# Appendices

## A Model Solution

### A.1 Final period: $t = M$

For both the parents and the child, the decision rules are solved using backward recursion beginning from the end of the development process, time $M$.

**Child's problem.** In period $M$, the child makes her choice of $\tau_{c,M}$ given the time and budget allocation of the parents, $a_{p,M}$, and, if the parents are using an Internal CCT, the contract specified by $\{r_M, b_M\}$. Then the child's period $M$ problem is

$$\begin{aligned} V_{c,M}(\Gamma_M|a_{p,M}, r_M, b_M) = \max_{\tau_{c,M}|a_{p,M},r_M,b_M} \quad & \lambda_1 \ln(\tilde{T}_M - \tau_{p,M} - \tau_{c,M}) + \lambda_2 \ln x_M + \lambda_3 \ln k_M \\ & + \beta_{c,M}\psi_{c,M+1} E(\ln k_{M+1}|\tau_{c,M}, a_{p,M}, r_M, b_M) \end{aligned}$$

where $\tilde{T}_M$ is the child's time endowment after subtracting exogenous school time $s_M$. We can substitute out these two components:

$$\begin{aligned} E(\ln k_{M+1}|\tau_{c,M}, a_{p,M}, r_M, b_M) &= \ln R_M + \delta_{1,M} \ln \tau_{1,M} + \delta_{2,M} \ln \tau_{2,M} + \delta_{3,M} \ln \tau_{12,M} \\ &\quad + \delta_{4,M} \ln e_M + \delta_{5,M} \ln \tau_{c,M} + \delta_{6,M} \ln k_M, \\ \ln x_M &= b_M + r_M \ln \tau_{c,M} \end{aligned}$$

Since we assume that all parameters, including Total Factor Productivity $R_M$, are known at the time of the period $M$ decisions, there is no uncertainty present in the production technology, allowing us to drop the expectation operator. The optimal decision of the child is given by

$$\begin{aligned} \tau^*_{c,M}(\tau_{p,M}, r_M) &= \frac{\lambda_2 r_M + \Delta_{c,M}}{\lambda_1 + \lambda_2 r_M + \Delta_{c,M}}(\tilde{T}_M - \tau_{p,M}) \\ &= \gamma_M(r_M)(\tilde{T}_M - \tau_{p,M}) \end{aligned} \tag{A-1}$$

where $\Delta_{c,M} \equiv \beta_{c,M}\psi_{c,M+1}\delta_{5,M}$. Given the properties of the production, utility and reward functions, the choice of time in investment is independent of all of the parents' decisions with the exception of (1) the total time they spend interacting with the children, $\tau_{p,M}$, the effect of which is to reduce the child's effective time endowment, and (2) the child's "wage" rate $r_M$, which corresponds to the elasticity of child consumption with respect to child study time. The fact that $b_M$ drops out will prove useful in deriving some of the results below. Note that when $r_M = 0$, this solution simplifies to the special case in which the parents make a fixed transfer of $x_M$ to the child that is not tied to the child's investment time. Clearly, the solution to the child's problem is increasing in $r_M$ and the child can be induced to spend virtually all of its time in investment as $r_M$

becomes arbitrarily large. However, this could never be optimal since (1) the child has an incentive compatibility (IC) constraint that must be satisfied whenever the parents use an ICCT scheme, and (2) even in the absence of an IC constraint, the parents would not want their child to have zero leisure as long as they are altruistic ($\varphi > 0$).

**Parents' problem.** Given the child's reaction function $\tau^*_{c,M}(\tau_{p,M}, r_M)$, the parents solve the following problem:

$$V_{p,M}(\Gamma_M) = \max_{a_{p,M}, r_M, b_M} \tilde{u}_p(l_{1,M}, l_{2,M}, c_M, k_M, l_{c,M}, x_M) + \beta_p \psi_{p,M+1} \ln(k_{M+1})$$
$$+\mu_M \Big( \lambda_1 \ln(l_{c,M}) + \lambda_2 \ln(x_M) + \lambda_3 \ln(k_M) + \beta_{c,M} \psi_{c,M+1} \ln(k_{M+1}) - V_{c,M}(\Gamma_M | a^0_{p,M}) \Big) \quad \text{(A-2)}$$

where $\mu_M \geq 0$ is the Lagrange multiplier on the child's IC constraint, and $V_{c,M}(\Gamma_M | a^0_{p,M})$ denotes the child's outside option, i.e. the indirect value function evaluated at the parents' choices in the absence of an ICCT. We can substitute out (1) $c_M$ for the period $M$ budget constraint, (2) $l_{1,M}$, $l_{2,M}$ and $l_{c,M}$ for the individual time constraints, (3) $\ln k_{M+1}$ for the production technology, and (4) $\tau_{c,M}$ for the child's optimal reaction function derived in the previous paragraph. In order to simplify the first order conditions with respect to the remaining choices $\{h_{1,t}, h_{2,t}, \tau_{1,t}, \tau_{2,t}, \tau_{12,t}, e_t, x_t, r_t, b_t\}$, note that the parents jointly choose the triple $\{x_M, r_M, b_M\}$ subject to the reward function. Rearranging this equation yields

$$\begin{aligned} b_M &= \ln(x_M) - r_M \ln(\tau^*_{c,M}(\tau_{p,M}, r_M)) \\ &= \ln(x_M) - r_M \ln(\gamma_M(r_M)) - r_M \ln(\tilde{T}_M - \tau_{p,M}) \end{aligned}$$

Conditional on $\{x_M, r_M, \tau_{1,M}, \tau_{2,M}, \tau_{12,M}\}$, this will pin down the optimal choice of $b_M$. Taking first order conditions with respect to $\{e_M, x_M\}$ and using the budget constraint yields the following solutions for the expenditures (conditional on labor supply choices and the multiplier on the child's IC constraint):

$$c^\star_M = \frac{\tilde{\alpha}_3}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,M+1} \delta_{4,M} + \mu_M(\lambda_2 + \beta_{c,M} \psi_{c,M+1} \delta_{4,M})} Y_M \quad \text{(A-3)}$$

$$e^\star_M = \frac{\beta_p \psi_{p,M+1} \delta_{4,M} + \mu_M \beta_{c,M} \psi_{c,M+1} \delta_{4,M}}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,M+1} \delta_{4,M} + \mu_M(\lambda_2 + \beta_{c,M} \psi_{c,M+1} \delta_{4,M})} Y_M \quad \text{(A-4)}$$

$$x^\star_M = \frac{\tilde{\alpha}_6 + \mu_M \lambda_2}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,M+1} \delta_{4,M} + \mu_M(\lambda_2 + \beta_{c,M} \psi_{c,M+1} \delta_{4,M})} Y_M \quad \text{(A-5)}$$

where $Y_M = w_{1,M} h_{1,M} + w_{2,M} h_{2,M} + I_M$ and $\mu_M \geq 0$. It is easy to see that the fraction of income spent on the parents' private consumption, $c_M$, is strictly decreasing in $\mu_M$. Conversely, the fraction spent on the child's consumption, $x_M$, is strictly increasing in $\mu_M$ under a weak condition on the primitives:

$$\frac{\partial x^*_M / Y_M}{\partial \mu_M} > 0 \iff \beta_p \psi_{p,M+1} - \varphi \beta_{c,M} \psi_{c,M+1} > \frac{-\tilde{\alpha}_3}{\delta_{4,M}}$$

Since the parents in our model are not perfectly altruistic ($\varphi < 1$), are more patient than children ($\beta_p > \beta_{c,t}$) and usually (but not always in our model) care more about future child quality than the child ($\psi_{p,t+1} > \psi_{c,t+1}$), the left-hand side of this inequality should typically be positive, making this condition satisfied.[31]

After substituting out the (conditional) optimal choices $\{c^\star_M, e^\star_M, x^\star_M\}$ and the child's reaction function in the value function $V_{p,M}$, and after dropping some constant terms, we are left with a "residual" maximization problem which can be decomposed into two separate maximization problems, conditional on the Lagrange multiplier $\mu_M$:

$$W_{p,M} = \max \nu_{1,M}(h_{1,M}, h_{2,M}, \tau_{1,M}, \tau_{2,M}, \tau_{12,M}; \mu_M) + \nu_{2,M}(r_M; \mu_M)$$

where the second component is given by

$$\nu_{2,M}(r_M; \mu_M) \equiv \left(\tilde{\alpha}_5 + \mu_M \lambda_1\right) \ln(1 - \gamma_M(r_M)) + \left(\Delta_{p,M} + \mu_M \Delta_{c,M}\right) \ln(\gamma_M(r_M)). \quad \text{(A-6)}$$

where $\Delta_{c,M} \equiv \beta_{c,M}\psi_{c,M+1}\delta_{5,M}$ and $\Delta_{p,M} \equiv \beta_p \psi_{p,M+1}\delta_{5,M}$.

**Unconstrained optimum.** Denote the unconstrained optimal time, budget and ICCT choices by the vector $\{a^{unc}_{p,M}, r^{unc}_M, b^{unc}_M\}$. If the multiplier $\mu_M$ equals 0, the parents' optimization problem becomes truly separable, since the two components, $\nu_{1,M}(a_{p,M})$ and $\nu_{2,M}(r_M)$, no longer have any common components. Therefore, none of the unconstrained optimal time and budget choices (summarized by $a^{unc}_M = \{h_{1,M},\ h_{2,M},\ \tau_{1,M},\ \tau_{2,M},\ \tau_{12,M},\ c_M,\ e_M,\ x_M\}$) will depend on the parents' choice of $r_M$. Indeed, under our functional form assumptions, it must be the case that

$$a^{unc}_{p,M} = a^0_{p,M} = \hat{a}_{p,M}$$

where $a^0_{p,M}$ denotes the optimal time and budget allocation in the no-ICCT Stackelberg equilibrium (where $r_M = 0$), and where $\hat{a}_{p,M}$ denotes the optimal time and budget allocation in the dictatorial model, where parents (hypothetically) choose the child's study time directly. Conditional on parental labor supply, we can find the optimal budget allocation by plugging in $\mu_M = 0$ into Equations (A-3)-(A-5). The remaining optimal choices can be found by maximizing the sub-function $\nu_{1,M}(a_{p,M})$ analytically.

