Optimal Parenting Styles: Evidence From a Dynamic Model with Multiple Equilibria *

Marco Cosconati‡
Research Department
Bank of Italy
marco.cosconati@gmail.com

Abstract
There is little consensus among social science researchers about the effectiveness of alternative parenting strategies in producing desirable child outcomes. Some argue that parents should set strict limits on the activities of their adolescent children, while others believe that adolescents should be given relatively wide discretion. In this paper, I develop and estimate a model of parent-child interaction in order to better understand the relationship between parenting styles and the development of human capital in children. Using data from the NLSY97, the estimates of the model indicate that the best parenting style depends on the stock of adolescent human capital. Setting strict rules increases the study time of children with low skills, but is detrimental for adult human of the more knowledgeable teenagers.

JEL Classification Code: J13, J18, J24.
KEYWORDS: Child development, parenting style, human capital, time use.

*PRELIMINARY AND INCOMPLETE - PLEASE DO NOT CITE OR QUOTE
‡I am grateful to Kenneth I. Wolpin and Antonio Merlo for their guidance and encouragement. I also thank Petra Todd for precious advise. I benefited from useful discussions with Melissa Tartari, Flavio Cunha and Jim Heckman. I thank the participants to the the seminars at La Pietra-Mondragone Workshop, UPenn, University of Bristol, University of Essex, University of Chicago, Tinbergen Institute, Bank of Italy and MOVE Workshop in Barcelona.


1 Introduction

When questioned about who decides on how late they can stay out at night, about 67% of youths between the ages of 12 and 13 surveyed in the 1997 youth cohort of the National Longitudinal Surveys (NLSY97) declared that parents decide, 30% reported that the decision is jointly made with their parents and 3% that they alone decided. The responses are more heterogenous for questions about who decides what TV shows the youth can watch or about who the youth’s friends can be. Although parents differ in the degree of self-regulation left to their children, there is no consensus on how effective alternative parenting strategies are in producing desirable child outcomes. Some argue that allowing children more discretion in making these kinds of choices is a better approach to parenting, while others believe that establishing strict rules is best at inducing good behavior. In the words of Bornstein (1991):

“Despite the fact that most people become parents and everyone who ever lived has had parents, parenting remains a mystifying subject about which everyone has opinions, but about few people agree. Freud once listed bringing up children as one of the three ‘impossible profession’-the other two being governing nations and psychoanalysis.”

The old debate on how to raise a child has recently been reinvigorated by the best seller *Battle Hymn of the Tiger Mother* by Amy Chua. As summarized by an article appeared on the *Economist*1, the author argues that “Chinese matriarchs are better because they are tougher, stricter and readier to be loathed for banning children from almost any form of fun, from play dates to -horrors!-computer games. Indulgent Western parents, cosseting their baa-lambs’ self-esteem and releasing them to play in the mud when they could be doing extra arithmetic or practicing scales, are condemning them to a life of underachievement.” Even president Obama, during his first presidential campaign, argued in favor of restricting children’s recreational activities:

We’re going to have to parent better, and turn off the television set, and put the video games away, and instill a sense of excellence in our children, and that’s going to take some time.

In this article I propose an economic model apt to address the issue of optimal parenting. The models gives a novel insight on why a “Chinese” approach to rear children might be undesirable2. To put it simply, I argue that if children act as a rational and

---

1The article is titled “Tiger cubs vs precious lambs” and appeared on the Economist on January 20th 2011.

2According to Chua the discrepancy between “Chinese” and “Western” approach to parenting is also due to the different concerns that punishment can have on children’s self-confidence:

“For example if a child come home with an A minus on a test, a Western parenting will most likely praise the child. The Chinese mother will gasp in horror and ask what went wrong. ...Chinese parents demand perfect grades because they believe that their child can get them. If their child doesn’t get them, the Chinese parent assumes it’s because the child didn’t work hard enough. That’s why the solution to substandard performances is always to excoriate, punish and shame the child. The Chinese parent believes that their child will be strong enough to take the shaming and to improve from it.
forward looking player and if in equilibrium poor performances are punished by strict rules - which I call PS equilibria- a too demanding parenting policy might not maximize children’s achievement. For children prefer loose limits, there is an incentive to work hard to fall “above the bar”. However, if parents set too high standards, the chance of reward in the future requires a disproportionate opportunity cost. Thus, once the choice problem of the child is allowed to be dynamic, there exists a trade-off related to an authoritarian parenting style. Setting strict limits is not necessarily an ex-ante superior policy for every type of children. In general the strategy which maximizes the child’s human capital makes use of both rewards and punishment.

To analyze the development of adolescent human capital as an equilibrium outcome of the parent-child interaction I propose a dynamic finite-horizon principal-agent model in which parties cannot commit to a course of actions. In the model parental rules diminish, on average, the time the child is allowed to enjoy leisure. The child, given the “induced” time constraint, decides the fraction of time dedicated to knowledge acquisition. Because the parent can only imperfectly determine the child’s effort, there is a moral hazard problem. Strict limits are also disliked by the parent because they require a higher monitoring/enforcement cost. The game is non-cooperative because parents and children, who are both rational and forward looking, attach different values to the child’s leisure and adult human capital. The dependence of the child’s behavior on his human capital stems from the endogenous formation of preferences: human capital increases the child’s discount factor. Using the standard methods of monotone comparative statics, I show that, under concavity of the players’ payoff functions, PS equilibria emerge.

The model is estimated by simulated maximum likelihood using five years of data (1997 to 2001) on all the cohorts of the NLSY97. The model is mapped to the data by dividing the school year in two semesters (fall and spring) starting from the 6th grade up to the 12th grade, which gives information on the parent-child interaction for 14 “periods”. The sample consists of youths who eventually graduate from high school with a straight grade progression, living with their parents during the school attendance period and with no co-residing sibling. I use all the responses available in the NLSY97 on who is the person setting the limit on curfew, friends and TV shows as observable inputs of the “discipline” production function; both the GPA achieved at the end of each academic year and the PIAT test scores are assumed to be informative about the child’s cognitive skills, while self-declared time spent doing homework measures the child’s effort. Children are assumed to differ only in an observable way through the observed academic performances and the test scores. Parents, on the contrary are allowed to be unobservably heterogeneous in terms of their monitoring costs. I deal with the potential multiplicity of PS equilibria arising from the possibility that multiple cutoffs are optimal without imposing an equilibrium selection rule. I adopt a two-step procedure: in the first step I recover the child’s preferences and the technologies by treating the cutoffs as estimable parameters. Given that the child’s best response is a function, in the first step I solve numerically for the equilibrium effort of the child and compute the associ-

---

3Empirical evidence on the importance of time use choices for human capital accumulation can be found in Lee (2012) and Stinebrickner and Stinebrickner (2008).

4Concavity is obtained through a standard assumption in the principal-agent model.
ated likelihood function. The estimation procedure also allows for classification error in the limits variable and measurement error in the time spent studying. The likelihood contribution for a given child-parents pair is the probability of observing the limits, the time spent studying in 1997 (the only year it is available), the PIAT test scores and GPA. Once this subset of primitives has been recovered in the second step I use the definition of the equilibrium cutoffs, as the basis of a GMM estimator which allows me to recover parental preferences.

I use the results from the first step to perform two thought experiments. In the first one I analyze the response to the most lenient parenting style, which I conceptualize as if the government were re-assigns the “property rights” on the decision on the limits from parents to children. The results indicate that the no child would benefit from such as policy: there are decreases in both the GPA and the PIAT test scores independently on the level of skills at the age of 12. However these effects are heterogenous: they are quite large(small) for children with low(high) “initial” human capital. In the second experiment I study the impact of laws which eliminate parents’ monitoring costs, such as mandatory curfew laws, bans and prohibitions. If a government does not internalize the loss suffered by teenagers subsequent to lower leisure availability and only seek to maximize children’s effort, what is the cost of passing these laws? The results indicate that the human capital of the least able would increase substantially while children with high level of adolescent skills would experience a decrease in both the GPA and the PIAT test scores. The intuition underlying this result, is that although they reduce the cost of setting limits for parents who would otherwise have done so, these public policies restrict the instruments available to parents to induce effort. In some families it could actually be optimal to increase parents’ monitoring/enforcement cost- this way it’s as if the government provides a credible commitment device.

This paper is organized as follows. In sections 1.1 I describe the relationship of this work to existing papers on parent-child interaction and on the technology of skills formation. In section 1.2 I provide some mathematical background which I use to show how PS equilibria can emerge. In section 2 the game is described and the conditions to obtain PS equilibria are spelled out. In section 3 I describe the data. In section

---

5 Curfew laws restrict the right of children to be outdoors or in public places during certain hours of the day. If a teenager breaks curfew, he or she can be temporarily detained by police and returned home. Currently, there is no state curfew. Such laws or ordinances are typically passed and enforced by local municipalities, cities and townships. Courts in California have generally upheld such laws as long as the local ordinance seeks to discourage loitering or remaining in certain places after certain hours. There are a few papers who estimate the impact of these kind of policies on teenagers’ outcomes. For example Kline (2012) finds sizable effects of juvenile curfew laws on crime involvement; Lee (2012) finds substantial impact of repealing Sunday closing laws on educational achievement.

6 Coleman and Roker (2001) illustrates some aspects of this discussion in UK:

“In the early days of the Labour government there was much discussion in the media about where the boundary lay between interferences and appropriate involvement. There are some who view the family as a private institutions and believe that, except in the extreme cases, it should be largely free from state interferences. On the other hand there are those who take the position that government does have a role in enhancing support for the family and in modifying attitudes about the importance of parenting.”
I describe the estimation algorithm. The results are described in section 5. I then conclude by summarizing and pointing at existing empirical and theoretical challenges that parent-child games pose to researchers of family economics.

1.1 Additional Background Literature

This paper is related to two strands of literature. The first is represented by the literature on skills formation (see Cunha and Heckman (2007), Cunha and Heckman (2008), Cunha, Heckman, and Schennach (2010) and Del Boca, Flinn, and Wiswall (2012)). These papers characterize the empirical properties of the technology of skill formation, such as self and cross productivity of skills, as well as the impact of parental investments. In these models the child has no active role in developing his own human capital—he is treated as a passive actor. Although I also analyze the relationship between parental influence and the development of human capital in children, I focus on parental incentives rather than school and home inputs. The main conceptual difference is that while the latter enter directly into the skills production function, the former only alter children’s behavior, which is an input of the human capital production function (indirect effect). This abstraction allows to analyze how adolescent human capital evolves as a result of the strategic interaction between parents and children, which is absent in the literature on skill formation.

The second strand of the literature related to this article, is represented by papers on parent-child interaction. The novelty here is that both players are forward looking; in particular the child recognizes the his behavior will affect the future play of the game. For example, Akabayashi (2006) analyzes how child maltreatment can emerge as a result of biased beliefs parents hold about the child’s ability; in his model the child is myopic. Weinberg (2001) analyzes a standard static principal-agent model in which the parent can use both monetary and non-monetary incentives. Hotz, Hao, and Jin (2008) study parent-child interaction to rationalize the existence of birth-order effects. The parent has an incentive to build a reputation of being tough by setting stricter limits on older children, yet the does not take into account how his behavior changes the parent’s actions. Bursztyn and Coffman (2012) provide interesting results on the information structure underlying parent-child interaction. They show that in poor urban Brazil parents prefer to link cash transfer to their children’s attendance, rather than receiving a higher amount unconditional on the child’s decisions. However parents prefer to obtain direct information about attendance rather than paying for conditionality. These results indicate that parents have little ability in predicting the child’s behavior—moral hazard—and prove the existence of miscongruence in preferences. Lizzeri and Siniscalchi (2008) develop a model of supervised learning. Parents optimally decide how much to shelter children from the consequences of their mistakes and how much to let them learn from experience. Because the child, being unaware of parental intervention, does not condition his action on the future play of the parent, they abstract from strategic aspects, which is the focus of this article.

---

7 The estimation of the skill technology problematic because of challenging econometric issues
8 The early literature, see Becker (1974) and Becker (1981), abstracts from this distinction.
The greatest body of research on parenting practices has been produced by sociologists and researchers of child development. Baumrind (1968) divides parents into three categories: authoritarian, permissive and authoritative. Authoritative parenting broadly refers to the method used by those parent who enforce their rules, set clear standards, provide consistent discipline that is not overly punitive. Authoritative parenting is also associated to the concept of “flexibility”. Parents are flexible when they alter their behavior when appropriate. In contrast authoritarian parenting style is sometimes referred to as the military parenting style. This type of parents puts an emphasis on obedience, and usually impose very strict family rules. Authoritarian parenting stifles intellectual growth and creativity. It also encourages children to either rebel against their parents. Permissive parents give up most control to their children; they few, if any, rules which are not consistently enforced. They do not set clear boundaries or expectations for their children’s behavior and tend to accept however the child behaves. It has been documented that socio-economic status is correlated with parenting methods (see Bornstein (1991)). In particular middle-class parents are more likely to choose an authoritative parenting method while in low class households authoritarian parenting is the most common.

1.2 Mathematical Background

I argue about the existence of PS equilibria using the standard tools of monotone comparative statics (MCS) (see Van Zandt (2002) for an introduction). MCS methods, introduced by Topkis (1978), allow to show how the optimal choices change in response to an increase of exogenous parameters, e.g. state variables, without having to assume continuity of the choice set, differentiability and/or concavity of the payoff functions. Because concavity has little to do with comparative statics, these method distinguish the role of diminishing returns in driving the economic arguments from their role in making the implicit function theorem applicable. I will start by providing a set of definitions.

**Definition 1.** A partially ordered set is a set endowed with a binary relation that is reflexive, antisymmetric and transitive. Such a set is a lattice if given any element $x_1$ and $x_2$, a sup (or join) $x_1 \lor x_2$ and an inf (or meet) $x_1 \land y$ of $x_1$ and $x_2$ are both in $X$. A subset $X'$ of a lattice $X$ is a sublattice of $X$ if it contains the sup and the inf of any pair of elements of $X'$.

**Definition 2.** Let $X$ be a lattice, $\Theta$ a partially ordered set. The function $f : X \times \Theta \to \mathbb{R}$ is supermodular (submodular) in $(x, \theta)$ if for all $x, y \in X$ we have

$$f(x) + f(y) \leq (\geq) f(x \lor y) + f(x \land y)$$

Moreover $f$ displays increasing (decreasing) differences in $(x, \theta)$ if for $x' \geq x$ $f(x', \theta) - f(x, \theta)$ is monotone non decreasing (increasing) in $\theta$.

For the special case of $X$ being a subset of $\mathbb{R}^m$ it is possible to establish the following relationship between supermodularity and increasing differences
Lemma 1. A function \( f : X \subset \mathbb{R}^m \to \mathbb{R} \) is supermodular (submodular) on \( X \) if and only if \( f \) has increasing (decreasing) differences on \( X \).

Interestingly for my purposes, as explained by Topkis (1998): “the set of all supermodular functions on a lattice \( X \) is a closed convex cone in the vector space for all real-valued functions on \( X \)”. This property is translated in the following lemma:

Lemma 2. Suppose that \( X \) is a lattice.

a) If \( f(x) \) is supermodular (submodular) on \( X \) and \( \alpha > 0 \), then \( \alpha f(x) \) is supermodular (submodular) on \( X \)

b) If \( f(x) \) and \( g(x) \) are supermodular (submodular) on \( X \), then \( f(x) + g(x) \) is supermodular (submodular)

c) If \( f_k(x) \) is supermodular (submodular) on \( X \) for \( k = 1, 2, \ldots \) and \( \lim_{k \to \infty} f_k(x) = f(x) \) for each \( x \) in \( X \), then \( f(X) \) is supermodular (submodular) on \( X \).

