331	0.323	0.312
Model	None	State FE	County FE	None	State FE	County FE

Note: Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. lnWage is log (inverse hyperbolic sine transformation) wage earnings at age 26. College is college attendance status at age 22. Additional controls include indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used at ages 8 through 14. Investment is estimated factor for cognitive stimulation scale for ages 8 through 14. AFQT is parent's AFQT score.

other groups, showing inconsistent interaction coefficient estimates. Table 4.18 excludes spanking in the construction of $Punish_i$ variable. The results are mostly the same as the baseline model from equation (4.4) but slightly weaker, suggesting the effect of spanking also depends on the parent's human capital. Table 4.19 includes a full set of interactions among investment, punishment and all the other household variables in the equation. The magnitudes of the coefficients do not change greatly and the punishment-AFQT interaction estimates remain significant for the college attendance model, but they decrease in magnitude and precision for the wage model. Investment now plays a much more important role in predicting wage outcome. The model seems to be a better explanation for the education outcome than for the wage outcome.[8]

Table 4.15: Child Outcomes, by Household Income

	(1)	(2)	(3)	(4)
	lnWage	lnWage	College	College
Punishment	0.127	-0.567	-0.253***	-0.005
	(0.436)	(0.675)	(0.063)	(0.106)
Pun.xAFQT	0.569	0.391	0.175**	0.213*
	(0.408)	(0.794)	(0.056)	(0.107)
AFQT	-0.424	-0.284	-0.107*	-0.005
	(0.425)	(0.755)	(0.054)	(0.101)
Observations	568	598	429	305
R^2	0.243	0.264	0.536	0.480
Adjusted R^2	0.026	0.059	0.365	0.189
Model	County FE	County FE	County FE	County FE
Income	Above Median	Below Median	Above Median	Below Median

Note: Standard errors are in parenthesis, clustered at the household level. + p¡.1 * p¡.05, ** p¡.01, *** p¡.001. lnWage is log (inverse hyperbolic sine transformation) wage earnings at age 26. College is college attendance status at age 22. Additional controls include log net household income, father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used at ages 8 through 14. Investment is estimated factor for cognitive stimulation scale for ages 8 through 14. AFQT is parent's AFQT score.

8. One possible reason is the high job mobility of young workers in the 20s. Wage measured after age 30 would be a better measure of life cycle outcome.

Table 4.16: Child Outcomes, by Gender

	(1) lnWage	(2) lnWage	(3) College	(4) College
Punishment	0.154	-0.445	-0.170	-0.249***
	(0.651)	(0.494)	(0.126)	(0.059)
Pun.xAFQT	0.153	0.723	0.181+	0.116+
	(0.566)	(0.479)	(0.101)	(0.062)
AFQT	-0.350	-0.394	-0.127	-0.058
	(0.549)	(0.533)	(0.092)	(0.065)
Observations	576	590	345	389
R^2	0.225	0.292	0.486	0.580
Adjusted R^2	0.014	0.090	0.241	0.390
Model	County FE	County FE	County FE	County FE
Sex	Male	Female	Male	Female

Note: Standard errors are in parenthesis, clustered at the household level. + p¡.1 * p¡.05, ** p¡.01, *** p¡.001. lnWage is log (inverse hyperbolic sine transformation) wage earnings at age 26. College is college attendance status at age 22. Additional controls include log net household income, father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used at ages 8 through 14. Investment is estimated factor for cognitive stimulation scale for ages 8 through 14. AFQT is parent's AFQT score.

Table 4.17: Child Outcomes, by Race

	(1) lnWage	(2) lnWage	(3) lnWage	(4) College	(5) College	(6) College
Punishment	0.313	-1.368	-0.673	-0.216**	0.075	-0.214
	(0.535)	(0.999)	(1.092)	(0.078)	(0.174)	(0.162)
Pun.xAFQT	0.122	0.337	-0.602	0.149*	0.253	0.004
	(0.456)	(1.203)	(1.225)	(0.068)	(0.253)	(0.158)
AFQT	0.315	-0.831	0.463	-0.102	-0.129	0.040
	(0.465)	(1.218)	(1.149)	(0.063)	(0.238)	(0.128)
Observations	582	355	229	414	195	125
R^2	0.254	0.336	0.443	0.540	0.581	0.550
Adjusted R^2	0.066	0.044	0.142	0.369	0.211	0.114
Model	County FE	County FE	County FE	County FE	County FE	County FE
Race	White	Black	Hispanic	White	Black	Hispanic

Note: Standard errors are in parenthesis, clustered at the household level. + p¡.1 * p¡.05, ** p¡.01, *** p¡.001. lnWage is log (inverse hyperbolic sine transformation) wage earnings at age 26. College is college attendance status at age 22. Additional controls include log net household income, father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used at ages 8 through 14. Investment is estimated factor for cognitive stimulation scale for ages 8 through 14. AFQT is parent's AFQT score.

Table 4.18: Child Outcomes, with Non-corporal Punishment

	(1) lnWage	(2) lnWage	(3) lnWage	(4) College	(5) College	(6) College
Non-corp. Pun.	-0.244	-0.122	-0.147	-0.147***	-0.156***	-0.182***
	(0.276)	(0.294)	(0.333)	(0.041)	(0.045)	(0.046)
Pun.xAFQT	0.629*	0.563	0.483	0.116**	0.133***	0.118**
	(0.281)	(0.303)	(0.324)	(0.036)	(0.039)	(0.044)
Investment	0.515***	0.531**	0.408*	0.133***	0.126***	0.136***
	(0.155)	(0.170)	(0.168)	(0.020)	(0.022)	(0.024)
Inv.xAFQT	-0.038	-0.097	-0.061	0.028	0.029	0.040*
	(0.146)	(0.157)	(0.173)	(0.015)	(0.017)	(0.018)
AFQT	-0.459	-0.443	-0.350	-0.034	-0.056	-0.064
	(0.310)	(0.328)	(0.331)	(0.035)	(0.038)	(0.041)
Observations	1299	1166	1166	807	734	734
R^2	0.090	0.128	0.176	0.343	0.374	0.429
Adjusted R^2	0.072	0.075	0.060	0.324	0.316	0.306
Model	None	State FE	County FE	None	State FE	County FE

Note: Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. lnWage is log (inverse hyperbolic sine transformation) wage earnings at age 26. College is college attendance status at age 22. Additional controls include log net household income, father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used at ages 8 through 14 other than spanking. Investment is estimated factor for cognitive stimulation scale for ages 8 through 14. AFQT is parent's AFQT score.

Table 4.19: Child Outcomes, with More Interactions

	(1) lnWage	(2) lnWage	(3) lnWage	(4) College	(5) College	(6) College
Punishment	-1.604	0.277	0.039	-0.385	0.000	0.078
	(5.672)	(7.128)	(7.774)	(0.761)	(.)	(0.942)
Pun.xAFQT	0.500	0.167	0.075	0.172**	0.200***	0.147*
	(0.376)	(0.430)	(0.433)	(0.052)	(0.055)	(0.063)
Investment	3.986*	4.113*	4.171	-0.270	-0.025	0.000
	(1.864)	(2.027)	(2.179)	(0.205)	(0.251)	(.)
Inv.xAFQT	0.128	0.020	0.110	-0.001	0.009	0.009
	(0.199)	(0.207)	(0.242)	(0.019)	(0.022)	(0.023)
AFQT	-0.381	-0.101	0.006	-0.080	-0.115*	-0.088
	(0.394)	(0.435)	(0.430)	(0.047)	(0.050)	(0.057)
Observations	1299	1166	1166	807	734	734
R^2	0.130	0.174	0.224	0.403	0.432	0.488
Adjusted R^2	0.082	0.089	0.078	0.358	0.347	0.341
Model	None	State FE	County FE	None	State FE	County FE

Note: Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. lnWage is log (inverse hyperbolic sine transformation) wage earnings at age 26. College is college attendance status at age 22. Additional controls include log net household income, father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used at ages 8 through 14. Investment is estimated factor for cognitive stimulation scale for ages 8 through 14. AFQT is parent's AFQT score. Both investment and punishment are interacted with AFQT, log household income, father's presence at home, race, child's gender, child's birthyear and parent's birthyear.

I also examine the effect of punishment and its interaction in the dynamic model of skill formation. For cognitive skill factor $\theta_{i,t}^{C}$ and behavioral skill factor $\theta_{i,t}^{B}$, I estimate

$$\theta_{i,t+1}^{k} = \beta_0 + \beta_1 AFQT_i + (\beta_2 + \beta_3 AFQT_i) \times Punish_{i,t} + X_{i,t}\gamma + \epsilon_{i,t},\ k \in \{C, B\} \quad (4.5)$$

using RE and FE models. The FE model is estimated using the Arellano-Bond GMM estimator to avoid bias in the dynamic fixed effect model. $X_{i,t}$ includes punishment variables, lagged cognitive and behavioral skill factors, indicators for race, child gender and the child's birth year, the parent's birth year, log household income, father's presence at home, parental investment and its interaction with AFQT score. Table 4.20 shows that the model prediction is confirmed for cognitive skill model. β_3 estimate is positive and significant, suggesting that the effect of punishment on cognitive skill development is positive for parents with a high AFQT score, and negative otherwise. The interaction coefficient estimate is positive but insignificant for the behavioral model. Interestingly, the coefficient estimate for punishment itself is negative and significant for the RE model, but *positive* and significant for the FE model. It seems that household-specific time-invariant unobserved characteristics play an important role in determining both more punishment and more problematic behavior in childhood. Accounting for the unobserved characteristics, parental punishment, on average, reduced problematic behavior of the child as perceived by the mother. The fact that this is not observed in the case of cognitive skill may mean BPI captures only a small part of child behavior that are affected by punishment and generating cognitive skill.

Tables 4.21, 4.22, and 4.23 replicate these results using different measures of punishment, different measures of the parent's human capital and different measures of cognitive and behavioral skill. Tables 4.24, 4.25 and 4.26 repeat the analysis for subsamples based on high and low levels of household income, child's sex, and race. The results are qualitatively similar across racial groups, but stronger for girls and for those who are below average in household income. Table 4.27 excludes spanking in the construction of $Punish_{i,t}$, and the estimates are

Table 4.20: Child Human Capital Development

	(1) PIAT	(2) PIAT	(3) BPI	(4) BPI
L.punish	0.012	0.022^{+}	-0.063^{***}	0.091^{***}
	(0.009)	(0.013)	(0.012)	(0.017)
L.Pun.xAFQT	0.024^{**}	0.030^{*}	0.003	0.011
	(0.009)	(0.012)	(0.012)	(0.016)
L.Inv.	0.039^{***}	-0.016^{+}	0.049^{***}	-0.010
	(0.005)	(0.008)	(0.008)	(0.011)
L.PIAT	0.566^{***}	0.323^{***}	0.051^{***}	-0.022
	(0.008)	(0.021)	(0.008)	(0.014)
L.BPI	0.045^{***}	0.001	0.511^{***}	0.181^{***}
	(0.006)	(0.010)	(0.010)	(0.026)
Observations	6689	4416	6777	4383
Model	RE	FE	RE	FE
r2_b	0.781		0.658	

Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. + p¡.1, * p¡.05, ** p¡.01, *** p¡.001. FE estimates Arellano-Bond GMM estimator. PIAT is latent factor estimate from Peabody Individual Achievement Test scores. BPI is latent factor estimate from Behavior Problems Index, scaled so that higher value indicates better behavior. Additional controls include lag of skills, lag of log (inverse hyperbolic sine transformation) household net income, lag of father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used . Investment is cognitive stimulation scale from HOME score. AFQT is parent's AFQT score.

very close to those from equation (4.5). Table 4.28 includes a full set of interactions among investment, punishment and all the other household variables in the equation. The results lose magnitude and statistical significance for the child behavior model. Both investment and punishment coefficients and their interactions lose precision in the behavior model, while only coefficients for lagged skill measures remain significant. However, the results are consistent with those in the baseline estimation for the human capital development model.