It is convenient that the second sub-function, $\nu_{2,M}(r_M)$ shown in (A-6) only depends on the reward elasticity, $r_M$, due to the functional form of the ICCT reward function. The parents' ability to implement an ICCT (assuming a slack incentive constraint for the child) allows them to perfectly align the child's incentives with their own altruistic preferences by implementing the following ICCT contract:

$$\begin{aligned} r^{unc}_M &= \frac{\Delta_{p,M} - \varphi \Delta_{c,M}}{\lambda_2 \varphi}, \\ b^{unc}_M &= \ln(x^{unc}_M) - r^{unc}_M \ln(\tau_{c,M}(\tau^{unc}_{p,M}, r^{unc}_M)) \end{aligned}$$

[31]Given our model estimates and random simulation draws, this condition always holds in our analysis.

where $r_M^{unc}$ follows from differentiating $\nu_{2,M}(r_M)$, and $b_M^{unc}$ follows from the ICCT reward function. Note that given this contract, the child's optimal response coincides with what the parents would choose themselves if they were the household dictator (i.e. $\tau_{c,M}^{unc} = \hat{\tau}_{c,M}$).

By complementary slackness, we should finally verify whether the child's IC constraint in the parents' problem (A-2) is indeed slack, given the parents' unconstrained choice vector. As stated in Proposition 1, we know that whenever the parents are using an ICCT with $r_t \neq 0$, the child's incentive constaint must be binding in equilibrium, thereby effectively ruling out the unconstrained equilibrium we have just derived. For completeness, we also provide a more formal proof of Proposition 1 for a general period $t \in \{1, ..., M\}$.

First, we define the child's outside option, $V_{c,t}(\Gamma_t | a_{p,t}^0)$, as her indirect value when the parents are not using an Internal CCT scheme, i.e. when $a_{p,t} = a_{p,t}^0$ and $r_t = 0$. We prove by contradiction, i.e. by assuming that the child's IC constraint will be slack in the Stackelberg equilibrium with ICCT and $\mu_t = 0$. From before, we know that the optimal reaction function of the child is given by Equation (A-1). After plugging in this reaction function and the parents' optimal choices, we can write the child's indirect value function as follows:

$$V_{c,t}(\Gamma_t | a_{p,t}^{unc}, r_t^{unc}, b_t^{unc}) = \lambda_1 \ln(l_{c,t}^{unc}) + \lambda_2 \ln(x_t^{unc}) + \lambda_3(k_t) + \beta_{c,t}\psi_{c,t+1} \ln(k_{t+1}(a_{p,t}^{unc}, r_t^{unc}))$$

where the child's leisure $l_{c,t}^{unc} = (1 - \gamma_t(r_t^{unc}))(\tilde{T}_t - \tau_{p,t}^{unc})$. Importantly, in this indirect utility function, the child's consumption level no longer depends on the child's study time. Indeed, even though the parents are offering the child an incentive scheme (or reward function) given by $x_t(\tau_{c,t}; r_t, b_t)$, the child realizes that irrespective of how much she studies, the parents can always implement their first-best value of child consumption (given by $x_t^{unc} = x_t^0 = \hat{x}_t$) by simply readjusting (or reneging on) the value of $b_t^{unc}$ after the child has chosen how much time to devote to studying. This lack of commitment on behalf of the parents would make the child unwilling to participate in the incentive scheme and deviate back to the no-ICCT Stackelberg equilibrium. Indeed, given our previous result that $a_{p,t}^{unc} = a_{p,t}^0$, we can simplify child's IC constraint as follows:

$$\begin{aligned}
& V_{c,t}(\Gamma_t | a_{p,t}^{unc}, r_t^{unc}, b_t^{unc}) \geq V_{c,t}(\Gamma_t | a_{p,t}^0) \\
\iff \quad & \lambda_1 \ln(l_{c,t}^{unc}) + \beta_{c,t}\psi_{c,t+1}\delta_{5,t} \ln(\tau_{c,t}^{unc}) \geq \lambda_1 \ln(l_{c,t}^0) + \beta_{c,t}\psi_{c,t+1}\delta_{5,t} \ln(\tau_{c,t}^0) \\
\iff \quad & \lambda_1 \ln(1 - \gamma_t(r_t^{unc})) + \Delta_{c,t} \ln(\gamma_t(r_t^{unc})) \geq \lambda_1 \ln(1 - \gamma_t(0)) + \Delta_{c,t} \ln(\gamma_t(0))
\end{aligned}$$

First, note that this inequality is binding if and only if $r_t^{unc} = 0$, which is only optimal in the knife-edge case where $\Delta_{p,t} = \varphi \Delta_{c,t}$.[32] Second, while the right-hand side of the

[32] Under the relatively weak assumption on the primitives that $\Delta_{p,t} > \varphi \Delta_{c,t}$, parents prefer to positively incentivize their children, i.e. to set $r_t > 0$. Although our model does not rule out that some parents may prefer to implement negative incentive schemes ($r_t < 0$), it is never the case given our parameter estimates and random simulation draws. Moreover, even in those cases where $r_t^{unc} < 0$, the child would still prefer to deviate back to the no-ICCT equilibrium by studying *more* than what the parents prefer.

above inequality does not depend on $r_t$, we can show that the left-hand side is strictly *decreasing* in $r_t$, by taking the partial derivative:

$$\begin{aligned}\frac{\partial V_{c,t}(\Gamma_t|a_{p,t}^{unc}, r_t, b_t)}{\partial r_t} &= \frac{\partial \gamma_t(r_t)}{\partial r_t}\Big[\frac{\Delta_{c,t}}{\gamma_t(r_t)} - \frac{\lambda_1}{1-\gamma_t(r_t)}\Big] \\ &= \frac{\partial \gamma_t(r_t)}{\partial r_t}\Big[\frac{-\lambda_2 r_t}{\gamma_t(r_t)}\Big]\end{aligned}$$

where we implicitly use the results that the parents' time and budget allocation ($a_{p,t}^{unc}$) does not vary with $r_t$, and that the child's action does not depend on $b_t$. Since $\gamma_t(r_t)$ is strictly increasing in $r_t$, this partial derivative is zero only when $r_t = 0$, strictly negative whenever $r_t > 0$, and strictly positive whevener $r_t < 0$. This implies that compared to the no-ICCT equilibrium where $r_t = 0$, the child's value function (evaluated at $a_{p,t}^{unc}$) is globally maximized at $r_t = 0$, and strictly decreases whenever $r_t \neq 0$. This means the child's IC constraint is violated whenever $\mu_t = 0$, which rules out the unconstrained ICCT equilibrium. Therefore, the IC constraint is always binding in the ICCT equilibrium.

**Constrained optimum.** Although we cannot solve for $r_M$ in closed form, we can find the optimal value conditional on $\mu_M$, by differentiating $V_{p,M}$ (see Equation (A-2)) with respect to $r_M$. Assuming for now that there will be an interior solution (i.e. $r_M \neq 0$), we obtain:

$$\frac{dV_{p,M}(\Gamma_M|a_{p,M}, r_M, b_M; \mu_M)}{dr_M} = \frac{\partial V_{p,M}}{\partial r_M} + \frac{\partial V_{p,M}}{\partial \mu_M}\frac{\partial \mu_M}{\partial r_M} = 0 \qquad \text{(A-7)}$$

where we have imposed that $\frac{\partial a_{p,M}}{\partial r_M} = \frac{\partial b_M}{\partial r_M} = 0$ due to the optimality principle. We know that $\frac{\partial V_{p,M}}{\partial \mu_M} \leq 0$, since the presence of the child's incentive constraint must decrease the parents' value relative to the unconstrained equilibrium, which coincides with the parents' first-best outcome. Moreover, from Proposition 1, we know that the incentive constraint becomes binding whenever $r_M \neq 0$. Since $\mu_M = 0$ only if $r_M = 0$, this implies that the partial derivative $\frac{\partial \mu_M}{\partial r_M}$ is positive when $r_M^{unc} > 0$, and negative when $r_M^{unc} < 0$. Thus, the second component in (A-7) must be negative whenever $r_M^* > 0$, and positive whenever $r_M^* < 0$. By optimality, the first component must have the opposite sign as the second component. Given our previous discussion of the parents' constrained problem and the expression given in Equation (A-6), we can derive this first component as follows:

$$\begin{aligned}\frac{\partial V_{p,M}}{\partial r_M} = \frac{\partial \nu_{2,M}}{\partial r_M} &= \frac{d\gamma_M(r_M)}{dr_M}\Big[\frac{\Delta_{p,M} + \mu_M \Delta_{c,M}}{\gamma_M(r_M)} - \frac{\tilde{\alpha}_5 + \mu_M \lambda_1}{1-\gamma_M(r_M)}\Big] \\ &= \frac{d\gamma_M(r_M)}{dr_M}\Big[\frac{\Delta_{p,M} - \varphi\Delta_{c,M} - \lambda_2 r_M(\varphi + \mu_M)}{\gamma_M(r_M)}\Big]\end{aligned}$$

where we used the fact that $\tilde{\alpha}_5 = \varphi\lambda_1$. Since $\gamma_M(r_M)$ is strictly increasing in $r_M$, we can derive a bound on the constrained optimal value $r_M^*$:

$$\frac{dV_{p,M}}{dr_M} = 0 \iff |r_M^*| \leq \frac{|\Delta_{p,M} - \varphi\Delta_{c,M}|}{\lambda_2(\varphi + \mu_M)} < |r_M^{unc}| = \frac{|\Delta_{p,M} - \varphi\Delta_{c,M}|}{\lambda_2\varphi}$$

where $\varphi$ is the parents' altruism parameter. Note that as $\mu_M$ approaches 0, $r_M^*$ converges to $r_M^{unc}$. By plugging this bound into the child's reaction function, the corresponding bounds on the child's optimal fraction of study time are:

$$r_M^* > 0 \iff \frac{\Delta_{c,M}}{\lambda_1 + \Delta_{c,M}} < \gamma_M(r_M^*) \leq \frac{\Delta_{p,M} + \mu_M\Delta_{c,M}}{\tilde{\alpha}_5 + \Delta_{p,M} + \mu_M(\lambda_1 + \Delta_{c,M}))} < \frac{\Delta_{p,M}}{\tilde{\alpha}_5 + \Delta_{p,M}}$$

where all inequality signs reverse for the (rare) cases where $r_M^* < 0$, i.e. when $\Delta_{p,M} < \varphi\Delta_{c,M}$. Without an explicit expression for $\mu_M$, we cannot characterize the constrained ICCT optimum any further.

