CESifo

AREA CONFERENCES 2017

Public Sector Economics

Munich, April 27–29 2017

Universal versus Targeted Preschools: A Mechanism-Design Perspective

Brant Abbott and Dominik Sachs

CESifo

Universal versus Targeted Preschools: A Mechanism-Design Perspective*

Brant Abbott
Queen's University

Dominik Sachs
European University Institute

April 20, 2017

Abstract

We study optimal sliding fee schedules for public preschools in the presence of private alternatives and endogenous parental earnings. The optimal fee schedule trades off public enrollment, crowding-out of private enrollment and parental incentives. We estimate our model of preschool choice using the Early Childhood Program Participation Survey, and achieve parameter identification through a policy-discontinuity implied by Head Start eligibility criteria. The current level of enrollment could be implemented with 5% less government spending by targeting subsidies more efficiently. Optimally implementing higher enrollment implies negative fees for the poor. Ignoring distortions from welfare programs like food stamps leads to significant policy errors.

JEL-**classification:** H21, H23 , I20, J13

Keywords: Preschool Policies, Mechanism Design, Early Childhood

*Contact: abbottbrant@gmail.com and dominik.sachs@eui.eu. For comments and suggestions, we are grateful to Árpád Ábrahám, Richard Blundell, Eric French, Nezih Guner, James Heckman, Andrea Ichino, Fabian Kosse, Dirk Krueger, Erzo F.P. Luttmer, Christopher Rauh, José-Victor Ríos-Rull, Stefanie Stantcheva, Aleh Tsyvinski as well as seminar participants at the $(Ce)^2$-Workshop in Warsaw, CREST, EUI, the NBER Summer Institute 2016 (Macro Public Finance) and the Barcelona Summer Forum 2016 (Micro and Macro Perspectives on Taxes).

1 Introduction

"One solution to these problems is to make the programs universal but to offer a sliding fee schedule based on family income." Heckman (2013, p. 36)

The importance of early human capital investments has recently received substantial attention from economists and policy makers. While much of the economics literature has focused on evaluations of preschool,[1] policy debates have focused on how these programs ought to be implemented, with particular attention given to the relative merits of universal and targeted pre-K programs.[2] Universal preschool programs ensure access for all children, but they crowd out enrollment in private preschools leading to inefficient use of public funds.[3] Targeted programs – which are usually means tested by family income – are considered a more cost effective way of increasing enrollment because poor families could not afford private alternatives. Targeting, however, disincentivizes parental labor supply because it acts like an additional tax rate on earnings. It seems unlikely that current programs (e.g. Head Start), which depend on ad-hoc cut-off rules for participation, target in an optimal way.

This paper studies optimal preschool targeting within a rich empirical model. We characterize the nonlinear fee schedule for public preschools that implements a certain level of preschool enrollment at minimal fiscal costs, taking into account crowding-out of private enrollment and endogenous parental earnings. This problem is deliberately non-welfarist because we do not want to confuse the goal of efficiently promoting preschool enrollment with redistributive goals. Expanding preschool enrollment through provision of public preschools has been stated as an explicit goal by the Obama administration.[4] The aim of this paper is to inform policy makers how such a goal can be achieved at minimal costs. To achieve this, we combine methods from mechanism-design with quasi-experimental and structural estimation methods.

We develop a model in which parents decide whether to to send their child to a private preschool, a public preschool or no preschool. The fee schedule for public preschool influences this decision. A decrease in public fees increases public enrollment through two channels: (i) some children are sent to public preschool that would otherwise have not attended any preschool. This constitutes an increase in enrollment. (ii) some children are sent to public preschool that would otherwise have attended a private preschool. This crowding-out of private investment does not imply an increase in enrollment. Furthermore, the gradient of the fee schedule affects the earnings of parents who send their children to public preschool. An

[1] See Cunha, Heckman, and Schennach (2010), Heckman, Moon, Pinto, Savelyev, and Yavitz (2010), Currie and Almond (2011) and Duncan and Magnuson (2013) or Elango, García, Heckman, and Hojman (2015).

[2] See Barnett, Brown, and Shore (2004) and Fitzpatrick (2008).

[3] See Cascio and Schanzenbach (2013) and Bassok, Miller, and Galdo (2016).