Table 4.21: Child Human Capital Development with Different Punishment Measures

	(1) PIAT	(2) PIAT	(3) BPI	(4) BPI
L.Moderate Pun.	0.011	0.017	-0.038**	0.079***
	(0.010)	(0.014)	(0.013)	(0.018)
L.MPxAFQT	0.036***	0.042**	0.009	0.017
	(0.010)	(0.014)	(0.014)	(0.018)
L.Corporal Pun.	-0.022	0.001	-0.085***	0.018
	(0.013)	(0.017)	(0.017)	(0.022)
L.CPxAFQT	-0.002	-0.027	-0.034	-0.038
	(0.014)	(0.018)	(0.018)	(0.023)
L.Inv.	0.042***	-0.014	0.046***	-0.017
	(0.006)	(0.010)	(0.009)	(0.012)
L.PIAT	0.632***	0.393***	0.060***	-0.013
	(0.009)	(0.029)	(0.009)	(0.017)
L.BPI	0.044***	0.002	0.513***	0.170***
	(0.006)	(0.011)	(0.011)	(0.029)
Observations	6270	3895	6333	3876
Model	RE	FE	RE	FE
R2	0.814		0.644	

Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. FE estimates Arellano-Bond GMM estimator. PIAT is latent factor estimate from Peabody Individual Achievement Test scores. BPI is latent factor estimate from Behavior Problems Index, scaled so that higher value indicates better behavior. Additional controls include lag of skills, lag of log (inverse hyperbolic sine transformation) household net income, lag of father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Moderate Punishment is indicator for and Corporal Punishment is indicator for spanking used . Investment is cognitive stimulation scale from HOME score. AFQT is parent's AFQT score.

Table 4.22: Child Human Capital Development with Parent's Education Level

	(1) PIAT	(2) PIAT	(3) BPI	(4) BPI
L.punish	-0.062	-0.068	-0.143	-0.077
	(0.054)	(0.075)	(0.081)	(0.095)
L.punishXHGCmom	0.006	0.007	0.007	0.014
	(0.004)	(0.006)	(0.006)	(0.007)
Parent's Yrs of Educ	0.040***		0.010	
	(0.004)		(0.006)	
L.Inv.	0.041***	-0.019*	0.044***	-0.011
	(0.006)	(0.009)	(0.008)	(0.011)
L.PIAT	0.579***	0.346***	0.051***	-0.011
	(0.009)	(0.022)	(0.008)	(0.015)
L.BPI	0.042***	-0.001	0.504***	0.183***
	(0.006)	(0.010)	(0.010)	(0.026)
Observations	5947	4070	6039	4029
Model	RE	FE	RE	FE
R2	0.787		0.652	

Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. FE estimates Arellano-Bond GMM estimator. PIAT is latent factor estimate from Peabody Individual Achievement Test scores. BPI is latent factor estimate from Behavior Problems Index, scaled so that higher value indicates better behavior. Additional controls include lag of skills, lag of log (inverse hyperbolic sine transformation) household net income, lag of father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Moderate Punishment is indicator for and Corporal Punishment is indicator for spanking used . Investment is cognitive stimulation scale from HOME score. AFQT is parent's AFQT score.

Prediction 3: The child's expectation of his own outcome increases in punishment as s_M increases

For this test, I use the measures in CNLSY that ask the children "Looking ahead, how far do you think you will go in school?" Responses include "leave high school before graduation", "graduate from high school", "get some college or other training", "graduate from college", and "take further training after college". Responses are collected for children between ages

Table 4.23: Child Human Capital Development with Different Human Capital Measures

	(1) Math	(2) Math	(3) Headstrong	(4) Headstrong
L.punish	0.004	0.036*	-0.130***	0.087***
	(0.007)	(0.014)	(0.014)	(0.019)
L.Pun.xAFQT	0.018**	0.018	-0.006	0.026
	(0.007)	(0.014)	(0.014)	(0.018)
L.Inv.	0.028***	-0.008	0.042***	-0.017
	(0.004)	(0.009)	(0.008)	(0.012)
L.Math	0.683***	0.688***	0.018**	0.007
	(0.005)	(0.007)	(0.006)	(0.009)
L.Headstrong	0.016***	-0.011	0.447***	0.121***
	(0.004)	(0.009)	(0.009)	(0.021)
Observations	6946	4870	7104	4976
Model	RE	FE	RE	FE
R2	0.790		0.603	

Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. FE estimates Arellano-Bond GMM estimator. Math is PIAT math test score, standardized. Headstrong is BPI subscale, standardized and rescaled so that higher value indicates better behavior. Additional controls include lag of skills, lag of log (inverse hyperbolic sine transformation) household net income, lag of father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used . Investment is cognitive stimulation scale from HOME score. AFQT is parent's AFQT score.

Table 4.24: Child Human Capital Development, by Household Income

	(1) PIAT	(2) PIAT	(3) BPI	(4) BPI
L.punish	0.025	0.056*	0.078***	0.133***
	(0.016)	(0.025)	(0.020)	(0.039)
L.Pun.xAFQT	0.017	0.084**	0.019	0.038
	(0.015)	(0.028)	(0.018)	(0.044)
Observations	3060	1356	3054	1329
Model	FE	FE	FE	FE
Income	Above Median	Below Median	Above Median	Below Median

Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. + p¡.1, * p¡.05, ** p¡.01, *** p¡.001. FE estimates Arellano-Bond GMM estimator. PIAT is latent factor estimate from Peabody Individual Achievement Test scores. BPI is latent factor estimate from Behavior Problems Index, scaled so that higher value indicates better behavior. Additional controls include lag of skills, lag of log (inverse hyperbolic sine transformation) household net income, lag of father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used . Investment is cognitive stimulation scale from HOME score. AFQT is parent's AFQT score.

Table 4.25: Child Human Capital Development, by Gender

	(1) PIAT	(2) PIAT	(3) BPI	(4) BPI
L.punish	0.009	0.034*	0.079**	0.103***
	(0.019)	(0.017)	(0.024)	(0.023)
L.Pun.xAFQT	0.025	0.036*	0.001	0.022
	(0.018)	(0.017)	(0.023)	(0.022)
Observations	2182	2234	2174	2209
Model	FE	FE	FE	FE
Sex	Male	Female	Male	Female

Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. + p¡.1, * p¡.05, ** p¡.01, *** p¡.001. FE estimates Arellano-Bond GMM estimator. PIAT is latent factor estimate from Peabody Individual Achievement Test scores. BPI is latent factor estimate from Behavior Problems Index, scaled so that higher value indicates better behavior. Additional controls include lag of skills, lag of log (inverse hyperbolic sine transformation) household net income, lag of father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used . Investment is cognitive stimulation scale from HOME score. AFQT is parent's AFQT score.

Table 4.26: Child Human Capital Development, by Race

	(1) PIAT	(2) PIAT	(3) PIAT	(4) BPI	(5) BPI	(6) BPI
L.punish	0.019	0.026	0.037	0.109***	0.050	0.040
	(0.021)	(0.029)	(0.035)	(0.025)	(0.046)	(0.043)
L.Pun.xAFQT	0.025	0.016	0.083*	0.015	-0.023	-0.051
	(0.018)	(0.033)	(0.041)	(0.021)	(0.052)	(0.050)
Observations	2284	1279	853	2307	1249	827
Model	FE	FE	FE	FE	FE	FE
Race	White	Black	Hispanic	White	Black	Hispanic

Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. + p¡.1, * p¡.05, ** p¡.01, *** p¡.001. FE estimates Arellano-Bond GMM estimator. PIAT is latent factor estimate from Peabody Individual Achievement Test scores. BPI is latent factor estimate from Behavior Problems Index, scaled so that higher value indicates better behavior. Additional controls include lag of skills, lag of log (inverse hyperbolic sine transformation) household net income, lag of father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used . Investment is cognitive stimulation scale from HOME score. AFQT is parent's AFQT score.

9 and 14, although not many measures are collected for children at age 9. The mother's expectation is available for child ages 5 through 14. Tables 4.29 and 4.30 show that most variation occurs when a respondent thinks the child will be able to graduate college. I estimate the equation

$$ChildExpectation_{i,t} = \beta_0 + \beta_1 AFQT_i + (\beta_2 + \beta_3 AFQT_i) \times Punish_t + X_{i,t}\gamma + \epsilon_{i,t} \quad (4.6)$$

where $ChildExpectation_{i,t}$ is a binary variable that equals 1 if the child's answer is "graduate from college" or "take further training after college". $X_{i,t}$ includes the same set of variables as in equation (4.5).

Table 4.31 shows that only the child's expectation responds to punishment-AFQT interaction. The estimate is positive and significant in the RE model, but not significant in the the FE model. They are close to zero and insignificant in the models for mother's ex-

Table 4.27: Child Human Capital Development with Non-corporal Punishment

	(1) PIAT	(2) PIAT	(3) BPI	(4) BPI
L.non-Corp. Pun.	0.010	0.022	-0.052***	0.089***
	(0.009)	(0.014)	(0.013)	(0.018)
L.Pun.xAFQT	0.033***	0.038**	0.001	0.010
	(0.009)	(0.014)	(0.013)	(0.017)
L.Inv.	0.042***	-0.012	0.048***	-0.016
	(0.006)	(0.010)	(0.009)	(0.012)
L.PIAT	0.632***	0.390***	0.060***	-0.014
	(0.009)	(0.029)	(0.008)	(0.017)
L.BPI	0.045***	0.001	0.516***	0.165***
	(0.006)	(0.011)	(0.011)	(0.028)
Observations	6299	3923	6366	3902
Model	RE	FE	RE	FE
R2	0.814		0.645	

Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. + p¡.1, * p¡.05, ** p¡.01, *** p¡.001. FE estimates Arellano-Bond GMM estimator. PIAT is latent factor estimate from Peabody Individual Achievement Test scores. BPI is latent factor estimate from Behavior Problems Index, scaled so that higher value indicates better behavior. Additional controls include lag of skills, lag of log (inverse hyperbolic sine transformation) household net income, lag of father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used other than spanking. Investment is cognitive stimulation scale from HOME score. AFQT is parent's AFQT score.