**Conditions for Costly ICCT use.** Since implementing an Internal CCT is, in our most general model, costly for the parents, it may be optimal to not use one, by setting $r_M = 0$. The parents will choose to use an ICCT when the welfare gain from the *constrained* equilibrium exceeds the utility cost $\omega_M$, i.e. under the following necessary and sufficient condition:

$$r_M^* \neq 0 \iff V_{p,M}(\Gamma_M | a_{p,M}^*, r_M^*, b_M^*) - V_{p,M}(\Gamma_M | a_{p,M}^0) \geq \omega_M$$

where $a_{p,M}^*$ and $a_{p,M}^0$ denote the parents' optimal time and budget allocations in the constrained ICCT equilibrium and the no-ICCT equilibrium, respectively. Since the child's incentive constraint in the ICCT equilibrium is always binding, the optimal parental choices will change whenever $r_M \neq 0$ (i.e. $a_{p,M}^* \neq a_{p,M}^0$), preventing us from simplifying this expression any further. In the empirical implementation, we use a numerical solver to evaluate this necessary and sufficient condition for every household at every child age.

We have previously argued that the *unconstrained* parents' problem is separable into two parts, where only the second component, $\nu_2(r_M, \mu_M)$ (see (A-6)) depends on the parents' ICCT parameter. This insight allows us to derive the following necessary (but not sufficient) condition for the parents' choice whether to use an ICCT:

$$\begin{aligned} r_M^* \neq 0 \quad &\implies \quad \nu_{2,M}(r_M = r_M^{unc}, \mu_M = 0) - \nu_{2,M}(r_M = 0, \mu_M = 0) \geq \omega_M \\ &\iff \quad \Delta_{p,M} \ln\Big(\frac{\gamma_M(r_M^{unc})}{\gamma_M(0)}\Big) \geq \tilde{\alpha}_5 \ln\Big(\frac{1 - \gamma_M(0)}{1 - \gamma_M(r_M^{unc})}\Big) + \omega_M \end{aligned}$$

where the closed form for $r_M^{unc}$ is known, and where $\gamma_M(0) = \frac{\Delta_{c,M}}{\lambda_1 + \Delta_{c,M}}$ is the child's optimal fraction of study time in the no-ICCT equilibrium, which is strictly smaller than the fraction of study time in the unconstrained ICCT equilibrium, $\gamma_M(r_M^{unc}) =$

$\frac{\Delta_{p,M}}{\tilde{\alpha}_5+\Delta_{p,M}}$. Consider the most common case where the parents would choose $r_M^{unc} > 0$. Then, the left-hand side of the second line can be interpreted as the parents' benefit of implementing an internal CCT which, by raising the child's study time, will increase child capital in the next period. The right-hand side can be interpreted as the parents' utility cost of implementing the ICCT, comprising both the direct cost, $\omega_M$, as well as a utility loss through the reduction in the child's leisure time, which is weighted by the parents' value of child leisure, $\tilde{\alpha}_5 = \varphi\lambda_1$. Conversely, in the case where $r_M^{unc} < 0$, the parents would like to reduce the child's study time, which now has a utility benefit in terms of leisure, and a utility cost in terms of lost capital. Intuitively, the inequality will be satisfied if the parents' optimal reward elasticity is sufficiently different from 0, either positively or negatively. However, since this necessary condition does not include the additional utility loss the parents must incur due to the child requiring some additional compensation in the constrained equilibrium, it is not sufficient.

**Optimal choices.** Given the functional form assumptions and the presence of the child's incentive constraint, we cannot find closed form solutions for any of the parental choices $\{h_{1,M},\ h_{2,M},\ \tau_{1,N},\ \tau_{2,M},\ \tau_{12,M},\ c_M,\ x_M,\ e_M,\ r_M,\ b_M\}$. In the computational exercise, we will use a numerical solver to find the optimal choice vector, taking into account the possible corner solutions for labor supply. Enforcing the child's incentive constraint involves first solving the unconstrained parents' problem, which (1) allows us to verify that, in accordance with Proposition 1, the child's IC is violated whenever $r_M^{unc} \neq 0$, and (2) provides us with a good initial guess before numerically solving the harder constrained problem where the incentive constraint is imposed at equality. Appendix C contains more details on this estimation procedure.

## A.2 Remaining periods: $t = 1, ..., M-1$

The solution has exactly the same characteristics in the general period $t$ case. The only adjustments to the solution occur with respect to the variables $\psi_{j,t},\ j = c, p$, which measure the future impacts of improvements in child quality in period $t$ and the remaining periods in the development process. The time-varying characteristics that appear in the solution include the production function parameters, the realizations of wages and non-labor income in period $t$, and the discount factor of the child, which is monotonically increasing in $t$. Thus the $t$-period solution is as follows:

**No-ICCT Stackelberg Equilibrium.** First, we solve the household problem assuming the parents are not using an ICCT, such that $r_t = 0$. We denote the total vector of optimal parental choices in the no-ICCT equilibrium as $a_{p,t}^0$.

1. Condition on a choice vector of $\{h_{1,t},\ h_{2,t},\ \tau_{1,t},\ \tau_{2,t},\ \tau_{12,t}\}$, including potential corners for labor supply. Given these values, the household income in period $t$ is $Y_t = w_{1,t}h_{1,t} + w_{2,t}h_{2,t} + I_t$. Total parental time is defined as $\tau_{p,t} = \tau_{1,t} + \tau_{2,t} + \tau_{12,t}$.

2. The optimal expenditures in the no-ICCT Stackelberg equilibrium (conditional on labor supply) are given by

$$c_t^0 = \frac{\tilde{\alpha}_3}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,t+1} \delta_{4,t}} Y_t$$
$$e_t^0 = \frac{\beta_p \psi_{p,t+1} \delta_{4,t}}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,t+1} \delta_{4,t}} Y_t$$
$$x_t^0 = \frac{\tilde{\alpha}_6}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,t+1} \delta_{4,t}} Y_t$$

where

$$\psi_{p,M+1} \equiv \xi_p \alpha_4,$$
$$\psi_{p,t} \equiv \tilde{\alpha}_4 + \beta_p \delta_{6,t} \psi_{p,t+1}, \ t = 1, ..., M.$$

The optimal study time of the child in the absence of an ICCT is given by:

$$\tau_{c,t}^0(\tau_{p,t}, r_t = 0) = \frac{\Delta_{c,t}}{\lambda_1 + \Delta_{c,t}} (\tilde{T}_t - \tau_{p,t})$$

where

$$\Delta_{c,t} \equiv \beta_{c,t} \psi_{c,t+1} \delta_{5,t}, \ t = 1, ..., M,$$
$$\psi_{c,M+1} \equiv \xi_c \lambda_3,$$
$$\psi_{c,t} \equiv \lambda_3 + \beta_{c,t} \delta_{6,t} \psi_{c,t+1}, \ t = 1, ..., M.$$

3. By using the time constraints and the production technology function, we find the leisure of each individual $(l_{1,t}^0, l_{2,t}^0, l_{c,t}^0)$ and future child capital, $k_{t+1}^0$. This allows us to define the parental value function:

$$V_{p,t}(\Gamma_t, a_{p,t}^0) = \tilde{u}_p(l_{1,t}^0, l_{2,t}^0, c_t^0, k_t, l_{c,t}^0, x_t^0) + \beta_p \psi_{p,t+1} \ln(k_{t+1}^0)$$

We use a numerical solver to maximize this function with respect to the remaining choices for which we cannot find closed form solutions: $\{h_{1,t}, \ h_{2,t}, \ \tau_{1,t}, \ \tau_{2,t}, \ \tau_{12,t}\}$.

4. Finally, we evaluate the child's value function at the no-ICCT Stackelberg equilibrium to define the child's outside option:

$$V_{c,t}(\Gamma_t, a_{p,t}^0) = \lambda_1 \ln(l_{c,t}^0) + \lambda_2 \ln(x_t^0) + \lambda_3 \ln(k_t) + \beta_{c,t} \psi_{c,t+1} \ln(k_{t+1}^0)$$

**Constrained ICCT Equilibrium.** Now, we solve the household's problem if the parents are using an ICCT, summarized by the reward function $x_t(\tau_{c,t}; r_t, b_t)$. If the parents choose a strictly positive reward elasticiticy ($r_t > 0$), we know by Proposition 1 that the child's incentive compatibility constraint must be binding. Although some parents in our model might theoretically prefer to set a negative reward elasticity (see above), this is never the case for our estimates and random simulation draws. Therefore, we can abstract from those cases in the empirical implementation. We denote the vector of optimal parental time, budget and ICCT choices in this equilibrium by $\{a^*_{p,t}, r^*_t, b^*_t\}$. In the constrained optimum, we no longer have closed form solutions for any of the parental choices. To simplify the numerical solution, we first find the unconstrained parents' optimum, which is almost identical to the no-ICCT equilibrium, except now (1) the parents choose $r^{unc}_t = \frac{\Delta_{p,t} - \varphi\Delta_{c,t}}{\lambda_2\varphi} > 0$, and (2) consequently, the child studies more. Using this as an initial guess, we then solve the constrained problem as follows:

1. Condition on a choice vector of $\{h_{1,t}, h_{2,t}, \tau_{1,t}, \tau_{2,t}, \tau_{12,t}, e_t, r_t\}$, including potential corners for labor supply, and restricting $r_t$ to be strictly positive. Given these values, the household income in period $t$ is $Y_t = w_{1,t}h_{1,t} + w_{2,t}h_{2,t} + I_t$. Total parental time is defined as $\tau_{p,t} = \tau_{1,t} + \tau_{2,t} + \tau_{12,t}$.