A useful property which is implied by the previous lemma is that the expected value of a supermodular (submodular) function is supermodular (submodular).

Lemma 3. If \( F(w) \) is a distribution function on a set \( W \) and \( g(x, w) \) is supermodular (submodular) in \( x \) on a lattice \( X \) for each \( w \) in \( W \), then \( \int_w g(x, w) dF(w) \) is supermodular (submodular) in \( x \) on \( X \).

Supermodularity is a cardinal property, i.e. it is not invariant to monotone transformations. The following property, introduced by Milgrom and Shannon (1994), is instead a cardinal one and serves a basis to draw monotone comparative statics.

Definition 3. Let \( X \subset \mathbb{R}^m \), \( \Theta \subset \mathbb{R}^n \) and \( f : X \times \Theta \to \mathbb{R} \). \( f \) has the single crossing property (SCP) in \( (x, \theta) \) if for any \( x'' > x' \) and \( \theta'' > \theta' \)

\[
f(x'', \theta'') - f(x', \theta'') \geq (>0) \Rightarrow f(x'', \theta') - f(x', \theta') \geq (>0)
\]

Because monotone comparative statics provides conditions on the primitives such that after an increase of the parameter(s) of interest the optimal choice increases, an order both on the constraint set and on the set of maximizers is required. Milgrom and Shannon (1994) adopt the strong set order.

Definition 4. The set \( S' \) dominates \( S \) in the strong set order if for any \( y \) in \( S' \) and \( x \) in \( S \), their supremum \((x \vee y)\) is in \( S' \) and their infimum \((x \wedge y)\) is in \( S \).

Milgrom and Shannon (1994) prove that SCP is necessary and sufficient to draw monotone comparative statics.

Theorem 1. Let \( X \subset \mathbb{R}^m \), \( \Theta \subset \mathbb{R}^n \), \( f : X \times \Theta \to \mathbb{R} \) and \( S \subset X \). Then \( \arg \max_{x\in S} f(x, \theta) \) is monotone nondecreasing in \((\theta, S)\) if and only if \( f \) has the SCP in \((x, \theta)\) and it is supermodular in \( x \).

\footnote{In the more general version of the theorem \( X \) is a lattice, \( \Theta \) is a partially ordered set and \( f \) is quasisupermodular.}
Their result is at the root of the *ordinal* approach to obtain MCS results. In this application I use a *cardinal approach* to derive my results, which rely on the increasing differences property. In fact if $f(x, \theta)$ has increasing differences in $(x, \theta)$, then $f$ has the single crossing property in $(x, \theta)$.

It is finally useful to provide a comparative statics theorem which covers the case of uncertainty in the payoff function. I will use a result of Van Zandt and Vives (2007) rather than the most commonly used theorems provided by Athey (2002), as it is slightly more general in terms of the restriction imposed on the density function. It is necessary to introduce the following definition.

**Definition 5.** Let $(\Theta, F)$ be a measurable space and let $\geq$ be a partial order on $\Theta$. Let $P_H$ and $P_L$ be two probability measures on $(\Theta, F)$. We say that $P_H$ first-order stochastically dominates $P_L$ if for all increasing functions $f : \Theta \to \mathbb{R}$ that are integrable with respect to $P_H$ and $P_L$, $\int_{\Theta} f(\theta) dP_H \geq \int_{\Theta} f(\theta) dP_L$.

Let $X$ be a partially ordered set and let $u : X \times \Theta \to \mathbb{R}$ be measurable in $\Theta$. Let $M$ be the set of probability measures on $(\Theta, F)$, partially ordered set by first-ordered stochastic dominance. Define $U : X \times M \to \mathbb{R}$ by $U(x, P) := \int_{\Theta} u(x, \theta) dP(\theta)$.

**Lemma 4.** Assume that $u$ has increasing (decreasing) differences in $(x, \theta)$. Then, on the domain of $U$, $U$ has increasing (decreasing) differences in $(x, P)$.

The previous lemma is be used to prove the following theorem which incorporates uncertainty in the payoff function of the decision maker.

**Theorem 2.** Assume that $u$ is supermodular in $x$ and has increasing (decreasing) differences in $(x, \theta)$. Then $\arg \max_{x \in X} U(x, P)$ is increasing (decreasing) in $P$.

### 2 A principal agent-model of parent-child interaction

There are two players, the *child*(he) and the *parent*(she). The game is dynamic and each *period* a Stackelberg game is played. Time is discrete and there are $T$ periods which can be thought as the child's ages. At each stage game the parent moves first, i.e. she acts as *leader* conditional on her information set. The child then responds and the payoffs of the players are realized. In what follows I will use the words human capital, skills and knowledge interchangeably. I now the describe the primitives of the game, which falls into the class of dynamic principal-agent model (henceforth PA) without commitment.

**Children's preferences** The *child* cares about his *leisure* and the human capital accumulated at the end of the game. The child chooses how to allocate his total available time (normalized to one) between leisure and acquiring knowledge. Knowledge acquisition time ($e$) does not provide direct utility and entails an opportunity cost represented by the available alternative recreational activities$^{10}$. As in the theory of patience forma-

---

$^{10}$Notice how the density of the output is not altered by the current stock of skills $k$. It will be clear later
tion of Becker and Mulligan (1997), I allow human capital to increase the child’s rate of time preferences: children with little knowledge behave more myopically\(^{11}\). Future utility is discounted according to the function \(\delta(k)\), which is strictly increasing in \(k\) (ID)\(^{12}\).

Let \(k_0\) to be the endowment of the child at the beginning of the game so that \(k_{t-1}\) is the stock of human capital of the child at the beginning of period \(t\). The payoff function of the child is given by:

\[
v(k_t, l_t, e_t) = \begin{cases} 
  u(l_t) - c(e_t) & \text{if } t < T + 1 \\
  \Xi(k_T) & \text{if } t = T + 1 \text{ (the game is over)} 
\end{cases}
\]

where \(l_t\) signifies leisure time and \(e_t\) is the time spent acquiring new knowledge. In the above formula \(k_T\) - the adult human capital of the child - maps into future earnings through an increasing function \(\Xi\). The felicity function \(u\) is increasing in its argument, \(c\) is an increasing function which measures the psychic cost associated to studying.

**Parenting styles and time constraints** The parent cares about the adult human capital of the child; she is forward looking and discounts the future at rate \(\beta \in (0, 1)\). The parent select a parenting rule \(\rho\) - a parenting style - among \(n\) mutually exclusive alternatives\(^{13}\), with \(n \geq 2\). The parent’s choice set is denoted by:

\[
R = \langle \rho_1, \rho_2, \ldots, \rho_n \rangle
\]

From the practical point of view a parenting style determines a lower (upper) bound \(\tau (1 - \tau)\) on the effective amount of learning time (leisure activities). The child solves his time allocation problem under the constraints

\[
1 = l_t + e_t \\
e_t \geq \tau_t
\]

The first constraint is the usual time constraint, with total available time being normalized to one. The second one represents the link between the parent and the child\(^{14}\).
Consistently with the existence of shocks to the monitoring technologies adopted by the parent\textsuperscript{15}, I think of \( \tau \)-the minimum amount of time the child has to spend studying-as a random variable whose distribution \( G(\tau | \rho) \) depends on the parenting rule \( \rho \). The non-degeneracy of \( \tau \) captures a \textit{moral hazard problem}: the parent cannot perfectly determine (and predict) the child’s behavior\textsuperscript{16}. Lower the variance of \( \tau \) conditional on \( \rho \), more effective is the parent in forcing the child to learn-less pervasive the moral hazard problem. This “reduced form” approach describes situations in which the parent is unable to monitor the child’s whereabouts and/or the intrinsic inability to fully draw the child’s attention to homework and other learning material.

Given that child dislikes parental monitoring (\( k_{1g} \tau \)) the following definition provides a basis to order \( \mathcal{R} \) in terms of strictness. I will say that a parenting style \( \rho_i \) is higher (stricter) than \( \rho_j \) and write \( \rho_i > \rho_j \) if and only if \( G(\tau | \rho_i) \) first order stochastically dominates \( G(\tau | \rho_j) \). As a convention higher the index more permissive the parenting rule.

**The Parent’s Preferences** The utility she enjoys at the end period \( t \) is given by the following formula

\[
 w(\rho_{t},k_{t}) = \begin{cases} 
 -\sum_{i=1}^{n} \kappa_i [\rho_{t} = \rho_i] & \text{if } t < T + 1 \\
 \Pi (k_T) & \text{if } t = T + 1 \text{ (the game is over)}
\end{cases}
\]

(4)

with \( \kappa_i > \kappa_j \) for any \( i > j \): stricter limits entail a higher monitoring/enforcement cost.

The terminal value function \( \Pi \) is strictly increasing in its argument: more knowledgeable is the child as he starts off his adult life, higher is the utility enjoyed by the parent. To the extent that the cost function incorporates not only the price paid for higher monitoring but also the discomfort the parent feels in setting a harsher parenting rule, the model can accommodate the traditional notion of altruism\textsuperscript{17}. According to this specification the parent cares about the child because she gets utility from better career prospects, as proxied by the adult human capital.

**Human Capital Production Function** When deciding how to allocate his available time the child faces a trade-off for learning time increases knowledge in the sense of \textit{first order stochastic dominance} (FOSD)\textsuperscript{18}. Let \( k \) denote the human capital produced at the end of the period and \( k \) the current stock of human capital. Let the support of \( k \) be \([k, \bar{k}]\) which is independent on the inputs, let \( F \) be the CDF of the conditional density. Effort and the stock of human capital increase future human capital in a first order stochastic dominance sense:

\[
 F(K|e'',k) \leq F(K|e',k) \text{ for any } e'' > e' \quad (1E)
\]

such an interpretation.

\textsuperscript{15}Allowing parental monitoring to be subject to random negative shocks is consistent with the malfunctioning of the monitoring devices purchased by the parent. On the contrary positive shocks can be conceptualized in terms of the unexpected help of other family members in disciplining the child.

\textsuperscript{16}Bursztyn and Coffman (2012) provides a evidence in favor of this issue.

\textsuperscript{17}To have a traditional model of altruism, such as the one proposed by Weinberg (2001), the flow utility of the child would have to be plugged into the payoff function of the parent and weighted by a parameter between zero and one.

\textsuperscript{18}This property is weaker than the commonly used MLRP.
and
\[ F(K|e, k'') \leq F(K|e, k') \text{ for any } k'' > k' \] (6)

I will explain later why it is convenient not to separate explicitly the “structural” shock from the inputs of the production function.

To summarize the link between preferences and the technology of skill formation, higher human capital induces the child to study more by increasing his “motivation”-discount rate; more effort will in turn increase his knowledge (cognitive skills) which will then affect the noncognitive skills\(^{19}\) in the following period. The endogenous choice of effort is then the mechanism through which the model reproduces the cross productivity of the skills\(^ {20}\). For the remainder of this paper I assume that the primitive functions \(F, \Pi, \Xi, c, \delta\) and \(u\) are continuous in their arguments.

**Strategies and Equilibria** The history of the game at any point in time is given by a sequence of actions and human capital realizations \(H^t = (e^t, \rho^t, \tau^t, k^t)\) where \(k^t = (k_0, \ldots, k_{t-1})\), \(e^t = (e_1, \ldots, e_t)\) and \(\rho^t\) is defined analogously. A time \(t\)-strategy for the child is a function \(\phi^c_t : H^{t,c} \to [0,1]\), where \(H^{t,c} = H^t \cup \tau_t\); a time-\(t\) strategy for the parent is a map \(\phi^p_t : H^t \to \mathbb{R}\). An optimal strategy profile for the child, \(\vec{\phi}^c = \langle \phi^c_t(H^{t,c}) \rangle_{t=1}^{T}\), is given by a sequence of strategies such that:

\[
\vec{\phi}^c \in \arg \max \mathbb{E} \left[ \sum_{t=1}^{T} v(\phi^c_t(H^{t,c})) | \vec{\phi}^p \right] \tag{7}
\]

where the expectation is taken over the random variables \(k\) and \(\tau\). Conditioning on the parent’s strategy also implicitly describes the sequence of constraints the child is subject to. An optimal strategy for the parent is such that:

\[
\vec{\phi}^p \in \arg \max \mathbb{E} \left[ \sum_{t=1}^{T} w(\phi^p_t(H^t)) | \vec{\phi}^c \right] \tag{8}
\]

A Nash equilibrium of the game is given by a pair of strategies \((\vec{\phi}^c, \vec{\phi}^p)\) such that (7) and (8) are satisfied. The existence of a Nash equilibrium in pure strategies is guaranteed by the sequential nature of the stage-game, the finite horizon and the smoothness of the payoff functions in both the state and the choice variables.

### 2.1 An Intuitive Equilibrium

Although common sense suggests that often parents *let the child go* even after evidence of misbehavior, I focus my attention on *monotone equilibria*- the parent penalizes...\(^{19}\)The noncognitive skills can be thought as a bundle of primitive economic parameters such as preferences for leisure, the discount factor, risk aversion and self-control. It would be unfeasible to include all of those in the model.

\(^{20}\)Cross productivity arises when each type of skills is an input to produce the other. See Cunha and Heckman (2007) for a model of skills production function which incorporates both self-productivity and dynamic complementarity. The existence of both has been documented by Cunha et al. (2010).
bad outcomes by setting stricter limits\textsuperscript{21}. Cut-off equilibria of this kind have been analyzed by Banks and Sundaram (1998) and in the political economy literature by Ferejohn (1986) and Fiorina (1977). I refer to such a strategy as to a (PS) strategy (punishing strategy) or as to a strategy with the (PS) property; analogously I call an equilibrium in which a (PS) strategy is played a (PS) equilibrium. The following definition characterizes a (PS) equilibrium when the choice set of the parent is a given by $n$ possible values of the parenting style:

**Definition 6.** A (PS) equilibrium is characterized by a parental strategy $\langle \rho_t(k_{t-1}) \rangle_{t=1}^{T}$ such that there exists a $(n + 1) \times T$ matrix $c = [\bar{c}_1, \bar{c}_2, \ldots, \bar{c}_T]$, whose generic $(n - 1$-dimensional) column $\bar{c}_t$ contains the equilibrium cut-offs faced by the child in period $t - 1$. The $i$-th element of $\bar{c}_t$ is denoted by $\bar{c}_{ti}$ and it is such that:

$$k = \leq \bar{c}_{1t} \leq \cdots \leq \bar{c}_{nt} \leq \bar{c}_{n-1t} = \bar{k}$$

$$\rho_t(k_{t-1}) = \rho_i \iff \bar{c}_{ti} < k_{t-1} \leq \bar{c}_{ti+1} \text{ for all } i = 0, \ldots, n \text{ and } t = 1, \ldots, T$$

The above definition, in which $\bar{c}_{0t} = k$ and $\bar{c}_{nt} = \bar{k}$, captures the idea that good performances are rewarded by setting a more permissive parenting style (higher index). The inequalities in the definition reflect an assumption grounded on Pareto optimality: when indifferent between two parenting styles the parent chooses the more permissive one, thereby making the child better off. The elements of the matrix $c$ depend on the initial condition $k_0$ and all the parameters of technologies and utility functions.