[4] `https://obamawhitehouse.archives.gov/the-press-office/2013/02/13/fact-sheet-president-obama-s-plan-early-education-all-americans`

increasing fee schedule acts like an additional marginal tax rate on earnings, and therefore reduces work incentives. This in turn increases costs for the government because it reduces tax revenue. This creates an interesting interaction between the optimal progressivity of the fee schedule and the progressivity of the existing tax-transfer system.

To arrive at the fee schedule that optimally deals with these trade-offs, we determine optimal incentive-compatible allocations, and the implied optimal fee schedule, using the random participation approach pioneered by Rochet and Stole (2002). Since the preschool decision is an extensive margin decision, we can easily deal with multiple dimensions of heterogeneity (see also Choné and Laroque (2011)). This allows us to explore optimal policies in an empirically plausible model with rich heterogeneity.

Very few structural models of preschool choice exist in the literature,[5] and to the best of our knowledge there are none that distinguish public and private preschool choices.[6] Thus, in order to carry-out the desired optimal policy exercise, we must first develop and estimate such a model. Our model includes heterogeneity in parental preferences for public and private preschool, where the distributions of these preferences vary with observable characteristics of the parent and child. We show that these preferences are a composite of parental marginal values of wealth, returns to preschool and deeper preschool tastes. Parents compare their preferences over the various preschool options to the associated costs, and choose the option that generates the largest net utility gain.

We estimate the parameters of the model by maximum likelihood. However, our model requires exogenous variation in the policy environment in order to be identified. We use variation in Head Start eligibility around the poverty line as our source of such variation. Our use of quasi-experimental variation to identify the parameters of a structural model follows similar work by Blundell and Shephard (2012) and Attanasio, Meghir, and Santiago (2012). The Head Start eligibility threshold was recently use by Carneiro and Ginja (2014) as a source of variation to study the effects of preschool on health and behavioral outcomes. We use data on 2344 four year-olds from the Early Childhood Program Participation survey to estimate the model.

We then quantitatively study the fee schedule that implements the same enrollment as current policies at minimal costs. This schedule starts at a monthly fee of \$100 and then increases smoothly, almost linearly, in family income with an average marginal fee of 2.5%. Compared to current policies, enrollment decreases below the poverty line and for family incomes above \$60,000. It increases for income levels inbetween. Public expenditures decrease by 5.3% due to

[5] One existing study is that of Heckman and Raut (2016). Those authors estimate a model with a single preschool option based on the CNLSY data.

[6] We must stress the point that this paper is not about child care or day care, but rather early childhood education. Some examples of models of day care decisions include Domeij and Klein (2013) and Bernal (2008). Unlike day care, publicly provided preschool does not appear to affect maternal labor supply (Cascio and Schanzenbach 2013, Fitzpatrick 2010).

optimal targeting even though the implied adverse effects on parental income imply a loss in tax revenue which is as large as 2% of the current spending on public preschools.

We then elaborate on the importance of the baseline tax-transfer system, taking into account effective marginal tax rates associated with other welfare programs like food stamps, TANF etc., which imply much higher distortions for low incomes. Taking these programs into account changes the shape of the fee schedule heavily for low incomes. The optimal fee schedule starts at $250 per month and then decreases until the poverty line is reached where it takes a value of $150 per month implying an average marginal fee of -10% in that income region. For higher incomes the shape of the optimal fee schedule is very similar to benchmark version.

Finally, we ask how to implement a higher level of enrollment in a cost-minimizing way. We choose the enrollment target that a universal preschool, i.e. zero fees for all income levels, would achieve. Naturally, the cost-minimizing schedule to achieve the same level of enrollment has to include negative fees for some families, unless, of course, universal preschool itself is optimal. We find that fees should be negative for income levels below $30,000. Thus, these families should not only have free access to public preschool but should even receive up to $200 per month as an incentive. This resembles conditional cash transfer programs that are typically implemented in poorer countries.

Related Literature. This paper is related to the optimal taxation literature following Mirrlees (1971). In contrast to most of that literature we do not treat redistribution as the social objective. Another example of a non-welfarist objective can be found in Kanbur, Keen, and Tuomala (1994), who consider poverty reduction as a social objective. Kanbur, Pirttilä, and Tuomala (2006) provide a survey of other related papers. Whereas that strand of literature shares the non-welfarist approach with our paper, our paper differs in that we are not studying income taxation, but rather use the mechanism design approach to study optimal preschool pricing.