2. The optimal reaction of the child is given by
$$\begin{aligned}\tau^*_{c,t}(\tau_{p,t}, r_t) &= \frac{\lambda_2 r_t + \Delta_{c,t}}{\lambda_1 + \lambda_2 r_t + \Delta_{c,t}}(\tilde{T}_t - \tau_{p,t}) \\ &= \gamma_t(r_t)(\tilde{T}_t - \tau_{p,t})\end{aligned}$$

3. By using the time constraints, we find each individual's leisure $(l^*_{1,t}, l^*_{2,t}, l^*_{c,t})$. Since we know all the inputs $\{\tau_{1,t},\ \tau_{2,t},\ \tau_{12,t},\ e_t,\ \tau_{c,t},\ k_t\}$, we can also find future child quality, $k^*_{t+1}$. This allows us to invert the child's binding IC constraint, to find the amount of child consumption needed to make the child indifferent:
$$\ln(x^*_t) = \frac{1}{\lambda_2}\Big(V_{c,t}(\Gamma_t, a^0_{p,t}) - \lambda_1\ln(l^*_{c,t}) - \lambda_3\ln(k_t) - \beta_{c,t}\psi_{c,t+1}\ln(k^*_{t+1})\Big)$$
Finally, parental consumption $c^*_t$ follows from the budget constraint, and $b^*_t$ fan be backed out from the ICCT reward function:
$$b^*_t = \ln(x^*_t) - r_t\ln(\tau^*_{c,t})$$

4. The parents' value (not including the ICCT cost) can then be defined as:
$$V_{p,t}(\Gamma_t, a^*_{p,t}, r_t, b^*_t) = \tilde{u}_p(l^*_{1,t}, l^*_{2,t}, c^*_t, k_t, l^*_{c,t}, x^*_t) + \beta_p\psi_{p,t+1}\ln(k^*_{t+1})$$
We use a numerical solver to to maximize this function with respect to the remaining choices for which we cannot find closed form solutions: $\{h_{1,t},\ h_{2,t},\ \tau_{1,t},\ \tau_{2,t},\ \tau_{12,t}, e_t, r_t\}$.

5. By construction, the child is indifferent between the two equilibria. The parents will implement the ICCT equilibrium if and only if
$$V_{p,t}(\Gamma_t, a^*_{p,t}, r^*_t, b^*_t) - \omega_t \geq V_{p,t}(\Gamma_t, a^0_{p,t})$$
where $\omega_t$ is the per-period utility cost of implementing the ICCT.

Appendix C contains more details on the estimation procedure.

# B Data Appendix

## B.1 Sample criteria

All our results are based on a selected sample of households that satisfy the following criteria:
(1) All households are intact over the observed period (i.e. only stable two-parent households).
(2) Households have either one or two children. We select only one child from each household (see below).
(3) All children are biological; no adopted children, no step-parents.
(4) All selected children are at least three years old in 1997, because we need a valid initial Letter Word (LW) score observation.
(5) All selected children have an observed LW score in 1997 and in 2002. Some of these also have an observed LW score in 2007 as well, although it is not required.
(6) If a household has two eligible siblings satisfying requirements (4) and (5), we select the youngest sibling by default. This has two potential advantages: parental labor supply is probably more responsive to the age of the youngest sibling than the age of the oldest sibling, and we also have a higher chance of observing the youngest sibling in 2007, which enriches the total sample.
(7) We only keep data rows for which the selected child's age is between 0 and 16.

This sample selection approach results in a final sample of $N = 247$ children or households. We have exactly 17 data rows per child, and we load the following variables after cleaning the data in Stata (not all of which are used in the code): (1) household identifier, (2) year, (3) number of child, (4) mother's age, (5) father's age, (6) family size, (7) mother's education, (8) mother's weekly labor, (9) mother's hourly wage, (10) father's weekly labor, (11) father's hourly wage, (12) weekly non-labor income, (13) child's age, (14) Letter Word raw score, (15) father's education, (16) joint parental active time, (17) mother's active time, (18) father's active time, (19) total school time, (20) regular school time, (21) other school time, (22) child's effective time endowment, (23) child's age in 1997.

## B.2 Censoring and truncation

**Actual data.** Obvious reporting errors in the parental wage and labor supply data were resolved in the following way. For a given spouse in a given year, we replace the reported labor income and labor supply by missing values if (1) the reported labor income is positive but the reported labor hours are 0, (2) if the reported labor hours are positive but the labor income is 0, or (3) if either reported labor hours or labor income is missing.

If the non-labor income in any given year (calculated as the residual yearly income after subtracting both spouses' labor income) was either negative or above 1000 dollars per week, we replace all the corresponding hourly wage, labor hours and non-labor income data by missing values for that year.

If the labor supply for a given spouse was above 80 hours per week, we truncate that observation at 80. If an hourly wage rate for a given spouse was either less than $5 per hour or more than $150 per hour, we replace that observation by a missing value. However, we keep all the other information pertaining to that household.

**Simulated data.** All simulated data are being censored in exactly the same way as the original data. Hence, if the original data contain a missing value or a censored observation for some variable at some child age, then the simulated data will have a missing value in the corresponding cell (i.e. in all $R$ corresponding cells, since we simulate $R > 1$ data sets). Similarly, whenever the simulations yields a corner solution for labor supply, we censor the corresponding simulated wage. However, we do not censor extreme simulated wage draws (i.e. below $5 or above $150 per hour).

Given our estimation procedure for the non-labor income process, simulated non-labor income draws cannot be negative. In the event that they exceed $1000 per week, we truncate that draw at $1000. Note that we cannot replace these extreme draws by a missing value (as we did for the actual data), since we always need a real-numbered (non-missing) value of non-labor income to simulate household choices in each period.

## B.3 School time

We believe the reported school time data from the CDS to be relatively noisy, as can be seen in Table B-1, which shows the distribution of reported school time at each child age $t$. Given the implausibly wide data range of these reported school times, we only use the median of these reported values (conditional on child age $t$), and use that as a measure to define the child's effective time endowment at age $t$ as $T_{c,t} = 112 - med(s_t)$. To construct school time $s_t$, we use combined CDS data from 1997, 2002 and 2007, and define total school time as the sum of "regular" school time and "other" school time. These two subcomponents were constructed based on the following CDS time categories:

1. Regular school time: All time use with activity code
    - 5090: Student (full-time); attending classes; school if full-time student.
    - 5091: Daycare/nursery school for children not in school.
    - 5092-5093: School field trips inside/outside of regular school hours.
2. Other school time: all activities taking place at school with activity code
    - 5190-5193: Other classes, courses, lectures, being tutored.
    - 5680: Daycare/nursery before or after school only.

- 6130-6138: Attending a before or after school club (math, science, drama, debate, band, ...).

Detailed descriptive statistics of these schooling components are available upon request. Finally, we note that time spent with babysitters, time spent at daycare before or after school, or time spent in home care from a non-household member (CDS activity code 4870) is not counted as school time.

Table B-1: Total School Time $s_t$ by Child Age

| | Mean | Std. | Min | P25 | **Median** | P75 | Max | NrZeros | NrObs |
|---|---|---|---|---|---|---|---|---|---|
| $t = 3$ | 11.250 | 17.866 | 0.000 | 0.000 | 0.000 | 26.042 | 47.083 | 6 | 9 |
| $t = 4$ | 9.845 | 16.152 | 0.000 | 0.000 | 0.000 | 15.833 | 55.000 | 23 | 36 |
| $t = 5$ | 13.725 | 16.830 | 0.000 | 0.000 | 0.000 | 29.583 | 56.250 | 21 | 40 |
| $t = 6$ | 24.534 | 17.404 | 0.000 | 0.000 | 32.500 | 37.083 | 47.083 | 9 | 34 |
| $t = 7$ | 31.739 | 9.956 | 0.000 | 32.500 | 33.333 | 35.000 | 45.417 | 1 | 23 |
| $t = 8$ | 28.274 | 14.201 | 0.000 | 30.833 | 33.458 | 35.000 | 48.333 | 6 | 38 |
| $t = 9$ | 31.842 | 12.055 | 0.000 | 30.833 | 34.167 | 37.812 | 50.000 | 5 | 59 |
| $t = 10$ | 31.719 | 10.948 | 0.000 | 31.667 | 33.750 | 36.250 | 47.500 | 4 | 62 |
| $t = 11$ | 30.629 | 14.027 | 0.000 | 30.771 | 34.583 | 38.750 | 56.833 | 6 | 53 |
| $t = 12$ | 32.826 | 16.841 | 0.000 | 33.333 | 38.333 | 42.500 | 53.750 | 4 | 22 |
| $t = 13$ | 31.948 | 12.409 | 0.000 | 32.500 | 34.583 | 37.604 | 45.833 | 4 | 37 |
| $t = 14$ | 37.560 | 14.888 | 0.000 | 35.000 | 37.125 | 43.750 | 74.833 | 5 | 54 |
| $t = 15$ | 35.908 | 15.003 | 0.000 | 34.375 | 37.500 | 43.750 | 60.833 | 6 | 56 |
| $t = 16$ | 33.638 | 17.220 | 0.000 | 31.042 | 37.917 | 45.000 | 65.000 | 8 | 49 |

# C Estimation, Identification, and Computation Details

## C.1 Computation of Model Solution

Given some vector of model parameters, we next describe the model solution algorithm for a given household $i$ in the dataset. For each household there is a vector of observable characteristics $X$, including parental age (at birth) and parental education levels. In addition, we observe a measure of the child's cognitive skills at some child initial age (where the age of the initial test score observation can vary across children).

