Intergenerational Long-Term Effects of Preschool - Structural Estimates from a Discrete Dynamic Programming Model*

James J. Heckman
and
Lakshmi K. Raut†

Abstract

This paper formulates a structural dynamic programming model of preschool investment choices of altruistic parents and then empirically estimates the structural parameters of the model using the NLSY79 data. The paper finds that preschool investment significantly boosts cognitive and non-cognitive skills, which enhance earnings and school outcomes. It also finds that a standard Mincer earnings function, by omitting measures of non-cognitive skills on the right-hand side, overestimates the rate of return to schooling. From the estimated equilibrium Markov process, the paper studies the nature of within generation earnings distribution, intergenerational earnings mobility, and schooling mobility. The paper finds that a tax-financed free preschool program for the children of poor socioeconomic status generates positive net gains to the society in terms of average earnings, higher intergenerational earnings mobility, and schooling mobility.

Keywords: Preschool Investment, Early Childhood Development, Intergenerational Social Mobility, Structural Dynamic Programming
JEL Classification Nos.: J24, J62, O15, I21

Submitted Draft: December 2009
Revised Draft: October 1, 2015

*We would like to thank the anonymous Associate Editor and two referees of the *Journal of Econometrics* for many valuable comments. An earlier draft was presented at the Centre for Development Studies, Institute of Economic Growth, Indian Statistical Institute, Indira Gandhi Institute of Development Research, Center for Development Studies, Nanyang Technological University, Periyar University, Singapore National University, University of Nevada at Las Vegas, Tokyo University, University of Southern California, the Western Economic Association Meeting 2003, and the Public Economic Theory (PET) 2006, Hanoi, Vietnam. Comments of the participants of these workshops, especially of Juan Pantano as a discussant of the Western Economic Association conference, Lien H. Tran for presenting and commenting on the paper at the PET 2006 conference, and comments from Han A.T. Raut and T.N. Srinivasan are gratefully acknowledged. This research was supported in part by the American Bar Foundation, the Pritzker Children's Initiative, the Buffett Early Childhood Fund, NICHD R37 HD065072, R01 HD054702, the Human Capital and Economic Opportunity Global Working Group - an initiative of the Becker Friedman Institute for Research in Economics - funded by the Institute for New Economic Thinking (INET), and an anonymous funder. The views expressed in this paper are those of the authors and not necessarily those of the funders or commentators mentioned here.

†**Corresponding author**. Raut is an Economist at the Social Security Administration (SSA). This paper was prepared prior to his joining SSA, and the analysis and conclusions expressed are those of the authors and not necessarily those of SSA.

James J. Heckman
Department of Economics
University of Chicago
1126 E. 59th Street
Chicago, IL 60637
jjh@uchicago.edu
(773) 702-3478

Lakshmi K. Raut
Social Security Administration
400 Virginia Avenue, SW, Suite 300
Washington, DC 20024
Lakshmi.Raut@ssa.gov
(202)358-6513

1 Introduction

This paper formulates and estimates an altruistic model of parental preschool investment decisions. In our model, preschool investments affect the cognitive and non-cognitive skills of the children, and hence their lifetime permanent earnings and school outcomes. Optimal choices by parents determine the equilibrium controlled Markov process, characterizing the equilibrium dynamics of earnings distributions within each generation, and the schooling and earnings mobility across generations. We also examine the effect of a social policy that provides free preschool to children of low socioeconomic status (SES) financed by taxing all parents in the population, on the distribution of earnings within generation and on intergenerational earnings and schooling mobility. We use the NLSY79 (National Longitudinal Survey of Youth, 1979) and the NLSY79 Children and Young Adults data containing information on a nationally representative sample of parent-child pairs of the US population. This paper extends Raut (2003) by incorporating unobserved heterogeneity and estimating the structural parameters. The paper utilizes the Rust (1987) nested fixed point maximum likelihood estimation procedure.

Two important building blocks of our model are: (1) The stochastic production processes of the cognitive and non-cognitive skills with early childhood investment as one of the inputs; (2) An augmented Mincer earnings function that adds non-cognitive skills to the standard Mincer earnings function. We estimate these relationships. We provide an estimate of the extent to which the rate of return to schooling in the standard Mincer earnings function is inflated because the schooling level in the standard Mincer earnings function embodies the effect of the omitted non-cognitive skills variables.

In the past three decades, the income gap between the rich and the poor and the wage gap between the college educated and the non-college educated workers in the US have been widening. Equalizing education is advocated as the main policy in the US to reduce poverty and income disparities. Many are, however, highly skeptical about a positive answer to the basic question: *"Can we conquer poverty through school?"*

There are many reasons for this skepticism. In the US, education through high school level is virtually free. Yet many children of poor SES (Socio Economic Status) do not complete high school and many of them perform poorly in schools. Gaps in test scores between rich and poor children are substantial, and unequal schooling does little to widen this gap (Carneiro and Heckman, 2003; Heckman, 2008). In spite of its positive effects on test scores and earnings, the effects of improved school quality on school dropout rates is marginal.

A growing consensus reached among educators, media writers (see for instance Traub, 2000), researchers in economics (see, for instance, Cameron and Heckman, 1998; Carneiro and Heckman, 2003; Cunha et al., 2006; Heckman, 2000, 2008; Keane and Wolpin, 1997, 2001) finds that children of poor SES are not prepared for college because they were not prepared for school to begin with. The most effective intervention for the children of poor SES should be introduced at the preschool stage so that these children are prepared for school and college. The question is, then: does preschool experience have long-term positive effects on school performance and labor market success? This is the main issue that we address in this paper, and our finding corroborates the evidence in Cameron and Heckman (1998); Cunha and Heckman (2007, 2009); Heckman et al. (2010a,b); Keane and Wolpin (1997, 2001) that early intervention is effective.

There are quite a few quantitative studies on this issue. One set of studies uses data on high cost high quality pilot preschool programs such as the High Scope/Perry Preschool Program (see Heckman et al., 2010a,b) and the North Carolina Abecedarian Study (Campbell et al., 2012). These studies find a substantial lasting effect of these programs on school performance and labor market outcomes. The participants in these programs are not representative of the US population.

Another set of studies estimates the production function for children's cognitive achievements, which is usually measured by scores in math and reading tests in early childhood.[1] Most of these studies do not explicitly examine the effect of the mother's employment or types of childcare on cognitive and non-cognitive skill formation of children. Blau (1999), however, uses the childcare data on the nationally representative full NLSY79 sample of parents, matched with the NLSY79 Children data. He finds that the childcare investments during the first three years have no significant effect, but an experience with better quality childcare during the next two years has a significant positive effect on the cognitive developments of children in early school years. Other studies (see Blau and Currie, 2006) find negligible or negative effects of maternal employment on child outcomes. When a mother works, maternal time input for child development is reduced, which may yield a negative effect. This negative effect might be offset by the positive effects of higher income and better quality childcare on child outcomes, yielding a net small or negative effect of maternal employment. Similarly, the negligible effect of childcare may be because the mothers may use childcare to be able to work, which reduces mother's time input on child development, offsetting the positive effect of childcare on child outcomes. The problem is that childcare and maternal employment are endogenous variables. The regression models that treat these variables as exogenous regressors will produce biased estimates of their effects on child outcomes. Bernal (2008) and Bernal and Keane (2011) formulate and estimate structural models in which these two are choice variables. Using the same dataset as in Blau (1999), they find significant negative effects of maternal employment and informal childcare (i.e., care by relatives) on test scores of children. These studies do not distinguish between preschool and daycare centers of various

[1] See Todd and Wolpin (2003) and Blau and Currie (2006) for earlier surveys and summaries of these studies, and Cunha and Heckman (2008); Todd and Wolpin (2007) for more recent studies and recent references.

qualities that the respondent uses. The results are for the restricted groups in the sample of single mothers (or mothers that do not cohabit with a male) during the first five years of the child's life or for mothers who live with the husband/male-partner during the first five years of the child's life. In both cases, the mothers do not have another child for at least five years. See Blau and Currie (2006) for a summary of similar findings on various other subgroups.[2]

The other set of studies uses data on the Head Start preschool program which is funded by the Federal Government. It is available to children whose parents earn incomes below the poverty line. Not all eligible children are, however, covered by the program. The quality of the program is very poor compared to the enriched pilot programs or most private preschool programs. Some studies find that the Head Start Preschool Program has no long-term effect on children's cognitive achievements and school performance, especially for black children. Currie and Thomas (1995) carry out a careful econometric investigation and conclude that the benefits disappear for black children because most of the Head Start black children attend low-quality public schools. But after controlling for school quality, they find significant positive effects of the Head Start Preschool Program. Studying two types of preschool is beyond the scope of this study; see the recent study by Deming (2009).

The above studies are not based on nationally representative samples of children. Most studies examine only the effect on school performance, such as test performance in early school years, grade retention, and high school and college graduation rates, and do not model parental choice of investing in their children's

[2]"The most consistent evidence of negative effects of maternal employment comes from families in which some or all of the following are true: the mother returns to work when the child is less than one year old; young children spend very long hours in care; the mother's employment does not raise family income (as in some households where families have been forced off welfare); there is a single parent with few family members to draw on so that time spent in employment cannot be compensated by drawing on the time of other family members either for child care or for housework; and/or the work itself is very stressful and reduces the resources the mother brings to parenting. Some studies of shift-work, for example, suggest that it may have this effect. Adolescents may also suffer more negative effects of maternal employment than younger children, particularly if they are left unsupervised." (Blau and Currie, 2006, pp. 1170-1171)

preschool. In this paper, we formulate a model of parental investment in preschool that is guided by economic incentives. We show that the preschool experience benefits children in acquiring many useful cognitive and non-cognitive skills, especially for the children of poor SES who live in poor home environments—a measure of family investment. We also show the importance of non-cognitive skills in improving school performance and life-time earnings of children, after controlling for their education level, innate ability, and family background. See Raut (2003) for earlier estimates of the effects of cognitive and non-cognitive skills on school performance and earnings along the line of this paper. Almlund et al. (2011); Duckworth, Almlund, Kautz, and Heckman (Duckworth et al.); Heckman and Kautz (2012) summarize the literature on the effects of non-cognitive traits on earnings.