As in the standard PA model, a natural question to ask is: under what conditions does the parent find optimal to reward good performances - choose a high $\rho$? In short the answer to the original question is: concavity of the payoff function of the players and complementarity between effort and human capital are sufficient (not necessary) to obtain a (PS) equilibrium. The way I achieve this result in the context of my dynamic game is different than the standard PA model in that I do not make use of the first order conditions\textsuperscript{22}. As explained by Conlon (2009), there are two different approaches to obtain the concavity of the objective function of the agent: they differ on how the payoff function of the agent is represented. In one case the researcher represents the output in terms of a parametric form of $e$ and a random shock $\epsilon$, for example through the production function $\phi$

$$H = \phi(e, h, \epsilon)$$

This is what Conlon (2009) calls the state space representation. The concavity of the objective function following a state-space representation, is achieved by imposing con-

\textsuperscript{21}According to Amy Chua “Chinese parents demand perfect grades because they believe that their child can get them. If their child does not get them, the Chinese parent assumes it’s because the child didn’t work enough. That’s why the solution to substandard performances is always to excoriate, punish and shame the child. The Chinese parent believes that their child will be strong enough to take the shaming and to improve from it.”

\textsuperscript{22}Ex-post the conditions I impose are such that the first order approach is valid. Such a procedure consists in replacing the incentive compatibility constraint of the agent with his first order condition. As originally pointed out by Mirrless (1999) this procedure is in general incorrect. The conditions under which the first-order approach is valid are such that the objective function of the agent is concave whenever the principal’s first order condition hold. Perhaps surprisingly those conditions turn out to be rather stringent. See page 200 of Laffont and Martimort (2002) for a general introduction and section 3 of Conlon (2009) for a technical summary of the available conditions.
ditions on $\phi$ and $\Xi$. Alternatively the researcher can work with the so called Mirrlees representation. In this case the research will impose some structure on $F(K|e,k)$, typically the (CDFC) condition. I adopted a Mirrlees representation. Such an approach turns out to be particularly convenient when working with a dynamic model. I now cast the problem in a recursive way assuming that a (PS) equilibrium is played.

2.1.1 Value Functions

Let’s start considering the problem of the child, who takes the best response function of the parent, $\rho(k)$, as given. From now on I omit the time indexes, although it should be kept in mind that the policy functions and the value functions depend on time. I define the value functions associated to a given equilibrium by proceeding backwards. Conditional on the realization of $\tau$ following the action prescribed by $\rho(k)$ and given a human capital $k$, the child wishes to maximize with respect to $e$ and $l$ the following objective function:

$$U(e,l,k) =: u(l) + \delta(k) \int_{k}^{\Xi} dF(K|e,k)$$

s.t. (2) and (3)

where the constraint set of the maximization problem reflects the degree of (realized) parental intervention. Let $\mathcal{P}(k,\tau)$ denote the above program where, as in the remainder of the section, I omit the time indexes of the endogenous variables. Let $e(k,\tau) \in \langle \arg\max \mathcal{P}(k,\tau) \rangle$. The last period’s value function associated to $e(k,\tau)$ is defined as

$$V(k|\tau) := \max_{e,l} U(e,l,k)$$

s.t (2) and (3)

\[23\] Such an approach has been followed by Jewitt (1988).
\[25\] As explained by Conlon (2009) because it does not separate out the production function $\phi$ from the density of $\epsilon$ multiple state-space representation can be associated to a given Mirrlees representation.
\[26\] On the contrary if the researcher works with the state-space representation to obtain concavity, she may have to impose conditions which yields concavity of the value functions While the application of the envelope theorem may facilitate the task in the case of the value function of the child, the same may not be true for the parent. Indeed the future component of her objective function depends on a “non-primitive” function such as $e(k,\tau)$ which also need to be concave. It is in turn possible to obtain this if the researcher is willing to impose conditions on the third derivative of the utility function $u$ and the terminal value function $\Xi$. Given these considerations, on net, placing structure on $e(k,e)$ seemed an easier way to get concavity, while leaving probably more flexibility in the choice of the parametric forms assumed for the payoff functions of the players.

\[27\] The set $\langle \arg\max \mathcal{P}(k,\tau) \rangle$ is non empty due to the Weistrass’ thereom, provided that $\int_{k}^{\Xi} dF(K|e_{n},k) \rightarrow$ for any sequence $e_{n} \rightarrow e$. This is the case if $\Pi$ is continuous and $F$ is continuous and of bounded variation, due to Helly’s convergence theorem. $\mathcal{P}(k,\tau)$ needs not to be a singleton as I did not assume that the payoff function in quasi-concave.
Using this definition it is then possible to write down the \( T - 1 \) objective function of the child as follows:

\[
U(e,l,k) = u(l) + \delta(k) \int_{0}^{1} \int_{0}^{k} V(K|\tau)dG(\tau|\rho(k))dF(K|e,k) \tag{12}
\]

s.t (2) and (3)

Let

\[
V(K|\rho_i) = \int_{0}^{1} V(K|\tau)dG(\tau|\rho_i) \tag{13}
\]

denote the expected value function the child computes, when the human capital realization at time \( T \) is \( k \) prior the a given constraint \( \tau \) realizes. Such a value function, associated to a given (PS) equilibrium, can be written as follows:

\[
\left\{ \begin{array}{ll}
V(K|\rho_1) & \text{if } c_1 \leq K < c_2 \\
V(K|\rho_2) & \text{if } c_2 \leq K < c_3 \\
\vdots & \vdots \\
V(K|\rho_i) & \text{if } c_i \leq K < c_{i+1} \\
\vdots & \vdots \\
V(K|\rho_n) & \text{if } c_n \leq K < c_{n+1} \\
\end{array} \right.
\]

The value function is increasing in the state variable \( k \) both because when a (PS) strategy is played the expected value of \( \tau \) is decreasing in \( k \), and because conditional on a parental strategy the expected utility is increasing in the stock of human capital. Given this definition the \( P_{T-1}(k,\tau) \) can be re-written as:

\[
\max_{e,l} \left\{ U(e,l,k) = u(l) + \delta(k) \sum_{j=1}^{c_{j+1}} V(K|\rho_j)dF(K|e,k) \right\} \tag{14}
\]

s.t (2) and (3)

The value functions associated to periods programs \( P_t(k,\tau) \), with \( t = T - 1, T - 2, \ldots, 1 \), are defined analogously to (11), completing the set-up of the child's problem in a recursive fashion.

I now turn to the problem of the parent proceeding backwards. At time \( T \) the parent anticipates \( e(k,\tau) \) which can be plugged into her objective func-
tion. She solves the following problem:

$$\max_{\rho \in \mathbb{R}} \left\langle U^p(k, \rho) = -\kappa(\rho) + \beta \int_0^k \Pi(K) dF(K|\epsilon(k, \tau), k) dG(\tau|\rho) \right\rangle$$ \hspace{1cm} (15)$$

Let $Q_T(k)$ denote the above program and $\rho(k) \in \arg \max Q(k)$ an element of the best response correspondence. The value function of the parent at time $T$ associated to $\rho(k)$- $V^p(k)$- is then defined as the value attained by $U^p(k, \rho)$ at $\rho(k)$.

If a (PS) strategy is played, the objective function of the parent at period $T - 1$, reads as:

$$U^p(k, \rho) = -\kappa(\rho) + \beta \int_0^k V^p(k) dF(K|\epsilon(k, \tau), k) dG(\tau|\rho)$$ \hspace{1cm} (16)$$

s.t. $V(K|\rho) \geq q$ \hspace{1cm} (17)$$

The value function associated to the $T - 1$ problem is then defined as usual. Proceeding backwards and using the characterization of a PS equilibrium, the problem solved by the parent at each node of the game is described in a recursive fashion.

### 2.1.2 Sufficient conditions for (PS) equilibria

The natural argument to sustain that a (PS) strategy is optimal for the parent requires to show that the hypothesis of the monotonicity theorem hold. This is the case if the choice set expands as $k$ increases and the objective functions of the parent, in each period, are submodular (or display decreasing differences - DD) in $(k, \rho)$. As it appears from (15) and (16) those objective functions incorporate a derived (non primitive) object, namely $\epsilon(k, \tau)$. This feature complicates the analysis. The issue I face can be phrased, in more general terms, as follows. Consider an objective function $g(x, y, z)$ defined on compact intervals. Let $x(y, z) \in \arg \max g(x, y, z)$. Under what conditions on $g$ is $x(y, z)$ supermodular in $(y, z)$? To the best of my knowledge, there is no result which answers this question. In order to deal with this issue I first show in the next lemma that under certain conditions $e(k, \tau)$- the best response of a child endowed with human capital $k$ subject to constraint $\tau$ - displays decreasing differences (henceforth DD) in $(k, \tau)$.

### Footnotes

28 Differently than a standard PA model, as in Weinberg (2001), I abstract from the child's participation constraint. I will explain in the proof of lemma 6 why it is convenient to do so.

29 The ensures that the payoff function satisfies the single crossing property in $(\rho, k)$ making possible to apply the monotonicity theorem of Milgrom and Shannon (1994).

30 If the first order condition is necessary and sufficient to characterize the optimal solution hold one may be tempted to adopt the following strategy. First apply the IFT to the FOC to compute $\nabla(\epsilon(k, \tau), k)$ and then differentiate again this expression with respect to $z$ to compute $\nabla_{yz}(\epsilon(k, \tau), k)$. While the first operation would be legitimate, the second would not because the IFT only provides a formula to compute the gradient of $\nabla$ with respect to $x$ and $y$. 
Lemma 5. If $e(k, \tau)$ is (strictly) increasing in $k$ at any given $\tau$ (I) and the “unconstrained” objective function of the child, i.e. when $\tau = 1$, has a unique global maximum (U), then $e(k, \tau)$ displays (strictly) decreasing differences (DD) in $(k, \tau)$.

The proof, relegated to the appendix, is trivial once we refer to picture A which depicts the best responses of the objective functions drawn in figure A. For any $k_H > k_L$, under (I) and (U) $e^H$-the (unique) maximum at $k_H$- is higher than $e^K$, the maximum at $K_L$. The best responses are constant at their unconstrained global maximum if $\tau$ is lower than such a value; when $\tau$ is higher the constraint is binding and the optimal effort equals $\tau$.

In light of this lemma one way to obtain (PS) equilibria is to find conditions such that 1) $e(k, \tau)$ has the DD property - (I) and (U) hold- and 2) DD is inherited by the payoff function of the parent, are satisfied simultaneously.

The following lemma addresses issue 2) by making use of the Rogerson’s convexity of the distribution function condition (CDFC) and imposing further structure on the distribution of the output.

Lemma 6. If $F(K|e,k)$ is separable in $e$ and $k$, i.e. it can be written as

$$F(K|e,k) = F_1(K|e) + F_2(K|k) \text{ (AD)}$$

with $F_1$ and $F_2$ being non decreasing in $e$ and $k$, $F_1$ is convex in $e$ (CDFC) and $e(k, \tau)$ has properties I and U, the best response of the parent has the PS property.

Proof. $\square$

The optimality of a (PS) strategy stems from the CDFC and the properties of the best response function. The CDFC’s role is twofold: under the separability assumption - AD - of lemma 6, it preserves the DD property of the best response function and allows to place all the assumption on the production function. One then needs not to work with the value function which would require to place further structure on the flow utilities and the terminal value functions$^{31}$. A remark is also due with respect to the key role played by the DD of $e(k, \tau)$. In fact, the nature of parental intervention in my model-stricter limits diminish on average the amount of available leisure- is crucial to obtain DD$^{32}$. Such a formulation allows, for any given $k$, to compute the slope of $e(k, \tau)$ with respect to $\tau$: either zero, when the constraint is not binding, or $\tau$ when it is binding$^{33}$.

$^{31}$This strategy at first sight seems less attractive, given that the derivative of the value function with respect to the state variable cannot be obtained by applying the envelope theorem, due to the dependence on $e(k, \tau)$.

$^{32}$An alternative formulation according to which stricter limits diminish the marginal utility of leisure (proportional tax), rather than its amount, would not allow to draw immediately such a conclusion. In my dissertation - Cosconati (2009)- I worked out a two period version of such a model. I found sufficient condition to have a unique PS equilibrium.

$^{33}$Another natural question is what would be the implications of including a participation constraint, such
I now address issue 1), e.g. showing that under the structure imposed so far -CDFC and AD- whenever the parent adopts a (PS) strategy it is optimal for the child to use a best response function with properties I and U in lemma 5.

U is satisfied if the payoff function of the child is single peaked in the effort, which is the case if CDFC holds and \( u \) is strictly concave in \( l \) (or viceversa that the latter is strictly concave and the former is weakly concave or that they are both strictly concave). I follows from the separability implied by AD: the lack of interaction between \( e \) and \( k \) implies that the optimal effort function depends on the child’s knowledge only through the impact the latter has the discount factor and not because of changes in the productivity of study time\(^{34}\). The following lemma formalizes these observations.

**Lemma 7.** If the parent plays a (PS) strategy, CDFC and AD hold, \( u(l) \) is strictly concave in \( l \) (SC), then the best response function of the child is characterized by I and U.

**Proof.** Consider the objective function of the child in a generic period. For period \( T \), such a function is given by (11); while for a generic period \( t < T \) it is written in (12) (the time constraint is plugged in). Because when a PS parenting rule is played the (terminal) value function is increasing in the state variable, using the Conlon’s approximation (see (34)) the objective functions can be approximated as:

\[
\max_{e \in \mathcal{T}} \left[ u(1-e) + \delta(k) \left[ a_0 + \sum_{j=1}^{a_j} [1 - F_1(b_j|e) - F_2(b_j|k)] \right] \right]
\]

AD implies that the above function has increasing differences in \((e,k)\) because \( F_1(\cdot|e) \) is decreasing in \( k \), implying I; CDFC and SC imply that U also holds, because concavity is closed under summation.

The following theorem summarizes the assumptions consistent with (PS) equilibria.

**Theorem 3.** Under CDFC, SC, AD, the equilibria of the game have the (PS) property.

**Proof.** Under these assumptions lemma 6 and 7 hold and the result follows immediately.

**Discussion on the assumptions** A few comments are in order. SC seems indisputable: it simply requires diminishing marginal utility of leisure. As it is well known in the PA literature CDFC is less intuitive than it seems. To explain why, Conlon cites Rogerson:

“If output is determined by a stochastic production function with diminishing returns to scale in each state of nature, the implied distribution function over output will not, in general, exhibit the CDFC.”

as:

\[
\mathcal{V}(k|\rho) \geq \mathcal{W}
\]

To apply the monotonicity theorem I would have to show that \( \mathcal{V}(k''|\rho) \) is higher than \( \mathcal{V}(k'|\rho) \) in the strong set order, which is not necessarily the case. As an alternative way of ordering set I could have used the \( C \)-flexible set order introduce by Quah (2007), which often allows to compare sets which are not comparable in the strong set order. However in this case the requirement of the \( C \)-flexible set order are the same of the strong set order.