This paper is also related to Blundell and Shephard (2012) in the sense that we study optimal policies using a structural empirical model. Like them, we connect the statistical uncertainty of our estimated model to uncertainty about the optimal policies and what they can achieve.

Some recent papers in the 'New Dynamic Public Finance' tradition study implementations of second-best efficient allocations, where government makes use of both income taxes and income-contingent repayment of student loans (Findeisen and Sachs 2015, Stantcheva 2015, Koeniger and Prat 2015). As in these papers, there are two different policy instruments that influence labor supply: Labor income taxes and income-contingent repayment of student loans in their case and payment of preschool fees in our case. Note, however, that our paper is conceptually different in that we take the labor income tax as given and focus on the design of one instrument in isolation. Whereas such an optimization of one policy instrument in isolation is naturally less ambitious in terms of the social objective, it may be of more immediate policy relevance if it is easier for governments to reform one policy instrument (preschool policies in our case) instead of changing many at the same time (which would be preschool policies and income taxes in our case). In line with our approach is the paper by Ho and Pavoni (2016) who study efficient child care subsidies in the tradition of Mirrlees (1971). Their focus, however, is on setting efficient incentives for maternal labor supply.

Finally, this paper is related to Cascio and Schanzenbach (2013) and Bassok, Miller, and Galdo (2016) who have stressed the crowding out of private preschool enrollment due to universal preschool. Relatedly, Bick (2015) studies public child care subsidies in Germany in a quantitative macro model that he calibrates to Germany and also emphasizes the crowding out of alternative child-care options.

This rest of this paper goes like this: In Section 2 we present the general structure of our model and use this to derive formulas describing the optimal policies. In Section 3 we describe the data and our maximum likelihood estimation of the model parameters. In section 4 we present our quantitative results, and in Section 5 we conclude.

2 The Model Environment

The model of preschool decisions that we estimate and use for policy analysis is one of static discrete choices. However, households make decisions in a dynamic environment. In what follows we first introduce a fully specified dynamic model, and then show that the one-time discrete preschool choices made at the beginning of that model are equivalent to those in the static discrete choice model, up to a first-order approximation. An important aspect of this result is that it shows that our estimation identifies preferences that have been normalized by marginal utilities from wealth, rather than deeper utility parameters.

2.1 Model Basics: Deriving the Linear Static Model

Our model includes a continuum of heterogeneous households indexed by i. Households have preferences over consumption, preschool enrollment, and effort in the labor market. Preschool enrollment decisions are denoted by $ps_i \in \{no, pu, pr\}$, where the elements of the choice set are *no preschool*, *public preschool* and *private preschool*, respectively. The current wealth of household i is a_{it} and their ability to earn income is ω_{it}. Each household chooses their consumption c_{it}, savings a_{it+1}, labour supply ℓ_{it}, and what type of preschool to enroll their child in, if any. Periodic utility from consumption and leisure has the form $u(c_{it} - v(\ell_{it}))$, and utility from the preschool choice arises from variation in the household's continuation value. The fee paid for preschool $F_i(ps_i, y_i)$ varies with the type of preschool, the household, and (possibly) household income. Variation in fees across households that is unrelated to their income may arise because of geographic variation in private preschool markets or public policy.

To consider all dynamic decisions of households beginning from the time of the preschool choice, we would write their value function at the time of preschool choice:

$$V_{it}(a_{it}, \omega_{it}) = \max_{a_{it+1}, c_{it}, \ell_{it}, ps_i} \left\{ u\left(c_{it} - v\left(\frac{y_{it}}{\omega_{it}}\right)\right) + \beta V_{it+1}(a_{it+1}, \omega_{it+1}, ps_i) \right\}$$
$$s.t. \;\; c_{it} = a_{it} + y_{it} - \frac{a_{it+1}}{1+r} - T(y_{it}) - F_i(ps_i, y_{it}).$$