For each household in the dataset, and starting at the initial child age, we draw $r = 1, \ldots, R$ wage offer and non-labor income shocks, test score measure shocks, preferences. For each simulation draw, the model solution takes the following steps:

1. Solve for the latent cognitive skills given the draw.

2. Parental labor supply falls into 1 of 4 possible cases:

   (a) $h_{1,t} > 0, h_{2,t} > 0$

   (b) $h_{1,t} = 0, h_{2,t} > 0$

   (c) $h_{1,t} > 0, h_{2,t} = 0$

   (d) $h_{1,t} = 0, h_{2,t} = 0$

   For each of the four labor supply cases, we numerically solve the optimal time allocation vector $(h_{1,t}, h_{2,t}, \tau_{1,t}, \tau_{2,t}, \tau_{12,t})$ and, for the ICCT model, also for $(e_t, r_t)$. For the case where both mother and father have positive labor hours, there are 5 free choice variables in the no-ICCT model, and 7 choice variables in the ICCT model. We use the Newton-Raphson algorithm to solve for the utility maximizing choices. We constrain each choice appropriately using the logit transformaton:

$$q_i = \frac{exp(p_i)}{1 + exp(p_i)} \in (0, 1),$$

   and search over the $p_i \in (-\infty, \infty)$ parameters for $i = 1, \ldots, 7$.

   For each $q_i$ point, we define the choice variables sequentially as

   (a) Total parental investment time $\tau_{p,t} = q_1(\tilde{T}_t - s_t)$.

   (b) Mother's active time $\tau_{1,t} = q_2 \tau_{p,t}$.

   (c) Father's active time $\tau_{2,t} = q_3(\tau_{p,t} - \tau_{1,t})$.

   (d) Mother's labor time $h_{1,t} = q_4(T - \tau_{1,t} - \tau_{12,t})$

   (e) Father's labor time $h_{2,t} = q_5(T - \tau_{2,t} - \tau_{12,t})$

(f) For the ICCT model: child expenditures $e_t = q_6 Y_t$, where $Y_t = w_{1,t}h_{1,t} + w_{2,t}h_{2,t} + I_t$. For the no-ICCT model, we have a closed-form solution for $e_t$ as a function of $(h_{1,t}, h_{2,t})$.

(g) For the ICCT model: reward elasticity $r_t = q_7 r_{max}$, where $r_{max} = 20$. For the no-ICCT model, we set $r_t = 0$.

This ensures that all parental time choices (including joint parental time $\tau_{12,t} = \tau_{p,t} - \tau_{1,t} - \tau_{2,t}$) are strictly positive and satisfy the time and budget constraints. For each time allocation choice, we compute $\tau_{c,t}(\tau_{p,t}, r_t)$ using the child's reaction function.

3. In the no-ICCT model, we find $x_t$, $c_t$ and $e_t$ using the closed form solutions derived above. After defining $k_{t+1}$, we can define the child's outside option, $V^0_{c,t}(\Gamma_t | a^0_{p,t})$. In the constrained ICCT model, we numerically solve for $e_t$, so we can (1) use the technology function to define $k_{t+1}$ conditional on all inputs, (2) find $x_t$ by inverting the child's binding incentive compatibility constraint conditional on the outside option (see also Appendix A), and (3) find $c_t$ through the budget constraint.

4. We solve for the utility maximizing choices for all possible labor supply cases and retain the highest utility choices for both the no-ICCT and ICCT models. In the benchmark model with endogenous costly ICCT choice, we retain those choices which maximize the parents' value (net of the ICCT cost $\omega_t$).

5. With the optimal choices computed, we use the $r$th measurement shock to compute the measure $\tilde{k}_{r,t+1}$. Then, we reiterate by updating $t$ to $t+1$ and latent capital $k_t$ to $k_{t+1}$.

## C.2 Identification

In this sub-section, we provide more details on several of the more involved identification issues.

**Production Technology: Measurement Error** In order to focus on key issues, consider a simplified version of our production technology, where $\ln k = \ln R + \delta \ln \tau$, with $k$ representing latent cognitive ability, $R$ is TFP, and $\tau$ is an observed input with associated parameter $\delta$. Consider the following conditional mean of the observed test score $k^*$, given some level of the observed input.

$$E(k^* | \tau) = NQ \frac{\exp(\lambda_0 + \lambda_1 \ln R + \lambda_1 \delta \ln \tau)}{1 + \exp(\lambda_0 + \lambda_1 \ln R + \lambda_1 \delta \ln \tau)} \tag{C-1}$$

We observe the left-hand side of this expression in the data, and the right-hand side is a function of the primitives we would like to identify.

It is clear from this expression that we cannot separately identify the production function primitives $(R, \delta)$ from the measurement parameters $(\lambda_0, \lambda_1)$. This is a generic problem of indeterminacy due to the fact that latent child quality/skill $k$ does not have any natural units. Identification requires some normalization to fix the location and scale of the latent variable. We normalize $\lambda_{0t} = 0$ and $\lambda_{1t} = 1$ for all $t$, and proceed to identify the production primitives up to this normalization.

First, consider evaluating this conditional expectation at the point $\tau = 1$, so that

$$E(k^*|\tau = 1) = NQ\frac{R}{1+R}$$

Given the number of test questions $NQ$, we identify the TFP term $R$. Next, we identify $\delta$ from the difference in mean test scores for two values of the input $\tau \in \{1, b\}$, for $b \neq 1$:

$$E(k^*|\tau = b) - E(k^*|\tau = 1) = NQ\{\frac{\exp(\ln R + \delta \ln b)}{1 + \exp(\ln R + \delta \ln b)} - \frac{R}{1+R}\}$$

We can extend this approach to any number of multiple observed inputs.

**Production Technology: Unobserved Expenditures** In our data, in contrast to the time inputs, child expenditures are not observed directly (the PSID-CDS data provides some expenditure data but is likely incomplete). To identify the productivity of the unobserved child expenditure input, we require a different identification strategy from the one we utilized for the observed time inputs. Consider two households with the same observed time inputs, but who differ in their household income (due to differences in labor or non-labor income). Given child expenditures are a normal good, this implies that the higher income household has larger expenditures on children. Expanding our simplified production function notation to include an expenditure input $e$ and observed household income $Y$, we can construct the following conditional moment of the observed test scores:

$$E(k^*|\tau, Y) = NQ\frac{\exp(\ln R + \delta_\tau \ln \tau + \delta_e E(\ln e|Y))}{1 + \exp(\ln R + \delta_\tau \ln \tau + \delta_e E(\ln e|Y))}$$

$E(\ln e|Y)$ is the expected (log) expenditure for a household of income $Y$. Building on the analysis above, comparing households with different observed incomes then allows us to identify this term $\delta_e E(\ln e|Y)$ for any $Y$ in the support of our data.

Our task is then to separately identify the productivity parameter $\delta_e$ from the unobserved average level of expenditure by income $E(\ln e|Y)$. We separately identify these two components using the model structure, in particular the restrictions implied by the budget constraint and from observed household choices. From the solution to our model, the optimal expenditure on children is given by $e = \Delta_e Y$, where $\Delta_e \in (0, 1)$ is the income share spent on children, a non-linear function of the primitive household preferences and technology. $\Delta_e$ is identified jointly with the other household parameters, with the key parameters comprising this share parameter (i.e. the household preference for consumption relative to the taste for child skills) identified from the observed household time allocation (i.e. time with children and labor supply).

**Production Technology: Latent Skill Distribution** We also need to identify the distribution of latent skills because they serve as input to the production process, not only as an output. For each child we observe measures of child quality for at least two different ages. We use the first measure of child quality as an initial condition. However, to solve the model and identify the production technology, we require an initial level of latent child quality $k_t$, not the measure $k_t^*$.

Given the measurement error assumptions, the probability of answering a question correctly $p$ is distributed according to the Beta distribution, with parameters $(1 + k_t^*, (NQ - k_t^*) + 1)$, where $k_t^*$ is the observed number of correct answers out of the $NQ = 57$ items. For any given realization of $p$ (given $k_t^*$), $p = \tilde{p}$, we then invert the normalized measurement equation (2) to obtain a realized value of latent child quality:

$$k_t = \frac{\tilde{p}}{1 - \tilde{p}}.$$

Repeatedly drawing from the Beta distribution given the observed measure then provides a simulated distribution of latent child quality values. From these initial values of $k_t$, we then begin the construction of each sample path, recursively substituting the latent $k_t$ values and other endogenous inputs determining latent $k_{t+1}$. When we get to the period of the second measurement, at which time the child is of age $t' > t$, the observed test score is a draw from a Binomial distribution with parameters $(NQ, p(k_{t'}))$, as described above.

## C.3 Estimator

For the same household $i$, this process is repeated $S$ times, so that in the end we have $S \times N$ sample paths. Using the simulated data set, we then compute the analogous simulated sample characteristics to those determined from the actual data sample. The characteristics of any simulated sample are determined by $\Omega$, the vector of all primitive parameters that characterize the model, and the actual vector of pseudo-random number draws made in generating the sample paths. Denote the simulated sample characteristics generated under the parameter vector $\Omega$ by $\tilde{M}_S(\Omega)$. The Method of Simulated Moments (MSM) estimator of $\Omega$ is then given by

$$\hat{\Omega}_{S,N,W} = \arg\min_{\Omega}(M_N - \tilde{M}_S(\Omega))' W_N (M_N - \tilde{M}_S(\Omega)),$$

where $W_N$ is a symmetric, positive-definite weighting matrix.[33] Given random sampling from the population of married households with a given number of children (one or two, in our case), we have $\text{plim}_{N\to\infty} M_N = M$. The weighting matrix, $W_N$, is simply the inverse of the covariance matrix of $M_N$, which is estimated by resampling the data.