The rest of the paper is organized as follows. Section 2 describes the intergenerational altruism model of parental preschool investment within a structural dynamic programming framework. Section 3 describes the estimation algorithm that we use. Section 4 provides the empirical specification of the production processes of various skills and reports the parameter estimates. Section 5 conducts a policy analysis. Section 6 concludes.

2 The Basic Framework

In this section we formulate an econometrically implementable model of preschool investment of altruistic parents in a structural dynamic programming framework. We describe how we compute the long-run equilibrium distributions of earnings and schooling within generations. A transition probability matrix of earnings or schooling levels provides information about the degree of intergenerational earnings or schooling mobility or if there is an intergenerational poverty cycle. We explain how we compute the mobility index from a transition matrix and how we compute the long-run equilibrium tax rate to finance free preschool for children of poor SES, and the net-gains or losses from introduction of such a free preschool program to the society in terms of welfare gains and losses of various groups, and in terms of change in the per capita disposable (i.e., after tax) earnings in the long-run equilibrium.

We assume a parthenogenetic mode of biological reproduction in our model and with due respect to both genders, all individuals are male gender. Parents of period t will be referred to as generation t. Each parent has one child. After parents of generation t die at the end of period t, their children become the parents of generation $t+1$ and make decisions for their children. The economy goes on in this recursive manner forever.

In each period, parents are characterized by a vector of observed characteristics x, and a vector of unobserved characteristics ε, which are described in detail below. We summarize these traits by a vector $z = (x, \varepsilon)$. When we need to be specific about his generation t or period t, we write him as $z_t = (x_t, \varepsilon_t)$. We assume that each component of x takes a finite number of values, thus x will be from a finite set $\mathcal{X}$ with m elements. We assume that the set $\mathcal{X}$ is ordered with elements $x_1, ..., x_m$ in it. For a parent-child pair, if v is a variable that refers to the parent, we use v' to denote the corresponding variable for his child.

An individual's lifetime consists of several stages during which important life-cycle events relevant to learning and earning occur. A parent invests in his child's preschool activities during ages [0-5), which help develop his child's school readiness and various cognitive and non-cognitive skills. Denote by a the preschool investment choice of a parent. At the end of the preschool period, the child acquires levels of cognitive skill τ, social skill σ, motivational skill μ, self-esteem skill η, and internal self-control skill ϕ.

During ages [5-17), the child goes to school. School performance at this stage depends on his level of τ,

σ, μ, η, and ϕ that the child has acquired during the previous stage. The school performance also depends on many other variables such as the quality of the school that he attends,[3] the quality of the neighborhood, and the parental home inputs.[4] [5]

During ages [17-26), the child decides the number of years of schooling to complete, which depends on his parent's family background, his own cognitive and non-cognitive abilities τ, σ, μ, η, and ϕ, and some random shocks ε_s. We denote its dependence on these factors by the function $s = s(\tau, \sigma, \mu, \eta, \phi, s, \varepsilon_s)$.[6]

During ages [26-], he works, forms his family, procreates one child, and chooses a preschool investment plan for his child. In Section 4.1, we describe in detail the components of the observed characteristics vector $x = (\tau, \sigma, \mu, \eta, \phi, s)$. In this section, we sequentially define the components of the vector of his unobservable characteristics ε.

The production sector of the economy uses a linear production function with labor (measured in efficiency units) as the only input. An individual with observed cognitive and non-cognitive skills $x = (\tau, \sigma, \mu, \eta, \phi, s)$ is assumed to be equivalent to the unit of labor in efficiency units $w(x) + \varepsilon_1$, where ε_1, given x, is a mean-zero random productivity shock, or it can be interpreted as mean-zero measurement error. The individual and the firm observe ε_1, but it is unobserved by others. Let $\pi(x)$ be the probability density function on the set of observable characteristic $\mathcal{X}$ in that period and let $g_1(d\varepsilon_1)$ be the probability

[3] Even differences in school qualities and parental choices of school quality in an altruistic dynamic programming framework can limit social mobility and lead to the intergenerational poverty trap. See Nishimura and Raut (2007) for such a model.

[4] Home inputs include amount of hours the parent spends with the child doing homework, amount of hours the child watches TV, type of programs watched, and how stable and stimulating the relationships among the family members are. Many of these are choice variables for the parent. The omission can lead to biases in the estimates. We cannot measure them in our dataset.

[5] See the studies by Cunha et al. (2010) and Cunha and Heckman (2008), as well as Mohanty and Raut (2009) and Boca et al. (2013).

[6] We have assumed a reduced form specification for the schooling level s. The schooling level s is, in fact, the equibrium outcome of a parent-child bargaining game. Raut and Tran (2005) derive and estimate a model of schooling investment s as a Nash equilibrium outcome of a child-parent bargaining game in a model with only two overlapping generations. In the present framework, with an infinite number of overlapping parent-child generations, it is more complex to derive such a solution and is not further explored in this paper.

density of the random shock ε_1, given x.[7] The aggregate output, which turns out to also be the per capita income or the average income of the economy in any period is

$$Y = \sum_{x \in \mathcal{X}} \left[\int w\left(x\right) + \varepsilon_1 \right] \pi\left(x\right) g_1\left(d\varepsilon_1\right) = \sum_{x \in \mathcal{X}} w\left(x\right) \pi\left(x\right)^{8} \tag{1}$$

An individual with skills x and productivity shock ε_1 ends up with the marginal product $w = w\left(x\right) + \varepsilon_1$ in the labor market. w is his annualized permanent earnings out of which he makes a preschool investment choice a for his child. The annual cost of his preschool investment choice a is $\tilde{\theta}\left(a\right) \equiv \theta\left(a\right) + \varepsilon_2\left(a\right)$, where θ is a constant function for all parents and ε_2 is an unobserved parent-specific variation in the cost, assumed to have zero mean. The rest of his earnings makes up his annualized permanent consumption $c \equiv w - \tilde{\theta}\left(a\right) = w\left(x\right) - \theta\left(a\right) + \varepsilon\left(a\right)$, where $\varepsilon\left(a\right) \equiv \varepsilon_1 - \varepsilon_2\left(a\right)$. We assume that parents with observed characteristics x have a finite number of feasible preschool investment choices, which is represented by the ordered set $A\left(x\right)$. The utility or reward of a parent $\left(x, \varepsilon\right)$ from a preschool investment choice $a \in A\left(x\right)$ is the sum of two components. The first component is the current payoff function with the form $u\left(x, \varepsilon, a\right) = u\left(x, a\right) + \varepsilon\left(a\right)$ where $u\left(x, a\right) \equiv w(x) - \theta\left(a\right)$. Note that ε has two elements, the wage shock and the childcare cost shock. We assume utility is linear in consumption, hence it is additive in these shocks. In the rest of the exposition, we assume a general form for $u\left(x, a\right)$. The parents also derive utility from child outcomes as described below. Finally, we define the components of the unobserved heterogeneity vector ε of an individual of observed characteristics x as $\varepsilon = \left(\varepsilon\left(a\right), a \in A\left(x\right)\right)$, where $\varepsilon\left(a\right)$ is defined above. Denote by $\mathcal{E}$ the set of all possible ε.

[7] We use the convention of denoting the probability density g of a continuous random variable ε by the notation $g\left(d\varepsilon\right)$ and for a discrete random variable x by $g\left(x\right)$ and for their joint density as $g\left(x, d\varepsilon\right)$.

[8] In a similar theoretical model, Raut (1995) includes an external total factor productivity multiplier that increases with an increase in the number of skilled workers in the economy. The paper shows that policies that lead to higher social mobility also leads to higher economic growth.

In any period for a parent $z = (x, \varepsilon)$ with preschool investment choice a, his child's vector of cognitive and non-cognitive skills and unobserved heterogeneity shocks (i.e., the vector $z' = (x', \varepsilon')$) is produced stochastically, which is characterized by the transition probability density function $p(x', d\varepsilon' | x, \varepsilon, a)$.

The preschool investment choice problem of the parent (x, ε) is given by the following Bellman equation:

$$V(x,\varepsilon) = \max_{a \in A(x)} u(x,\varepsilon,a) + \beta \sum_{x' \in X} \int V(x',\varepsilon')\, p(x', d\varepsilon' | x, \varepsilon, a) \tag{2}$$

where $V(x,\varepsilon)$ is his maximized welfare (i.e. the value function), and $u(.)$ is the utility he derives from his own consumption. The utility he derives from his child's welfare is the expected maximized welfare $V(x', \varepsilon')$ of the child, discounted by β, the degree of parental altruism towards the child. His influence over his child's wellbeing is through his preschool investment choice a, which affects his child's cognitive and non-cognitive skill formations as reflected in the transition probability density function $p(x', d\varepsilon' | x, \varepsilon, a)$. Under general regularity conditions on $u(.)$, $p(x', d\varepsilon' | x, \varepsilon, a)$ and β, the measurable value function $V(x, \varepsilon)$ and measurable optimal decision rule $a(x, \varepsilon)$ exist.[9]

An equilibrium in the model is a controlled Markov process with a given initial distribution of parent population $\mu_0(x, d\varepsilon)$ on $\mathcal{X} \times \mathcal{E}$ in period $t = 0$, a family of optimal preschool investment decisions $a(x, \varepsilon)$, $x \in X$ and $\varepsilon \in \mathcal{E}$, and the stationary transition probability density function $p(x', d\varepsilon' | x, \varepsilon, a(x, \varepsilon))$. These variables determine the equilibrium dynamics of earnings distribution, the degrees of intergenerational earnings and college mobilities, and how these are affected by a public policy as described below.