We proceed by defining a related value function $\widetilde{V}_{it}(\cdot)$ that is the solution of the continuous part of the problem above, conditional on a particular preschool choice and the optimal labor supply. To do this we define a net-wealth state variable $a_{it} - F_i(ps_i, y^*_{it})$, where y^*_{it} is the optimal labor supply rule. Given these, the conditional value function is:

$$\widetilde{V}_i(a_{it} - F_i(ps_i, y^*_{it}), \omega_{it}|ps_i) = \max_{a_{it+1}, c_{it}} \{u\left(c_{it} - v\left(y^*_{it}/\omega_{it}\right)\right) + \beta V_{it+1}(a_{it+1}, \omega_{it+1}, ps_i)\}$$
$$s.t. \;\; c_{it} = a_{it} - F_i(ps_i, y^*_{it}) + y^*_{it} - \frac{a_{it+1}}{1+r} - T(y_{it}).$$

Implicitly, y^*_{it} solves the optimal labor supply first order condition:

$$1 - T'(y^*_{it}) - F'_i(ps_i, y^*_{it}) - v'(y^*_{it}/\omega_{it})(1/\omega_{it}) = 0. \tag{1}$$

Given the above definitions, we can determine the conditions under which public preschool would be chosen:

$$\begin{aligned} ps_i = pu \;\; \text{if} \quad & \widetilde{V}_i(a_{it} - F_i(pu, y^*_{it}), \omega_{it}|pu) > \widetilde{V}_i(a_{it}, \omega_{it}|no) \\ \text{and} \;\; & \widetilde{V}_i(a_{it} - F_i(pu, y^*_{it}), \omega_{it}|pu) > \widetilde{V}_i(a_{it} - F_i(pr, y^*_{it}), \omega_{it}|pr). \end{aligned}$$

These conditions simply say that public preschool earns greater net utility than either no preschool or private preschool, respectively. However, $\widetilde{V}_i(\cdot)$ is an unknown function, which makes it impossible to proceed directly. Instead, we replace $\widetilde{V}_i(\cdot)$ in the above conditions with a Taylor series expansion around the point where net-wealth is equal to assets. Using $\partial\widetilde{V}_i/\partial a_{it}$ is to denote the marginal value of net-wealth for household i, we can re-write the above public preschool choice conditions as:

$$ps_i = pu \;\; \text{if} \quad F_i(pu, y^*_{it}) < \frac{\widetilde{V}_i(\cdot|pu) - \widetilde{V}_i(\cdot|no)}{\partial\widetilde{V}_i/\partial a_{it}} \tag{2}$$

$$\text{and} \;\; F_i(pr, y^*_{it}) - F_i(pu, y^*_{it}) > \frac{\widetilde{V}_i(\cdot|pr) - \widetilde{V}_i(\cdot|pu)}{\partial\widetilde{V}_i/\partial a_{it}}. \tag{3}$$

By dividing through by the marginal value of net-wealth, we effectively monetize utility from preschool (at an individual level), which is why monetary costs appear on the left-hand sides. To simplify notation we denote monetized utility from preschool as follows:

$$\theta_i = \frac{\widetilde{V}_i(\cdot|pu) - \widetilde{V}_i(\cdot|no)}{\partial \widetilde{V}_i / \partial a_{it}} \tag{4}$$

$$\theta_{pr,i} = \frac{\widetilde{V}_i(\cdot|pr) - \widetilde{V}_i(\cdot|pu)}{\partial \widetilde{V}_i / \partial a_{it}}, \tag{5}$$

with the interpretation that θ_i is the monetary value of public preschool for i, and $\theta_{pr,i}$ is the additional value that i garners from choosing private preschool (over public).

Using the notation for monetary values of preschool choices more broadly, we can fully specify conditions for each of the three possible preschool choices:

$$\begin{array}{llll} ps_i = pu & \text{if} \quad F_i(pu, y^*_{it}) < \theta_i & \text{and} & F_i(pr, y^*_{it}) - F_i(pu, y^*_{it}) \geq \theta_{pr,i} \\ ps_i = pr & \text{if} \quad F_i(pr, y^*_{it}) < \theta_i + \theta_{pr,i} & \text{and} & F_i(pr, y^*_{it}) - F_i(pu, y^*_{it}) < \theta_{pr,i} \\ ps_i = no & \text{if} \quad F_i(pu, y^*_{it}) \geq \theta_i & \text{and} & F_i(pr, y^*_{it}) \geq \theta_i + \theta_{pr,i} \end{array} \tag{6}$$

Clearly, these choices are now characterized as derived from a linear and static model.