[33]Simulation in our context is used to solve the computationally intensive integration problem. Our choice of MSM vs. an alternative simulation estimator, for example simulated maximum likelihood (SMLE) is due the greater flexibility that the MSM estimator offers in combining data from multiple sources with different sampling schemes.

We computed the $M_N^g$ vector for each of $Q$ resamples of the original $N$ data points, and the covariance matrix of $M_N$ is given by

$$W_N = \left( Q^{-1} \sum_{g=1}^{G} (M_N^g - M_N)(M_N^g - M_N)' \right)^{-1} .$$

The number of draws, $Q$, was set at 200.

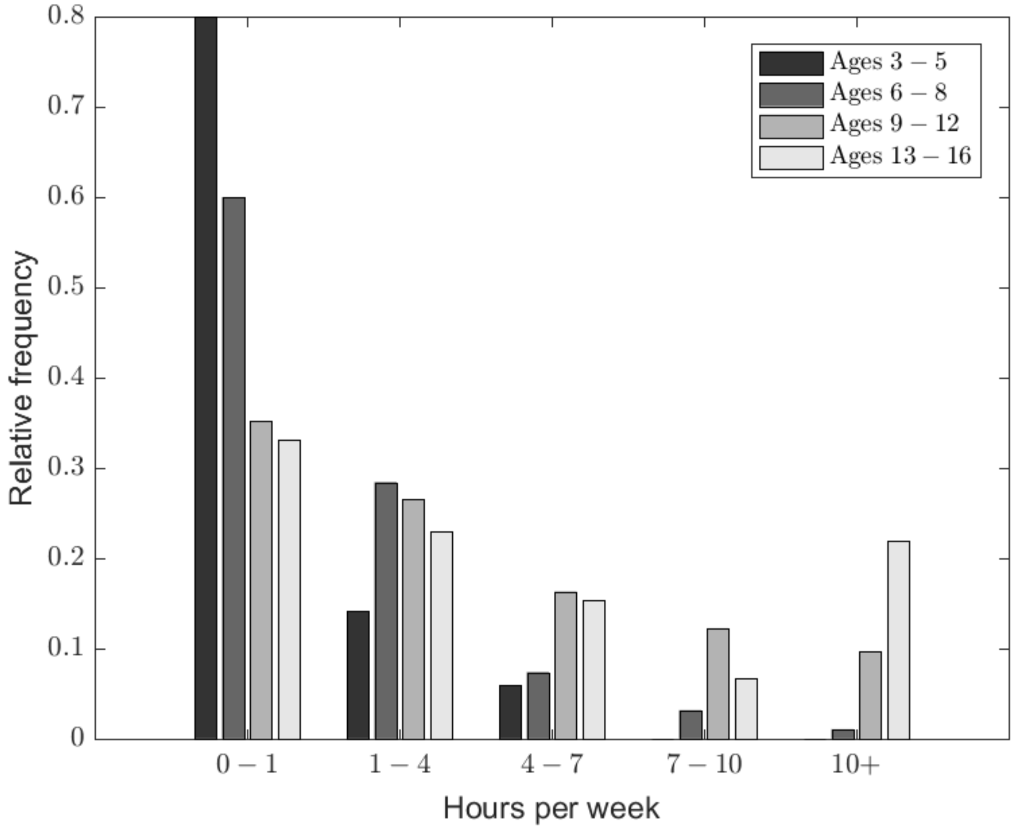
Given that the simulated moments are non-linear functions of the simulated draws so that $\tilde{M}_S$ is biased for fixed $S$, for consistency of the MSM estimator we require that $S$ also grow indefinitely large. Let the true value of the parameter vector characterizing the model be denoted by $\Omega_0$. Then $\text{plim}_{S\to\infty} \tilde{M}_{S,N}(\Omega_0) = M_N(\Omega_0)$. Given identification and these regularity conditions,

$$\underset{N\to\infty,S\to\infty}{\text{plim}} \tilde{\Omega}_{S,N,W} = \Omega \text{ for any positive definite } W.$$

Since $W_N$ is positive definite by construction, our estimator $\Omega_{S,N,W_N}$ is consistent as well. We have not utilized the asymptotically optimal weighting matrix in this case due to the computational cost and issues regarding the differentiability of the objective function given the crude simulator we use. This does not seem to be a major concern since virtually all of the parameters are precisely estimated with the exception of those which we know from our earlier discussion to be tenuously identified in a data set that is the size of ours.

# D Additional Tables and Figures

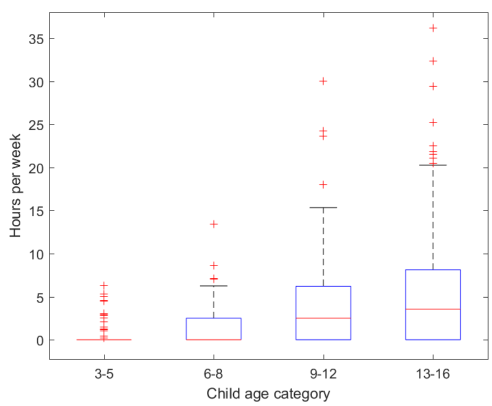
Figure D-1: Distribution of Child Self-investment Time by Age



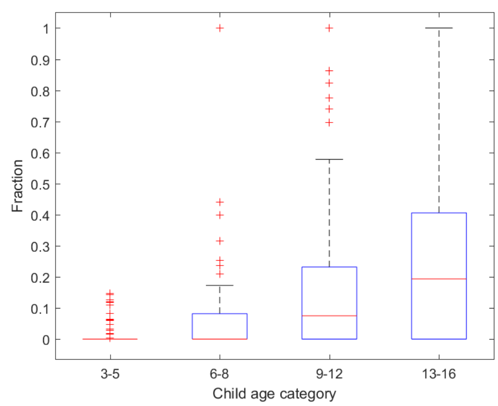

**Source:** PSID-CDS combined sample from 1997, 2002 and 2007 interviews.
**Notes:** Within each child age category, the vertical bars represent the fraction of households whose reported child self-investment time was between 0 – 1 hours, 1 – 4 hours, 4 – 7 hours, 7 – 10 hours, or more than 10 hours per week.

Figure D-2: Boxplots of Child Self-investment Time by Age

(a) Hours per week

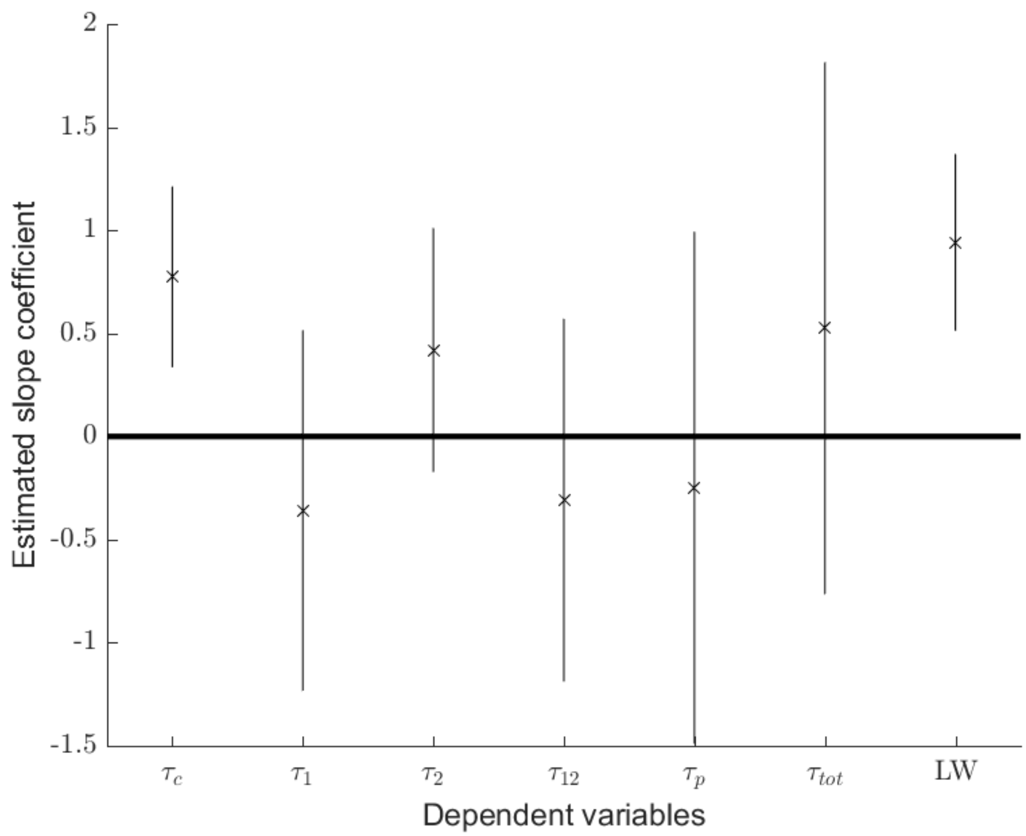



(b) Fraction of total investment time




**Source:** PSID-CDS combined sample from 1997, 2002 and 2007 interviews.
**Notes:** The left panel plots the distribution of the reported weekly child study time for each child age category. The right panel shows child study time as a fraction of total investment time, defined as the sum of child study time and all active time with either or both of the parents.