\(^{34}\)As an alternative/additional modeling assumption able to generate a positive correlation between human capital and effort one could assume that the marginal psychic cost of effort depends negatively on \( k \).
Conlon concludes: “Thus, neither Rogerson’s one-signal results nor Sinclair-Desgagne’s multisignal results build on the usual economic notion of diminishing marginal returns.” Although self-productivity of cognitive skills is allowed, AD rules out the interaction between the inputs. For instance a linear production function \( H = h + e + \epsilon \) would satisfy AD as long as the distribution of the shock \( e \) is such that CDFC holds. The issue is then whether one can find distributions which satisfy AD and CDFC simultaneously. Licalzi and Spaeter (2003) provide two classes of distributions which satisfy the monotone likelihood ratio property (MLRP)\(^{35}\) and the CDFC. The following lemma accommodates these classes so that \( F \) is also a function of \( k \) and generalized as follows:

**Lemma 8.** Let \( K \in S = [0, 1] \). The distribution function

\[
F(K|e,k) = K + \beta(k)[\gamma(e) + \eta(k)]
\]

satisfies AD and CDFC if

i) \( \beta(k) \) is a positive function on \( S \) such that \( \lim_{H \to 0} \beta(k) = \lim_{H \to 1} = 1 \) and \( |\beta'(k)| \leq 1 \)

ii) \( \gamma(e) \) is decreasing and convex on \( S \), \( \eta(k) \) is decreasing on \( S \), \( |\gamma(e) + \eta(k)| < 1 \).

**Proof.** The proof follows immediately from the proof of Proposition 1 of Licalzi and Spaeter (2003). \(\square\)

The above class is slightly more general than the one proposed in Proposition 1 by Licalzi and Spaeter (2003) as \( \beta \) is not restricted to be concave because FOSD is less restrictive than MLRP.

The following lemma adapts the second class of Licalzi and Spaeter (2003).

**Lemma 9.** Let \( K \in S = [0, 1] \). The distribution function

\[
F(K|e,k) = \delta(K)\left[\frac{\exp[\beta(K)\gamma(e)]}{2} + \frac{\exp[\beta(K)\eta(k)]}{2}\right]
\]

satisfies AD and CDFC if

i) \( \beta(K) \) is a nonconstant, negative, increasing such that \( \lim_{K \to 1} \beta(K) = 0 \)

ii) \( \gamma(e) \) and \( \eta(k) \) are strictly positive and increasing, \( \gamma(e) \) is strictly convex

iii) \( \delta(K) \) is positive, strictly increasing, \( \lim_{K \to 0} \delta(K) = 0 \), \( \lim_{K \to 1} \delta(K) = 1 \)

**Proof.** First, for \( F \) to be a CDF it needs to be the case that \( \lim_{K \to 1} F(K|e,k) = 1 \). Secondly \( \lim_{K \to 0} F(K|e,k) = 0 \), as it can be verified. Third, it needs to be strictly increasing:

\[
F(K|a,k) = \delta(k)\left[\frac{\gamma(e)}{2} + \frac{\eta(k)}{2}\right] + \delta'(k)\left[\cdot\right] > 0
\]

\(^{35}\)MLRP is more restrictive than FOSD.
Finally it can be verified that the CDFC holds (see the proof of Proposition 2 of Licalzi and Spaeter (2003)).

As pointed out by the authors the second class is richer than the first one. They provide the following example, adapted to incorporate $k$:

$$F(K|a,k) = K\left[e^{e(K)} + e^{k(K)}\right]$$

### 2.2 Uniqueness

The result in lemma 3 does not ensure that the set of SPNE equilibria of the game is a singleton. In particular there could be multiple cut-offs which are a best response to the child’s effort strategy. To see why, consider the last period and for simplicity let $R = \{\rho_l, \rho_h\}$, with $\rho_h > \rho_l$. Let the cost associated to $\rho_h$ equal to $\kappa$, while the cost of $\rho_l$ is normalized to zero. An equilibrium cut-offs $c$ of the last period is such that:

$$\Delta(k) |_c = \kappa$$

where:

$$\Delta(k) = \int \Pi(K) dF(k|e(k,\tau),k) [dG(\tau|\rho_h) - dG(\tau|\rho_l)]$$

Notice that $\Delta(k)$ might not be strictly decreasing in $k$ because $e(k,0)$ - the child’s “unconstrained” best response- might not be strictly increasing, e.g. a corner solution might be optimal given high/low values of $k$. In particular for a given configurations of parameters there exists $k^*$ and $k^{**}$ in $[k, \bar{k}]$, with $k^* \leq k^{**}$ and an increasing function $\tau(k)$, such that

$$e(k,0) = \begin{cases} 0 & \text{if } k \leq k^* \\ \tau(k) & \text{if } k^* < k < k^{**} \\ 1 & \text{if } k \geq k^{**} \end{cases}$$

which implies that there exist $\Delta > \bar{\Delta}$ and a strictly decreasing function $\hat{\Delta}(k)$ such that

$$\Delta(k) = \begin{cases} \Delta & \text{if } k \leq k^* \\ \hat{\Delta}(k) & \text{if } k^* < k < k^{**} \\ \bar{\Delta} & \text{if } k \geq k^{**} \end{cases}$$

Multiplicity arises if, at the true parameter values $c = \Delta$, in which case any value in $[0, k^*]$ is an optimal cut-off. This event has measure zero which is why the equilibrium of the $T$-period stage game is generically unique. More precisely, even if $c = \Delta$, if one perturbs the set of parameters exogenously the probability of multiple optimal cutoff is zero. Notice that $\Delta(k)$ is not a function of $c$ because the child only responds to current incentives, i.e. the realized $\tau$. However, when we go back to $T - 1$, the effort of the child is a function of the (unique) $T$-period equilibrium cutoff which is itself a function
Thus at $T - 1$ also the function $\Delta(k)$ depends on $c$ and shifts in response to an $\epsilon$ perturbation of $c$\textsuperscript{36}. Because an upward shift of exactly $\epsilon$ cannot be ruled out ex-ante, multiplicity can arise in the multi-stage game\textsuperscript{37}.

3 Data and Patterns

The National Longitudinal Study of Youth 1997 (NLSY97) provides information about youths living in the United States and born during the years 1984 through 1980. The original sample contains 8,984 individuals living in 6,819 households. The interviews were conducted annually since 1997. Blacks and Hispanics are over represented in the sample, which consists of roughly 50% of boys. In each survey round, the Youth Questionnaire is administered. The topics covered range from child’s schooling to employment activities. A rich set of information on the respondent’s family background is also available. In the first round the Parent questionnaire collected information from one of the child’s biological parent if present, or from the child’s guardian/parent figure\textsuperscript{38}. The NLSY97 contains variables that can be used to proxy for the state and choice variables of the theoretical model. I now describe them in more detail.

Parenting Rules The Autonomy/Parental control section of the NLSY97 Youth Questionnaire asks in the first three survey years to children born in 1983 and 1984 about the person or people who make decisions concerning three of their activities: i) how late they may stay out at night ii) the kind of TV shows or movies they may watch, and iii) who they are allowed to “hang out” with. Three alternative responses are available:

1. Parent or Parents Set Limits
2. My Parents and I Decide Jointly
3. Parents Let Me Decide

That is for each limit it is reported if the decision is made only by the respondent, by the the child jointly with her parents or only by the parents\textsuperscript{39}. ALTHOUGH THESE ARE NOT PERFECT MEASURES....

Table 2 shows the distribution of the grade attended by children for whom the information on parental autonomy is available. COMMENT Table 10 shows the correlation between the race of the child and parental monitoring

\textsuperscript{36} Gaudet and Walant (1991) provide conditions for uniqueness in multi-stage games of quantity competition. The issue tackled in their paper is analogous to the problem faced here.
\textsuperscript{37} Uniqueness can be achieved by ruling out corner solutions through Inada conditions. However as it appears from table 3 a non-negligible proportions of children choose zero as optimal time spent doing homework- my measure of effort. Because the model needs to generate a mass at 0, the estimation procedure needs to accommodate potential multiple equilibria.
\textsuperscript{38} For detailed information about the topics covered in the Parent’s questionnaire and the Youth Questionnaire see the NLSY97 guide
\textsuperscript{39} From now on I will refer to the limits of type i) as the curfew-limit, of type ii) as the TV limits and to limits in iii) as to the friends limits.
conditional on the age of the child. For all these ethnic groups we observe an intuitive age-pattern: older the child more likely it is he(she) has some role in the decisions on the three activities. This is not surprising as enforcing limits is typically more costly and/or less beneficial for older teenagers. Black parents are the ones who are more likely not to leave any role to the child, while white parents are the ones who leave the most often complete discretion on the rule to the child. Hispanic parents are the ones who negotiate the most often with their children on these type of decisions. It is interesting to see that the difference between races in the approach to parenting varies with the the type of rule parents decide on. Namely there is quite enough difference in the decisions on the curfew and in the decisions on the peers while there is less distance in the degree of monitoring on the TV shows. It is also indicative to inspect the dynamic behavior of parental decisions on limits, namely the transitions of the limits variables documented in table 11. On top of reflecting the age pattern mentioned before, the table documents the existence of the opposite type of switches: parents that did not set limits or decide them together at a given age of the child find it optimal to set limits when he(she) is one year older. For example for what concerns the curfew limits almost 35% of the youths that were allowed to decide and 41% of those that decided together with parents on how late to stay out at night, have no longer a role in deciding about this matter when they age one year. Moreover about 63% percent of children between the ages of 13 and 14 who have some or all the power in deciding on how late to stay out at night, report that the parent decides on the curfew when 14 or 15 years old. The ‘loss of power situation’ happens in about 18% and 19% of the times for the TV and for the friends limits when the child is 13 or 14 years old and 15%-16% of the times when she is 14 or 15 years old. How “objective” are the responses to the these questions? Because in the Parent questionnaire, responding parents of children born in 1983 and 1984 are asked the same questions on the limits, we can learn the extent to which the parenting style adopted by the parent is perceived differently by the child. Table 12 is similar to a transition matrix, whose rows sum to 100. The table reveals that there is usually consensus whenever the child declares that the parent(s) decide, with an “agreement” rate ranging from 91% for the curfew to 55% for the limits on the TV shows. There is more disagreement when the child reports that he/she is the person in charge of the decision. Moreover the agreement is not uniform across type of limits: the most homogeneous responses are the ones on the curfew, while there is the least consensus on who is the person deciding the TV shows. Not surprisingly the highest degree of ambiguity arises when the child declares that the decision is made jointly with the parent; for example in the case of curfew 80% of the times in which the child declares that the decision is taken together, the parent believes that he(she) is the one deciding.

Assuming that the responses to the questions on the person setting the limit do reflect an actual constraint the parent imposes on the child, it makes sense to wonder whether parents generally enforce the limits they set on the child’s recreational activities. The

---

40 Most likely parents are of the same race of the respondent. There is roughly 4% of households in which the race of the child differ from the ones of the parent.

41 One explanation for this fact might be that, because they are more likely to live in dangerous neighborhoods, Black and Hispanic parents have more incentives to impose those type of limits that prevent the interaction with those environments. Such an explanation is consistent with the fact the distribution of responses on the TV limits are similar across ethnic groups.
data give an opportunity to address this issue, as they contain information on the (self-declared) number of times the child breaks the limit in the month prior to the interview. Table 13 shows that the majority of children do not violate the curfew in the first round survey, but the propensity to do so increases with age. Moreover children are more likely to break limits on the "allowed" friends and TV shows. It is still the case that the majority of children violates the rules only once in the month prior to the interview date. Overall it seems that parents enforce the limits although not always successfully.\(^\text{42}\)

**Time Allocation** In the first survey round it is asked how much time it is spent doing homework (only to children enrolled in school) and watching TV in a typical week. Table 3 shows the proportion of children who dedicate a positive amount of time doing homework conditional on race and age. White children have a 10% probability more their black peers of dedicating some time to study. Hispanic children are half-way between their white and black peers. Table 4 shows, at each age, some statistics of the time spent doing homework of black, white and hispanic children. Black children are the ones who study the least: slightly less than 4 hours per week when they are 12 years old. Interestingly, at any at the age covered by the data, Hispanic children are the ones who study the most: about 5 hours per week on average, while the time dedicated by white children is about 4 hours and 45 minutes per week.\(^\text{43}\) This ranking does not hold as the child ages if one looks at the median time. The median of hispanic children is still the highest, but the ranking between white and black children changes as a function of age.\(^\text{44}\) Pooling together the responses of all the age groups, the order hispanic-white-black is respected both for the mean and the median time dedicated to homework. Such a variable is correlated with different measures of educational attainment. Table 6 shows the correlation, for each age group, of the extensive margin with the probability of dropping out from high school, of obtaining a high school diploma (without enrolling into college) and with the likelihood of and of enrolling into some college, after graduating from high school. At the age of twelve, at which the most commonly attended grade is the eighth grade, there is a very small difference between high school graduate and drop outs: about nine minutes a week. On the other hand the difference between simple high school graduates and those who will enroll into college is of 27 minutes. At the age of thirteen, when the child typically attends the 9th grade, these differences increase: children who will enroll into college study three quarters of an hour more than those who will only graduate. This difference among these two groups increases slightly with age: at the age of fourteen (fifteen) the former study on average an hour (an hour and

\(^{42}\) If the respondent reports that there are some kind of limits a follow up question asks about the number of times the respondent broke the rules in the last month and what kind of consequence (if any) there would be if parents find out that the rules have been violated, and who would implement the (eventual) punishment. From table 14, we can see that there are two most common consequences: 'Discuss it calmly' and 'Take away a privilege, ground you, or give you a chore'. For each type of limit these two alternatives are chosen in more than 80% of the cases. It thus seems that when rules are broken roughly half of children expect their parents to approach the situation using confrontation and discussion, while the other half believes that some sort of tangible punishment will be implemented.

\(^{43}\) Given that the test scores of hispanic are on average lower than the ones of white children, this statistics indicates that initial conditions play an important role in the skill formation process.

\(^{44}\) The pattern of the median value of black children resembles a U-shape while the one of the white children has more the one turning point.
6 minutes) more than the latter. The median difference is constant at one hour, with the exception of the age of fourteen, at which the difference peaks at 1.5 hours per week. Time spent studying increases monotonically with age for high school graduates (whether or not they enroll into college). The age pattern of high school drop outs is instead non-monotonic probably due to the endogeneity of the enrollment status, which determines whether the question is asked or not to the child. These differences in behavior across educational attainment groups are also reflected in the extensive margin: among children between the ages of twelve and fifteen interviewed in the first round, 5% of those who will eventually enroll into college do not dedicate any time to studying, while such a proportion for high school graduates (drop outs) is 9% (20%).

Moving to the time spent weekly amount of TV watching, as it appears from table 5, there are substantial differences among racial groups. For the age group 12-15 years old on average black children watch four and half hours per week than hispanic children and about eight hours more than white children. The differences in terms of the median values are about five and eight hours respectively\(^{45}\). There are also marked differences in terms of the age patterns of means and medians. White children’s averages go from 16.72 at the age of twelve to 16.25 at the age of fourteen without changing much during the intermediate ages; the median is flat at 14 hours per week. On the contrary black children experience an increase in TV watching of about an hour for each additional year, after the decrease at 15 years old which is common to hispanic and white children. The median of black children goes from 21 (at 12 years old) to 25 (at 15 years old) and decreases at 18 for children of fifteen years old; the median of hispanic also takes an inverted U-shape ranging from 16 hours (at the age of 12) to 17 hours at the age of fifteen\(^{46}\).