This level of generality makes the estimation of the model computationally intractable. We are more interested in studying the equilibrium dynamics over the observable states $\mathcal{X}$. Since $\mathcal{X}$ is finite, the equilib-

[9] See, Bhattacharya and Majumdar (1989, Theorem 3.2).

rium dynamics over it is a Markov chain, determined by the initial distribution π_0 of population over $\mathcal{X}$ and the transition probability matrix $\Pi = [\Pi(x, x')]_{x,x' \in \mathcal{X}}$. We derive π_0 and Π from the above equilibrium controlled Markov process, $\mu_0(.), a(.)$ and $p(.|.)$. A stationary or long-run equilibrium in this reduced set-up is a probability density function π over the observable states $\mathcal{X}$, such that $\pi = \pi\Pi$ (i.e., an invariant distribution π of the transition probability matrix Π).

Given π_0 and Π, we can examine how the population distribution π_t on $\mathcal{X}$ changes over time t. The structure of Π can tell us if a unique invariant distribution exists and whether the equilibrium population distribution π_t over time t converges to the invariant distribution as t becomes large. A sufficient condition for both is $\Pi(x, x') > 0$ for all $x, x' \in \mathcal{X}$. If the equilibrium transition matrix of Π exhibits a block-diagonal structure (after reordering the states in $\mathcal{X}$ if necessary), then the economy would exhibit an intergenerational poverty cycle. However, our empirical estimates of Π have all elements strictly positive. Hence, we do not have intergenerational poverty cycles. The unique invariant distribution is the long-run equilibrium distribution that the economy will converge to, starting at any initial distribution π_0.

A number of mobility measures have been proposed in the literature for the Markov process determined by a transition matrix. Sommers and Conlisk (1979) argues that $1 - \lambda_{\max}$ is the most appropriate measure of social mobility, where $\lambda_{\max}$ is the second highest positive eigenvalue of the transition probability matrix (it is well-known that the highest positive eigenvalue of a transition probability matrix is always 1). We use this measure for earnings or college mobility and use the Gini coefficient of average earnings over the observable states (i.e., the Gini coefficients of earnings distribution $(w(x), \pi(x), x \in \mathcal{X})$) to compare the effects of our public preschool policy.

2.1 Public preschool policy

We consider the effect of introducing a publicly provided, free preschool to children of poor SES, financed by taxing all parents. Given the type of information available in our dataset, choice variable a takes two values: a value 0 if no preschool and a value 1 if preschool. The cost of preschool as a function of preschool choices will now on be taken as $\theta(a) = \theta a$, where $\theta > 0$ is the cost of preschool. In any period, we define parents of observable state x to fall in the poor SES category if $w(x) \leq 0.7\bar{w}$, where $\bar{w} = \sum w(x)\pi(x)$ is the average or per capita earnings. Our public preschool program makes free preschool participation compulsory for each child of poor SES. Denote by $\mathcal{X}_p$ the set of observable characteristics of the parents of poor SES. The equilibrium tax rate τ_{ax} is then given by $\tau_{ax} = \theta \sum_{x\in\mathcal{X}_p} \pi(x) / \sum_{x\in\mathcal{X}} w(x)\pi(x)$. Once such policy is introduced, a new set of optimal preschool investment decision rules and a new transition matrix will emerge. This will affect the invariant distribution, degree of earnings and schooling inequalities within generations, and the degree of social and college mobilities between generations. We estimate these effects empirically.

2.2 The Econometric Methodology

We follow Rust's (1987, 1994) approach to estimation of dynamic discrete choice model. He introduces the following three assumptions to convert the choice problem in Equation (2) into a random utility model.

Assumption 1 *For $u(x,\varepsilon,a) = u(x,a) + \varepsilon(a)$, the support of $\varepsilon(a)$ is the real line for all $a \in A(x)$.*

Assumption 2 *The transition probability $p(x', d\varepsilon' | x, \varepsilon, a) = f(x'|x,a)\ g(d\varepsilon'|x')$ for some twice continuously differentiable density function g with finite first moment.*

Assumption 3 *The components of ε are independently and identically distributed as extreme value distribution with location parameter 0 and scale parameter 1.*

Assumption 2 means that there are no persistent unobserved heterogeneities across parents and children. It also means that cognitive and non-cognitive skills and the schooling levels of children depend on their parents' skills and schooling levels as well as preschool investment choices.[10] Assumption 3 implies that there are no common unobservables across alternative choices; since in our case we have only one choice, this is not relevant. Let $\Omega(x,a) = \{\varepsilon | \text{for individual } (x,\varepsilon), \text{the choice } a \text{ is optimal}\}$. The conditional choice probabilities are defined as $P(a|x) = \int_{\Omega(x,a)} g(d\varepsilon|x)$. Denote the vector of conditional choice probabilities by $\mathcal{P} = \{P(a|x), a \in A(x), x \in \mathcal{X}\}$. Let Δ be the set of all possible vectors of conditional probabilities. Under the above assumptions, the transition probability matrix Π and the average welfare of individuals in the observable characteristics group can be computed solely with the conditional choice probabilities. Furthermore, the computation of the conditional choice probabilities becomes a simpler iterative fixed point computation of a map Ψ on the finite dimensional compact set Δ as given below.

$$\Pi(x,x') = \sum_{a \in A(x)} f(x'|x,a) P(a|x). \tag{3}$$

[10]This assumption is made for computational simplicity. However, random variable ε represents the unobserved heterogeneity, the omitted factors that are important for the production of skills, and the measurement errors of the included observed variables that could be correlated with the included input variables and correlated across generations. Generally one uses exclusion restrictions, instrumental variables or includes random or fixed effects in microeconometric studies to handle these problems (for example, see Keane et al., 2011; Keane and Wolpin, 2009; Todd and Wolpin, 2006). In our set-up, given the nature of the available data, it is not clear how to utilize these econometric procedures in a multi-generational equilibrium model. See, Arcidiacono and Miller (2011) for some examples of how to incorporate correlated shocks and time-invariant unobserved heterogeneity and then use the EM algorithm to estimate the parameters in related models.

The average welfare of the group with observable state x has the form:

$$v(x) \equiv \int V(x,\varepsilon)\, g(d\varepsilon|x) = \sum_{a\in A(x)} P(a|x)\left[u(x,a) + e(x,a) + \beta F(x,a)\cdot v\right] \tag{4}$$

where $v = [v(x_1), ..., v(x_m)]'$ is a column vector, $e(x,a) = \int_{\Omega(x,a)} \varepsilon g(d\varepsilon)$, and $F(x,a) = [f(x_1|x,a),$ $\ldots,$ $f(x_m|x,a)]$, a m dimensional row vector. Recall that m is that number of ordered discrete states in each period. $F(x,a)$ is the row vector of transition probabilities of the m states that x' can take in the next period given the current state x and choice a. The column vector v contains the values of these states in the next period. Thus, $F(x,a)\cdot v$ is the expectation of the next period's value function conditional on this period's state x and choice a.

Under Assumptions 1 and 2, Rust (1987) shows that the problem in Equation (2) becomes a random utility model. Using the McFadden result where a random utility model under Assumption 3 has a Logit representation, Rust shows that the conditional choice probabilities have the following Logit representation,

$$\begin{aligned} P(a|x) &= \frac{e^{\tilde{v}(x,a)}}{\sum_{a'\in A(x)} e^{\tilde{v}(x,a')}}, \text{ where} \\ \tilde{v}(x,a) &= u(x,a) + \beta F(x,a)\left[I_m - \beta\bar{F}\right]^{-1}[\bar{u} + \bar{e}] \end{aligned} \tag{5}$$

where I_m is a $m\times m$ identity matrix, $\bar{F}$ is an $m\times m$ matrix with the element in the (x,x') position is $\sum_{a\in A(x)} f(x'|x,a)\, P(a|x)$; $\bar{u} = [\bar{u}(x_1), ..., \bar{u}(x_m)]'$, and $\bar{e} = [\bar{e}_1(x_1), ..., \bar{e}_m(x)]'$ are m dimensional column vectors with elements $\bar{u}(x) = \sum_{a\in A(x)} u(x,a)\, P(a|x)$ and $\bar{e}(x) = \sum_{a\in A(x)} e(x,a)\, P(a|x)$, $x \in \mathcal{X}$.

Given our data, how do we estimate the structural parameters and hence choose a particular model to study all the policy issues? This is explained in the next section.