2.1.1 Limitations of Approximation

The linear static model is exactly equivalent to the preschool choice dimension of the dynamic model only if preschool fees have no effect on the marginal value of wealth. To the extent that this is not true the linear model is only a first-order approximation. Our belief is that preschool fees only amount to a small fraction of lifetime resources even for lower income families. The benefits of focusing on the static model include an exact derivation of the likelihood function, which clarifies parameter identification, and also an exact theoretical formulation of the optimal policy schedule, which clarifies the nature of quantitative optimal policy results.

Further, the scope of the policy analysis that our model is suited to is not too limited by this first-order approximation. We consider only consider changes in the preschool fee and these fees are only paid for two years, so the effect of a $1,000 increase in a preschool fee is very different from a $1,000 dollar increase of a tax because a tax is paid every year. In other words, preschool fees are relatively small as compared to permanent earnings and therefore changes in the fee should have only small effects on the marginal utility of consumption.

2.2 Aggregation

The social objective we consider are functions of aggregate preschool enrollment rates, therefore we must characterize the distribution of households over the parameter space. The distributions of the preference parameters will be allowed to depend on parental ability, ω_i, and a vector of other household characteristics, X_i (henceforth we repress time subscripts as we are working in a static environment). Let $G_\theta(\theta_i|\omega_i, X_i)$ and $g_\theta(\theta_i|\omega_i, X_i)$ be the CDF and PDF of θ_i, respectively. Also, let $G_{pr}(\theta_{pr,i}|\omega_i, X_i)$ and $g_{pr}(\theta_{pr,i}|\omega_i, X_i)$ be the CDF and PDF of $\theta_{pr,i}$, respectively. The model also includes variation in preschool fees, even after conditioning on income. Let the PDF of public preschool fees be $\Xi_{pu}(F_i(pu, y_i^*)|\omega_i)$ and the PDF of private preschool fees be $\Xi_{pr}(F_i(pr, y_i^*)|\omega_i)$.

Note that we write the distribution of fees as conditional on ω_i, when in reality fees will depend on parental income y_i. We are able to exchange the conditioning variable in this way because the optimal labor supply rule in equation 1 is a one to one mapping between the two. When we analyze optimal policies the household labor supply rule will be an incentive compatibility constraint on the policy choice, and thus the variables will continue to be exchangeable by construction.

A very useful measure for defining and analyzing policy is the share of children whose parents have a particular value of ω and make each of the possible preschool choices. We define $s_j(\omega)\ \forall j \in \{pu, pr, no\}$ to be functions defined over the space of ω that report these shares. When deriving optimal policy formulas in the next section, we derive them in terms of $s_j(\omega)\ \forall j \in \{pu, pr, no\}$ because this highly simplifies notation without limiting the economic intuition. $s_j(\omega)\ \forall j \in \{pu, pr, no\}$ is just a reduced form and we provide the formulas in terms of primitives of the model in the appendix.[7]

2.3 Optimal Policy Problem

Our policy analysis considers the problem of a policy maker (planner) who has a given preschool enrollment target and wants to determine the lowest cost way to achieve it. Alternatively, one could consider the dual problem of maximizing preschool enrollment subject to a fixed budget. To achieve this cost minimization the planner optimally chooses the public preschool fee schedule $F_{pu}(\cdot)$. The planner can choose any fee schedule: No ex-ante restrictions are placed on the functional form. Commonly analyzed targeted or universal programs are within the choice set, as are more flexible blends of these programs. In making their choice the planner will consider how various programs affect incentives to enroll children in public versus private preschool, and how any progressiveness would distort parental labor supply.

[7] These shares are also related to the likelihood function present in the estimation section below.

Heterogeneity in ω_i, θ_i and $\theta_{pr,i}$ is not observed by the policy maker, but the planner does observe household income and knows what the joint distribution of heterogeneity is. Naturally, observing income means that the planner will be able to screen households ω_i types. However, the planner will not be able to screen out heterogeneity in θ_i and $\theta_{pr,i}$, and thus approaches the problem as one of random participation random participation (see Rochet and Stole (2002)).