**Human Capital** The GPA achieved at the end of each academic year of high school and the scores of the Peabody Individual Achievement Test (PIAT) are the two measures that can be used to proxy for human capital of the child as they are available for multiple survey rounds\(^{47}\). This test measures academic achievement of children of ages five and over and it is among the most widely used brief assessments of academic achievement\(^{48}\). One of the PIAT subtests, the Mathematics Assessment, was given in round 1 respondents not yet enrolled in the 10th grade and in rounds 2 through 6 to respondents who were age 12 as of December 31, 1996, and who were in the 9th grade or lower in round 1. For grades 9-12, the transcript data survey report the academic GPA achieved at the end of each academic year. Each of these measures presents pros and cons. On the one...
hand parents are more likely to know the child's grades then the scores of tests such as the PIAT, which are not shown to them after being administered. On the other hand academic grades are not standardized. For instance if parents believe that grades are a poor measure of the child's skills because the grading system is too generous, they may not condition their actions on academic performances. The number of observations for which each information is available for the 1983 and 1984 cohorts is described in table 1. As it appears from the table there are children born in 1984 for whom both the PIAT scores and the GPA are available in all the survey years. The universe of children satisfying this requirement is small for survey year 1 and 2 when few children born in 1984 attend high school. In particular in the third survey year (1999) there are enough children with information on parental limits, GPA and PIAT.

\footnote{In particular the second column refers to the GPA achieved at the end of the academic year preceding the survey year. For example for the 1984 cohort the are only 3 children with information on the GPA of academic year 1996-1997}

INSERT: GPA and PIAT age pattern condit. on race and educational attainment

3.1 Correlations Among Endogenous Variables

**Time allocation choices, parenting rules and children's human capital** In order to show how the child's behavior is related to his human capital and to parenting styles, I regress the time spent studying (both intensive and extensive margin) on dummies for whether the parent sets alone the limits, controlling for the test scores and the race of the child. To take into account the forward looking behavior of the child I not only include the decision taken in the first survey round (PAR_CURFEW_1, PAR_FRIENDS_1, PAR_TV_1) but also future decisions. The first two columns of table 7 show little correlation of the intensive margin with the "contemporaneous" decisions on the limits, although there exists an association of the decision on the friends taken in the following survey round: if the parent is the only person who will likely decide on the limit, less time is dedicated to study. On the other hand not letting the child decide on the curfew is associated with an higher probability of spending a positive amount of time to study. The weak correlation between limits and the homework-time is not surprising. Setting strict limits on recreational activities has probably an effect on the former only through the lesser time dedicated to the latter.

The third column in table 7 shows the correlation between TV viewership (intensive margin) and the responses on the person setting the limits. There exists a negative correlation with the dummy which takes value one if the parent sets the curfew and if the child does not decide on the friends is spends time with, indicating a possible substitution between outdoor activities and TV watching. On the other hand, as one would expect, when no role is left to the child on the type of TV shows the child can watch, less time is spent watching TV which is consistent with the mechanism postulated by the model. Such a correlation holds also for the decision on the TV shows in subsequent rounds, indicating that the child anticipates the future moves of the parent. Interestingly the correlation between cognitive skills, as measured by test scores, and children's time allocation choices is consistent with a PS equilibrium.
higher the human capital of the child, more (less) time is dedicated to homework (TV watching).

The fourth column of the table 7 examines the relationship between time spent doing homework and TV viewership. The estimates reveal a negative and statistically significant association between TV viewership and the extensive margin of time spent doing homework (the dummy STUDY), while there appears to be no correlation between the intensive margin of TV watching and homework doing.

**Parenting rules and children's human capital** Table 8 shows the correlation between the propensity not to let the child decide on the rules in each of the three activities and my two measures of human capital, controlling for several family and child time-invariant characteristics which might shift parental preferences/constraints. The parameters attached to PIAT, the child’s PIAT test scores, is significant and negative in all the regressions, which is consistent with the comparative statics implied by PS equilibria: higher the skills of the child more likely it is that the parent sets permissive limits. When using the GPA as a proxy for the underlying state variable, the correlations differ as a function of the type of activity considered. It is still negative for the curfew but it is positive for the rules on TV shows, while there appears to be no association between GPA and the limits on friends.

**Human capital production function** Although the primary goal of this paper is not to assess how time allocation choices impact children's cognitive skills, it is useful to have a sense of their impact on the PIAT test scores and the GPA. To this end, given the data limitation I face, I adopt a value added specification (VA). I estimate an OLS regressions in which the dependent variable is either the GPA or the PIAT test scores reported in the second survey round and the regressors are the time spent doing homework and watching TV, with age-specific coefficients. As it appears from table 9, TV watching has a negative and small effect on test scores: an increase of an hour of TV watching per week decreases the test scores by 0.05% of a standard deviations. The effect of study time is bigger compared to the negative impact of TV watching, yet quite small in absolute terms: at the age of thirteen an increase of an hour spend doing a homework per week increases the PIAT test

---

50The dummies BLACK and HISPANIC take value 1 if the child belongs to those racial groups, MOTHDROP (FATHDROP) and MOTHHS (FATHHS) takes value one if the mother (father) is a high school dropout (graduate), BOY takes value one if the child is a boy, SINGLOMOTH is dummy for whether the child is reared by a single mother, NETWORTH is a proxy for family income which measures the household net worth values, NSIB gives the number of siblings living with the respondent child at the time of the interview. 51 Todd and Wolpin (2003) and Todd and Wolpin (2007) clarify the identifying assumptions underlying different specifications of the test scores production function. Because the questions on time spent doing homework and watching TV are asked only in the first survey round, a VA specification seems the best alternative given the data limitations I face. 52 The specification I adopt is also consistent with the skills production function estimated by Cunha and Heckman (2008). The dummies capturing the responses on the limits did not have an impact on test scores. As one would expect their direct impact is negligible. 53 Results on the impact of TV watching in children's outcomes are mixed. Fiorini and Keane (2011) adopt different specifications to assess the impact of time allocation on cognitive and non-cognitive skills. They find that time allocation matters and that time spending using media, such as TV, has positive effect on cognitive skills. Using a sample of children between the ages of 6-9 Haung and Lee (2010) finds that the effect of TV watching is non linear and depends on the child's age. They estimate a negative but small effect. Gentzkow and Shapiro (2008) use a exogenous variation of TV availability in the 60's to assess the impact of TV viewership on adolescent test scores. They find small and positive effects of TV exposure.
scores by 0.016% of a standard deviation, while there is no return for effort exerted at the age of fourteen\textsuperscript{54}. To put these numbers is prospective PIAT test scores recorded in the first survey round, which should capture the impact of past inputs, increase the test scores of the second survey round by 0.63 – 0.72 of a standard deviation, depending on the age of the child.

The results for GPA production function are somehow different: increasing study time of an hour per week at the age of fourteen increases of 1.2 of a standard deviation the GPA of the following academic year, while there is no effect of homework doing when the child is thirteen. Time spent watching TV has no effect on academic performances, while current GPA of the current year and 

3.2 Sample and Data Issues

As it appears from table 1 the information contained in the NLSY97 present several limitations which render the map of the model to the data problematic and pose potential obstacles to the identification of the parameters. They can be summarized as follows:

L\textsubscript{1} the PIAT test scores are available only for the born in 1984, the youngest cohort interviewed in the NLSY97

L\textsubscript{2} The GPA is available only for children enrolled in high school

L\textsubscript{3} The information on the person setting the limit are available only for children born in 1983 and 1984

L\textsubscript{4} The time allocation choices are elicited only in the first survey round

Because of L\textsubscript{3} and L\textsubscript{1} the data on the cohort of 1984 contains the most information on the relationship of interest. In fact, if one had to proxy human capital using only the PIAT test scores, the 1984 would suffice. On the one hand using the data on the 1983 cohort allows to expand the age pattern of the child rearing practices, which is a target of the model. On the other, because I also want to use the GPA as a measure of the underlying state variable of the model, such a choice is problematic for two reasons. First, to identify the parameters of the production function using the GPA as measures of human capital one needs enough children enrolled in high school in the first two survey years, due to L\textsubscript{2}. As it appears from table 2 few children born in those two years attend in the first two survey years high school. Thus the information provided by 9\textsuperscript{th},10\textsuperscript{th} and 11\textsuperscript{th} graders born in 1980 and 1981 interviewed in the first two rounds, are extremely useful. The cohort of 1982 does not provide extra information, yet it helps to reconstruct the age pattern of the time allocation choices, which the model needs to match. Given these considerations I use the information for all the cohorts and model parent-child interaction starting from 6\textsuperscript{th} up to 12\textsuperscript{th} grade. I start from the 6\textsuperscript{th} grade because it is the earliest grade for which I have enough children interview in the first round born in 1983

\textsuperscript{54} The evidence on the impact of study time on achievement is scarce. Stinebrickner and Stinebrickner (2008) found a sizable effect on academic performances among college students.
and 1984 (76 and 605, respectively). Stopping at the 12th grade is also natural as many teenagers will leave home after graduating from high school. I use data for the cohorts I indicated contained in the first five rounds: the fifth round is the latest one for which some information on the parenting rules is available. Thus the terminal value of the theoretical model can naturally interpreted as the payoff from obtaining a high school diploma with a certain level of “knowledge”. I divide the year into two semesters: fall and spring, which implies that the theoretical model is estimated using 14 “periods”.

Sample

4 Estimation

The goal of the structural estimation exercise is to recover the parameters of the payoff and production functions using the information on the person(s) deciding the curfew, the rules on TV viewership, curfew and friends, as well as from the self-declared time allocated to homework/TV watching and from the PIAT test scores.

The estimation of games with multiple equilibria is a complicated matter, due to the well known indeterminacy (or incompleteness) problem, which makes the map from the structural parameters to the likelihood of observed events a correspondence rather than a function. There are two prevalent approaches. In the context of discrete games, both static and dynamic, the approach developed by Hotz and Miller (1993) can be fruitfully applied to games, as done by Bajari, Benkard, and Levin (2007). The convenience of this approach relies on the fact that there is no need to solve for the equilibria of the game at any given parameter trial. Therefore the typical inner loop of the standard full-solution method, is not performed. In the context of static games, exploiting the particular structure of his game, Moro (2003) develops a two step estimator which also does not require to compute all the equilibria of the game. The starting point of these papers is that computing all the equilibria can be undoable for large games. Both approaches hinge on the assumption that a unique equilibrium is played across markets/families. Bajari, Hong, and Ryan (2010) develop an algorithm which relaxes this assumption and allows the data generating process to be a mixture of likelihood functions, each associated to a specific equilibrium.

To tackle the multiplicity problem arising in the class of (PS) equilibria, I adopt a two-step

---

55. In the first round there are 1231, 1744 and 14 children who are 12, 13 and 14 at the date of the interview, respectively.

56. As it appears from table 1 it is known who is the person setting limits on the friends.

57. Interviews are conducted in the summer only in the first round. Fall goes from the 1st of September to 31st of December; spring starts from the 1st of January to the 31st of June of year \( t + 1 \).

58. See Tamer (2003) for an explanation of the issue of incompleteness and Bajari, Hong, and Nekipelov (2010) for a survey of the methods in empirical IO.

59. See Moro, Bisin, and Topa (2011) for an application to social interaction.

60. They estimate a game of entry in auctions by computing all the pure and mixed strategy equilibria. This is achieved by using the software Gambit, developed by which solves systems of polynomial equations and inequalities. The authors provide conditions for the semi-parametric identification of the payoff function and of the selection's mechanism.
algorithm, similar to the one proposed by Moro (2003). Recall that the source of multiplicity in the class of (PS) equilibria is due to the best response of the parent, which is a correspondence rather than a function. The converse however is not true: given a sequence of cutoffs set by the parent, a unique time allocation strategy is optimal, given (SC). Therefore, by treating the matrix of cutoffs as a set of estimable parameters, one can obtain information on the technologies and on the child's preferences using the data on the child's time allocation choices/performances and on the person setting the limits on the three activities. Once these parameters have been estimated, the estimates of $c$ can be used to identify the parameters of the parent’s preferences.

To better illustrate the method it is useful to partition the parameter space as follows:

$$\Theta = (\Theta^p, \Theta^c, \Theta^f, \Theta^m)$$

where $\Theta^p$ and $\Theta^c$ are the set of parameters entering into the parent’s and child’s payoff function in (1) and (4); $\Theta^f$ are the parameters entering into the technologies $F(K|e,k)$ and $G(\tau|\rho)$ and $\Theta^m$ is a set measurement error parameters entering into the measurement error equations. In the first step I estimate $(\Theta^c, \Theta^f, \Theta^m, c)$, while in the second I recover $\Theta^p$.

4.1 First Step

The goal of the first step of the estimation procedure is to recover $(\Theta^c, \Theta^f, \Theta^m, c)$. This is accomplished by solving the child’s problem, for each initial condition $k_0$. Measurement error is then used to reconcile the model’s predictions with the observed events using the estimation method proposed by Keane and Wolpin (2001). In particular given a matrix $c$, a probability distribution on the observed events needs to be placed. Let $l_{t,i}$ denote the time allocation choices of a child $i$ in period $t$, with $t = 1, \ldots, 14$; let $\rho_{t,i}$ be a 3-dimensional vector which describes the observed parenting styles decisions. As a convention the first, second and third element contain the information on the person(s) deciding the curfew, the type of TV shows the child $i$ can watch and friends he/she is allowed to spend time with, respectively. To conserve on parameters I pool together the responses "jointly"/"child alone". Therefore $\rho_{t,i}^1(1) = 1$ signifies the the parent of child $i$ at time $t$ sets the curfew alone, while $\rho_{t,i}^1(3) = 0$ means the the limits on the friends of child $i$ are not decided by the parent alone. Finally let $(PIAT_{t,i}, GPA_{t,i})$ denote the PIAT math test and the GPA achieved by the child at the end of period $t$. Ignoring for now that there are two measures for the human capital of the child, at any equilibrium there exists a determinist map between the underlying state variable and the players’ actions: given a matrix of cutoffs $c$ each player at any given state $k$ chooses an optimal action $(l_t, \rho_t)$ with probability one. As a result the contribution to the likelihood function of a generic pair child-parent is generically zero, in particular for the endogenous variables which are continuous. Given a matrix $c$ which is consistent with a (PS) equilibrium, the method proposed by Keane and Wolpin (2001) can applied to my first stage as follows:

1) Draw an initial condition $k_0$
2) Solve for the optimal sequence \( \langle l_t \rangle_{t=1}^T \)

3) Using the randomness of the human capital production function simulate \( N \) outcome history \( \hat{O}^n = (l^n_t, \rho^n_t, k^n_t)_{t=1}^T \)

4) Repeat these steps \( N \) times for \( N \) fictitious parent-child pairs

Let \( O^n_i = (PIAT_{i,t}^n, GPA_{i,t}^n, e^n_{i,t}, \rho^n_{i,t}) \) with \( t = 1, \ldots, T \) denote an outcome history observed for child \( i \), which may contain missing elements. Using the simulated histories one can construct an unbiased estimator for the probability that the observed history \( O^n_i \) is generated by a simulated history \( \hat{O}^n \)

\[
\hat{P}_N(O^n_i) = \frac{1}{N} \sum_{i=1}^N P(O^n_i | \hat{O}^n)
\]

The convenience of this method is that it allows to construct the likelihood without having to compute the probability of an action conditional on the state variable \( k_{t-1} \). Given that my measures of \( k_{t-1} \) are not always available this method does not require to integrate out \( k_{t-1} \) which can be computationally impractical; the densities of the measurement/classification errors are used to reconcile \( O^n_i \) and \( \hat{O}^n \).