3 Econometric implementation

For each vector of structural parameters, we need to compute the optimal choice probabilities $\mathcal{P} = \{P(a|x), a \in A(x), x \in \mathcal{X}\}$ and use them to compute the likelihood of the sample and the maximum likelihood estimates of the structural parameters. To that end, Rust (1987) uses a fixed point algorithm on the set of functions to compute the value function and uses the value function to compute the optimal choice probabilities. We use the fixed point algorithm on choice probabilities and used these choice probabilities to compute the value function and to estimate the structural parameters as explained below.[11]

Based on what is known in the child-development literature, we specify the stochastic production functions for cognitive and non-cognitive skills as follows (recall that τ denotes cognitive skill, σ, μ, φ denote social skills and s denotes schooling):

$$\begin{aligned} f_\gamma(x'|x,a) &= q_\tau(\tau'|\tau,s,a) \times q_\sigma(\sigma'|\tau',\tau,\sigma,\mu,\eta,\phi,s,a) \\ &\quad \times q_\mu(\mu'|\tau',\tau,\sigma,\mu,\eta,\phi,s,a) \times q_\eta(\eta'|\tau',\tau,\sigma,\mu,\eta,\phi,s,a) \\ &\quad \times q_\varphi(\varphi'|\tau',\tau,\sigma,\mu,\eta,\phi,s,a) \times q_s(s'|\tau',\sigma',\mu',\eta',\phi',s,a) \end{aligned} \tag{6}$$

where each component probability density function is further specified as a Logit model with the regressors as the conditioning variables of the component. In our model, τ' is the innate ability of the child. We assume that τ' depends only on parent's schooling level s, innate ability τ and preschool investment a. The details

[11] For other estimation procedures, see a recent survey of the literature by Aguirregabiria and Mira (2010).

of the production process of the non-cognitive skills are discussed in Section 4.4. Denote by γ the vector of all of these regression parameters, which together determine the transition probabilities $f_\gamma\left(x'|x,a\right)$. Denote the parameters of the reward function, θ and the altruism parameter β by the vector $\xi=(\theta,\beta)$.

We have data of the type (x_i,a_i,x_i'), $i=1,...,n$, on n parent-child pairs. The problem is to estimate the structural parameters $\zeta=(\xi,\gamma)$ using this data.

Note that for fixed ζ, Equation (5) defines a map $\Psi:\Delta\rightarrow\Delta$ since the right hand side of the equation is a function of conditional probabilities. The fixed point of which is the set of conditional choice probabilities of the dynamic programming problem in Equation (2). It can be shown that for each structural parameter ζ, the iterative process $\mathcal{P}_n=\Psi\left(\mathcal{P}_{n-1}\right)$, starting from any initial $\mathcal{P}_0$, always converges to a unique fixed point $\mathcal{P}_\zeta=\left(P_\zeta\left(a|x\right),a\in A\left(x\right),x\in\mathcal{X}\right)$. We use $\mathcal{P}_\zeta$ to calculate the log-likelihood of the sample in the following procedure.

The likelihood can be split up into the parameters of payoffs and the parameters that govern the laws of motion of the state variables. To see this, note that $\Pr\left(a,x'|x\right)=\Pr\left(a|x\right).\Pr\left(x'|x,a\right)=P_\zeta\left(a|x\right).f_\gamma\left(x'|x,a\right)$. The log-likelihood function for the sample is then given by $L\left(\xi,\gamma\right)=L_1\left(\xi,\gamma\right)+L_2\left(\gamma\right)$, where $L_1\left(\xi,\gamma\right)=\sum_{i=1}^n\ln P_\zeta\left(a_i|x_i\right)$ and $L_2\left(\gamma\right)=\sum_{i=1}^n\ln f_\gamma\left(x_i'|x_i,a_i\right)$. The full information maximum likelihood estimation procedure requires maximization of the full likelihood function $L\left(\xi,\gamma\right)$, which involves numerous parameters. The maximization algorithm for such objective functions does not always converge and this is true in our case.

We follow a two-step procedure instead: In the first step, we compute a consistent estimate $\hat{\gamma}$ by maximizing the conditional likelihood function $L_2\left(\gamma\right)$, which given the recursive structure in Equation (5), is equivalent to estimating the individual Logit models constituting the parts of $f_\gamma\left(x'|x,a\right)$. In the second step, we estimate ξ by maximizing $L_1(\xi, \hat{\gamma})$.

Denote this two-step estimate by $(\hat{\xi}_r, \hat{\gamma}_r)$ and the full information maximum likelihood estimate by $(\hat{\xi}_f, \hat{\gamma}_f)$. How close is the estimate $\hat{\xi}_r$ to $\hat{\xi}_f$? How precise is the estimate of the standard error $\Sigma_{\xi\xi\cdot\gamma}$ of $\hat{\xi}_r$ obtained from the restricted maximum likelihood procedure by fixing a value of $\gamma = \hat{\gamma}_r$?

We use the bootstrap with 300 replications to calculate the variance-covariance matrix of our parameter estimates, as this accounts for the two-step nature of our estimation procedure.

4 Empirical Findings

4.1 The Dataset and the Variables

We use the NLSY79 and NLSY79 Children and Young Adults data. The NLSY79 dataset contains a nationally representative sample of 12,686 young men and women who were 14-22 years old when they were first surveyed in 1979 (i.e., these sampled individuals represent a population born in the 1950s and 1960s, and living in the United States in 1979). These individuals are interviewed annually. The dataset has records of school and labor market experiences of these individuals and also information on their cognitive and non-cognitive traits. We, however, need information on most of these variables for the parents of the respondents, but this dataset does not have much information on them. We have linked this dataset with the NLSY79 Children and Young Adults dataset. The child survey dataset includes longitudinal assessments of each child's cognitive, attitudinal, social, motivational, academic and labor market experiences. We generate separate observations, one for each child for families with multiple children, and treat such parent-child pairs as independent observations. We construct the variables of our study as follows:[12]

[12] We only describe these for the parent sample; the same cut-off points are used for the children in the children sample.

Early childhood inputs and home environment: We use parent's education levels to measure the child's family background. The NLSY dataset has poor measures of respondent's early childhood inputs. It has only a binary variable containing information on whether the respondent had preschool (not including Head Start) experience or not. We treat individuals with Head Start experience as having no preschool in our analysis. Notice that this may lead to underestimation of the effect of preschool investment. We use the AFQT score to measure the innate ability.

Socialization skill (σ): Each respondent is asked how social he/she felt towards others at age 6, expressed on a scale of 1 to 4, the highest number representing most social. We create a binary sociability variable by assigning the value 1 if a respondent reported a number 3 or 4 and assigned 0 otherwise.

Motivational skill (μ): We measure motivational skill as the job aspiration of the respondents in the main NLSY79 sample. For the children sample, the average of the various motivation measures is taken at a young age of the child and assigned the value 1 if the average is greater than 3.75 and the value 0 otherwise.

Rosenberg measure of self-esteem skill (η): We measure the positiveness with which individuals regard themselves in society (i.e., a positive sense of self). Six questions were taken from the classic Rosenberg (1965) scale in the NLSY surveys. There is, however, no well-accepted definition of adequate self-esteem. Based on the distribution, we divided the 25-point scale by treating a score of 20 or greater to indicate a high self-esteem, assigning the value 1 to η and the value 0 to η otherwise.

Pearlin mastery scale of internal self-concept (ϕ): This measures to what extent an individual believes that his life chances are under his own control (Pearlin et al., 1981). This is similar to the Rotter scale of self-control. The respondents are asked 7 questions yielding scores ranging from 0 to 28. We assign the value 1 to represent a high sense of self-control to respondents with a score between 23 and 28 inclusive, otherwise we assign the value 0.[13]

4.2 An Augmented Earnings Function - The Role of cognitive and non-cognitive skills

Non-cognitive traits are important determinants of both earnings and learning.[14] We carry out a rudimentary analysis in this section to emphasize the importance of these traits for earnings. We estimate an augmented Mincer earnings function by adding measures of non-cognitive skills such as social, motivational, self-esteem and internal self-concept skills in the standard Mincer earnings function that includes only cognitive skills such as innate ability and the number of years of schooling. The schooling level variable is correlated with the omitted non-cognitive skill variables. Thus the schooling level variable captures the effects of non-cognitive skills in the standard Mincer earnings function estimation, producing an over-estimate of the rate of return to schooling. As a by-product of our analysis, we provide an estimate of this upward bias.

Mincer (1958) shows that if foregone earnings is the only cost of schooling and the effect of an extra year of schooling on earnings is proportional and constant, then the log-earnings is a linear function of the number of years of schooling. He later extends this model by allowing experience (measured by the square of the age of the worker) to affect earnings over the life-cycle as follows:

$$\ln w = \alpha_0 + \alpha_1 s + \alpha_2 \text{age} + \alpha_3 \text{age}^2 + \varepsilon$$

This basic Mincer earnings function has been estimated using various datasets. It has been given many interpretations by deriving it from various models of schooling choice.[15] We estimate the basic model by

[13] For further discussion of these measures, see Duckworth, Almlund, Kautz, and Heckman (Duckworth et al.).

[14] For surveys of the effect of non-cognitive traits on earnings, see Borghans et al. (2008) and Almlund et al. (2011).

[15] See, for instance, Card (1999); Heckman et al. (2006, 2008); Raut and Tran (2005); Weiss (1995).

taking w as the annual earnings of the respondents in the NLSY79 dataset. The heteroskedasticity adjusted estimates for this basic model are reported in the second column (under heading Basic) in Table 1. Our estimate for α_1 is 11.12 percent, which is close to what is found in other studies.[16]

What exactly is the role of education in the production of earnings? Does an extra year of education have any intrinsic value in the production of the output? Or is it a surrogate for other factors, such as innate ability, hence the estimated returns to education is higher than its actual worth in production?[17]

We include the AFQT score variable (a widely used measure of ability) as a regressor together with other standard variables used in the literature, such as family background measured by the parents' education levels, and a dummy variable for the female gender. These are reported in the third column (under heading Extended) in Table 1. The estimate for the schooling coefficient drops to 6.94 percent. This estimate is corrected for ability bias or gender bias in the estimated returns to schooling and is close to what is found in other studies (see Card (1999)). We now add to it our four measures of non-cognitive skills to see how much of the above estimate of the returns to education is biased upward because it captures the effects of the omitted non-cognitive skills. The estimates are shown in the fourth column of Table 1 (under heading Augmented). We see that all of the four non-cognitive skill variables have significant positive effects on earnings, and the rate of returns to education has dropped by about 1 percentage point. By looking at the R^2 values, we see that about 1 percent variation in earnings is explained by the inclusion of the non-cognitive skills in the standard Mincer earnings function. Note that adding the non-cognitive skills leads to much less improvement in fit than adding the cognitive skill.