Table 4.28: Child Human Capital Development with More Interactions

	(1) PIAT	(2) PIAT	(3) BPI	(4) BPI
L.punish	-0.138^{+}	-0.079	-0.035	-0.008
	(0.074)	(0.114)	(0.113)	(0.146)
L.Pun.xAFQT	0.019	0.027^{+}	-0.005	0.003
	(0.012)	(0.016)	(0.015)	(0.020)
L.Inv.	0.068^{+}	-0.025	-0.024	-0.012
	(0.038)	(0.057)	(0.053)	(0.072)
L.Inv.xAFQT	-0.009	-0.010	-0.012	0.014
	(0.007)	(0.011)	(0.010)	(0.014)
L.PIAT	0.560***	0.329***	0.032***	-0.028*
	(0.009)	(0.022)	(0.008)	(0.014)
L.BPI	0.039***	0.001	0.495***	0.171***
	(0.006)	(0.010)	(0.010)	(0.026)
Observations	6185	4183	6196	4091
Model	RE	FE	RE	FE
R2	0.783		0.674	

Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. + p¡.1, * p¡.05, ** p¡.01, *** p¡.001. FE estimates Arellano-Bond GMM estimator. PIAT is latent factor estimate from Peabody Individual Achievement Test scores. BPI is latent factor estimate from Behavior Problems Index, scaled so that higher value indicates better behavior. Additional controls include lag of skills, lag of log (inverse hyperbolic sine transformation) household net income, lag of father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used . Investment is cognitive stimulation scale from HOME score. AFQT is parent's AFQT score. Both investment and punishment are interacted with AFQT, log household income, father's presence at home, race, child's gender, child's birthyear and parent's birthyear.

pectation compared to the models for child's expectation. Although the interaction term is insignificant in the FE model, comparing columns 2 and 4 shows there is a difference between the interaction coefficients for the mother and the child.

Table 4.29: Child's Expectation

Child's Age	9-10	11-12	13-14	Total
less than HS	35	115	104	254
HS grad	323	1,004	934	2,261
some college	171	498	558	1,227
college grad	803	2,846	2,569	6,218
more than college	465	1,790	2,071	4,326
Total	1,797	6,253	6,236	14,286

Source: CNLSY.

Table 4.30: Parent's Expectation

Child's Age	5-6	7-8	9-10	11-12	13-14	Total
less than HS	14	26	56	97	117	310
HS grad	184	363	652	1,194	1,111	3,504
some college	245	430	797	1,217	1,161	3,850
college grad	993	1,459	2,024	2,717	2,279	9,472
more than college	311	480	705	893	768	3,157
Total	1,747	2,758	4,234	6,118	5,436	20,293

Source: CNLSY.

3 Alternative Explanations

The signaling model says that while the parent may choose to punish or not, the effect of punishment is determined by the quality of punishment, which is measured by the harshness and inconsistency in parenting. The quality of punishment is determined by parenting skill, which is the parent's ability to manage parent-child conflict so that signaling noise is

Table 4.31: Subjective Expectations

	(1) Mother	(2) Mother	(3) Child	(4) Child
L.Pun.xAFQT	0.011	-0.001	0.023**	0.014
	(0.008)	(0.011)	(0.009)	(0.014)
L.punish	0.009	0.011	-0.000	-0.007
	(0.009)	(0.011)	(0.010)	(0.016)
AFQT	0.087***	0.000	0.027**	0.000
	(0.009)	(.)	(0.008)	(.)
Observations	11155	11155	9371	9371
Model	RE	FE	RE	FE
R2	0.302	0.026	0.154	0.068

Note: Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. Outcome variable: subjective expectation that child will at least graduate 4 yr college. Additional controls include lagged cognitive and behavior factors, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used at ages 8 through 14. Investment is estimated factor for cognitive stimulation scale for ages 8 through 14. AFQT is parent's AFQT score.

minimized. The model predicts that punishment, used consistently in response to child misbehavior, should have a better effect on child outcomes than punishment used inconsistently. This section presents alternative models of parent-child interaction. I show that modeling parental punishment as costly incentive often leads to predictions that are inconsistent with the empirical evidence presented in this paper.

A model of guided learning The signaling model presented in this paper can be considered an extension of the model by [50]. In that work, the authors analyzed a model of guided learning, where the child's goal was to learn the mean parameter of a normal random variable through Bayesian updating from quadratic loss function. The child lives for multiple periods, where in each period he draws from i.i.d random variables $X_t \sim \mathcal{N}\left(M,\ p_X^{-1}\right)$. The child's period utility is $-\left(X_t + \bar{a}_t - \bar{b}_t\right)^2$, where the child chooses action $\bar{b}_t$ and the parent chooses action $\bar{a}_t$ in each period. The child does not know the value of M but has a prior belief, described by normal distribution. Child's belief is updated in each period according to the realization of X_t.

$\bar{a}_t$ is the 'guide' that the parent provides for a finite period of time, lessening the pain of learning of the child. The child is assumed to disregard the parent's influence and act as if $\bar{a}_t = 0$ for all t. Then, knowing the true value of M, the parent can choose to forego learning efficiency by appropriately choosing $\bar{a}_t$ to 'shelter' the child from the pain of learning. Optimal parenting policy is shown to be between full sheltering, where period utility is maximized, and 'boot camp' parenting, where the child's learning is maximized.[9]

My model extends their model by three ways. First, human capital channel and belief channel are separated. Some policies target child's human capital, while others target child's behavior and beliefs. Their effects and potential interactions can be investigated in this model. Second, parenting behaviors are also separated, between parental investment and parental punishment. In particular, the model allows measures of parental punishment and punishment consistency to be interpreted in the economic framework. Third, the model pinpoints a policy-relevant parameter that is proven to be malleable. Evaluation of the parent training program shows that parenting skill can be effectively improved, leading to behavioral improvement of the children.

Principal-agent model with monetary and non-monetary incentives [66] was one of the first to model the parent's use of punishment in the context of parent-child interaction. The parent-principal chooses between monetary and non-monetary incentives, where a non-monetary incentive is interpreted as corporal punishment. Binding budget constraint leads the parent to use corporal punishment. A key implication of the model is that non-monetary incentive such as corporal punishment is more likely to be used in low-income households.

9. It can be shown that parent's sheltering is greater when the child's initial prior is farther away from the value of the mean observed by the parent. Some geneticists, cited in their paper, showed that twins reared apart are just as similar to each other as twins reared together, and argued that 'nature' dominates 'nurture'. Lizzeri and Siniscalchi shows that this evidence can be generated by their model, where the parent increases sheltering when the children have different characteristics than her own. Since the twin reared by different foster parents would retain many of their original behavior, it is no surprise that they end up being 'similar'.

Table 4.6 confirms this implication, showing that household income is negatively associated with punishment use in the CNLSY data. Table 4.32 shows that this implication remains true for each of the punishment strategies measured.

Table 4.32: Descriptive Model with Punishment Measures

	(1) Spanking	(2) Send-to-room	(3) Privilege	(4) Grounding	(5) Allowance
Log Net HH income	-0.015***	-0.014***	-0.007**	-0.010***	-0.003*
	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)
Father presence at home	0.028***	0.009	-0.002	-0.036***	-0.014**
	(0.008)	(0.010)	(0.009)	(0.008)	(0.005)
AFQT	-0.070***	-0.010	-0.003	-0.063***	-0.027***
	(0.005)	(0.007)	(0.006)	(0.005)	(0.003)
firstborn	0.030***	0.037***	0.023***	0.014*	0.002
	(0.005)	(0.007)	(0.006)	(0.006)	(0.003)
Child is female	-0.051***	-0.057***	-0.072***	-0.052***	-0.013***
	(0.006)	(0.008)	(0.007)	(0.007)	(0.004)
BLACK	0.085***	-0.037*	0.011	0.037**	0.041***
	(0.012)	(0.015)	(0.014)	(0.014)	(0.009)
NON-BLACK, NON-HISPANIC	0.040**	0.010	-0.042**	-0.025*	-0.006
	(0.012)	(0.015)	(0.013)	(0.013)	(0.007)
Constant	0.375***	0.585***	0.386***	0.389***	0.108***
	(0.031)	(0.031)	(0.028)	(0.029)	(0.016)
Observations	8680	8348	8355	8358	8336
R2	0.066	0.015	0.021	0.086	0.058

Note: Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. Punishment models are based on linear probability model. and columns indicate privilege restriction and allowance restrictions, respectively. Baseline racial category is the Hispanic.

Weinberg's principal-agent model, however, does not allow for the possibility that parental punishment can have a negative effect on child outcome. Parent's use of punishment is always a positive incentive for the child to increase effort. The model is also unsuited to explain the effect of the randomized parenting intervention on child outcomes. One could argue that the training increases the magnitude of utility transferred to the child when the parent uses non-monetary incentive. However, the intervention describes discipline techniques that actually *lowers* the parent's over-reactivity in punishment, which would reduce the overall disutility the child receives as a result of misbehavior. While it is very plausible that the choice of parental punishment is partly driven by the parent's inability to use more costly parenting, principal-agent model with different incentive costs is not suited to explain the empirical evidence.

Principal-agent model with imperfect monitoring The parent-child interaction model of [4] also relies on the principal-agent framework. The principal-parent in his model uses the history of human capital accumulation of the child (such as school grades) as noisy signals of the underlying child effort in each. Suppose, by luck, the child's human capital accumulation is exceptionally high in one period. Because the child cannot be lucky every period, the child's human capital accumulation may well be lower than what the parent expected in the next period. Then, the parent mistakenly believes the child's effort to be lower than it actually is, and punishes the child by choosing low level of utility transfer. 'Maltreatment equilibrium' may arise in this case, where the parent punishes repeatedly without realizing the intended effect on child effort.[10]

This model can account for the observed evidence in CNLSY if the parent's human capital is positively correlated with the parent's monitoring of the child's behavior. It might be the case that a parent is better able to deduce the child's underlying effort if the parent herself has higher human capital.[11] However, this advantage in monitoring ability should be strong enough to compensate for the fact that the parent with more human capital also tend to work longer hours, and therefore would have less hours available to monitor the child. Also, the model does not account for the parent's punishment responses to the child's tantrum or

10. Model implications are consistent with the evidence in [13]. [13] evaluates a field experiment where low-income mothers are given text messages on the child's participation in school. When the parents initially overestimate the child's effort, better monitoring technology leads the parents to *increase* the use of punishment strategies such as privilege restriction.

11. If parental punishment based on poor monitoring leads to parent-child conflict, monitoring ability is nested as a part of parenting skill in the signaling model in Section 3.

resisting the parent's requests, where monitoring is not an issue.

Finally, although the curriculum of parent training in Section 1 includes identifying the child's problematic behavior and ignoring petty resistance from the child, the program primarily focuses on the parent's behavioral techniques in actually implementing punishment and maintaining the parent's control of the parent-child interaction. Signaling model provides a more plausible interpretation of the experimental evidence than principal-agent model with imperfect monitoring.