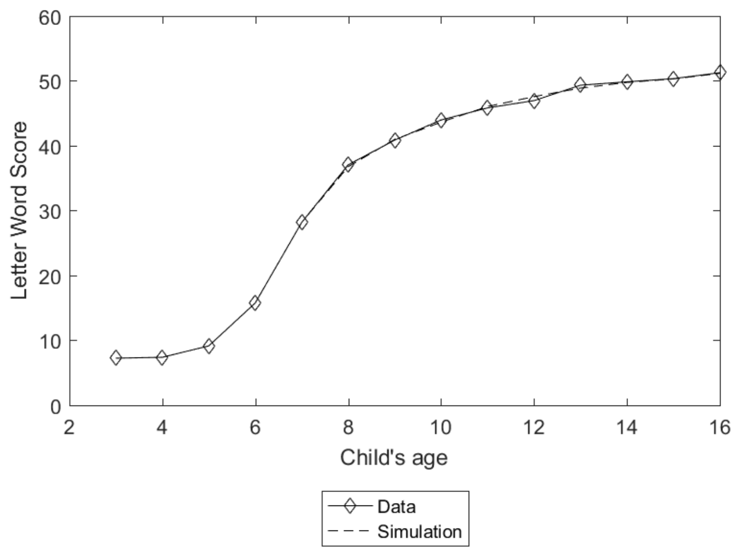
Figure D-3: The Effect of Household Income on Productive Time Inputs and Test Scores




**Source:** PSID-CDS combined sample from 1997, 2002 and 2007 interviews and PSID core data between 1986 and 2010.

**Notes:** We regress various weekly time inputs and test scores on weekly household income (in thousands of dollars, averaged across all observed years). All regressions also include child age fixed effects. We plot the estimated slope coefficients on income and their corresponding 95% confidence intervals. The dependent variables are (from left to right): (1) the child's self-investment time, $\tau_c$, (2) mother's active time, $\tau_1$, (3) father's active time, $\tau_2$, (4) joint parental time, $\tau_{12}$, (5) total parental time, $\tau_p = \tau_1 + \tau_2 + \tau_{12}$, (6) total investment time, $\tau_{tot} = \tau_c + \tau_p$ and (7) the child's raw Letter Word score, LW.

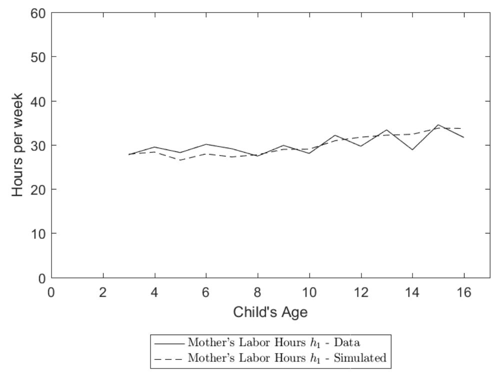
Figure D-4: Simulated and Actual Average Child's Letter Word Score



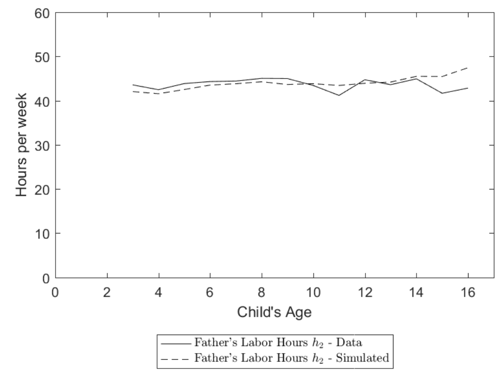

**Notes**: Data is actual data. Simulated is the model prediction at estimated parameters given above.
**Source**: PSID-CDS combined sample from 1997, 2002 and 2007 interviews and PSID core data between 1986 and 2010.

Figure D-5: Parental Labor Supply and LFP by Child Age

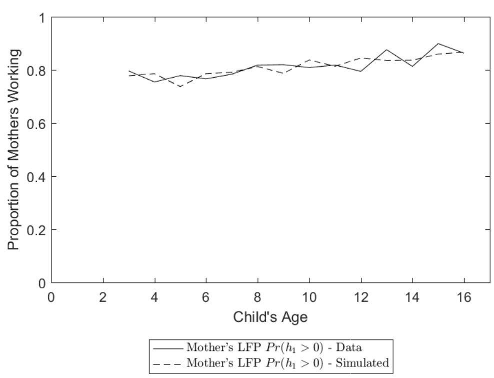
(a) Working Mother's Labor Supply




(b) Working Father's Labor Supply




(c) Mother's Labor Force Participation




(d) Father's Labor Force Participation

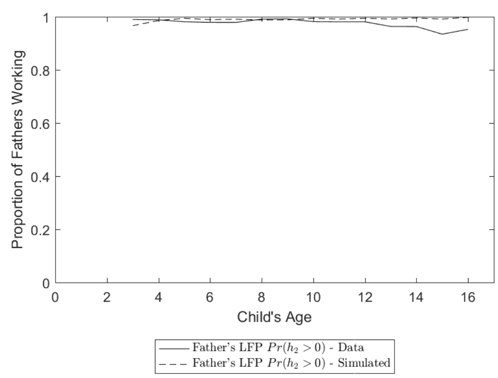
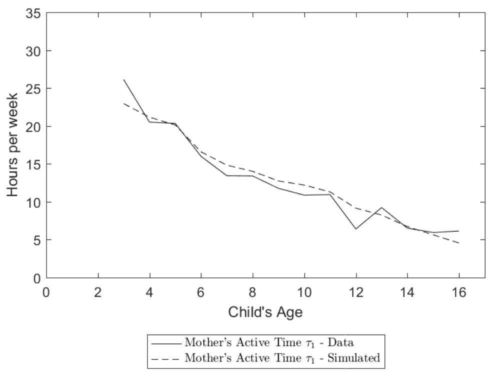
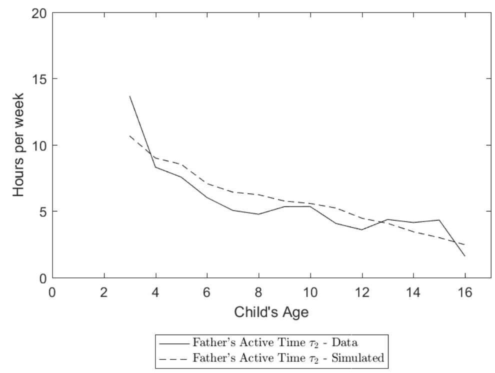


Figure D-6: Productive Time Inputs by Child Age

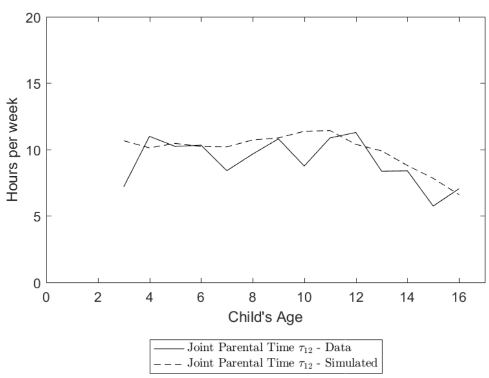
(a) Mother's Active Time




(b) Father's Active Time




(c) Joint Parental Time




(d) Child's Self-investment Time

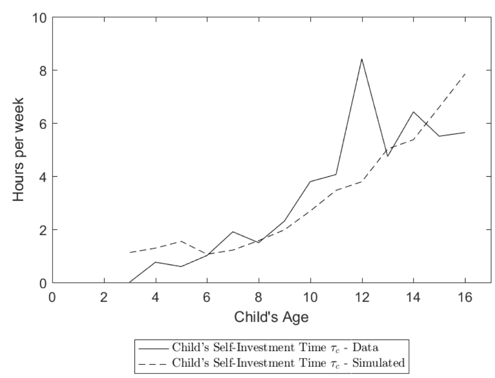
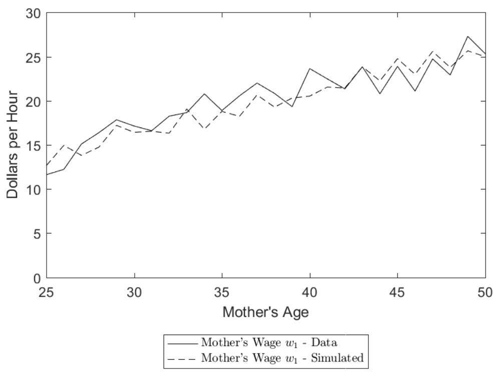
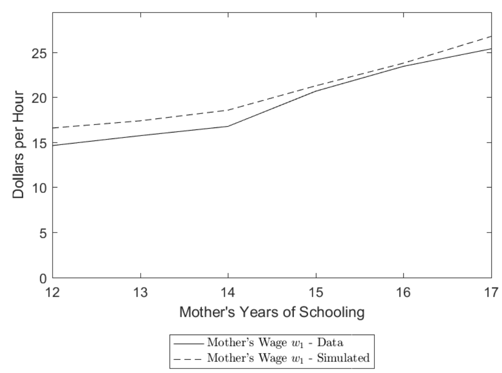


Figure D-7: Parental Hourly Wages by Parental Age and Education

(a) Mother's Hourly Wage, by Age




(b) Mother's Hourly Wage, by Education




(c) Father's Hourly Wage, by Age


(d) Father's Hourly Wage, by Education

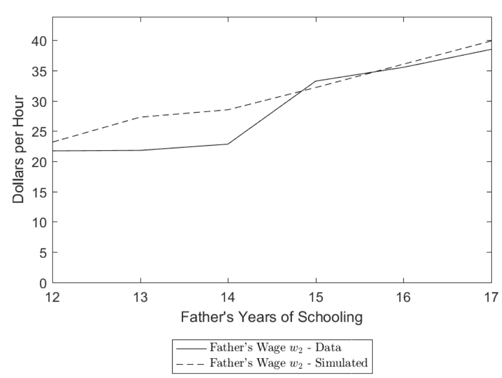
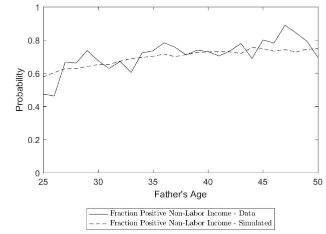


Figure D-8: Weekly Non-Labor Income by Parents' Age and Education

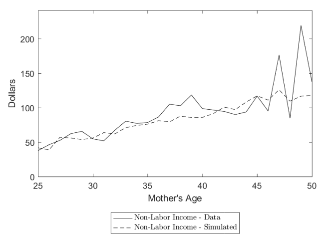
(a) All, by Father's Age


(b) Positives, by Father's Age


(c) Fraction > 0, by Father's Age

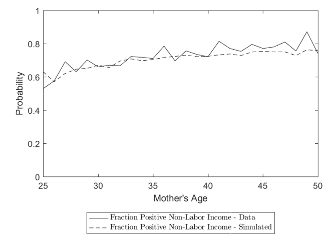



(d) All, by Mother's Age




(e) Positives, by Mother's Age

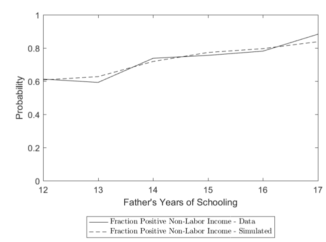

(f) Fraction > 0, by Mother's Age



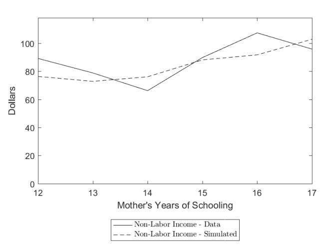

(g) All, by Father's Educ.