In writing the planner's objective it is helpful to simplify notation to some extent. Let $y^*(\omega)$ be the optimal income of a household that does not enroll their child in public preschool, and let $y^*_{pu}(\omega)$ be the optimal income of a household that does. Also, we repress individual i subscripts when analyzing policy. The planner's cost objective is given by:

$$\mathcal{C}(F_{pu}(\cdot)) = -\int_{\Omega} s_{pu}(\omega)\left[T(y^*_{pu}(\omega) - T(y^*(\omega)) + F_{pu}(y^*_{pu}(\omega)) - \mathcal{C}_{pu}\right] dF(\omega).$$

Families that do not send their child to public preschool do not show up because the way they affect public expenditure is exogenous. Thus, for public expenditure, we aggregate over those families only that do send their children to public preschool. The fee schedule influences this objective in three ways: First of all, the obvious direct fiscal impact is $F_{pu}(y^*_{pu}(\omega)) - \mathcal{C}$: Parents pay the fee $F_{pu}(y^*_{pu}(\omega))$ but imply costs of $\mathcal{C}$. Second, the share of children in public preschool $s_{pu}(\omega)$ is endogenous w.r.t. the fee. Thirdly, the progressivity of the fee schedule influences the income choice of the parents. If the fee schedule is increasing, parents will work less if they send their child to public preschool which implies a reduction in tax revenue.

The government's objective is to ensure a certain enrollment level $\bar{\mathcal{E}}$ with as little public spending as possible. Thus, the problem reads as:

$$\min_{F_{pu}(\cdot)} \mathcal{C}(F_{pu}(\cdot)) \quad \text{subject to} \quad \mathcal{E}\left(F_{pu}(\cdot)\right) \geq \bar{\mathcal{E}} \tag{7}$$

where

$$\mathcal{E}\left(F_{pu}(\cdot)\right) = \int_{\Omega} \left(s_{pr}(\omega) + s_{pu}(\omega)\right) dF(\omega) \tag{8}$$

The government also has to take into account individual optimization behavior when choosing the fee schedule. First, the government is constrained by individually optimal preschool enrollment decisions that are summarized in equation 6. Second, for those parents that choose to send their child to a public preschool, the government has to take into account that these parents choose their labor supply endogenously with respect to the fee schedule. Formally, this means that the labor supply of households with children in public preschool will satisfy:

$$1 - T'(y^*_{pu}(\omega)) - F'_{pu}(y^*_{pu}(\omega)) - v'(y^*_{pu}(\omega)/\omega)(1/\omega) = 0.$$

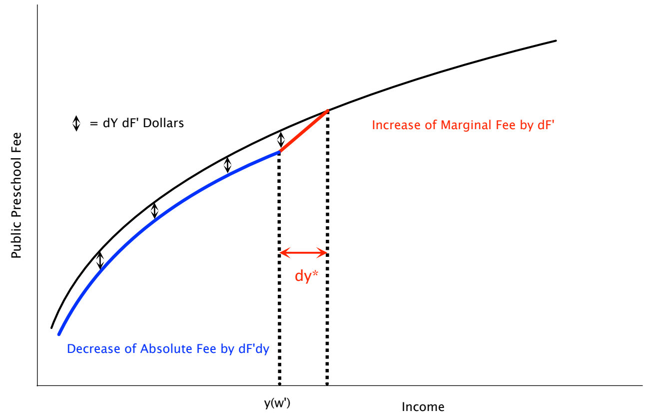


Figure 1: Perturbation of Public Preschool Fee Schedule

In Section 2.4 we solve this problem by optimizing the function $F_{pu}(\cdot)$ using a perturbation approach. In Appendix A, we show that this problem can also be solved using a mechanism-design approach in the tradition of the optimal income taxation literature following Mirrlees (1971).

2.4 The Optimal Sliding Fee Schedule

In the following, we use an intuitive perturbation approach in the spirit of Piketty (1997) and Saez (2001). We show that it is equivalent to the mechanism-design solution in Appendix A. The value of using a fee perturbation approach is that is shows in a more intuitive way where the various components of the optimal fee schedule come from. In what follows we list these components one by one, and then combine them into a formula for the optimal fee schedule.

Assume that the black bold line in Figure 1 presents the optimal preschool fee schedule. Then, slightly lowering the fee for incomes below $y^*_{pu}(\omega')$, as illustrated, should have no first-order effect on the planner's objective. Note that the perturbation is such that the marginal preschool fee is slightly increased by dF'_{pu} in an interval around $y^*_{pu}(\omega')$ with length dy. We think of dy and dF'_{pu} being infinitesimal. This small reform will have direct effects on the planner's objective, as well as indirect effects that work through the government budget constraint. If the initial fee schedule is optimal, the sum of these effects has to be zero.