Principal-agent model with child heterogeneity Cosconati's model ([19], and later [20]) focuses on the heterogeneity in the preference of children for future outcome. Some children are assumed to care more about their own future outcome, and therefore choose higher level of effort. 2009 model assumed different types of children defined by the baseline level of human capital. 2013 model linked the child's human capital to his discount rate, as in [10]. Cosconati's model avoids the awkward issue of having to think about the implicit contract between a parent and a child in the household by considering the average effect of the parent's attempt at controlling the child's leisure activities on the child's utility. Knowing that the parent would try to impose restrictions if the child performance was bad, the child has an incentive to avoid this consequence by working hard. Cosconati structurally estimated the model, where parental restrictions were represented by whether the parent decides how the child uses leisure time. The key implication of this model is that a blanket restriction ("curfew") on the child's leisure time use may help children with initially low human capital level, but hurt others. The children with initially high level of human capital can no longer lift the restrictions through good performance, thereby losing the dynamic incentive for more effort. In other words, the children with initially high level of human capital benefit more from conditional use of incentive, while those with low level of human capital may benefit from blanket restrictions.

It is not clear how Cosconati's model, which focuses on the heterogeneity in children's characteristics, can be applied to interpret the experiment evidence in Section 1, which changes parental characteristics. Also, the parental punishment empirically analyzed in Section 2 does not necessarily play the same role as parental restriction in Cosconati's model. A parent could allow considerable leeway in the child's time use, reserving punishment only for serious misbehavior. Nonetheless, suppose that parental punishment plays the same role as parental restriction of leisure activities by children. Then, frequent punishment would have similar effect as blanket restriction. Always-punishing parent would hurt the children with high baseline human capital and benefit children with low baseline human capital. As shown in Table 4.7, low-skilled parents used punishment more often and indiscriminately while high-skilled parents used punishment less often and more conditional on the child's misbehavior. Were Cosconati's model applicable to this case, punishment by low human capital parents would be more effective for children with low baseline human capital. Then, the sign of β_3 in equations (4.4) and (4.5) should be negative for low baseline human capital children. Tables 4.33 and 4.34 estimates equations (4.4) and (4.5) respectively for a sample of CNLSY households whose children had below sample-average human capital at child ages 5-6. The signs of β_3 estimates are positive in all the cases considered, lending support to the signaling interpretation of parental punishment, rather than the interpretation based on the principal-agent model with child heterogeneity.

Principal-agent model with reputation formation The model by [42], later extended by [46], assumes that parents are heterogeneous in their preference for incentive use. In their model, some parents, even if they value child outcomes, do not wish to use incentive to motivate child effort.[12] When there are multiple siblings who do not know the type of the parent, the parent has an incentive to punish the misbehavior of the older children, in order to establish reputation as a 'tough' (transferable-utility type) parent to younger children. In equilibrium, parents of all types are more likely to punish the misbehavior of earlier-born

12. The authors in [42] assume that some family preferences satisfy transferable utility, while others do not.

Table 4.33: Child outcomes

	(1) Wage	(2) Wage	(3) Wage	(4) College	(5) College	(6) College
Punishment	0.892	1.788	-0.574	-0.027	-0.024	-0.378
	(1.187)	(1.072)	(1.197)	(0.109)	(0.169)	(0.397)
Pun.xAFQT	1.472	2.332*	0.694	0.326***	0.429**	0.204
	(1.235)	(0.955)	(1.337)	(0.095)	(0.147)	(0.277)
AFQT	-1.195	-1.878*	0.146	-0.282***	-0.362**	-0.248
	(1.259)	(0.931)	(1.553)	(0.082)	(0.115)	(0.247)
Observations	201	184	184	116	106	106
R^2	0.258	0.380	0.594	0.493	0.663	0.819
Adjusted R^2	0.152	0.107	0.115	0.367	0.345	0.207
Model	None	State FE	County FE	None	State FE	County FE

Note: Sample consists of below-sample-average child human capital at age 5-6. Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. lnWage is log (inverse hyperbolic sine transformation) wage earnings at age 26. College is college attendance status at age 22. Additional controls include log net household income, father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used at ages 8 through 14. Investment is estimated factor for cognitive stimulation scale for ages 8 through 14. AFQT is parent's AFQT score.

Table 4.34: Child skills

	(1)	(2)	(3)	(4)
	PIAT	PIAT	BPI	BPI
L.punish	0.012	0.040^{+}	-0.087^{***}	0.105^{**}
	(0.019)	(0.022)	(0.024)	(0.038)
L.Pun.xAFQT	0.025	0.061^{**}	-0.010	0.012
	(0.021)	(0.023)	(0.026)	(0.038)
L.Inv.	0.022^{*}	-0.016	0.035^{**}	-0.001
	(0.010)	(0.013)	(0.013)	(0.022)
Observations	1569	1115	1570	1094
Model	RE	FE	RE	FE
R2	0.669		0.661	

Note: Sample consists of below-sample-average child human capital at age 5-6. Observations are number of parent-child pairs. Standard errors are in parenthesis, clustered at the household level. + p¡.1, * p¡.05, ** p¡.01, *** p¡.001. FE estimates Arellano-Bond GMM estimator. PIAT is latent factor estimate from Peabody Individual Achievement Test scores. BPI is latent factor estimate from Behavior Problems Index, scaled so that higher value indicates better behavior. Additional controls include lag of skills, lag of log (inverse hyperbolic sine transformation) household net income, lag of father's presence at home, indicators for race, child's birth year, parent's birth year and child's gender. Punishment is indicator for any punishment used . Investment is cognitive stimulation scale from HOME score. AFQT is parent's AFQT score. ChildHC is child's average PIAT score between ages 8 and 14.

children. Anticipating the parent's strategy, earlier-born children are more likely to choose good behavior. Their empirical analysis using NLSY79 data[13] confirms that the earlier-born children are less likely to engage in risky adolescent behavior, and the parents are more likely to punish misbehavior when the children are earlier-born. The punishment measures are based on the parent's decision whether to finance half of annual living expenses after child age 18, and whether to allow the child to live with the parent after child age 18. [46] finds similar results for academic achievement and parental monitoring using CNLSY data.

It is not clear how the reputation model can be used to interpret the empirical evidence in this paper. Their model is silent on the underlying nature of the parental type differences, focusing instead on the implications of different beliefs the children have about the type of the parent. Additionally, the presence of other siblings is necessary for the reputation formation mechanism, while the signaling between a parent and a child can occur regardless of the child's siblings. It could still be argued that parenting skill is one of the elements that define parental types, since a parent with low parenting skill is limited in her ability to transfer utility to the child as intended. Then, the signaling model and the experimental evidence in this paper present a particular way in which the parental type can be changed.

Principal-agent model with principal's private information The model in this paper bears important resemblance to that of [12], in that the principal-parent has a private information about the returns to effort of the agent-child. In their model, the parent knows the type of the child, while the child draws private signal regarding his own type. When the child is a high-productivity type, the parent anticipates that the child will likely draw a high signal and voluntarily choose high effort, regardless of parental incentives. Since incentive provision is costly to the parent, the parent of the high-productivity type child would not announce the incentive, and the child will not receive incentive transfer even when high effort is chosen. In an equilibrium with two productivity types, parents of high-type children never

13. They use the main respondents of NLSY79 as the 'children'. My study uses CNLSY, where the respondents of NLSY79 are the parents.

announce incentive to their children, while parents of low-type children randomizes between announcement and no announcement. Although announcing incentive has a short-run effect of encouraging the child's effort, the child infers from the announcement that his long-run productivity must be low.

The cost associated with the parent's incentive provision plays an essential role in deriving the equilibrium behavior in Benabou-Tirole model. If the incentive did not have any cost, all parents would announce the incentive, and the trade off between short-run positive effect and long-run negative effect would vanish. In the model of parental punishment, the child is rewarded for effort by the parental withholding punishment. Although sometimes a parent may have to 'hold herself down', it would be strange to assume that withholding punishment imposes direct utility cost to the parent. It is much more natural to assume that punishing the child imposes direct utility cost to the parent, based on the parent's use of time and effort. In this case, the equilibrium analysis in the Benabou-Tirole model is reversed. If a child is of high type, then the parent would always announce the punishment (incentive), since she expects the punishment to be never used. A parent of a low-type child also wants to announce punishment to encourage the child's effort, but now the parent would have to pay the cost of punishment if the child *fails*. In this case, parents of low-type children would randomize between announcement and no announcement.[14] The model implies that parents of low-type children would use less punishment on average. Assuming that a child's type is positively correlated with household socioeconomic status, this prediction is inconsistent with the stylized facts (and Table 4.7) that punishment use is more frequent in low SES households. Thus Benabou-Tirole model is not suitable to explain the empirical evidence on parental punishment. Additionally, similar to the Cosconati's model, this model is based on the heterogeneity of the characteristics of children. It is not clear how to interpret the experimental evidence in this paper using their model.

14. Suppose both types of parents pool into not announcing punishment. Low-type parent does not want to deviate when she anticipates that the child is likely to fail, in which case she would have to pay the cost of punishment. Knowing this, high-type parent can deviate and signal to the child that he is of high type.

CHAPTER 5

CONCLUSION

What is the effect on a child's life cycle outcomes of of parental punishment during childhood? The effect is positive when the punishment effectively communicates the parent's message regarding optimal behavior for the child. The effect is negative when it does not. When the parent-child interaction surrounding the punishment is characterized by the parent's emotional over-reaction, punishment is implemented harshly and inconsistently with respect to the underlying child behavior. Then, the child receives a low-quality message regarding the behavior expected of him, which also informs him of the returns to his own effort in the long-run. Short-run and long-run efforts are connected through the role of human capital in generating life cycle outcomes.[1] The child's long-run outcome suffers from imprecise learning, because the now-independent child makes misinformed efforts choice towards life-cycle outcomes.

The model generates clear implications for policies that target households with children. Improving child outcomes through direct interaction with children in early childhood can be difficult and costly. Effects of direct interventions on young children would depend on household and parental characteristics that are hard to monitor and control. However, experimental data shows that the parents can learn to use punishment in a calm and consistent manner through education and training. The success of the randomized intervention program suggests that harshness and inconsistency of punishment are examples of behavioral skill, or *parenting skill* as it is referred to in this paper. Interpreting a behavior as a skill implies it can be changed through skill investment more so than by other means.[2] Parents' responses to policy interventions may be constrained by their parenting skills, but, as this paper demonstrates, parenting skills can be improved by external investment.

1. Parental investment and punishment in $t = 1$ and child effort in $t = 2$ are both dependent on the value of long-run productivity of human capital R in generating child outcomes.

2. See [47] for an example of a program targeting behavioral skills of adolescents.

These results are based on a highly stylized model with restrictive assumptions of how the child interprets signals from parental investment and punishment. The model also does not investigate intergenerational implications, where the beliefs held by the parent about the state of the world are transmitted to the child through parenting. In particular, the intergenerational belief transmission can guide the formation of family and community culture regarding schooling and work. Relaxing the restrictive assumptions of the model and extending the model to intergenerational setting constitute interesting avenues of research.