To deal with unobserved heterogeneity let \( \pi_m | k_0 \) the probability that a given parent-child pair is of type \( m \in \{1, \ldots, M\} \), given the initial condition \( k_0 \). In this case:

\[
\hat{P}_N(O^n_i) = \frac{1}{N} \sum_{m=1}^M \sum_{i=1}^{N/M} P(O^n_i | \hat{O}^n) \frac{\pi_m | k_0}{N/M}
\]

The following section describes the details of the measurement error equations.

4.2 Second Stage

Given an initial condition \( k_0 \), an associated matrix of estimated cutoffs \( \hat{c} = (\hat{c}_1, \ldots, \hat{c}_T) \) estimated in the first step can be used as a basis to form a system of moment inequalities. In particular given a vector \( \hat{c}_t \) suppose that \( \hat{c}^{i-2}_t < \hat{c}^{i-1}_t < \hat{c}^i_t \). This means that at \( \hat{c}^i_t \) the parent is indifferent between \( \rho_{i-1} \) and \( \rho_i \) (\( \rho_{i-2} \)). If instead \( \hat{c}^{i-1}_t = \hat{c}^i_t < \hat{c}^{i+1}_t \) the parent never finds optimal to use \( \rho_{i-1} \) because the differential benefit of \( \rho_i \) versus \( \rho_{i-1} \) is always higher than the relative cost. Let \( S1 \) denote the set of interior cutoffs -the ones belonging to the first case- and \( S2 \) the one who fall in the case of the second example. For any \( i = 1, \ldots, n-1 \), at any given \( \theta^p \in \Theta^p \), the following system of equalities/inequalities holds:

\[
\Delta^i_t(\hat{c}^i_t) = \kappa_{i-1} - \kappa_i \quad \text{if} \quad \hat{c}^i_t \in S1
\]

\[
\Delta^i_t(\hat{c}^i_t) < \kappa_{i-1} - \kappa_i \quad \text{if} \quad \hat{c}^i_t \in S2
\]

\[62\] As the authors explain “in this construction, missing elements of \( O^n_i \) are counted as consistent with any entry in the corresponding element of \( \hat{O}^n \).”

\[62\] Recall that by convention \( c^n_0 = \bar{k} \) and \( c^n_T = \bar{k} \)
where the expected benefit of choosing a stricter parenting style $\rho_i$ versus a more lenient one $\rho_{i-1}$ is defined on the support of the state variable $k_{t-1}$ as:

$$
\Delta_i^t(k_{t-1}) = \int_{k_{t-1}} \frac{1}{k} W_t(k_{t-1}, \epsilon_t(k_{t-1}, \tau), k) \left[ dG(\tau | \rho_{i-1}) - dG(\tau | \rho_i) \right] dG(\tau | \rho_{i-1}) - dG(\tau | \rho_i)
$$

for $i = 1, \ldots, n$ \hfill (20)

$W$ denotes the value function of the parent in the following period. At any given $\theta^p$ the quantity in (20) is known because in the first step $e_t(k_{t-1}, \tau)$ has been computed. To point identify $\theta^p$ one could use only the set of moments provided by the system in (20) provided that the number of equations exceeds/equals the cardinal of $\Theta^p$. If this is not the case using also the inequalities in (20) only gives set identification of the true parameters $\theta^p$.

### 4.2.1 Measurement Error Equations

**Parenting Rules** I assume that the responses on the person(s) setting the limit are classified with error, i.e. there is a positive probability of misreporting. As in Keane and Wolpin (2001) and Keane and Sauer (2009), the classification error is unbiased, which means that the probability of reporting a given choice for any activity $j$, with $j = 1, 2, 3$ equals the true probability that such an option is selected:

$$
Pr(\rho^o(j) = 1 | \rho(j) = 1) = E_j + (1 - E_j) Pr(\rho(j) = 1)
$$

for $j = 1, 2, 3$. $E_j \in [0, 1]$ is the “base” classification error parameter which I estimate and assumed to differ across the responses on the three activities. In the above formulation as $Pr(\rho(j) = 1)$ goes to one, the probability of observing option $j$ goes to one so that unbiasedness is preserved; on the contrary as $Pr(\rho(j) = 1)$ goes to zero $Pr(\rho^o(j) = 1 | \rho(j) = 1)$ goes to $E_j$.

**Measures of human capital** As in Cunha and Heckman (2008) I specify a linear factor model to relate $k_t$ to $GPA_t$ and $PIAT_t$:

$$
GPA_t = n_0 + n_1 k_t + \epsilon^G_t
$$

$$
PIAT_t = m_0 + m_1 k_t + \epsilon^P_t
$$

where the $\epsilon^G_t$ and $\epsilon^P_t$ are measurement errors, which are uncorrelated across time and individuals, uncorrelated among themselves and normally distributed with mean 0 and variances $\sigma^2_G$ and $\sigma^2_P$. As postulated by Cunha and Heckman, I also assume that there exists no “fixed effect”, in other words all the unobserved heterogeneity in “mental ability” is captured, up to measurement error, by the PIAT test scores and the GPA. This is due to limitation L4: permanent unobserved heterogeneity shifting $F(K | e, k)$ would not be identified because only one cross-section of data on $e$ is available. To conserve on parameters I assume, differently than Cunha and Heckman (2008), that the factor loadings
(m and n parameters) are not time-varying. For identification purposes I set $n_1 = 1$.63

**Time allocation choices** I adopt an hurdle model to deal with the cases in which the child declares not to study/watch TV, e.g. the zeros. Let $e^o$ denote an observed value of the child’s effort. The hurdle model requires to specify the probability of observing a zero, conditional on an effort level predicted by the model:

$$Pr(e^o = 0 | e) = \rho_e = \frac{1}{1 + \exp(\gamma_e e)}$$

as well as a conditional probability of observing a strictly positive measure of effort. I adopt a truncated normal distribution $g(\cdot | e)$ with mean $\lambda_{0,e} + \lambda_{1,e} e$ and variance $\sigma_e^2$. The measurement error model is written as follows:

$$g(e^o) = \begin{cases} 
\rho_a & \text{if } e^o = 0 \\
\frac{1-\rho_e}{1-G(0|e)} g(e^o | e) & \text{if } e^o > 0
\end{cases}$$

where $G$ denotes the CDF of the density $g(\cdot | a)$.

An analogous model is postulated for the observed time spent watching TV. Summing up there are 17 parameters coming from the measurement error equations:

$$\Theta^m = \langle E, m_0, n_0, m_1, \sigma^2_e, \sigma^2_P, y_e, \lambda_{0,e}, \lambda_{1,e}, \sigma^2_e, y_{TV}, \lambda_{0,TV}, \lambda_{1,TV}, \sigma^2_{TV} \rangle$$

### 4.3 Parametrizations

In order to structurally estimate the model it is necessary to parametrize its primitives. The following functional forms were adopted.

**Payoff Functions**

The child’s preferences are given by

$$v(\alpha_t) = \begin{cases} 
TV_t^{\eta_1}(1 - TV_t - e_t)^{\eta_2} & \text{if } t < T + 1 \\
\Xi(k_T) & \text{if } t = T + 1 \text{ (the game is over)}
\end{cases}$$

with $0 < \eta_1 + \eta_2 \leq 1, \eta_j$ for $j = 1, 2$, which ensures that the child’s flow utility function is globally concave. I parametrize the terminal value function through the flexible power risk aversion utility function (PRA)64

$$\Xi(k_T) = \frac{1}{\xi} \left\{ 1 - \exp \left[ - \xi \left( \frac{k_T^{1-\sigma} - 1}{1 - \sigma} \right) \right] \right\}$$

---

64This class of utility functions has been introduced by Xie (2000). An even more flexible class, the three parameter utility function (TFP), has been proposed by Conniff (2007).
with $\sigma$ and $\xi$ greater or equal to zero. Such a flexible parametrization nests a constant or decreasing absolute risk aversion (CARA and DARA) as well as increasing, constant or decreasing relative risk aversion (IRRA, CRRA and DRRA).

As pointed by Akabyashi (2006) a sensitive parametrization for the **discount rate function** is given by

$$\delta(k) = \frac{\theta_0}{\theta_0 + \exp(-\theta_1 k)}$$

with $\theta_0, \theta_1 > 0$. This parametrization/restrictions ensures that the discount factor is increasing in $k$ and has an upper bound at one.

The **parent’s preferences** are given by

$$w(\rho_t, k_t) = \begin{cases} -\sum_{j=1}^{3} \kappa_j \mathbb{1}[\rho_t(j) = 1] & \text{if } t < T + 1 \\ \Pi(k_T) & \text{if } t = T + 1 \text{ (the game is over)} \end{cases}$$

(27)

Given my convention $\kappa_1$ is the cost associated to deciding the curfew while $\kappa_2(\kappa_3)$ denote the monitoring cost subsequent to not making the child decide the rules on the TV shows (friends). To save on parameters the cost function does not include interaction effects so that the cost of setting a given type of limit is independent on the decisions on the other activities.

I also adopt a PRA utility function to parametrize the parent’s terminal value:

$$\Pi(k_T) = \frac{1}{\pi} \left\{ 1 - \exp \left[ -\pi \left( \frac{k_T^{1-\sigma} - 1}{1 - \sigma} \right) \right] \right\}$$

(28)

To save on parameters I assume that that the miscongruence in preferences with the child arises solely because $\pi$ is allowed to be different than $\xi$.

**Production Functions**

The **discipline production function** is parametrized as a beta distribution$^{65}$

$$G(\tau|\lambda) = \text{Beta}[d(\rho_t), \lambda_1] \text{ with } \tau \in [0,1], \lambda_1 > 0$$

(29)

I parametrize $d(\rho_t)$ as follows

$$d(\rho_t) = \frac{\lambda_0 \rho_t}{1 + \lambda_0 \rho_t}$$

Such a functional form coupled with the restriction that $\lambda_0 \in R^3_+$ guarantees that, stricter the limits, higher the probability the a high $\tau$ will realize$^{66}$. The components of the vector $\lambda_0$ measure the relative effectiveness of not letting the child decide on each

---

$^{65}$Beta distributions which are typically used to model random variables with bounded support. See Johnson and Kotz (1970).

$^{66}$If $\rho = 0$ the distribution will be degenerate and the child will enjoy complete autonomy, i.e. $\tau = 0$ with probability one. I do not impose any ranking among the elements of $\lambda_0$ letting the data tell which type of restriction has the greatest effect on study time.
of the leisure activities elicited in the data.

The knowledge production function is also parametrized as the sum of power density distributions:

\[ F(K|e,k) = \frac{K^a(e)}{2} + \frac{K^b(k)}{2} \quad \text{with} \quad K \in [0, 1] \]  

(30)

In order for \( F \) to be a distribution it is required that \( a(e) \) and \( b(K) \) are strictly positive, which also implies that effort and current knowledge increase future knowledge in a first order stochastic dominance; for the CDFC to hold I need \( a''(e) < 0 \). Moreover for the distribution to be well defined it is required that \( a(e) \) and \( b(K) \) are in \((0, 1)\). The following parametrizations satisfies these requirements:

\[ a(e) = \frac{\theta_0 + e^{\theta_1}}{1 + \theta_0 + e^{\theta_1}} \]  

(31)

\[ b(k) = \frac{\exp(\theta_2 k)}{1 + \exp(\theta_2 k)} \]  

(32)

under the restriction \( \theta_0 > 0, \theta_1 \in (0, 1) \) and \( \theta_2 > 0 \).

**Initial Conditions and Unobserved Heterogeneity** The initial condition of the game is given by the stock of human capital, the child is endowed with at time 0. Given that support of \( k_t \) is \([0, 1]\) it is natural to assume that \( k_0 \sim \text{Beta}[a_0, b_0] \) with support also in \([0, 1]\). I also allow parents to differ in their monitoring costs \( \kappa_i \). I postulate the existence of \( M \) finite number of parental types which are characterized by a cost-shifter parameter \( \kappa_i^j \), with \( i = 1, \ldots, M \). As a result the cost associated to parenting style \( j \) suffered by a parent of type \( i \) is \( \kappa_i^j \kappa_i \), where I normalize \( \kappa_i^1 \) to one and order parental types accordingly: \( \kappa_i^j > \kappa_i^{j-1} \), for \( j = 2, \ldots, M \). Notice how some type of heterogeneity is needed to rationalize, without measurement error, observations in which parents of children endowed with same human capital, choose different parenting styles. After some experimentation letting \( M = 2 \) turned out to be sufficient to fir the data. To take into account the initial condition problem - I only observe the players’ behavior only starting from the 6th grade - I allow the probability the parent is of type \( I \) (the lenient type), which I call \( \pi \), to be correlated with the initial human capital \( k_0 \) according to a multinomial logit distribution:

\[ \pi = \frac{\exp(k_0 c_0)}{1 + \exp(k_0 c_0)} \]

with cutoff \( c_0 \). Summing up the parameters coming from the initial condition distribution are \( (a_0, b_0, c_0, c_0) \).

**5 Results from the First Stage**

At the first stage I estimate 17 structural parameters and 17 parameters which enter into the measurement error equations. Moreover, for each of the two parental types, I need to estimate the matrix \( c \) which consists of 98 parameters, which gives an additional
196 parameters for a total of 132 parameters which I estimate in the first step.
ADD TABLES AND ESTIMATES
A Omitted Proofs

Proof of lemma 5

Let \( k_L < k_H \) denote two levels of human capital which are realized in the last period. Assumptions I and U imply that there exists \( e^L \) and \( e^H \) with \( e_L \leq e_H \) and \( \tau_H \geq \tau_L \) such that

\[
e(k_L, \tau) = \begin{cases} 
e_L & \text{if } 0 \leq \tau \leq \tau_L \\
\tau & \text{if } \tau > \tau_L
\end{cases}
\]

and

\[
e(k_H, \tau) = \begin{cases} 
e_H & \text{if } 0 \leq \tau \leq \tau_H \\
\tau & \text{if } \tau > \tau_H
\end{cases}
\]

Let \( \Delta_1 = e(\tau'', k') - e(\tau', k') \) and \( \Delta_2 = e(\tau'', k'') - e(\tau', k''). \) I argue that for any \( \tau'' > \tau', \Delta_1 \geq \Delta_2 \), that is the function \( \Delta(k) = e(\tau'', k) - e(\tau', k) \) is decreasing in \( k \). Six cases need to be considered:

1. \( \tau' < \tau'' < \tau_L < \tau_H \). This implies that \( \Delta_1 = e^L - e^L = 0 = \Delta_2 = e^H - e^H \)
2. \( \tau' < \tau_L < \tau'' < \tau_H \). This implies that \( \Delta_1 = \tau'' - \tau' > \Delta_2 = e^H - e^H > 0 \)
3. \( \tau_L < \tau' < \tau'' < \tau_H \). This implies that \( \Delta_1 = \tau'' - \tau' > \Delta_2 = e^H - e^H > 0 \)
4. \( \tau_L < \tau' < \tau_H < \tau'' \). This implies that \( \Delta_1 = \tau'' - \tau' > \Delta_2 = \tau'' - e^H \)
5. \( \tau' < \tau_L < \tau_H < \tau'' \). This implies that \( \Delta_1 = \tau'' - \tau' > \Delta_2 = \tau'' - e^H \)
6. \( \tau_L < \tau_H < \tau' < \tau'' \). This implies that \( \Delta_1 = \tau'' - \tau' = \Delta_2 = \tau'' - \tau' \)

Proof of lemma 6

The result follows if one can show that the objective functions of the parent displays in each period DD in \( (k, \rho) \).