Mechanical Revenue Effect First of all, each *infra-marginal* family that enrolls their child in public preschool and who has $\omega < \omega'$, will now pay $dF'_{pu}dy$ dollars less in preschool fees. We call these families infra-marginal because they would enroll their children in public preschool regardless of this small reform. The impact on the government's objective from this effect is

$$\Delta_M(\omega') = -dF'_{pu}dy \times \int_{\underline{\omega}}^{\omega'} s_{pu}(\omega)dF(\omega).$$

Labor Supply Effect All public preschool parents with income $y^*_{pu}(\omega')$ now face a higher (implicit) marginal tax rate. We follow Piketty (1997) and Saez (2001) on how to formalize this: These parents will change their behaviour according to[8]

$$\frac{\partial y^*_{pu}(\omega')}{\partial \tau} dF'_{pu} = -\varepsilon_{y,1-\tau} \frac{y^*_{pu}(\omega')}{1-\tau(\omega')} dF'_{pu}.$$

Whereas this labor supply response of the parents has no first-order effect on their utility because of the envelope theorem, it does have a first-order effect on the government's objective. At the margin, of each additional dollar earned, the government obtains $\tau_{pu}(\omega') = T'(y^*_{pu}(\omega')) + F'_{pu}(y^*_{pu}(\omega'))$. Per individual, the effect on the government's objective is given by:

$$-\tau(\omega')\varepsilon_{y,1-\tau}(\omega') \frac{y^*_{pu}(\omega')}{1-\tau(\omega')} dF'_{pu}. \tag{9}$$

What is the mass of individuals that change their labor supply? It reads as

$$s_{pu}(\omega') \times f(\omega') = s_{pu}(\omega') \times f(\omega') \frac{dy}{\varepsilon_{y,\omega}(\omega')} \frac{\omega'}{y(\omega')} .d\omega$$

Multiplying the effect per individual (9) with the mass of individuals yields the overall effect on preschool enrollment of this implied labor supply change:

$$\Delta_{LS}(\omega') = -\frac{\tau(\omega')}{1-\tau(\omega')} \varepsilon_{y,1-\tau} \omega' \frac{1}{\varepsilon_{y,\omega}} \times s_{pu}(\omega') f(\omega') \times dy dF'_{pu}.$$

Note that this negative fiscal effect is proportional to the overall labor wedge $\tau(\omega) = T'(y_{pu}(\omega)) + F'_{pu}(y_{pu}(\omega))$, which shows the importance of the progressivity of the tax-transfer system.

Enrollment Effect Public preschool enrollment will rise because of the fee reduction. Some of the newly enrolled children will have otherwise not been enrolled in preschool at all. The measure of such children is the what we term the 'enrollment effect'. That is, some parents with $\omega < \omega'$ have been indifferent between sending their children to public preschool or to no preschool at all. However, at the same time some parents with $\omega < \omega'$ have been indifferent between sending their children to public preschool or to private preschool and now switch to public preschool because of the fee reduction. This shift from private to public enrollment is not an increase in enrollment and has to be subtracted. The overall enrollment effect on the government's objective is therefore given by

[8] $\varepsilon_{y,1-\tau}$ is the elasticity of labor income w.r.t. ones minus the labor wedge. Relatedly $\varepsilon_{y,\omega}$ is the elasticity of labor income w.r.t. to the wage.

$$\Delta_E(\omega') = \Lambda \int_{\underline{\omega}}^{\omega'} \left(\frac{\partial s_{pu}(\omega)}{\partial F_{pu}(y^*_{pu}(\omega))} + \frac{\partial s_{pr}(\omega)}{\partial F_{pu}(y^*_{pu}(\omega))} \right) d\omega, \tag{10}$$

where we denote by Λ the Lagrange multiplier on the enrollment constraint (8) that translates the enrollment increase into public funds.