[16] See, for instance, the survey by Card (1999), and the analyses of Heckman et al. (2006, 2008) and Raut and Tran (2005).

[17] See Borghans et al. (2011); Heckman and Kautz (2012) for limitations of this measure.

Table 1: Determinants of earnings – role of cognitive and non-cognitive skills (from the parent sample)

Variables	Basic	Extended	Augmented
Intercept	1.7137 (26.89)	2.3440 (34.43)	1.6978 23.80
Grade	0.1112 (81.79)	0.0694 (38.33)	0.0595 (32.28)
Age	0.3363 (79.93)	0.3277 (73.85)	0.3279 (73.51)
Age Square	-0.0040 (59.74)	-0.0039 (55.07)	-0.0039 (54.86)
Mother's Grade		-0.0022 (1.59)	-0.0050 (3.56)
Father's Grade		0.0079 (6.83)	0.0065 (5.57)
Dummy Variable for Female		-0.5187 (80.31)	-0.5137 (78.88)
Dummy Variable for Non-Black and Non-Hispanic		0.0545 (7.18)	0.0794 (10.31)
τ : AFQT Score		0.0059 (37.21)	0.0048 (29.37)
σ : Socialization			0.0111 (1.68)
μ : Motivation - Job Aspiration			0.0261 (3.49)
η : Self-Esteem (Rosenberg Scale)			0.0193 (18.18)
ϕ : Internal Self-Control (Pearlin Scale)			0.0251 (22.33)
n	118,477	95,253	93,166
R^2	0.3083	0.3752	0.3839

Notes: Absolute values of t-statistics based on the heteroskedasticity adjusted standard errors are in parentheses.

4.3 Estimation of Schooling Function

Consider two specifications of the schooling function, $s(\tau', \sigma', \mu', \eta', \phi', a, \varepsilon')$. In the first specification, assume that the schooling level is a continuous variable and the function $s(\tau', \sigma', \mu', \eta', \phi', a, \varepsilon')$ is linear. Assume that variable ε' constitutes an additive error term with zero mean and possibly heteroskedastic variances. We include our measures of cognitive and non-cognitive skills and family background. The parameter estimates of this model with the *t-statistics* based on the heteroskedasticity adjusted standard errors are shown in Table 2.

In the second specification, consider only two levels of schooling: $s = 1$ for completed college or more, and $s = 0$ otherwise. Assume that $s(\tau', \sigma', \mu', \eta', \phi', a, \varepsilon')$ is a Logit model. The parameter estimates from this model are shown in Table 2.

It is clear from the estimates that the most significant determinant of schooling is the innate ability measured by the AFQT score. Moreover, even after controlling for family background, we find that all non-cognitive skills have significant positive effects on schooling level.

4.4 Production of non-cognitive skills

As established in the cited literature, non-cognitive skills are important determinants of earnings and learning. In this section we estimate the production process of these skills and estimate the effect of preschool experience on the development of these non-cognitive skills. Childhood investment is the most crucial input for the development of cognitive and non-cognitive skills.[18]

We create the binary variable τ, assigning the value 1 to denote an individual as highly talented if his

[18] See Cunha and Heckman (2007, 2009); Heckman et al. (2008); Raut (2003)

Table 2: Determinants of grade and College completion – role of cognitive and non-cognitive skills (from the parent sample)

Variables	OLS model of years of completed schooling	Logit model of completing college
Intercept	9.1570 (353.41)	-7.9304 (117.45)
Mother's Grade	0.0817 (32.44)	0.1145 (23.76)
Father's Grade	0.0430 (21.60)	0.0705 (19.59)
Preschool	0.4999 (34.62)	0.5800 (24.72)
τ : AFQT Score	0.0384 (165.38)	0.0472 (104.15)
σ : Socialization	0.0776 (7.00)	0.1332 (6.80)
μ : Motivation - Job Aspiration	0.4890 (43.04)	0.9446 (34.09)
η : Self-Esteem (Rosenberg Scale)	0.3551 (20.07)	0.3781 (14.66)
ϕ : Internal Self-Control (Pearlin Scale)	0.4399 (32.67)	0.7299 (20.62)
n	108,565	108,636
R^2 *	0.4263	0.3436

* Notes: The R^2 in the second column is McFadden's R^2.

AFQT score is 70 or higher (on a scale of 0 to 100), and assigning the value $\tau = 0$ otherwise. For the children sample, we take the average of available multiple cognitive test scores (on a scale of 0 to 100) and assign the value $\tau' = 0$ if the average score is less than 70. Otherwise we assign the value $\tau' = 1$.[19] Other binary skill variables are described earlier. We estimated the Logit models for each of the cognitive and non-cognitive skills-types in the children sample. These parameter estimates constitute the components of the parameter vector γ of the transition probability function $f_{\gamma}(x'|x,a)$. We report the parameter estimates in Table 3 for the specifications of each components of x and a. These are used in the two-step estimation procedure to estimate $\xi = (\theta, \beta)$ given the parameters γ of the transition probability function fixed at these estimates. To compare the sensitivity of our estimates and inference of the structural parameters, we estimated another specification in which we included only those regressors that are significant.

From Table 3, it is clear that after controlling for parents' grade, preschool has a significantly positive effect on socialization skill and on the levels of talent and schooling, but it has no direct effect on Pearlin measure of internal self-concept and the Rosenberg measure of self-esteem. The estimates in the table also show that the level of talent has strong positive effects on the formation of all skills.

4.5 Optimal Parental Preschool Investment Decision

We assume that the state variables $s, \tau, \sigma, \mu, \eta$ and ϕ are all binary (i.e., the number of states is $m = 2^6 = 64$) and the components of the random variable ε are continuous. Recall that preschool investment choice a is a binary variable assigned value 1 if the parent decides to invest in preschool and assigned the value 0 otherwise. For many children in our sample there are two parents alive, but in our model we have assumed one-parent families. We use both parents' information to create a synthetic parent as follows. We construct

[19] Alternatively, we could have taken the first component of the Principal Component Analysis of these cognitive test scores.

Table 3: Logit model of cognitive and non-cognitive skills.

Variables	τ'	σ'	μ'	η'	ϕ'	s
Intercept	-2.8005	-1.1219	-0.8990	-2.5222	-2.7063	-3.9698
	(41.76)	(20.80)	(17.02)	(32.42)	(32.61)	(33.60)
τ	1.4300	0.1508	-0.0713	-0.5082	-0.4989	2.1359
	(23.99)	(2.47)	(1.19)	(6.99)	(6.69)	(26.38)
τ'		0.9459	1.2590	0.2423	0.1800	
		(16.78)	(22.85)	(4.18)	(3.04)	
σ		0.2414	0.1940	0.1209	0.1044	0.3041
		(5.64)	(4.62)	(2.54)	(2.14)	(3.92)
μ		0.1005	-0.0211	-0.0449	-0.0312	0.7126
		(2.26)	(0.48)	(0.89)	(0.61)	(6.78)
η		0.2581	0.2577	0.2863	0.2542	0.5727
		(5.82)	(5.91)	(5.90)	(5.13)	(7.31)
ϕ		-0.0177	-0.0466	0.1294	0.1333	0.6198
		(0.41)	(1.11)	(2.66)	(2.68)	(7.72)
s	0.8456	0.5096	0.4588	1.5443	1.6694	1.4013
	(11.92)	(10.64)	(9.60)	21.21	(21.38)	(15.49)
a : Preschool	0.8766	0.7972	0.0496	-0.0731	-0.0647	0.6569
	(16.75)	(18.58)	(1.16)	(1.53)	(1.33)	(7.13)
n	11,428	11,428	11,428	11,428	11,428	7,732
McFadden's R^2	0.109	0.0911	0.0623	0.0681	0.0705	0.2205

Notes: A variable x without a $'$ refers to the parent and with a $'$ refers to his child.

τ : AFQT Score
σ : Socialization
μ : Motivation - Job Aspiration
η : Self-Esteem (Rosenberg Scale)
ϕ : Internal Self-Control (Pearlin Scale)

The schooling level in column one refers to parents' schooling level in all models. While for other models the attributes Socialization, Motivation, Internal Self-Control (Pearlin) and Self-Esteem (Rosenberg) in the first column are parents' attributes, for schooling model s, these attributes in column one are the individual's own attributes. Variable s in column one corresponds to parents' education level and this model is estimated using the 1979 youth sample.

a parent's binary schooling variable s to have value 1 if either parent has 16 or more years of education, otherwise $s = 0$.

The two-step maximum likelihood estimates of the structural parameters $\xi = (\theta, \beta)$ are shown in Table 4 with two sets of specifications of transition probabilities $f_\gamma(x'|x, a)$. The first column contains estimates from the specification in which only the significant conditioning variables are included and the second column contains the estimates of the parameters in which all conditioning variables are included. The remainder of the paper uses the parameter estimates in the second column of Table 4.

An estimate of $\hat{\theta} = 1.224$ in the table means that the cost per year during the first 5 preschool years is $\$6,120$. This results from us having annualized earnings and costs over 25 years of a parent's life-cycle. Thus, the total preschool cost over the entire life-cycle is $\$1,224 \times 25 = \$30,600$. This total amount is actually spent over the first 5 preschool years of the child's life, giving us an estimated preschool cost of $\$6,120$ per year. Schweinhart et al. report an estimate of the average yearly preschool cost to be $\$6,178$ using the actual preschool cost. Our maximum likelihood estimate of the cost is very close to their direct estimate of cost.