The starting observation is that any increasing and concave (convex) transformation of a function which displays DD and is increasing in its arguments, has itself DD. Ignoring for the time being that the parent is not choosing directly \( \tau \), I wish to show that, under CDFC, the future component of the utility has DD in \( (k, \tau) \). Starting from the last period and imposing AD, consider the future component of the last period’s payoff:

\[
\int_\tau^\pi \Pi(K)dF_1(k|e(k, \tau))dK + \int_\tau^\pi \Pi(k)dF_2(k|k)dK \tag{33}
\]

Because supermodularity is preserved under summation I can focus on the first integral in (33), call it \( \phi(k, \tau) \). Following Conlon (2009) it can be approximated by a sum of nondecreasing step functions:

\[
\phi(k, \tau) \approx a_0 + \sum_{j=1}^\pi a_j \int_\tau^\pi x(k, b_j)dF_1(K|e(k, \tau))dK = a_0 + \sum_{j=1}^\pi a_j[1 - F_1(b_j|e(k, \tau))] \tag{34}
\]

\[67^\text{For instance, consider the function } f(g(x, y)), \text{ with } f \text{ increasing and } g \text{ increasing in } x \text{ and } y. \text{ Let } f \text{ and } g \text{ be differentiable. It is easy to check that } f_{\gamma x} \leq 0 \text{ whenever } f'' \leq 0 \text{ and } g_{xy} \leq 0.\]
where the $a_j$'s are positive for $j \geq 1$ and $x(k, b)$ is the step function which is 0 for $K < b$ and 1 for $K \geq b$.

By the CDFC this function is concave in $e(k, \tau)$. Moreover because it is increasing in $e(k, \tau)$, it follows that $\phi(k, \tau)$ has DD in $(k, \tau)$, due to the initial observation of my proof and to the fact that $e(k, \tau)$ has DD in $(k, \tau)$ due to $l$ and $U$ (lemma 5). To conclude that a PS policy is optimal at $T$, I need to show that the function

$$\xi(\rho, k) = \int_0^1 \phi(k, \tau)dG(\tau|\rho)d\tau$$

has DD in $(\rho, k)$. Consider now the function

$$\Delta \hat{\phi}(\tau) = -[\phi(k'', \tau) - \phi(k', \tau)]$$

which displays increasing differences in $(k, \tau)$. The function

$$\Delta \hat{\xi}(\rho) = \int_0^1 \Delta \hat{\phi}(\tau)dG(\tau|\rho)d\tau$$

is increasing in $\rho$ because $G(\tau|\rho'')$ FOSD $G(\tau|\rho')$ for any $\rho'' > \rho'$.

Therefore the function:

$$\hat{\xi}(\rho, k) = \int_0^1 \hat{\phi}(\tau, k)dG(\tau|\rho)$$

has increasing differences in $(k, \rho)$ (see lemma 9 of Van Zandt and Vives (2007)). This is equivalent to state that $\xi(k, \rho)$ has DD in $(k, \rho)$. The flow utility has (trivially) DD in $(k, \rho)$ because DD is preserved under summation. Therefore the overall objective function has DD in $(k, \rho)$ or, equivalently, ID in $(k, \hat{\rho})$, where $\hat{\rho} = -\rho$. The constraint set is not altered by changes in $k$, therefore the monotonicity theorem can be applied to conclude that a PS strategy is optimal at $T$. Notice now that also $V_p\rho(k)$ is increasing in $k$ due to $l$, $IE$ and $IH$.

Proceeding backwards consider the last-but-one period’s payoff of the parent, as written in (16). Because $V_p\rho(k)$ is increasing in $k$, the Conlon’s approximation made for the $T$-period problem can be replicated by replacing $\Pi(k)$ in (33) with $V_p(k)$. The steps which lead to conclude that the $T - 1$ period payoff has DD in $(k, \rho)$ are identical to $T$-period case. Also at $T - 1$ the constraint is expanding because, once the parent plays a PS strategy at $T$, at $T - 1$ $V(K''|\rho) \leq V(K''|\rho)$ due to $IH$ and ID. The monotonicity theorem implies again that the optimal parenting policy at $T - 1$ has the PS property.

Iterating this argument backwards completes the proof. ☐
Figure 1: Objective functions

\[ V(e,k_L) \]
\[ V(e,k_H) \]
Figure 2: Properties of Best Responses
References


Cunha, F. and A. Badev (2011). Do parents know the technology of skill formation.


Moro, A., A. Bisin, and G. Topa (2011, January). The empirical content of models with multiple equilibria in economies with social interaction.


### Table 1: Data Availability

<table>
<thead>
<tr>
<th>Survey Round</th>
<th>PIAT</th>
<th>GPA</th>
<th>CAT-ASVAB</th>
<th>Time Study</th>
<th>Curfew</th>
<th>TV Limits</th>
<th>Friends Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>1984 Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1665</td>
<td>3</td>
<td>1413</td>
<td>1660</td>
<td>1738</td>
<td>1739</td>
<td>1737</td>
</tr>
<tr>
<td>2</td>
<td>1602</td>
<td>64</td>
<td>0</td>
<td>0</td>
<td>1611</td>
<td>1611</td>
<td>1609</td>
</tr>
<tr>
<td>3</td>
<td>1614</td>
<td>681</td>
<td>0</td>
<td>0</td>
<td>1568</td>
<td>1572</td>
<td>1572</td>
</tr>
<tr>
<td>4</td>
<td>1516</td>
<td>1049</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1462</td>
<td>1052</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1417</td>
</tr>
<tr>
<td>6</td>
<td>406</td>
<td>984</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Survey Round</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1983 Cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1600</td>
<td>37</td>
<td>1685</td>
<td>1685</td>
<td>1768</td>
<td>1771</td>
<td>1737</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>644</td>
<td>0</td>
<td>0</td>
<td>1599</td>
<td>1600</td>
<td>1597</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1017</td>
<td>0</td>
<td>0</td>
<td>1566</td>
<td>1566</td>
<td>1566</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1022</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>947</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1265</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>383</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Number of observations for which the information is available
Table 2: Distribution of Grade Attended in Survey Round 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2&lt;sup&gt;nd&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3&lt;sup&gt;rd&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5&lt;sup&gt;th&lt;/sup&gt;</td>
<td>71</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6&lt;sup&gt;th&lt;/sup&gt;</td>
<td>605</td>
<td>76</td>
<td>11</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>7&lt;sup&gt;th&lt;/sup&gt;</td>
<td>1023</td>
<td>659</td>
<td>122</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>8&lt;sup&gt;th&lt;/sup&gt;</td>
<td>64</td>
<td>998</td>
<td>677</td>
<td>116</td>
<td>21</td>
</tr>
<tr>
<td>9&lt;sup&gt;th&lt;/sup&gt;</td>
<td>1</td>
<td>63</td>
<td>955</td>
<td>725</td>
<td>181</td>
</tr>
<tr>
<td>10&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0</td>
<td>1</td>
<td>74</td>
<td>944</td>
<td>615</td>
</tr>
<tr>
<td>11&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>61</td>
<td>786</td>
</tr>
<tr>
<td>12&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>65</td>
</tr>
<tr>
<td>13&lt;sup&gt;th&lt;/sup&gt;</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; Year College</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>ungraded</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>N</td>
<td>1768</td>
<td>1807</td>
<td>1841</td>
<td>1871</td>
<td>1690</td>
</tr>
</tbody>
</table>

Table 3: Proportion of Children Spending Some Time Doing Homework in a Typical Week

<table>
<thead>
<tr>
<th>Age</th>
<th>Black Children</th>
<th>Hispanic Children</th>
<th>White Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>88.51 (296)</td>
<td>85.42 (240)</td>
<td>92.81 (612)</td>
</tr>
<tr>
<td>13</td>
<td>86.47 (414)</td>
<td>85.17 (344)</td>
<td>91.5 (871)</td>
</tr>
<tr>
<td>14</td>
<td>82.25 (462)</td>
<td>83.14 (344)</td>
<td>90.22 (941)</td>
</tr>
<tr>
<td>15</td>
<td>82.71 (133)</td>
<td>82.68 (127)</td>
<td>83.71 (264)</td>
</tr>
</tbody>
</table>

Number of observations in parenthesis

Table 4: Time Spent Doing Homework in a Typical Week

<table>
<thead>
<tr>
<th>Age</th>
<th>Black Children</th>
<th>Hispanic Children</th>
<th>White Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-16</td>
<td>4.34</td>
<td>3.75</td>
<td>3.71</td>
</tr>
<tr>
<td>12</td>
<td>4.53</td>
<td>4</td>
<td>3.23</td>
</tr>
<tr>
<td>13</td>
<td>5.10</td>
<td>4.5</td>
<td>3.53</td>
</tr>
<tr>
<td>14</td>
<td>5.29</td>
<td>4.5</td>
<td>2.29</td>
</tr>
<tr>
<td>15</td>
<td>5.79</td>
<td>5</td>
<td>3.65</td>
</tr>
</tbody>
</table>

Number of observations in parenthesis
Means and Medians are calculated on non-zero observations
Units: hours per week
Table 5: Time Spent Watching TV in a Typical Week

<table>
<thead>
<tr>
<th>Age</th>
<th>Black Children</th>
<th>Hispanic Children</th>
<th>White Children</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>S.D</td>
</tr>
<tr>
<td>12-15</td>
<td>24.49</td>
<td>22</td>
<td>14.84</td>
</tr>
<tr>
<td>12</td>
<td>23.39</td>
<td>21</td>
<td>14.19</td>
</tr>
<tr>
<td>13</td>
<td>24.38</td>
<td>21</td>
<td>15.5</td>
</tr>
<tr>
<td>15</td>
<td>22.53</td>
<td>18</td>
<td>15.34</td>
</tr>
</tbody>
</table>

Number of observations in parenthesis
Means and Medians are calculated on non-zero observations
Units: hours per week

Table 6: Homework Doing and Educational Attainment: Intensive Margin

<table>
<thead>
<tr>
<th>Time Spent Doing Homework</th>
<th>Enrolled into College</th>
<th>H.S. Graduate</th>
<th>H.S. Drop Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Mean</td>
<td>Median</td>
<td>N</td>
</tr>
<tr>
<td>12</td>
<td>5.43</td>
<td>5</td>
<td>294</td>
</tr>
<tr>
<td>13</td>
<td>5.73</td>
<td>5</td>
<td>461</td>
</tr>
<tr>
<td>14</td>
<td>6.12</td>
<td>5.5</td>
<td>481</td>
</tr>
<tr>
<td>15</td>
<td>6.14</td>
<td>5</td>
<td>153</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time Spent Watching TV</th>
<th>Enrolled into College</th>
<th>H.S. Graduate</th>
<th>H.S. Drop Out</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Mean</td>
<td>Median</td>
<td>N</td>
</tr>
<tr>
<td>12</td>
<td>14.9</td>
<td>12.5</td>
<td>319</td>
</tr>
<tr>
<td>13</td>
<td>15.3</td>
<td>13</td>
<td>502</td>
</tr>
<tr>
<td>14</td>
<td>15.2</td>
<td>13</td>
<td>524</td>
</tr>
<tr>
<td>15</td>
<td>15.69</td>
<td>14</td>
<td>168</td>
</tr>
</tbody>
</table>

Statistics computed using NLSY97 custom weights only on positive observations
Units: hours per week
<table>
<thead>
<tr>
<th></th>
<th>Study Time (Int. Margin)</th>
<th>Study Time (Ext. Margin)</th>
<th>TV Viewership</th>
<th>est4</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAR_CURFEW_1</td>
<td>0.027 (0.042)</td>
<td>0.173** (0.082)</td>
<td>0.072** (0.036)</td>
<td>0.072** (0.032)</td>
</tr>
<tr>
<td>PAR_CURFEW_2</td>
<td>-0.067 (0.042)</td>
<td>0.012 (0.082)</td>
<td>-0.020 (0.035)</td>
<td></td>
</tr>
<tr>
<td>PAR_CURFEW_3</td>
<td>-0.023 (0.042)</td>
<td>-0.131 (0.082)</td>
<td>0.038 (0.036)</td>
<td></td>
</tr>
<tr>
<td>PAR_FRIENDS_1</td>
<td>0.075 (0.051)</td>
<td>-0.126 (0.096)</td>
<td>0.088** (0.043)</td>
<td>0.045 (0.039)</td>
</tr>
<tr>
<td>PAR_FRIENDS_2</td>
<td>-0.169** (0.067)</td>
<td>-0.002 (0.127)</td>
<td>-0.047 (0.057)</td>
<td></td>
</tr>
<tr>
<td>PAR_FRIENDS_3</td>
<td>0.078 (0.073)</td>
<td>0.088 (0.145)</td>
<td>-0.004 (0.063)</td>
<td></td>
</tr>
<tr>
<td>PAR_FRIENDS_5</td>
<td>0.003 (0.065)</td>
<td>-0.100 (0.122)</td>
<td>0.045 (0.054)</td>
<td></td>
</tr>
<tr>
<td>PAR_TV_1</td>
<td>0.020 (0.044)</td>
<td>0.039 (0.089)</td>
<td>-0.133*** (0.037)</td>
<td>-0.132*** (0.035)</td>
</tr>
<tr>
<td>PAR_TV_2</td>
<td>0.031 (0.057)</td>
<td>0.014 (0.114)</td>
<td>-0.109** (0.049)</td>
<td>-0.119*** (0.044)</td>
</tr>
<tr>
<td>PAR_TV_3</td>
<td>-0.037 (0.063)</td>
<td>0.229* (0.136)</td>
<td>-0.159*** (0.054)</td>
<td>-0.122** (0.049)</td>
</tr>
<tr>
<td>PIAT_1</td>
<td>0.003*** (0.001)</td>
<td>0.010*** (0.002)</td>
<td>-0.003*** (0.001)</td>
<td>-0.003*** (0.001)</td>
</tr>
<tr>
<td>STUDY</td>
<td></td>
<td></td>
<td>-0.071 (0.052)</td>
<td></td>
</tr>
</tbody>
</table>

First column's dependent variable = log of weekly hours spent doing homework in a typical week
Second column's dependent variable = 0 if the child does not spend any time studying and 1 otherwise
Third and fourth column dependent variable: log of weekly hours spent watching TV
Age, sex and race dummies included
S.E. in parenthesis