This paper extends the literature on the analysis of child human capital production. Both the model and the empirical analysis presented here suggest parental punishment itself should not be interpreted as an input in human capital production function. It is difficult to justify, *a priori*, restriction on the sign of punishment's effect on child outcomes when the effect may be either positive or negative. On the other hand, the harshness or inconsistency of punishment has an unambiguously negative effect on the child's human capital development and life cycle outcomes. These should be treated as measures of parenting skill, which is an input to human capital production in children. An interesting avenue of research would be to examine the interaction of parenting skill with other human capital production inputs and other life cycle choices of parents in the context of child human capital development.

REFERENCES

[1] T. M. Achenbach. *Manual for the Youth Self-Report and 1991 Profile*. Burlington: University of Vermont, Department of Psychiatry, 1991.

[2] Thomas M Achenbach. *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. Department of Psychiatry, University of Vermont Burlington, VT, 1991.

[3] Francesco Agostinelli and Matthew Wiswall. Estimating the technology of children's skill formation. Technical report, NBER Working Paper No. 22442, 2016.

[4] Hideo Akabayashi. An Equilibrium Model of Child Maltreatment. *Journal of Economic Dynamics and Control*, 30(6):993–1025, 2006.

[5] David S. Arnold, Susan G. O'Leary, Lisa S. Wolff, and Maureen M. Acker. The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment*, 5(2):137–144, 1993.

[6] Bernd Bäse. *Die sozialräumliche Gliederung der Stadt Braunschweig: Methodik und Durchführung sozialgeographischer Analyse im städtischen Wohnumfeld auf der Grundlage des Zensus von 1987 [The social-ecological structure of the city of Brunswick. Method and procedure of socio-geographical analyses in urban living areas based upon the census of 1987]*. PhD thesis, Technische Universität Braunschweig, 1995.

[7] Diana Baumrind. Effects of Authoritative Parental Control on Child Behavior. *Child Development*, 37(4):887–907, 1966.

[8] Diana Baumrind. The influence of parenting style on adolescent competence and substance use. *The Journal of Early Adolescence*, 11(1):56–95, 1991.

[9] Diana Baumrind, Robert E Larzelere, and Philip A Cowan. Ordinary physical punishment: is it harmful? comment on gershoff (2002). *Psychological Bulletin*, 128(4):580–589, 2002.

[10] Gary S Becker and Casey B Mulligan. The endogenous determination of time preference. *The Quarterly Journal of Economics*, 112(3):729–758, 1997.

[11] Gary S Becker and Nigel Tomes. An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy*, 87(6):1153–1189, 1979.

[12] Roland Benabou and Jean Tirole. Intrinsic and extrinsic motivation. *The Review of Economic Studies*, 70(3):489–520, 2003.

[13] Peter Bergman. Parent-child information frictions and human capital investment: Evidence from a field experiment. 2017.

[14] James Berry. Child control in education decisions an evaluation of targeted incentives to learn in india. *Journal of Human Resources*, 50(4):1051–1080, 2015.

[15] Leonardo Bursztyn and Lucas C Coffman. The schooling decision: Family preferences, intergenerational conflict, and moral hazard in the brazilian *Favelas*. *Journal of Political Economy*, 120(3):359–397, 2012.

[16] Peter Burton, Shelley Phipps, and Lori Curtis. All in the family: A simultaneous model of parenting style and child conduct. *American Economic Review*, 92(2):368–372, 2002.

[17] Bettye M. Caldwell and Robert H. Bradley. *Home Observation for Measurement of the Environment*. University of Arkansas at Little Rock Little Rock, 1984.

[18] Magnus Carlsson, Gordon B Dahl, Björn Öckert, and Dan-Olof Rooth. The effect of schooling on cognitive skills. *Review of Economics and Statistics*, 97(3):533–547, 2015.

[19] Marco Cosconati. *Parenting Style and the Development of Human Capital in Children*. PhD thesis, University of Pennsylvania, 2009.

[20] Marco Cosconati. Optimal Parenting Styles: Evidence from a Dynamic Model with Multiple Equilibria. Technical report, Technical report, Bank of Italy, 2013.

[21] Flávio Cunha, Irma Elo, and Jennifer Culhane. Eliciting maternal expectations about the technology of cognitive skill formation. Technical report, NBER Working Paper No. 19144, 2013.

[22] 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.

[23] Flavio Cunha, James J Heckman, and Susanne M Schennach. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931, 2010.

[24] Nancy Darling and Laurence Steinberg. Parenting style as context: An integrative model. *Psychological Bulletin*, 113(3):487–496, 1993.

[25] Daniela Del Boca, Christopher Flinn, and Matthew Wiswall. Household Choices and Child Development. *Review of Economic Studies*, 81(1):137–185, 2014.

[26] Thomas J. Dishion and Gerald R. Patterson. *Developmental Psychopathology*, volume 3, chapter The Development and Ecology of Antisocial Behavior in Children and Adolescents, pages 503–541. 2 edition, 2015.

[27] Matthias Doepke and Fabrizio Zilibotti. Parenting with Style: Altruism and Paternalism in Intergenerational Preference Transmission. *Econometrica*, 85(5):1331–1371, 2017.

[28] Martin Dooley and Jennifer Stewart. Family Income, Parenting Style and Child Behavioural-Emotional Outcomes. *Health Economics*, 16(2):145–162, 2007.

[29] Sanford M Dornbusch, Philip L Ritter, P Herbert Leiderman, Donald F Roberts, and Michael J Fraleigh. The relation of parenting style to adolescent school performance. *Child development*, pages 1244–1257, 1987.

[30] Orla Doyle, Colm Harmon, James J. Heckman, Caitriona Logue, and Seong Hyeok

Moon. Early skill formation and the efficiency of parental investment: A randomized controlled trial of home visiting. *Labour Economics*, 2016.

[31] Nancy Eisenberg, Tracy L. Spinrad, and Ariel Knafo-Noam. *Handbook of Child Psychology and Developmental Science*, chapter Prosocial Development, pages 1–47. John Wiley & Sons, Inc., 2015.

[32] Sneha Elango, Jorge Luis García, James J Heckman, and Andrés Hojman. Early Childhood Education. Technical report, NBER Working Paper No. 21766, 2015.

[33] Stephen A Erath, Mona El-Sheikh, J Benjamin Hinnant, and E Mark Cummings. Skin Conductance Level Reactivity Moderates the Association between Harsh Parenting and Growth in Child Externalizing Behavior. *Developmental Psychology*, 47(3):693, 2011.

[34] Christopher J. Ferguson. Spanking, corporal punishment and negative long-term outcomes: A meta-analytic review of longitudinal studies. *Clinical Psychology Review*, 33(1):196–208, 2013.

[35] Mario Fiorini and Michael P Keane. How the Allocation of Children's Time Affects Cognitive and Non-Cognitive Development. *Journal of Labor Economics*, 32(4):787–836, 2014.

[36] Sebastian Galiani, Matthew Staiger, and Gustavo Torrens. When children rule: Parenting in modern families. Technical report, NBER Working Paper No. 23087, 2017.

[37] Jorge Luis García, James J Heckman, Duncan Ermini Leaf, and Maria José Prados. Quantifying the Life-Cycle Benefits of a Prototypical Early Childhood Program. Technical report, NBER Working Paper No. 23479, 2017.

[38] Elizabeth T Gershoff, Kierra MP Sattler, and Arya Ansari. Strengthening Causal Estimates for Links Between Spanking and Children's Externalizing Behavior Problems. *Psychological Science*, 29(1):110–120, 2018.

[39] Elizabeth Thompson Gershoff. Corporal punishment by parents and associated child behaviors and experiences: A meta-analytic and theoretical review. *Psychological Bulletin*, 128(4):539, 2002.

[40] Florence Laura Goodenough. *Anger in young children*. Westport, CT: Greenwood Press., 1975 (1931).

[41] Erik Grönqvist, Björn Öckert, and Jonas Vlachos. The intergenerational Transmission of Cognitive and Noncognitive Abilities. *Journal of Human Resources*, 52(4):887–918, 2017.

[42] Lingxin Hao, Joseph V. Hotz, and Ginger Z. Jin. Games Parents and Adolescents Play: Risky Behaviour, Parental Reputation and Strategic Transfers. *Economic Journal*, 118(528):515–555, 2008.

[43] Robert Haveman and Barbara Wolfe. The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature*, 33(4):1829–1878, 1995.

[44] Nina Heinrichs and Amanda Jensen-Doss. The effects of incentives on families' long-term outcome in a parenting program. *Journal of Clinical Child & Adolescent Psychology*, 39(5):705–712, 2010.

[45] George W. Holden, Paul A. Williamson, and Grant W. O. Holland. Eavesdropping on the Family: A Pilot Investigation of Corporal Punishment in the Home. *Journal of Family Psychology*, 28(3):401, 2014.

[46] V Joseph Hotz and Juan Pantano. Strategic parenting, birth order, and school performance. *Journal of Population Economics*, 28(4):911–936, 2015.

[47] Tim Kautz and Wladimir Zanoni. Measuring and fostering non-cognitive skills in adolescence: Evidence from chicago public schools and the onegoal program. *University of Chicago*, 2014.

[48] Jun Hyung Kim, Wolfgang Schulz, Tanja Zimmermann, and Kurt Hahlweg. Parent-child interactions and child outcomes: Evidence from randomized intervention. *Labour Economics*, 54:152–171, 2018.

[49] Robert E Larzelere, Marjorie Lindner Gunnoe, Mark W Roberts, and Christopher J Ferguson. Children and parents deserve better parental discipline research: Critiquing the evidence for exclusively "positive" parenting. *Marriage & Family Review*, 53(1):24–35, 2017.

[50] Alessandro Lizzeri and Marciano Siniscalchi. Parental Guidance and Supervised Learning. *Quarterly Journal of Economics*, 123(3):1161–1195, 2008.

[51] Nicholas Long, Mark C Edwards, and Jayne Bellando. Parent training interventions. In *Handbook of Childhood Psychopathology and Developmental Disabilities Treatment*, pages 63–86. Springer, 2017.

[52] Viara R Mileva-Seitz and Alison S Fleming. How being mothered affects the development of mothering. In *Parental Stress and Early Child Development*, pages 235–259. Springer, 2017.

[53] Sebastian Naumann, Heike Bertram, Annett Kuschel, Nina Heinrichs, Kurt Hahlweg, and Manfred Döpfner. Der erziehungsfragebogen (efb). *Diagnostica*, 2010.

[54] Seollee Park, Yaeeun Han, and Hyuncheol Bryant Kim. Knowledge, Food Vouchers, and Child Nutrition: Evidence from a Field Experiment in Ethiopia. 2018.

[55] Gerald R Patterson, John B Reid, and Thomas J Dishion. *Antisocial Boys*, volume 4. Castalia, 1992.