5 Economic Benefits from Public Provision of Preschool

We have shown that investment in preschool enhances certain skills that are important for learning and earning. We define parents to fall in the poor SES if their earnings are less than 70 percent of the average earnings in the economy. From the empirical estimates of the optimal choice, we find that very few parents of poor SES invest in their children's preschool. We consider a public policy of providing preschool to children of poor socioeconomic status (SES) in all periods. This will impose a tax burden on all parents, but

Table 4: Maximum likelihood parameter estimates of $\xi = (\theta, \beta)$ and other derivd macroeconomic parameters, given two different estimates of $f_\gamma(x'|x,a)$

	Given estimates of $f_\gamma(x'|x,a)$ with	
	only significant x	all x
Cost ($\hat{\theta}$) of preschool (in '000 dollars)	1.222	1.224
t-stat	(15.79)	(15.16)
Degree of altruism: $\hat{\beta}$	0.443	0.486
t-stat	(2.24)	(2.64)
Long-run Equibrium Tax Rate: τ (in percent)	5.94	5.83
Percent of population in poor SES:		
Before the policy introduction ($\tau = 0$)	36.22	35.71
After the policy introduction	29.64	29.14
Per capita after tax annual earnings:		
Before the policy introduction ($\tau = 0$)	5621.85	5640.08
After the policy introduction	5734.93	5759.38
gains in per capita income	113.09	119.30
log-likelihood	-7424.97	-7429.575

such a policy may also improve the social mobility, reduce the earnings inequality and eventually may lead to a higher level of per capita earnings in the long-run. We examine if the gain from per capita earnings can outpace the cost of providing such a social insurance program. We also look at its within-generation effects on earnings, and on the intergenerational effects on earnings and college mobility. It is important to note that the magnitude of the effect of publicly provided preschool will depend on if the social protection will be available to all future generations or if it is just a onetime policy.[20] In our model, it is clear that, if social protection is given only once, its effect will wear out in the long-run, although it may have significant effect during the transition to the long-run equilibrium.

Table 4 reports the estimates of the percent of parents falling into the poor SES status in the long-run before and after the introduction of the public policy; the tax rate τ_{ax} that finances the public preschool policy in the long-run equilibrium; and the long-run disposable (i.e., after tax) average yearly earnings of

[20] See Heckman and Masterov (2004); Raut (2003).

workers before and after the introduction of the social contract policy.[21]

5.1 Intergenerational Earnings Mobility

To examine how the introduction of a public policy providing free preschool to children of poor SES affects earnings mobility between generations, we compute the mobility index of a stationary transition probability matrix of an equilibrium Markov process of earnings distributions over time.[22] Our estimate of the measure of earnings mobility before the introduction of the social contract is 0.5945. After the introduction of the public preschool program it is 0.6468.

It is difficult to compare our estimate of the mobility index with previous studies, because there is no commonly agreed upon measure of earnings mobility.[23]

5.2 College Mobility

Denote by $Q^s = [q_{ij}], i, j = 1, 2$ the intergenerational college mobility matrix in which state 1 represents no college, and state 2 represents college or more. The element q_{ij} represents the probability that a child of a parent of college education status i will move to college education status j, for all i and $j = 1, 2$. We report below the estimated college mobility matrices, the corresponding invariant distributions, and the estimates of the mobility measure before and after the introduction of the social contract. These estimates indicate that the introduction of the social contract will increase college enrollment from 6.71 percent to 9.45 percent (i.e. a 2.74 percentage point increase for a child of non-college parent). The percentage of college-educated

[21] One can calculate consistent confidence intervals for these policy effects (and for those that follow) using the approach in Woutersen and Ham (2013), but computational constraints prevented us from doing this here.

[22] We are assuming that parents will send their children to preschool when preschool is offered free of cost. We assume with the preschool policy the extra children who will attend preschool will not to change the preschool cost (i.e., the estimated cost of preschool has priced in the cost of preschool buildings, teachers and preschool materials).

[23] For a survey of various measures of mobilities and their properties, see Geweke et al. (1986).

population will increase in the long-run from a rate of 10.16 percent without a social contract to a higher rate of 13.76 percent with a social contract. In the long run there will be about a 3.6 percentage point increase in college enrollment.

College mobility statistics before introduction of social contract:

$$Q_b^s = \begin{bmatrix} 0.93287 & 0.06713 \\ 0.59380 & 0.40620 \end{bmatrix}, p_b^s = \begin{bmatrix} 0.8984 & 0.1016 \end{bmatrix}, 1 - \lambda^s_{\max,b} = 0.6609$$

College mobility statistics after introduction of social contract:

$$Q_a^s = \begin{bmatrix} 0.90553 & 0.09447 \\ 0.59184 & 0.40816 \end{bmatrix}, p_a^s = \begin{bmatrix} 0.8624 & 0.1376 \end{bmatrix}, 1 - \lambda^s_{\max,a} = 0.6863$$

5.3 Lifetime Earnings Inequality

Preschool investment would increase the income of children from poor SES families and thus, presumably reduce the income gap between the rich and poor. Using the Gini-coefficient to measure income inequality, we would expect that income inequality will improve over time after the public preschool program is introduced. The long-run income distribution observed is the invariant distribution. We compute the Gini-coefficient of income inequality for the invariant income distribution *before* the introduction of the public policy, and compare it with the Gini-coefficient for the invariant income distribution *after* the introduction of the policy. The estimated Gini-coefficients of average lifetime earnings are, respectively, 0.2363 without the social contract, and 0.2335 with the social contract. The estimated Gini coefficient of the current generation from our data is 0.2291. Thus, our estimates show that income inequality of future generations will rise. However, the social contract of publicly providing preschool to children of poor SES produces a lower inequality of long-term earnings than the inequality without the social contract.

5.4 The Tax Burden of the Social Policy

Suppose the government provides preschool to the children of poor SES perpetually. We know that the size of the population of poor SES will change over time. Thus, the resource needs of the program will become smaller, and the tax revenues will become higher over time. We can study the stream of these costs and benefits to society and then compute the average per period costs and benefits to calculate the tax-burdens of the social contract. Applying the Ergodic Theorem, this boils down to computing the costs and benefits of the invariant distribution that will result after the introduction of the social contract. Our computations below are based on the long-run equilibrium.

For the current generation, 31.13 percent of the population falls in the poor SES. Without a public policy, approximately 35.71 percent of the population in the long-run will fall in the poor socioeconomic status, by our definition of poor SES. The introduction of the public policy will reduce the population in the poor SES to 29.14 percent. From Table 5, we see that while the welfare of the income groups that have publicly provided preschool will be higher, the welfare of the rest of the population will be lower. It is difficult to estimate the net effect of the policy on social welfare since there is no universally agreed upon aggregation rule for social welfare. We use average yearly disposable earnings over the life-cycle to compare the net gain or loss to the society. These estimates in Table 4 show that the average yearly disposable earnings of the society in the long-run are higher by $113 after the introduction of the policy. Based on this, we conclude that there is a net gain to the society by introducing a publicly provided preschool program for the children of poor SES.

Our benefit calculation does not take into account other public savings that will result due to the policy, such as savings from welfare assistance programs, savings to the criminal justice system, and potential victims of crimes. If we incorporate these, the returns will be even higher. Using data from the High/Scope Perry Preschool Program, Heckman et al. (2010b) estimate a total benefit of 7 percent per annum from all these sources for each dollar spent on the preschool program, even counting the social costs of taxation.

6 Conclusion

This paper formulates an altruistic model of preschool investment choices of parents in a structural dynamic programming framework. It uses NLSY79 and NLSY79 Children and Young Adult data to estimate the structural parameters.

The paper estimates the production processes of two types of cognitive skills - the IQ score and the schooling level, and four types of non-cognitive skills - the socialization skill, the motivational skill, the Rosenberg measure of self-esteem skill and the Pearlin mastery scale of internal self-concept skill. The paper finds that preschool boosts significantly both types of cognitive skills and only the socialization skills among the four measures of non-cognitive skills. Moreover, all of these cognitive and non-cognitive skills have significant positive effects on level of schooling and labor market earnings of individuals.

The paper estimates the structural parameters and then uses those to carry out policy analysis for this economy to examine the effect of a publicly provided preschool to economically disadvantaged children and financing it by taxing all parents. Taking into account the within generation and between generation effects of such a policy, the paper finds that the introduction of such a public policy: (a) improves the intergenerational earnings mobility from 0.5945 to 0.6468, measured on a scale of 0 to 1, (b) improves the college mobility from 0.6609 to 0.6863, measured on a scale of 0 to 1, (c) increases the college completion rate of the children of non-college educated parents from 6.71 percent to 9.45 percent (i.e. a 2.74 percentage point increase), and percent of college educated population increases from 10.16 percent to 13.76 percent (i.e., a 3.6 percentage point increase), (d) reduces the within-generation earnings inequality measured by the Gini coefficient from 0.2363 to 0.2335 on a scale of 0 to 1, and (e) results in a net gain (net of taxes) in the long-run per capita earnings.

The effects that we report in this paper may be underestimates for many reasons. First, we have treated Head Start children the same as children without preschool. Second, the preschool programs that the respondents attended were the ones that existed during the 1960's. The quality of preschool programs since then has improved significantly and thus the effects of current preschool programs may be much higher than the estimates that we have. The positive effects of the public preschool policy may be even higher in reality because we have used the estimated benefits from the lower quality preschool programs that existed in the 1960s. Furthermore, if there is a positive externality in the aggregate production function created by the size of the skilled labor as it is assumed in the endogenous growth models, the gains from a public preschool policy could be even higher. There are, however, other sources of bias in our empirical estimates, such as omitted variables in the skill production functions, persistent unobserved heterogeneity across generations and across life stages, and failure of the independence of irrelevant alternatives (IIA) assumption within periods. Given these factors, it is probably impossible to sign any bias.