N 2086 2307 2213 2648
<table>
<thead>
<tr>
<th></th>
<th>Curfew</th>
<th>Limits on TV</th>
<th>Limits on Friends</th>
<th>Curfew</th>
<th>Limits on TV</th>
<th>Limits on Friends</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIAT</td>
<td>-0.051** (0.022)</td>
<td>-0.066*** (0.023)</td>
<td>-0.170*** (0.023)</td>
<td>0.013 (0.130)</td>
<td>0.277*** (0.094)</td>
<td></td>
</tr>
<tr>
<td>BLACK</td>
<td>0.236*** (0.056)</td>
<td>-0.096 (0.060)</td>
<td>0.286*** (0.059)</td>
<td>0.260*** (0.093)</td>
<td>0.097 (0.141)</td>
<td></td>
</tr>
<tr>
<td>HISPANIC</td>
<td>0.052 (0.060)</td>
<td>-0.061 (0.065)</td>
<td>0.163** (0.066)</td>
<td>0.047 (0.098)</td>
<td>0.073 (0.112)</td>
<td></td>
</tr>
<tr>
<td>SINGLEXMOTH</td>
<td>-0.048 (0.048)</td>
<td>-0.288*** (0.053)</td>
<td>-0.094* (0.052)</td>
<td>-0.291*** (0.082)</td>
<td>-0.064 (0.114)</td>
<td></td>
</tr>
<tr>
<td>MOTHDROP</td>
<td>0.214*** (0.063)</td>
<td>-0.088 (0.068)</td>
<td>0.111 (0.074)</td>
<td>0.258** (0.103)</td>
<td>0.286*** (0.094)</td>
<td></td>
</tr>
<tr>
<td>MOTHHS</td>
<td>0.141** (0.057)</td>
<td>-0.091 (0.062)</td>
<td>0.112 (0.069)</td>
<td>0.097 (0.095)</td>
<td>0.027 (0.131)</td>
<td></td>
</tr>
<tr>
<td>FATHDROP</td>
<td>0.056 (0.106)</td>
<td>-0.125 (0.115)</td>
<td>-0.026 (0.108)</td>
<td>-0.096 (0.173)</td>
<td>-0.276 (0.237)</td>
<td></td>
</tr>
<tr>
<td>FATHHS</td>
<td>-0.031 (0.051)</td>
<td>-0.044 (0.055)</td>
<td>-0.006 (0.055)</td>
<td>-0.191** (0.084)</td>
<td>0.067 (0.111)</td>
<td></td>
</tr>
<tr>
<td>NETWORTH</td>
<td>-0.000*** (0.000)</td>
<td>-0.000 (0.000)</td>
<td>-0.000*** (0.000)</td>
<td>-0.000** (0.000)</td>
<td>-0.000 (0.000)</td>
<td></td>
</tr>
<tr>
<td>BOY</td>
<td>-0.002 (0.040)</td>
<td>0.207*** (0.043)</td>
<td>0.214*** (0.045)</td>
<td>0.004 (0.067)</td>
<td>0.194** (0.091)</td>
<td></td>
</tr>
<tr>
<td>NSIB</td>
<td>0.042** (0.017)</td>
<td>0.098*** (0.017)</td>
<td>0.037** (0.018)</td>
<td>0.079*** (0.028)</td>
<td>0.170*** (0.035)</td>
<td></td>
</tr>
<tr>
<td>GPA</td>
<td>-0.001** (0.001)</td>
<td>0.003*** (0.001)</td>
<td>-0.001 (0.001)</td>
<td>-0.001 (0.001)</td>
<td>-0.001 (0.001)</td>
<td></td>
</tr>
</tbody>
</table>

Dependent variable = 1 if the parent sets the limit alone and 0 otherwise
Probit Regression using pooled data from all the survey years
Age dummies included
Estimated coefficients are the marginal effects
Table 9: Limits and Human Capital

<table>
<thead>
<tr>
<th></th>
<th>PIAT</th>
<th>GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>STUDYAGE1</td>
<td>0.016**</td>
<td>(0.007)</td>
</tr>
<tr>
<td>STUDYAGE2</td>
<td>0.011</td>
<td>(0.009)</td>
</tr>
<tr>
<td>TVAGE1</td>
<td>-0.006***</td>
<td>(0.002)</td>
</tr>
<tr>
<td>TVAGE2</td>
<td>-0.006**</td>
<td>(0.003)</td>
</tr>
<tr>
<td>PIATAGE1</td>
<td>0.720***</td>
<td>(0.021)</td>
</tr>
<tr>
<td>PIATAGE2</td>
<td>0.681***</td>
<td>(0.024)</td>
</tr>
<tr>
<td>PIATAGE3</td>
<td>0.634***</td>
<td>(0.047)</td>
</tr>
<tr>
<td>STUDYAGE3</td>
<td>-0.166</td>
<td>(0.875)</td>
</tr>
<tr>
<td>TVAGE3</td>
<td>-1.328***</td>
<td>(0.217)</td>
</tr>
<tr>
<td>GPAAGE1</td>
<td>0.323*</td>
<td>(0.175)</td>
</tr>
<tr>
<td>GPAAGE2</td>
<td>0.114</td>
<td>(0.155)</td>
</tr>
<tr>
<td>GPAAGE3</td>
<td>0.146***</td>
<td>(0.024)</td>
</tr>
</tbody>
</table>

N 1393 648

Dependent variable = 1 if the parent sets the limit alone and 0 otherwise
Probit Regression using pooled data from all the survey years
Age dummies included
PIAT test scores and GPA have been standardized using the standard deviation.
<table>
<thead>
<tr>
<th></th>
<th>Age 12-13</th>
<th>Age 13-14</th>
<th>Age 14-15</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parents</td>
<td>Child</td>
<td>Jointly</td>
</tr>
<tr>
<td><strong>Curfew</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>67.45</td>
<td>3.94</td>
<td>28.61</td>
</tr>
<tr>
<td>Black</td>
<td>75.52</td>
<td>3</td>
<td>21.48</td>
</tr>
<tr>
<td>White</td>
<td>64.55</td>
<td>2.37</td>
<td>33.08</td>
</tr>
<tr>
<td><strong>Friends’ Limits</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>28.16</td>
<td>43.36</td>
<td>28.16</td>
</tr>
<tr>
<td>Black</td>
<td>34.87</td>
<td>36.49</td>
<td>28.64</td>
</tr>
<tr>
<td>White</td>
<td>17.35</td>
<td>51.08</td>
<td>31.57</td>
</tr>
<tr>
<td><strong>TV Limits</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>35.96</td>
<td>35.17</td>
<td>28.87</td>
</tr>
<tr>
<td>Black</td>
<td>35.33</td>
<td>35.33</td>
<td>29.33</td>
</tr>
<tr>
<td>White</td>
<td>37.5</td>
<td>27.8</td>
<td>34.7</td>
</tr>
</tbody>
</table>
### Table 11: Transitions

#### Transitions of the Curfew Limits

<table>
<thead>
<tr>
<th>Age 12-13</th>
<th>Age 13-14</th>
<th>Age 14-15</th>
<th>Age 14-15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parents</td>
<td>63.48</td>
<td>31.97</td>
<td>4.55</td>
</tr>
<tr>
<td>Jointly</td>
<td>41.36</td>
<td>53.09</td>
<td>5.54</td>
</tr>
<tr>
<td>Child</td>
<td>34.78</td>
<td>41.30</td>
<td>23.91</td>
</tr>
<tr>
<td>Parents</td>
<td>60.61</td>
<td>5.42</td>
<td>33.96</td>
</tr>
<tr>
<td>Jointly</td>
<td>27.09</td>
<td>8.86</td>
<td>64.05</td>
</tr>
<tr>
<td>Child</td>
<td>35.00</td>
<td>22.50</td>
<td>42.50</td>
</tr>
<tr>
<td>Parents</td>
<td>52.35</td>
<td>6.81</td>
<td>40.84</td>
</tr>
<tr>
<td>Jointly</td>
<td>33.41</td>
<td>8.30</td>
<td>58.3</td>
</tr>
<tr>
<td>Child</td>
<td>41.86</td>
<td>18.60</td>
<td>39.53</td>
</tr>
</tbody>
</table>

#### Transitions of the Friends Limits

<table>
<thead>
<tr>
<th>Age 12-13</th>
<th>Age 13-14</th>
<th>Age 14-15</th>
<th>Age 14-15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parents</td>
<td>26.70</td>
<td>31.06</td>
<td>42.23</td>
</tr>
<tr>
<td>Jointly</td>
<td>11.98</td>
<td>39.46</td>
<td>48.55</td>
</tr>
<tr>
<td>Child</td>
<td>5.56</td>
<td>20.33</td>
<td>74.12</td>
</tr>
<tr>
<td>Parents</td>
<td>30.65</td>
<td>38.71</td>
<td>30.65</td>
</tr>
<tr>
<td>Jointly</td>
<td>10.38</td>
<td>49.44</td>
<td>40.18</td>
</tr>
<tr>
<td>Child</td>
<td>4.50</td>
<td>78.04</td>
<td>17.45</td>
</tr>
<tr>
<td>Parents</td>
<td>19.06</td>
<td>48.07</td>
<td>32.87</td>
</tr>
<tr>
<td>Jointly</td>
<td>10.09</td>
<td>57.68</td>
<td>32.24</td>
</tr>
<tr>
<td>Child</td>
<td>4.66</td>
<td>77.23</td>
<td>18.11</td>
</tr>
</tbody>
</table>

#### Transitions of the TV Limits

<table>
<thead>
<tr>
<th>Age 12-13</th>
<th>Age 13-14</th>
<th>Age 14-15</th>
<th>Age 14-15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parents</td>
<td>32.14</td>
<td>32.33</td>
<td>35.53</td>
</tr>
<tr>
<td>Jointly</td>
<td>12.8</td>
<td>33.8</td>
<td>53.40</td>
</tr>
<tr>
<td>Child</td>
<td>5.74</td>
<td>14.13</td>
<td>80.12</td>
</tr>
<tr>
<td>Parents</td>
<td>38.92</td>
<td>30.14</td>
<td>30.94</td>
</tr>
<tr>
<td>Jointly</td>
<td>12.92</td>
<td>45.35</td>
<td>41.53</td>
</tr>
<tr>
<td>Child</td>
<td>4.27</td>
<td>83.57</td>
<td>12.16</td>
</tr>
<tr>
<td>Parents</td>
<td>24.16</td>
<td>45.35</td>
<td>30.5</td>
</tr>
<tr>
<td>Jointly</td>
<td>9.17</td>
<td>62.95</td>
<td>28.24</td>
</tr>
<tr>
<td>Child</td>
<td>4.56</td>
<td>84.49</td>
<td>10.95</td>
</tr>
<tr>
<td></td>
<td>Parents’ Responses on Curfew</td>
<td>Parents’ Responses on Friends’ Limits</td>
<td>Parents’ Responses on TV Limits</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------------------</td>
<td>---------------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td></td>
<td>Parent</td>
<td>Child</td>
<td>Jointly</td>
</tr>
<tr>
<td>Parent</td>
<td>91.26</td>
<td>0.68</td>
<td>8.06</td>
</tr>
<tr>
<td>Child</td>
<td>81.16</td>
<td>0.84</td>
<td>18.00</td>
</tr>
<tr>
<td>Jointly</td>
<td>80.00</td>
<td>2.35</td>
<td>7.65</td>
</tr>
</tbody>
</table>
Table 13: Breaking Limits in the Last Month

<table>
<thead>
<tr>
<th>Age</th>
<th>Never</th>
<th>Once</th>
<th>More than once</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>69.97</td>
<td>8.79</td>
<td>21.24</td>
<td>1172</td>
</tr>
<tr>
<td>13</td>
<td>70</td>
<td>8.29</td>
<td>21.71</td>
<td>1749</td>
</tr>
<tr>
<td>14</td>
<td>67.56</td>
<td>10</td>
<td>22.44</td>
<td>2081</td>
</tr>
<tr>
<td>15</td>
<td>65</td>
<td>9.31</td>
<td>25.69</td>
<td>2867</td>
</tr>
<tr>
<td>16</td>
<td>59.29</td>
<td>10.26</td>
<td>30.45</td>
<td>1382</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Never</th>
<th>Once</th>
<th>More than once</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>71.01</td>
<td>8.69</td>
<td>20.3</td>
<td>852</td>
</tr>
<tr>
<td>13</td>
<td>73.2</td>
<td>7.79</td>
<td>19.01</td>
<td>1129</td>
</tr>
<tr>
<td>14</td>
<td>74.3</td>
<td>6.28</td>
<td>19.42</td>
<td>1082</td>
</tr>
<tr>
<td>15</td>
<td>73.37</td>
<td>6.93</td>
<td>19.7</td>
<td>1168</td>
</tr>
<tr>
<td>16</td>
<td>77.85</td>
<td>6.14</td>
<td>16.01</td>
<td>456</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Never</th>
<th>Once</th>
<th>More than once</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>80.93</td>
<td>5.26</td>
<td>13.81</td>
<td>666</td>
</tr>
<tr>
<td>13</td>
<td>79.4</td>
<td>6.97</td>
<td>13.63</td>
<td>918</td>
</tr>
<tr>
<td>14</td>
<td>75.7</td>
<td>7.23</td>
<td>17.07</td>
<td>884</td>
</tr>
<tr>
<td>15</td>
<td>76.8</td>
<td>6.98</td>
<td>16.22</td>
<td>1073</td>
</tr>
<tr>
<td>16</td>
<td>71.26</td>
<td>6.7</td>
<td>22.04</td>
<td>522</td>
</tr>
<tr>
<td>17</td>
<td>69.7</td>
<td>6.37</td>
<td>23.93</td>
<td>675</td>
</tr>
<tr>
<td>18</td>
<td>71</td>
<td>6.96</td>
<td>22.04</td>
<td>431</td>
</tr>
<tr>
<td>Table 14: Punishment by Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Punishment for Children Between the Ages of 12 and 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Curfew Punishment</td>
<td>TV Shows Punishment</td>
<td>Friends Limits Punishment</td>
<td></td>
</tr>
<tr>
<td>Discuss it calmly with you</td>
<td>41.8</td>
<td>53.62</td>
<td>54.28</td>
<td></td>
</tr>
<tr>
<td>Ignore it, pretend that it didn't happen or let you get away with it</td>
<td>2.02</td>
<td>3.37</td>
<td>1.82</td>
<td></td>
</tr>
<tr>
<td>Sulk, pout, or give you the silent treatment</td>
<td>0.59</td>
<td>0.76</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Take away a privilege, ground you, or give you a chore</td>
<td>45.12</td>
<td>32.91</td>
<td>30.41</td>
<td></td>
</tr>
<tr>
<td>Make threats that won't be kept</td>
<td>1.72</td>
<td>1.85</td>
<td>2.03</td>
<td></td>
</tr>
<tr>
<td>Yell, shout, or scream at you</td>
<td>6.36</td>
<td>5.13</td>
<td>6.53</td>
<td></td>
</tr>
<tr>
<td>Use physical punishment</td>
<td>2.38</td>
<td>2.36</td>
<td>4.07</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>1682</td>
<td>1188</td>
<td>934</td>
<td></td>
</tr>
</tbody>
</table>

| Punishment for Children Between the Ages of 13 and 14 |
| | Curfew Punishment | TV Shows Punishment | Friends Limits Punishment |
| Discuss it calmly with you | 39.22 | 54.22 | 56.38 |
| Ignore it, pretend that it didn't happen or let you get away with it | 3.89 | 4.92 | 3.04 |
| Sulk, pout, or give you the silent treatment | 0.73 | 1.13 | 1.06 |
| Take away a privilege, ground you, or give you a chore | 43.18 | 28.75 | 28.27 |
| Make threats that won't be kept | 2.97 | 2.52 | 1.22 |
| Yell, shout, or scream at you | 8.64 | 7.19 | 8.21 |
| Use physical punishment | 1.38 | 1.26 | 1.82 |
| N | 1517 | 793 | 658 |

| Punishment for Children Between the Ages of 14 and 15 |
| | Curfew Punishment | TV Shows Punishment | Friends Limits Punishment |
| Discuss it calmly with you | 41 | 55.41 | 53.07 |
| Ignore it, pretend that it didn't happen or let you get away with it | 3.25 | 6.37 | 3.97 |
| Sulk, pout, or give you the silent treatment | 0.9 | 0.8 | 1.26 |
| Take away a privilege, ground you, or give you a chore | 41.41 | 27.87 | 29.24 |
| Make threats that won't be kept | 4.57 | 2.55 | 3.79 |
| Yell, shout, or scream at you | 7.96 | 5.57 | 7.22 |
| Use physical punishment | 0.9 | 1.43 | 1.44 |
| N | 1444 | 628 | 554 |