[56] Martin Pinquart. Associations of parenting styles and dimensions with academic achievement in children and adolescents: A meta-analysis. *Educational Psychology Review*, 28(3):475–493, 2016.

[57] Rodrigo Pinto. Selection Bias in a Controlled Experiment: The Case of Moving to Opportunity. Technical report, Ph. D. Thesis, University of Chicago, Department of Economics, 2015.

[58] Joseph H Pleck and Sandra L Hofferth. Mother involvement as an influence on father involvement with early adolescents. *Fathering*, 6(3):267, 2008.

[59] Rebecca M Ryan, Ariel Kalil, Kathleen M Ziol-Guest, and Christina Padilla. Socioeconomic Gaps in Parents' Discipline Strategies from 1988 to 2011. *Pediatrics*, page e20160720, 2016.

[60] Matthew R. Sanders, James N. Kirby, Cassandra L. Tellegen, and Jamin J. Day. The Triple P-Positive Parenting Program: A systematic review and meta-analysis of a multi-level system of parenting support. *Clinical Psychology Review*, 34(4):337–357, 2014.

[61] Matthew R Sanders, Carol Markie-Dadds, and Karen MT Turner. *Theoretical, Scientific and Clinical Foundations of the Triple P-Positive Parenting Program: A Population Approach to the Promotion of Parenting Competence*, volume 1. Parenting and Family Support Centre, The University of Queensland Queensland, 2003.

[62] William Schneider, Michael MacKenzie, Jane Waldfogel, and Jeanne Brooks-Gunn. Parent and Child Reporting of Corporal Punishment: New Evidence from the Fragile Families and Child Wellbeing Study. *Child Indicators Research*, 8(2):347–358, 2015.

[63] Christopher Spera. A Review of the Relationship among Parenting Practices, Parenting Styles, and Adolescent School Achievement. *Educational Psychology Review*, 17(2):125–146, 2005.

[64] Murray A Straus, Emily M Douglas, and Rose Anne Medeiros. *The Primordial Violence: Spanking Children, Psychological Development, Violence, and Crime*. Routledge, 2014.

[65] Christopher R Walters. Inputs in the production of early childhood human capital: Evidence from head start. *American Economic Journal: Applied Economics*, 7(4):76–102, 2015.

[66] Bruce a. Weinberg. An Incentive Model of the Effect of Parental Income on Children. *Journal of Political Economy*, 109(2):266–280, 2001.

APPENDIX A

PROOF FOR THEOREM 1

1 Child's problem

Assume that the child believes $\mu = \lambda_M R$ and $\varphi = \lambda_I R$.

Child's objective function is

$$\max_{a_1} -\frac{va_1^2}{2} - E_{\mathcal{H}_1}\left[(M - a_1)^2\right] + \delta E_{\mathcal{H}_1}\left[R(\theta_2 + \beta_\tau a_2) - \frac{\tau a_2^2}{2}\right]$$

Note that

$$E_{\mathcal{H}_1}[\rho_2] = E_{\mathcal{H}_1}\left[\frac{s_1\rho_1 + s_M\lambda_M^2 R + s_I\lambda_I^2 R}{s_2}\right] = \rho_1$$

In other words, the restriction placed in the main model is consistent with the assumption that the child believes his own belief will remain the same in expectation. Then child's expected utility in the second period from the child's perspective is

$$E_{\mathcal{H}_1}\left[R(\theta_2 + \beta_\tau a_2) - \frac{\tau a_2^2}{2}\right] = \rho_1 E_{\mathcal{H}_1}[\theta_2] + E_{\mathcal{H}_1}\left[R\beta_\tau a_2 - \frac{\tau a_2^2}{2}\right]$$

The second term $E_{\mathcal{H}_1}\left[R\beta_\tau a_2 - \frac{\tau a_2^2}{2}\right]$ depends entirely on ρ_1, which does not depend on the choice of a_1. Expected value of the punishment is

$$E_{\mathcal{H}_1}\left[(M - a_1)^2\right] = s_M^{-1} + (\lambda_M R - a_1)^2$$

due to the child's belief $w = 1$. Then the first order condition is

$$va_1 + \frac{\partial}{\partial a_1} E_{\mathcal{H}_1}\left[(\lambda_M R - a_1)^2\right] = \delta\rho_1\beta_a$$

$$\rightarrow (v - 2)\,a_1 + 2\lambda E_{\mathcal{H}_1}[R] = v\lambda_M\rho_1$$

$$\rightarrow a_1 = \lambda_M\rho_1$$

for $\lambda_M \equiv \frac{\delta\beta_1}{v}$, $E_{\mathcal{H}_1}[R] = \rho_1$.

1.1 Child's problem at higher level of sophistication

Now suppose the child's is at some higher level of sophistication. $k_M \neq 0$, $k_I \neq 0$ or both.

Key difference from the $k = 0$ model is that

$$E_{\mathcal{H}_1}\left[(M - a_1)^2\right] = s_M^{-1} + \left(w_{k_M}\lambda_M R + \left(1 - w_{k_M}\right) a_1 - a_1\right)^2$$

and

$$E_{\mathcal{H}_1}[\rho_2] = \frac{1}{s_2}\left[s_1\rho_1 + s_M w_{k_M}^2 \lambda_M^2 R + s_I \lambda_I^2 R\right] = \rho_1$$

Since the child still believes that he can back out the true value of R, his best guess of future belief is still his current belief. He knows how confident he will be with his future belief, which is determined by $w_{k_M}^2$.

Child's first order condition is

$$\begin{aligned} va_1 + \frac{\partial}{\partial a_1} E_{\mathcal{H}_1}\left[\left(w_{k_M}\lambda_M R + \left(1 - w_{k_M}\right) - a_1\right)^2\right] &= \delta\rho_1\beta_a \\ \rightarrow \left(v - 2w_{k_M}\right) a_1 + 2w_{k_M}\lambda_M\rho_1 &= v\lambda_M\rho_1 \\ \rightarrow a_1 &= \lambda_M\rho_1 \end{aligned}$$

Somewhat surprisingly, at any level of sophistication, the child still makes initial period decision solely based on the information available to him at the time of the decision. The simplicity is due to the rigid timing structure of the model–when the child decides on the effort choice, the only information available to him on the value of R is described by his prior belief. Sophistication is relevant when he extracts information from the signals once they are made available, but not before.

2 Parent's problem

Assume that child is at $k_M = k_I = 0$ belief, so child believes $\mu = \lambda_M R$ and $\varphi = \lambda_I R$.

Parent's objective function is

$$\max_{w,\varphi} - \frac{c\varphi^2}{2} - \alpha_1 E\left[(M - a_1)^2\right] + \alpha_2 \delta E\left[R\left(\theta_2 + \beta_\tau a_2 - \frac{\tau a_2^2}{2}\right)\right]$$

Restating the expression,

$$\max_{w,\varphi} - \frac{c\varphi^2}{2} - \alpha_1\left(s_M^{-1} + w^2\lambda_M^2\left(R - \frac{a_1}{\lambda_M}\right)^2\right) + \alpha_2\delta\left(R\theta_2 + \frac{\beta_\tau^2}{2\tau}\left(2R\rho_2^e - (\rho_2^e)^2\right) - \frac{\beta_\tau^2}{2\tau} var\left(\rho_2\right)\right)$$

2.1 Investment choice

First order condition for investment is

$$-c\varphi + \alpha_2\delta\left[R\beta_I + \frac{\beta_\tau^2}{\tau}\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)\left(R - \rho_2^e\right)\right] = 0 \tag{A.1}$$

$R\beta_I$ is the direct effect on child's adult utility through human capital. $\frac{\beta_\tau^2}{\tau}\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)\left(R - \rho_2^e\right)$ term is the change in the child's adult utility through the change in the child's expectation. Increasing φ increases ρ_2^e, so the second term is positive as long as $\rho_2^e < R$. Change in the expected value of child's belief is

$$\frac{\partial}{\partial\varphi}\rho_2^e = \frac{s_I\lambda_I^2}{s_2\lambda_I}$$

Following algebraic manipulation derives the optimal investment choice:

$$\begin{aligned}
R-\rho_2^e &= R-\frac{1}{s_2}\left(s_1\rho_1+s_M\lambda_M^2\frac{\mu}{\lambda_M}+s_I\lambda_I^2\frac{\varphi}{\lambda_I}\right)\\
&=\frac{1}{s_2}\left(s_1\left(R-\rho_1\right)+s_M\lambda_M^2\left(R-wR-\left(1-w\right)\rho_1\right)+s_I\lambda_I^2\left(R-\frac{\varphi}{\lambda_I}\right)\right)\\
&=\frac{1}{s_2}\left(s_1+s_M\lambda_M^2\left(1-w\right)\right)\left(R-\rho_1\right)+\frac{1}{s_2}s_I\lambda_I^2\left(R-\frac{\varphi}{\lambda_I}\right)\\
&=\frac{1}{s_2}\left(s_1+s_M\lambda_M^2\left(1-w\right)\right)\left(R-\rho_1\right)+\frac{1}{s_2}s_I\lambda_I^2\frac{1}{\lambda_I}\left(\lambda_I R-\varphi\right)\\
&=\frac{1}{s_2}\left(s_1+s_M\lambda_M^2\left(1-w\right)\right)\left(R-\rho_1\right)+\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)\left(\varphi^\dagger-\varphi\right)
\end{aligned}$$

First order condition for investment, restating (A.1):

$$\begin{aligned}
\varphi &= \frac{\alpha_2\delta}{c}\left[R\beta_I+\frac{\beta_\tau^2}{\tau}\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)\left(R-\rho_2^e\right)\right]\\
&=\frac{\alpha_2\delta}{c}R\beta_I+\frac{\alpha_2\delta}{c}\frac{\beta_\tau^2}{\tau}\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)\frac{1}{s_2}\left(s_1+s_M\lambda_M^2\left(1-w\right)\right)\left(R-\rho_1\right)\\
&+\frac{\alpha_2\delta}{c}\frac{\beta_\tau^2}{\tau}\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)\left(\varphi^\dagger-\varphi\right)
\end{aligned}$$

Collecting terms for φ,

$$\begin{aligned}
\left(1+\frac{\alpha_2\delta}{c}\frac{\beta_\tau^2}{\tau}\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)^2\right)\varphi &= \frac{\alpha_2\delta}{c}R\beta_I\\
&+\frac{\alpha_2\delta}{c}\frac{\beta_\tau^2}{\tau}\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)\frac{1}{s_2}\left(s_1+s_M\lambda_M^2\left(1-w\right)\right)\left(R-\rho_1\right)\\
&+\frac{\alpha_2\delta}{c}\frac{\beta_\tau^2}{\tau}\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)^2\varphi^\dagger
\end{aligned}$$

Note that $\varphi^\dagger \equiv \lambda_I R \equiv \frac{\alpha_2 \delta \beta_I}{c} R$. Then

$$\begin{aligned}
\varphi^* &= \varphi^\dagger + \frac{\frac{\alpha_2\delta}{c}\frac{\beta_\tau^2}{\tau}\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)\frac{1}{s_2}\left(s_1 + s_M\lambda_M^2\left(1-w\right)\right)\left(R-\rho_1\right)}{1+\frac{\alpha_2\delta}{c}\frac{\beta_\tau^2}{\tau}\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)^2} \\
&= \varphi^\dagger + \frac{\frac{\alpha_2\delta}{c}\frac{\beta_\tau^2}{\tau}\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)^2}{1+\frac{\alpha_2\delta}{c}\frac{\beta_\tau^2}{\tau}\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)^2}\frac{1}{\frac{\partial}{\partial\varphi}\rho_2^e}\frac{1}{s_2}\Big(s_1 + s_M\lambda_M^2\left(1-w\right)\Big)\left(R-\rho_1\right) \\
&= \varphi^\dagger + \eta\Big(s_1 + s_M\lambda_M^2\left(1-w\right)\Big)\left(R-\rho_1\right)
\end{aligned}$$

for $\eta \equiv \frac{\mathcal{B}}{1+\mathcal{B}}\frac{1}{\Delta_I}\frac{1}{s_2} = \frac{\mathcal{B}}{1+\mathcal{B}}\frac{1}{s_I\lambda_I}$, $\mathcal{B} \equiv \frac{\alpha_2\delta}{c}\frac{\beta_\tau^2}{\tau}\left(\Delta_I\right)^2$, $\Delta_I \equiv \frac{1}{s_2}s_I\lambda_I^2\frac{1}{\lambda_I} = \frac{\partial}{\partial\varphi}\rho_2^e$.