Due to data limitations and to avoid computational complexities, the multi-generational equilibrium model of our paper makes two simplifications. First, the aggregate output in the economy is produced with a linear production function with aggregate labor measured in efficiency units as the input and without any external effect from the aggregate skilled labor. This is equivalent to assuming that skill prices are fixed, which could be justified for a small open economy in a globalized world. But in a large or a closed economy, introduction of public preschool policy will change the supplies of various skills produced by preschool, and hence their prices and net benefits of the public policy may be lower. However, if there is a positive externality from the number of skilled workers in the production of aggregate output, as it is assumed in endogenous growth models, it is not clear whether the general equilibrium skill prices will fall or rise after the public preschool policy is introduced. Second, the production functions for the cognitive and non-cognitive skills do not include maternal time as one of the inputs. With maternal time input included in the production of skills, introduction of public preschool policy will have positive income effects, a negative substitution effect, and the net effect is undetermined and needs be empirically determined. The positive income effect will accrue because the parents will use the free preschool program as a daycare, enabling them to work outside the home. Those who are already using a daycare as a means to work outside the home will switch to free preschool program for their children. Both of these effects will lead to a gain in family income, leading to the parent's ability to buy more of the market inputs that are important for skill production. Positive preschool effects will occur because a preschool will increase the production of the cognitive and non-cognitive skills of the children. The negative substitution effect will occur because free preschool will increase maternal employment and thus, will reduce the maternal time input for skill production. Empirical evidence on these effects are limited and more work in this area will be useful. For our multi-generational equilibrium model, we do not have data on labor supply of the respondents' parents in the main NLSY sample, so we have assumed a simplified specification of these skill productions. Future work with better data can shed more light on these issues.

Bibliography

Aguirregabiria, V. and P. Mira (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics 156*(1), 38–67. 11

Almlund, M., A. Duckworth, J. J. Heckman, and T. Kautz (2011). Personality psychology and economics. In E. A. Hanushek, S. Machin, and L. Wößmann (Eds.), *Handbook of the Economics of Education*, Volume 4, pp. 1–181. Amsterdam: Elsevier. 1, 14

Arcidiacono, P. and R. A. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica 79*(6), 1823–1867. 10

Bernal, R. (2008, Nov). The effect of maternal employment and child care on childrens cognitive development. *International Economic Review 49*(4), 1173–1209. 1

Bernal, R. and M. P. Keane (2011, Jul). Child care choices and children's cognitive achievement: The case of single mothers. *Journal of Labor Economics 29*(3), 459–512. 1

Bhattacharya, R. N. and M. Majumdar (1989). Controlled semi-markov models: The discounted case. *Journal of Statistical Planning and Inference 21*(3), 365–381. 9

Blau, D. and J. Currie (2006). Chapter 20 pre-school, day care, and after-school care: Whos minding the kids? *Handbook of the Economics of Education*, 11631278. 1, 2

Blau, D. M. (1999). The effect of child care characteristics on child development. *The Journal of Human Resources 34*(4), pp. 786–822. 1

Boca, D. D., C. J. Flinn, and M. Wiswall (2013). Household choices and child development. *Review of Economic Studies, Forthcoming*. 5

Borghans, L., A. L. Duckworth, J. J. Heckman, and B. ter Weel (2008, Fall). The economics and psychology of personality traits. *Journal of Human Resources 43*(4), 972–1059. 14

Borghans, L., B. H. H. Golsteyn, J. J. Heckman, and J. E. Humphries (2011). Identification problems in personality psychology. *Personality and Individual Differences 51*(3: Special Issue on Personality and Economics), 315–320. 17

Cameron, S. V. and J. J. Heckman (1998, Apr). Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of American males. *Journal of Political Economy 106*(2), 262. 1

Campbell, F., G. Conti, J. Heckman, S. Moon, and R. Pinto (2012). The long-term health effects of early childhood interventions. Under review, *Economic Journal*. 1

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 5, pp. 1801–1863. New York: North-Holland. 15, 4.2, 16

Carneiro, P. and J. J. Heckman (2003). Human capital policy. In J. J. Heckman, A. B. Krueger, and B. M. Friedman (Eds.), *Inequality in America: What Role for Human Capital Policies?*, pp. 77–239. Cambridge, MA: MIT Press. 1

Cunha, F. and J. J. Heckman (2007, May). The technology of skill formation. *American Economic Review 97*(2), 31–47. 1, 18

Cunha, F. and J. J. Heckman (2008, Fall). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources 43*(4), 738–782. 1, 5

Cunha, F. and J. J. Heckman (2009, April). The economics and psychology of inequality and human development. *Journal of the European Economic Association 7*(2-3), 320–364. Presented as the Marshall Lecture, European Economics Association, Milan, Italy, August 29, 2008. 1, 18

Cunha, F., J. J. Heckman, L. J. Lochner, and D. V. Masterov (2006). Interpreting the evidence on life cycle skill formation. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Preschool 12, pp. 697–812. Amsterdam: North-Holland. 1

Cunha, F., J. J. Heckman, and S. M. Schennach (2010, May). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931. 5

Currie, J. and D. Thomas (1995, June). Does Head Start make a difference? *American Economic Review 85*(3), 341–364. 1

Deming, D. (2009, July). Early childhood intervention and life-cycle skill development: Evidence from Head Start. *American Economic Journal: Applied Economics 1*(3), 111–134. 1

Duckworth, A., M. Almlund, T. Kautz, and J. J. Heckman. The Relevance of Personality Psychology for Economics. Forthcoming, *Handbook of the Economics of Education*. 1, 13

Geweke, J., R. C. Marshall, and G. A. Zarkin (1986). Mobility indices in continuous time markov chains. *Econometrica 54*(6), 1407–1423. 23

Heckman, J. J. (2000, March). Policies to foster human capital. *Research in Economics 54*(1), 3–56. 1

Heckman, J. J. (2008, July). Schools, skills and synapses. *Economic Inquiry 46*(3), 289–324. 1

Heckman, J. J., F. Flyer, and C. Loughlin (2008, January). An assessment of causal inference in smoking initiation research and a framework for future research. *Economic Inquiry 46*(1), 37–44. 18

Heckman, J. J. and T. Kautz (2012, August). Hard evidence on soft skills. *Labour Economics 19*(4), 451–464. Adam Smith Lecture. 1, 17

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006). Earnings equations and rates of return: The Mincer equation and beyond. In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Volume 1, Preschool 7, pp. 307–458. Amsterdam: Elsevier. 15, 16

Heckman, J. J., L. J. Lochner, and P. E. Todd (2008, Spring). Earnings functions and rates of return. *Journal of Human Capital 2*(1), 1–31. 15, 16

Heckman, J. J. and D. V. Masterov (2004, September). The productivity argument for investing in young children. Preschool Working Paper No. 5, Committee on Economic Development. 20

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Q. Yavitz (2010a, August). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative Economics 1*(1), 1–46. 1

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Q. Yavitz (2010b, February). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics 94*(1-2), 114–128. 1, 5.4

Keane, M. P., P. E. Todd, and K. I. Wolpin (2011). *The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications*. Elsevier. 10

Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *Journal of Political Economy 105*(3), pp. 473–522. 1

Keane, M. P. and K. I. Wolpin (2001). The effect of parental transfers and borrowing constraints on educational attainment. *International Economic Review 42*(4), pp. 1051–1103. 1

Keane, M. P. and K. I. Wolpin (2009, Jan). Empirical applications of discrete choice dynamic programming models. *Review of Economic Dynamics 12*(1), 1–22. 10

Mincer, J. (1958, August). Investment in human capital and personal income distribution. *Journal of Political Economy 66*(4), 281–302. 4.2

Mohanty, L. L. and L. K. Raut (2009). Home ownership and school outcomes of children: Evidence from the PSID Child Development Supplement. *American Journal of Economics and Sociology 68*(2), 465–489. 5

Nishimura, K. and L. K. Raut (2007). School choice and the intergenerational poverty trap. *Review of Development Economics 11*(2), 412–420. 3

Pearlin, L. I., E. G. Menaghan, M. A. Lieberman, and J. T. Mullan (1981, December). The stress process. *Journal of Health and Social Behavior 22*(4), 337–356. 4.1

Raut, L. K. (1995). Signalling equilibrium, intergenerational mobility and long-run growth. Preschool 9603002, EconWPA. Presented at the Seventh World Congress of the Econometric Society, Tokyo, Japan. 8

Raut, L. K. (2003). Long term effects of preschool investment on school performance and labor market outcome. Preschool 0307002, EconWPA. 1, 18, 20

Raut, L. K. and L. H. Tran (2005, Aug). Parental human capital investment and old-age transfers from children: Is it a loan contract or reciprocity for Indonesian families? *Journal of Development Economics 77*(2), 389–414. 6, 15, 16

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press. 4.1

Rust, J. (1987, September). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica 55*(5), 999–1033. 1, 2.2, 2.2, 3

Rust, J. (1994). Structural estimation of Markov decision processes. In R. Engle and D. McFadden (Eds.), *Handbook of Econometrics, Volume*, pp. 3081–3143. New York: North-Holland. 2.2

Schweinhart, L. J., H. V. Barnes, and D. Weikart (1993). *Significant Benefits: The High-Scope Perry Preschool Study Through Age 27*. Ypsilanti, MI: High/Scope Press. 4.5

Sommers, P. M. and J. Conlisk (1979). Eigenvalue immobility measures for Markov chains. *Journal of Mathematical Sociology 6*(2), 253–276. 2