(h) Positives, by Father's Educ.

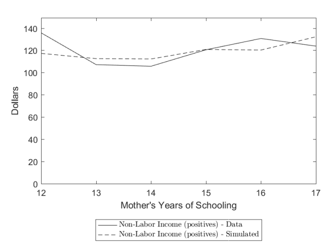

(i) Fraction > 0, by Father's Educ.




(j) All, by Mother's Educ.




(k) Positives, by Mother's Educ.




(l) Fraction >0, by Mother's Educ.

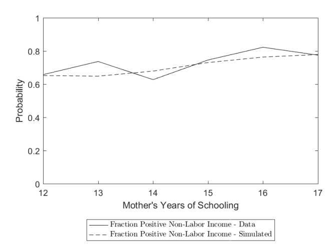
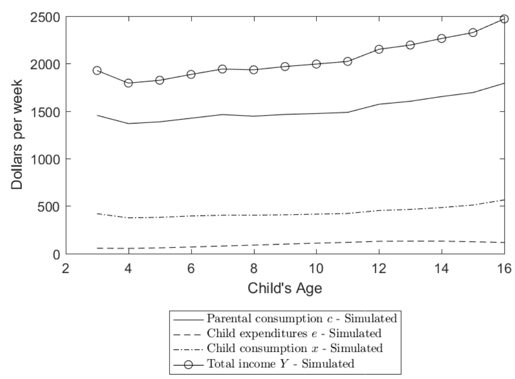
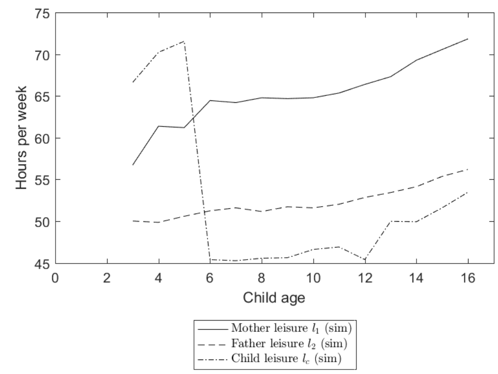


Figure D-9: Expenditures, Leisure and Internal CCT Use by Child Age

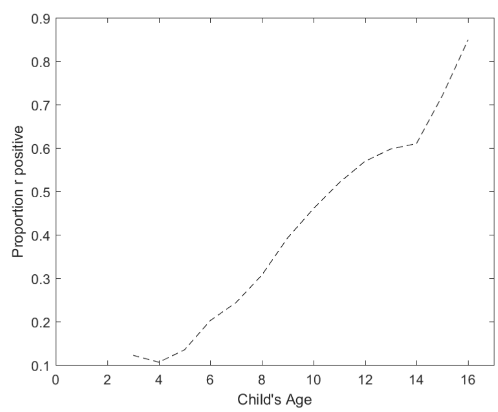
(a) Consumption, Expenditures and Income



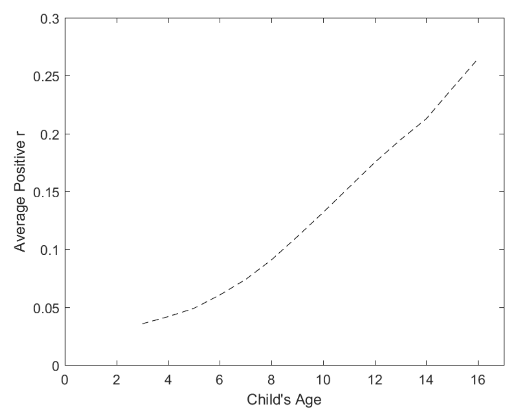

(b) Household Leisure Time



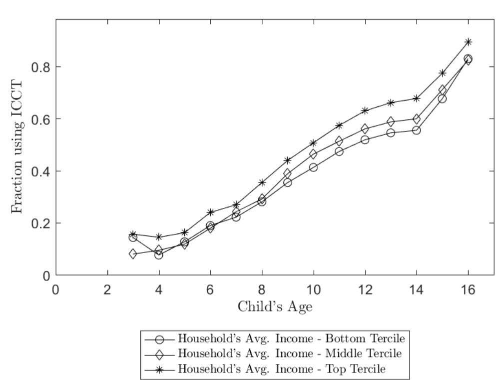

(c) Fraction of Households using Internal CCT




(d) Avg. Reward Elasticity $r$ (if $r > 0$)




(e) Fraction using ICCT by Income




(f) Avg. Reward Elasticity by Income

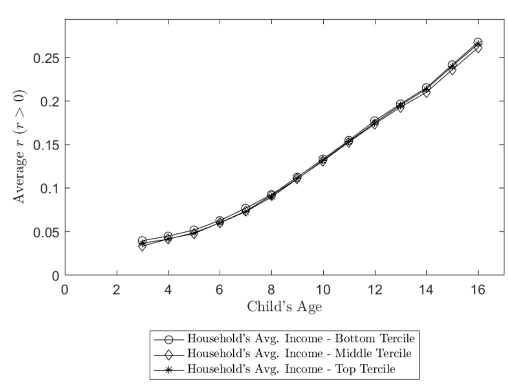


Table D-1: Data Correlations

| **Child Ages** | **9-12** | **13-16** |
|---|---|---|
| Letter Word Score, Child time | 0.264 | 0.128 |
| | (0.069) | (0.071) |
| Letter Word Score, Mother's Educ. | 0.245 | 0.265 |
| | (0.068) | (0.068) |
| Letter Word Score, Father's Educ. | 0.301 | 0.342 |
| | (0.067) | (0.066) |
| Letter Word Score, HH Income | 0.325 | 0.287 |
| | (0.076) | (0.077) |
| Child time, Mother's Educ. | 0.078 | 0.160 |
| | (0.072) | (0.071) |
| Child time, Father's Educ. | 0.095 | 0.283 |
| | (0.071) | (0.069) |
| Child time, HH Income | 0.145 | 0.280 |
| | (0.082) | (0.079) |

**Source**: PSID-CDS combined sample from 1997, 2002 and 2007 interviews and PSID core data between 1986 and 2010. To alleviate the missing data problem at young child ages, "HH income" is defined as the average total household income within each relevant child age bin. Standard Errors of the correlations are between brackets, and are defined as $SE_r = \sqrt{\frac{1-r^2}{n-2}}$.

Table D-2: Technology Parameter Estimates

| | | **Estimate** | **SE** |
|---|---|---|---|
| Mother's Active Time ($\delta_1$) | Intercept $\gamma_{1,0}$ | -0.244 | (0.03502) |
| | Slope $\gamma_{1,1}$ | -0.252 | (0.00683) |
| | Mother's Educ. $\gamma_{1,2}$ | -0.001 | (0.00007) |
| Father's Active Time ($\delta_2$) | Intercept $\gamma_{2,0}$ | -1.662 | (0.04973) |
| | Slope $\gamma_{2,1}$ | -0.239 | (0.00846) |
| | Father's Educ. $\gamma_{2,2}$ | 0.042 | (0.00276) |
| Joint Parental Time ($\delta_3$) | Intercept $\gamma_{3,0}$ | -1.259 | (0.06380) |
| | Slope $\gamma_{3,1}$ | -0.133 | (0.00127) |
| | Mother's Educ. $\gamma_{3,2}$ | 0.020 | (0.00219) |
| | Father's Educ. $\gamma_{3,3}$ | 0.018 | (0.00100) |
| Child Expenditures ($\delta_4$) | Intercept $\gamma_{4,0}$ | -4.219 | (0.17291) |
| | Slope $\gamma_{4,1}$ | -0.053 | (0.00118) |
| Child's Self-Investment Time ($\delta_5$) | Intercept $\gamma_{5,0}$ | -7.930 | (0.13529) |
| | Slope $\gamma_{5,1}$ | 0.249 | (0.00942) |
| Last Period's Child Quality ($\delta_6$) | Intercept $\gamma_{6,0}$ | -1.644 | (0.01502) |
| | Slope $\gamma_{6,1}$ | 0.264 | (0.00170) |
| Total Factor Productivity ($R_t$) | $\gamma_{7,0}$ | 0.47365 | (0.00677) |
| | $\gamma_{7,1}$ | 1.01128 | (0.00414) |
| | $\gamma_{7,2}$ | 1.44493 | (0.13486) |
| | $\gamma_{7,3}$ | 8.24483 | (0.10487) |

**Notes**: Productivity parameters take the form $\delta_{i,t} = 0.01 + 0.99\frac{\exp(\gamma_{i,0}+\gamma_{i,1}(t-1))}{\exp(\gamma_{i,0}+\gamma_{i,1}(t-1))}$, for all $i = 1, ..., 6$ and $t = 1, ..., 16$. Total Factor Productivity parameters take the form $R_t = \gamma_{7,0} + \frac{\gamma_{7,1}-\gamma_{7,0}}{1+\exp(-\gamma_{7,2}(t-\gamma_{7,3}))}$. SEs are standard errors computed using a cluster bootstrap sampling each household with replacement.