We have

$$\varphi^* = \lambda_I R + \eta \times \Big[s_1 + s_M\lambda_M^2\left(1-w^*\right)\Big]\left(R-\rho_1\right)$$

for $\mathcal{B} \equiv \left(\frac{\alpha_2\delta}{c}\frac{\beta_\tau^2}{\tau}\right)\left(\frac{\partial}{\partial\varphi}\rho_2^e\right)^2$ and $\eta \equiv \frac{\mathcal{B}}{1+\mathcal{B}}\frac{1}{s_I\lambda_I}$. $\frac{\partial}{\partial\varphi}\rho_2^e$ equals zero when $s_I = 0$, which means that φ^* approaches $\lambda_I R = \varphi^\dagger$ from above, the perfect-information solution. As s_I goes to infinity, φ^* approaches $\varphi^\dagger$ yet again. For any other value of $s_I > 0$, $\eta > 0$. So $\varphi^* \geq \varphi^\dagger$ if $R \geq \rho_1$ and $\varphi^* < \varphi^\dagger$ otherwise. The difference is proportional to $R - \rho_1$, so that the parent has additional incentive to invest on the child if the child's belief is sub-optimal, and a disincentive to invest if the child's belief is above optimal level. An interesting implication is that due to the effect on child belief, for some $s_I \geq s_I^*$, φ^* could decrease as s_I increases. This implication is unlikely to be relevant for low-SES households who are often targets of policies, as their low SES implies low s_I.

2.2 *Discipline choice*

Discipline choice is solved easier by plugging in the expression for the investment choice.

First order condition for w is

$$\alpha_1 2w\lambda_M^2\left(R-\frac{a_1}{\lambda_M}\right)^2=\alpha_2\delta\frac{\beta_\tau^2}{\tau}\rho_2^{e\prime}\left(R-\rho_2^e\right) \tag{A.2}$$

where

$$\begin{aligned}
\rho_2^e&=\frac{1}{s_2}\left(s_1\rho_1+s_M\lambda_M^2\frac{\mu}{\lambda_M}+s_I\lambda_I^2\frac{\varphi}{\lambda_I}\right)\\
\frac{\mu}{\lambda_M}&=wR+(1-w)\frac{a_1}{\lambda_M}\\
\rho_2^{e\prime}&\equiv\frac{\partial}{\partial w}\rho_2^e\\
&=\frac{1}{s_2}s_M\lambda_M^2\left(R-\frac{a_1}{\lambda_M}\right)\\
s_2&=s_1+s_I\lambda_I^2+s_M\lambda_M^2
\end{aligned}$$

In solving the first order condition, following expressions are convenient: $\mathcal{A}\equiv\delta\frac{1}{2}\frac{\beta_\tau^2}{\tau}$, $\frac{\partial}{\partial\varphi}\rho_2^e=\frac{s_I\lambda_I^2}{s_Is_2}$, $R-\rho_2^e=\frac{1}{s_2}\left(s_1\left(R-\rho_1\right)+s_M\lambda_M^2\left(R-\rho_1\right)-ws_M\lambda_M^2\left(R-\rho_1\right)+s_I\lambda_I^2\left(R-\frac{\varphi}{\lambda_I}\right)\right)$

Then equation (A.2) is

$$\begin{aligned}
w\alpha_1 2\lambda_M^2\left(R-\frac{a_1}{\lambda_M}\right)^2&=\alpha_2\delta\frac{\beta_\tau^2}{\tau}\rho_2^{e\prime}\left(R-\rho_2^e\right)\\
\rightarrow w2\lambda_M^2\left(R-\frac{a_1}{\lambda_M}\right)^2&=\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\rho_2^{e\prime}\left(R-\rho_2^e\right)\\
&=\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\frac{a_1}{\lambda_M}\right)\\
&\times\frac{1}{s_2}\left(s_1\left(R-\rho_1\right)+s_I\lambda_I^2\left(R-\frac{\varphi^*}{\lambda_I}\right)+s_M\lambda_M^2\left(R-\frac{a_1}{\lambda_M}\right)(1-w)\right)\\
&=\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right)\\
&\times\frac{1}{s_2}\left(s_1\left(R-\rho_1\right)+s_M\lambda_M^2\left(R-\frac{a_1}{\lambda_M}\right)(1-w)\right)\\
&+\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right)\\
&\times\frac{1}{s_2}s_I\lambda_I^2\left(R-\frac{\varphi^*}{\lambda_I}\right)
\end{aligned}$$

$$
\begin{aligned}
w2\lambda_M^2 \left(R-\rho_1\right)^2 = & \frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right) \\
& \times \underbrace{\frac{1}{s_2}\Big(s_1\left(R-\rho_1\right)+s_M\lambda_M^2\left(R-\rho_1\right)\left(1-w\right)\Big)}_{\mathbb{A}} \\
& + \frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right) \\
& \times \underbrace{\frac{1}{s_2}s_I\lambda_I^2\left(R-\frac{1}{\lambda_I}\Big(\lambda_I R+\eta\times\Big[s_1+s_M\lambda_M^2\left(1-w^*\right)\Big]\left(R-\rho_1\right)\Big)\right)}_{\mathbb{B}}
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{A} = & \frac{1}{s_2}\Big(s_1\left(R-\rho_1\right)+s_M\lambda_M^2\left(R-\rho_1\right)\left(1-w\right)\Big) \\
= & \frac{1}{s_2}\Big(s_1+s_M\lambda_M^2\Big)\left(R-\rho_1\right) \\
& - \frac{1}{s_2}s_M\lambda_M^2\left(R-\rho_1\right)w
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{B} = & \frac{1}{s_2}s_I\lambda_I^2\left(R-\frac{1}{\lambda_I}\Big(\lambda_I R+\eta\times\Big[s_1+s_M\lambda_M^2\left(1-w\right)\Big]\left(R-\rho_1\right)\Big)\right) \\
= & \frac{1}{s_2}s_I\lambda_I^2\left(-\frac{\eta}{\lambda_I}\Big[s_1+s_M\lambda_M^2\left(1-w\right)\Big]\left(R-\rho_1\right)\right) \\
= & -\frac{1}{s_2}s_I\lambda_I^2\frac{\eta}{\lambda_I}\Big[s_1+s_M\lambda_M^2\Big]\left(R-\rho_1\right) \\
& + \frac{1}{s_2}s_I\lambda_I^2\frac{\eta}{\lambda_I}\Big[s_M\lambda_M^2\Big]\left(R-\rho_1\right)w
\end{aligned}
$$

$$
\begin{aligned}
w2\lambda_M^2 \left(R-\rho_1\right)^2 = & \frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right)\frac{1}{s_2}\left(s_1+s_M\lambda_M^2\right)\left(R-\rho_1\right) \\
& -\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right)\frac{1}{s_2}s_M\lambda_M^2\left(R-\rho_1\right)w \\
& -\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right)\frac{1}{s_2}s_I\lambda_I^2\frac{\eta}{\lambda_I}\left[s_1+s_M\lambda_M^2\right]\left(R-\rho_1\right) \\
& +\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right)\frac{1}{s_2}s_I\lambda_I^2\frac{\eta}{\lambda_I}\left[s_M\lambda_M^2\right]\left(R-\rho_1\right)w
\end{aligned}
$$

$$
\frac{1}{s_2}s_I\lambda_I^2\frac{\eta}{\lambda_I} = \frac{\partial}{\partial\varphi}\rho_2^e\frac{\mathcal{B}}{1+\mathcal{B}}\frac{1}{s_I\lambda_I}
$$

Collecting terms with w,

$$
\begin{aligned}
& w2\lambda_M^2\left(R-\rho_1\right)^2 \\
& +\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right)\frac{1}{s_2}s_M\lambda_M^2\left(R-\rho_1\right)w \\
& -\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right)\frac{1}{s_2}s_I\lambda_I^2\frac{\eta}{\lambda_I}\left[s_M\lambda_M^2\right]\left(R-\rho_1\right)w \\
= & w2\lambda_M^2\left(R-\rho_1\right)^2 \\
& +\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right)\left(\frac{1}{s_2}-\frac{1}{s_2}s_I\lambda_I^2\frac{\eta}{\lambda_I}\right)s_M\lambda_M^2\left(R-\rho_1\right)w \\
= & w2\lambda_M^2\left(R-\rho_1\right)^2 \\
& +\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right)\left(\frac{1}{s_2}\frac{1}{1+\mathcal{B}}\right)s_M\lambda_M^2\left(R-\rho_1\right)w
\end{aligned}
$$

$$\begin{aligned}
&\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}(R-\rho_1)\frac{1}{s_2}\left(s_1+s_M\lambda_M^2\right)(R-\rho_1)\\
&-\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}(R-\rho_1)\frac{1}{s_2}s_I\lambda_I^2\frac{\eta}{\lambda_I}\left[s_1+s_M\lambda_M^2\right](R-\rho_1)\\
=&\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}(R-\rho_1)\left(s_1+s_M\lambda_M^2\right)(R-\rho_1)\frac{1}{s_2}\left(1-s_I\lambda_I^2\frac{\eta}{\lambda_I}\right)\\
=&\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}(R-\rho_1)\left(s_1+s_M\lambda_M^2\right)(R-\rho_1)\frac{1}{s_2}\frac{1}{1+\mathcal{B}}
\end{aligned}$$

$$\rightarrow w=\frac{\mathbb{C}}{\mathbb{D}}$$

where

$$\begin{aligned}
\mathbb{C}&=\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}(R-\rho_1)\left(s_1+s_M\lambda_M^2\right)(R-\rho_1)\frac{1}{s_2}\frac{1}{1+\mathcal{B}}\\
&=s_1\times\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}(R-\rho_1)^2\frac{1}{s_2}\frac{1}{1+\mathcal{B}}\\
&+s_M\lambda_M^2\times\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}(R-\rho_1)^2\frac{1}{s_2}\frac{1}{1+\mathcal{B}}
\end{aligned}$$

$$\mathbb{D}=2\lambda_M^2(R-\rho_1)^2+s_M\lambda_M^2\times\underbrace{\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}(R-\rho_1)^2\frac{1}{s_2}\frac{1}{1+\mathcal{B}}}_{\mathbb{E}}$$