Todd, P. E. and K. I. Wolpin (2003). On the specification and estimation of the production function for cognitive achievement*. *The Economic Journal 113*(485), F3–F33. 1

Todd, P. E. and K. I. Wolpin (2006). Assessing the impact of a school subsidy program in mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *The American Economic Review 96*(5), pp. 1384–1417. 10

Todd, P. E. and K. I. Wolpin (2007, Dec). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital 1*(1), 91–136. 1

Traub, J. (2000). What no school can do. *The New York Times Magazine January 16*(Section 6), 52–57. 1

Weiss, A. (1995). Human capital vs. signalling explanations of wages. *Journal of Economic Perspectives 9*(4), 133–154. 15

Woutersen, T. and J. Ham (2013). Confidence sets for continuous and discontinuous functions of parameters. Technical report, Working Paper. 21

7 APPENDIX

Table 5: Equilibrium Solution

State	PV Wage	obsd freq	$P_b(a=1\|x)$	$P_a(a=1\|x)$	optV_b	optV_a	$p{*}_b$	$p{*}_a$
[0,0,0,0,0,0]	3.0993	9.5730	33.8937	100.0000	8.5885	8.8587	32.5119	26.1168
[0,1,0,0,0,0]	3.5662	3.6839	34.2812	100.0000	9.0979	9.3356	0.8377	0.9192
[0,0,1,0,0,0]	3.5977	17.8684	33.9587	100.0000	9.0866	9.3284	2.3604	1.9432
[0,1,1,0,0,0]	4.0646	6.4491	34.3223	33.7294	9.5959	9.3847	0.1404	0.1589
[0,0,0,0,1,0]	4.4821	3.4739	33.8578	33.2821	9.9837	9.7555	2.8119	2.2938
[0,1,0,0,1,0]	4.9490	1.2776	34.2534	33.6584	10.4946	10.2290	0.1549	0.1739
[0,0,1,0,1,0]	4.9805	7.2454	33.9235	33.3450	10.4812	10.2241	0.2520	0.2337
[0,0,0,1,0,0]	5.0917	2.7739	34.3792	33.7795	10.6518	10.3746	0.0401	0.0484
[1,0,0,0,0,0]	5.2129	0.2450	46.8940	45.5622	10.9709	10.6463	14.2778	11.9679
[0,1,1,0,1,0]	5.4474	4.1740	34.2954	33.6988	10.9919	10.6974	0.7858	0.8910
[0,1,0,1,0,0]	5.5586	1.1201	34.7075	34.0914	11.1672	10.8509	1.0646	0.9269
[0,0,1,1,0,0]	5.5902	5.1628	34.4179	33.8169	11.1491	10.8427	0.1431	0.1679
[1,1,0,0,0,0]	5.6799	0.0875	47.4014	46.0409	11.4798	11.1192	1.2668	1.0921
[1,0,1,0,0,0]	5.7114	1.5051	46.9130	45.5788	11.4709	11.1167	0.1570	0.1829
[0,1,1,1,0,0]	6.0571	2.8176	34.7170	34.1007	11.6641	11.3187	0.1208	0.1197
[1,1,1,0,0,0]	6.1783	0.3150	47.3865	46.0251	11.9796	11.5895	0.0440	0.0545
[0,0,0,1,1,0]	6.4746	2.6689	34.3569	33.7527	12.0510	11.6897	16.5757	19.1858
[1,0,0,0,1,0]	6.5958	0.2275	46.8482	45.5145	12.3604	11.9540	0.7195	1.0225
[0,1,0,1,1,0]	6.9415	1.5926	34.6966	34.0748	12.5682	12.1678	1.2792	1.4759
[0,0,1,1,1,0]	6.9730	5.6965	34.3964	33.7908	12.5475	12.1573	0.1441	0.1962
[0,0,0,0,0,1]	7.0009	0.1050	52.0818	50.3491	13.4421	12.9084	1.5090	1.7322
[1,1,0,0,1,0]	7.0627	0.1663	47.3710	46.0073	12.8703	12.4279	0.1559	0.2125
[1,0,1,0,1,0]	7.0942	2.0039	46.8676	45.5315	12.8599	12.4241	0.1579	0.1882
[1,0,0,1,0,0]	7.2054	0.1138	47.5753	46.2025	13.0277	12.5739	0.0486	0.0641
[0,1,1,1,1,0]	7.4399	4.0340	34.7068	34.0849	13.0644	12.6350	7.4653	8.8418
[0,1,0,0,0,1]	7.4678	0.1838	52.4400	50.6829	14.0018	13.4268	0.7117	1.0156
[0,0,1,0,0,1]	7.4993	0.7088	52.0526	50.3226	13.9299	13.3686	0.5989	0.7082
[1,1,1,0,1,0]	7.5611	1.0151	47.3563	45.9920	13.3697	12.8979	0.1546	0.2107
[1,1,0,1,0,0]	7.6723	0.0350	47.9744	46.5770	13.5407	13.0497	0.7052	0.8298

Table 6: Equilibrium Solution (continued)

[1,0,1,1,0,0]	7.7038	0.4638	47.5571	46.1838	13.5269	13.0437	0.1668	0.2276
[0,1,1,0,0,1]	7.9662	0.2275	52.3699	50.6172	14.4889	13.8863	0.0783	0.0956
[1,1,1,1,0,0]	8.1707	0.1925	47.9216	46.5251	14.0396	13.5192	0.0536	0.0710
[0,0,0,0,1,1]	8.3837	0.0700	52.1615	50.4140	14.8749	14.2534	0.9937	1.0713
[1,0,0,1,1,0]	8.5882	0.1313	47.5553	46.1782	14.4201	13.8842	0.2684	0.3582
[0,1,0,0,1,1]	8.8506	0.0613	52.5352	50.7627	15.4375	14.7746	0.0756	0.0806
[0,0,1,0,1,1]	8.8821	0.4375	52.1318	50.3871	15.3619	14.7128	0.0503	0.0663
[0,0,0,1,0,1]	8.9933	0.2800	52.6654	50.8863	15.6096	14.9342	0.0959	0.1016
[1,1,0,1,1,0]	9.0551	0.2100	47.9729	46.5700	14.9346	14.3615	0.0586	0.0772
[1,0,1,1,1,0]	9.0866	1.1376	47.5371	46.1598	14.9189	14.3536	0.0088	0.0098
[1,0,0,0,0,1]	9.1146	0.0613	63.9675	61.7066	16.0457	15.3176	0.0149	0.0197
[0,1,1,0,1,1]	9.3490	0.4550	52.4642	50.6964	15.9237	15.2333	1.3296	1.4389
[0,1,0,1,0,1]	9.4603	0.0175	52.8843	51.0862	16.1762	15.4594	0.7621	1.0180
[0,0,1,1,0,1]	9.4918	0.3150	52.5898	50.8156	16.0957	15.3927	0.0970	0.1043
[1,1,1,1,1,0]	9.5535	0.3763	47.9197	46.5180	15.4330	14.8305	0.1413	0.1868
[1,1,0,0,0,1]	9.5815	0.0350	100.0000	100.0000	15.8182	15.0181	0.1235	0.1319
[1,0,1,0,0,1]	9.6130	0.5250	63.8780	61.6182	16.5365	15.7801	0.1651	0.2178
[0,1,1,1,0,1]	9.9587	0.4113	52.7707	50.9790	16.6615	15.9172	0.0112	0.0128
[1,1,1,0,0,1]	10.0799	0.2888	63.8606	61.5975	17.0876	16.2938	0.0421	0.0558
[0,0,0,1,1,1]	10.3762	0.1313	52.7758	50.9807	17.0488	16.2852	1.4851	2.0297
[1,0,0,0,1,1]	10.4974	0.0438	63.9837	61.7144	17.4679	16.6548	0.7157	1.0742
[0,1,0,1,1,1]	10.8431	0.0788	53.0088	51.1943	17.6178	16.8130	0.1179	0.1576
[0,0,1,1,1,1]	10.8746	1.0063	52.6993	50.9093	17.5341	16.7430	0.1525	0.2180
[1,1,0,0,1,1]	10.9643	0.0788	63.9986	61.7251	18.0234	17.1726	0.1478	0.1973
[1,0,1,0,1,1]	10.9958	0.9450	63.8950	61.6272	17.9577	17.1164	0.1753	0.2514
[1,0,0,1,0,1]	11.1070	0.0263	100.0000	100.0000	17.4154	16.5192	0.0153	0.0202
[0,1,1,1,1,1]	11.3415	1.0326	52.8937	51.0860	18.1024	17.2702	0.0493	0.0679
[1,1,1,0,1,1]	11.4627	0.5163	63.8846	61.6133	18.5121	17.6332	1.9466	2.6493
[1,1,0,1,0,1]	11.5739	0.0263	63.9065	61.6305	18.7516	17.8514	2.0323	3.0223
[1,0,1,1,0,1]	11.6054	0.3850	63.9250	61.6539	18.6824	17.7919	0.1504	0.2004
[1,1,1,1,0,1]	12.0723	0.3325	63.7741	61.5009	19.2389	18.3107	0.4324	0.6130
[1,0,0,1,1,1]	12.4898	0.0438	64.0692	61.7869	19.6239	18.6758	0.1891	0.2515
[1,1,0,1,1,1]	12.9567	0.1750	63.9402	61.6554	20.1848	19.1988	0.4979	0.7079
[1,0,1,1,1,1]	12.9882	1.6188	63.9548	61.6750	20.1113	19.1353	0.0199	0.0265
[1,1,1,1,1,1]	13.4552	1.5401	63.8076	61.5259	20.6711	19.6573	0.1407	0.1933