$\mathbb{E}=\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}(R-\rho_1)^2\frac{1}{s_2}\frac{1}{1+\mathcal{B}}=\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M}{s_2^2}\frac{1}{1+\mathcal{B}}\lambda_M^2(R-\rho_1)^2$

Finally,

$$
\begin{aligned}
w &= \frac{\left(s_1 + s_M\lambda_M^2\right)\frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M}{s_2^2}\frac{1}{1+\mathcal{B}}\lambda_M^2\left(R-\rho_1\right)^2}{2\lambda_M^2\left(R-\rho_1\right)^2 + s_M\lambda_M^2 \times \frac{\alpha_2}{\alpha_1}\delta\frac{\beta_\tau^2}{\tau}\frac{s_M\lambda_M^2}{s_2}\left(R-\rho_1\right)^2\frac{1}{s_2}\frac{1}{1+\mathcal{B}}} \\
&= \frac{\left(s_1 + s_M\lambda_M^2\right)\delta\frac{\beta_\tau^2}{2\tau}\frac{s_M}{s_2^2}\frac{1}{1+\mathcal{B}}}{\frac{\alpha_1}{\alpha_2} + s_M\lambda_M^2 \times \delta\frac{\beta_\tau^2}{2\tau}\frac{s_M}{s_2}\frac{1}{s_2}\frac{1}{1+\mathcal{B}}} \\
&= \frac{\left(s_1 + s_M\lambda_M^2\right)\Psi}{\frac{\alpha_1}{\alpha_2} + s_M\lambda_M^2\Psi} \qquad \text{(A.3)}
\end{aligned}
$$

for $\Psi = \delta\frac{\beta_\tau^2}{2\tau}\frac{s_M}{s_2^2}\frac{1}{1+\mathcal{B}}$.

APPENDIX B

OTHER IMPLICATIONS OF THE BASELINE MODEL.

1 Accuracy of parent's expectation

The parent does not observe the realized investment and punishment when she implements them, but her expectation of their values becomes more accurate as the signaling precision increases. Parent's expectation of child outcome is

$$RE\left[\theta_2\right] + R\beta_\tau E\left[a_2\right] - \frac{\tau E\left[a_2^2\right]}{2}$$

where expectation is taken over the realization of M and I. Child, on the other hand, observes the realization of inputs so his expectations are affected the same way outputs are affected. Child's expectation of outcome is where expectation is taken over R.

Tables B.1 and B.2 show that parent's expectation is more accurate when the parent has more skill. The interaction of child's expectation and parent's AFQT is small and insignificant. Parent's expectation is more accurate than child's expectation for college outcome, while child's expectation is more accurate than parent's expectation for high school graduation outcome. This finding is inconsistent with the idea that the child has more private information about long-run prospect of the child. Child seems to have more information about the short-run outcome of the child, while the parent has more information about the long-run outcome of the child.

Table B.1: HS Graduation and Parent's Expectation

	(1)	(2)	(3)
	hsgrad20	hsgrad20	hsgrad20
Mother expects at least HS grad	0.277	0.323	0.253
	(0.223)	(0.250)	(0.180)
Mother.xAFQT	0.252	0.286	0.405*
	(0.244)	(0.255)	(0.178)
Child expects at least HS grad	0.707***	0.694***	0.776***
	(0.055)	(0.060)	(0.062)
Child.xAFQT	0.162	-0.015	-0.005
	(0.136)	(0.086)	(0.106)
AFQT	-0.389	-0.245	-0.383*
	(0.274)	(0.233)	(0.190)
Observations	869	786	786
R^2	0.158	0.189	0.282
Adjusted R^2	0.149	0.134	0.149
Model	RE	State FE	County FE

Note: Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. Expectation: child will at least graduate HS. Additional controls include lagged cognitive and behavior factors, indicators for race, child's birth year, parent's birth year and child's gender. AFQT is parent's AFQT score.

Table B.2: College Attendance and Parent's Expectation

	(1)	(2)	(3)
	somecol22	somecol22	somecol22
Mother expects at least 4yr college	0.256***	0.263***	0.243***
	(0.032)	(0.034)	(0.036)
Mother.xAFQT	0.086**	0.081*	0.078*
	(0.031)	(0.033)	(0.035)
Child expects at least 4yr college	0.097**	0.076	0.077
	(0.035)	(0.040)	(0.043)
Child.xAFQT	0.028	0.034	0.011
	(0.038)	(0.044)	(0.048)
AFQT	0.007	0.009	0.018
	(0.036)	(0.043)	(0.045)
Observations	1040	936	936
R^2	0.357	0.382	0.434
Adjusted R^2	0.352	0.345	0.345
Model	RE	State FE	County FE

Note: Standard errors are in parenthesis, clustered at the household level. * p¡.05, ** p¡.01, *** p¡.001. Expectation: child will at least graduate college. Additional controls include lagged cognitive and behavior factors, indicators for race, child's birth year, parent's birth year and child's gender. AFQT is parent's AFQT score.

APPENDIX C

LITERATURE SURVEY ON THE PREVALENCE OF PARENTAL PUNISHMENT

Without direct measurement of parental punishment in the data, empirical analysis relies entirely on the self-report of the parents. To gain some understanding of the validity of self-reported punishment measures, I survey the literature on the prevalence of punishment in psychological literature. Much of the literature focus on the parental use of corporal punishment.

[64] (chp. 2) report the prevalence of corporal punishment based on the nationally representative US survey they conducted in 1995. Their definition of corporal punishment includes spanking on the bottom with and, slapping on the hand or limb, pinching, shaking (for age 3 or above), hitting on the bottom with an object, and slapping on the face or head. These are legal forms of corporal punishment provided that they do not lead to injury. In page 25, they show that based on a nationally representative 1995 survey, above 80% percent of parents of children at age 7 or 8 report that they used corporal punishment in the past 12 months. The percentage remains above 50% for ages between 1 and 12. More than 30% of parents with children below age 1 and more than 10% of parents with children at age 17 report having used corporal punishment. For children between ages 6 and 14, the frequency of corporal punishment in the past year for those used it is between 4 and 16 (page 26). The authors of the study believe that these are severe underestimates because spanking children is such a routine, taken-for-granted, and unremarkable event that few parents realize how often they have done it over the course of a year. Similar result is demonstrated in an older study which found that the frequency of spanking recalled during an interview is only one-sixth of the frequency recorded in a parenting diary ([40]).

[62] compare the parent and child (9 y.o.) report of corporal punishment in Fragile and Child Well Being study. Across White, Hispanic and Black ethnicities and child gender, at

least 50% of the mothers reported having used corporal punishment in the past 12 months. Children are slightly less likely to report that corporal punishment took place last year. The predictors of corporal punishment use were similar whether it was reported by the mother or the child—poverty status and child's externalizing behavior were important predictors. [33] surveyed a sample of 251 households (mean child age 8 y.o.) in a small town in southeastern US. Children were from two-parent families and did not have significant physical or mental illnesses. In that sample corporal punishment was reported by 95% of the parents and 89% of the children.

[59] surveyed 4 nationally representative data sets for years 1988-2011, focusing on families with 5 to 7 y.o. children. The include datasets are NLSY79 Child Supplement, PSID CDS, ECLS-K 1998/1999, and ECLS-K 2010/2011. While the support for corporal punishment declined across all SES groups over time, even in 2011, 12% of the parents in 90th-percentile household income endorsed the use of physical punishment, and 10% of them reported having used it in the past week. Support for using time-out increased over time, and between 70 to 80% of the parents in all SES groups endorsed its use.

[45] report the frequency of corporal punishment in the household based on audio recording device. The sample is based on 33 mothers of 2-to-5-year-old children recruited from day care centers in a large southwestern US city. They wore recording device for up to 6 evenings, between 5pm and 10pm. They found that after corporal punishment, children were misbehaving within 10 minutes after 73% of the incidents. Mother's self reports corresponded with audio recordings 81% of the times. Most of the discrepancy were in mother self-reporting corporal punishment but without audio evidence. The frequency of corporal punishment was much higher than what is reported in the literature. Finally, the recordings reveal that most parents were responding either impulsively and/or emotionally, rather than instrumentally and intentionally."

These studies suggest that recall bias is a significant concern, because the use of punishment is considered such a routine activity in the household. Child's externalizing behavior is an important predictor of parent's punishment use regardless of the identity of the reporter. Although the parent's endorsement of corporal punishment has been declining over time, studies conducted in as late as 2011 report surprisingly high rate of self-reported use of corporal punishment use across ethnic groups, child gender and child age. Report of corporal punishment by the parent and the child do not seem to differ in a systematic way.

I unfortunate was not able to find out whether the validity of self reports were correlated with parent's SES status. However, corporal punishment use is less likely for parents with high income ([66]). [59] report that 80% of the parents in high-income household endorse the use of non-corporal punishment. My analysis on punishment use shows that the proportion of households using any form of punishment is not significantly different across households. While caution is necessary, in this study I maintain the assumption that parents truthfully report their corporal punishment use to the best of their knowledge.

APPENDIX D

SELECTED SURVEY MEASURES OF PARENTAL HARSHNESS

Table D.1: Selected Measures of Parental Harshness

	Triple P ([48])
over-reactive	If my child is naughty or behaving inappropriately, I am so angry and frustrated that my child knows it.
	When there is a problem with my child, things get out of control and I do things I wouldn't want.
inconsistent	If I had to do something, because my child has misbehaved, I'll stick with what I have said, ignoring his protests.
	If my child does something I do not like, I let it pass easily.
	If we are not at home, I let my child get away with more than at home.
	I threaten with things that I know I won't do.
	LSAC ([35])
over-reactive	How often are you angry when you punish this child?
	How often do you feel you are having problems managing this child in general?
inconsistent	When you give this child an instruction or make a request to do something, how often do you make sure that he/she does it?
	How often does this child get away with things that you feel should have been punished?
	When you discipline this child, how often does he/she ignore the punishment?
	NLSCY ([28])
over-reactive	How often does parent get angry when punishing child?
	Does punishment depend on parent's mood?
inconsistent	Does parent follow through with threatened punishment?
	How often does child get away with behavior that should be punished?
	How often does child avoid punishment?
	How often does child ignore punishment?