Cost Effect The increase in enrollment due to the cost decrease of course has direct fiscal consequences. The effect on the government budget for one child with parental ability ω is

$$\Delta T(\omega) = T(y^*_{pu}(\omega)) - T(y^*_j(\omega)) + F_{pu}(y^*_{pu}(\omega)) - \mathcal{C}$$

with $j = no, pr$. Thus, the overall effect on the government budget is:

$$\Delta_C(\omega') = \int_{\underline{\omega}}^{\omega'} \Delta T(\omega) \frac{\partial s_{pu}(\omega)}{\partial F_{pu}(y^*(\omega))} d\omega. \tag{11}$$

By comparing the enrollment effect (10) and the cost effect (11), we can see clearly the crowding out effect. The share $-\frac{\partial s_{pr}(\omega)}{\partial F_{pu}(y^*_{pu}(\omega))}$ switch from private to public preschool. They *do not* increase enrollment which is why this term is subtracted in (10). They *do* increase government spending because so they are not subtracted in (11).

Optimality If the fee schedule is optimal, the sum of the effects on the social objective has to be zero, thus the necessary conditions for an optimal public preschool fee schedule are:

$$\forall \omega \in \Omega : \Delta_M(\omega) + \Delta_{LS}(\omega) + \Delta_E(\omega) + \Delta_C(\omega) = 0. \tag{12}$$

Rearranging (12) yields the following proposition:[9]

Proposition 1. *Optimal labor wedges for parents who send their children to public preschool are given by:*

$$\frac{\tau_{pu}(\omega')}{1-\tau_{pu}(\omega')} = \left(1 + \frac{1}{\varepsilon_{y,1-\tau}(\omega')}\right) \frac{\mu(\omega')}{s_{pu}(\omega') f(\omega') \omega'}, \tag{13}$$

where

$$\mu(\omega') = \int_{\underline{\omega}}^{\omega'} \left(-s_{pu}(\omega) f(\omega) + \Delta T(\omega) \frac{\partial s_{pu}(\omega)}{\partial F_{pu}(y^*_{pu}(\omega))} - \Lambda \left(\frac{\partial s_{pu}(\omega)}{\partial F_{pu}(y^*_{pu}(\omega))} + \frac{\partial s_{pu}(\omega)}{\partial F_{pu}(y^*_{pu}(\omega))} \right) \right) d\omega.$$

and Λ is implicitly defined by $\mu(\overline{\omega}) = 0$.

This condition for the optimal labor wedge trades-off the labor supply margin, the public/private preschool margin and the public/no preschool margin. It is formally similar to

[9]One also has to use that $\varepsilon_{y,\omega}(\omega) = \varepsilon_{y,1-\tau}(\omega)$.

conditions for optimal redistributive taxes such as in the pioneering papers of Diamond (1998) and Saez (2001) and even closer to those papers that also studied an extensive margin in addition. The latter include Saez (2002) and Jacquet, Lehmann, and Van der Linden (2013) who consider the labor force participation margin, Lehmann, Simula, and Trannoy (2014) who consider migration and Scheuer (2014) who considers the occupational choice margin. The motives for labor supply distortions here, however, are very different. The usual redistribution motive coming from differences in marginal utility of income is shut down here as we deliberately choose this non-welfarist objective.

Note that Proposition 1 does not tell us explicitly what the optimal fee schedule is, but only about the optimal labor supply distortion (wedge). The way that this translates into the steepness of the preschool fee schedule depends on the pre-existing marginal tax rates:

Corollary 1. *The optimal public preschool fee schedule is described by*

$$F'_{pu}(y^*_{pu}(\omega)) = \tau_{pu}(\omega) - T'(y^*_{pu}(\omega)).$$

These theoretical derivations have shed light on the different forces at work and how they should optimally be traded-off. The interesting question is now what these results imply quantitatively. How steep should the optimal preschool fee schedule be? Or should it maybe even be decreasing because pre-exisiting labor supply distortions from the tax-transfer system are already very large? And how much better do optimal policies compare to current policies?

3 Model Estimation

3.1 Data: Early Childhood Program Participation Survey

A suitable data set for estimating our model must include data on costs of preschool attendance and must allow us to distinguish between public and private preschool enrollment. To this end we use data from the Early Childhood Program Participation Survey (ECPP), which is conducted by the Nation Center for Education Statistics (NCES). Alternative data sets with information on preschool enrollment that we considered include the Current Population Survey (CPS) October Supplement (used by Cascio and Schanzenbach (2013)) and the Children of the National Longitundinal Survey of Youth (CNLSY) (used by Carneiro and Ginja (2014)). The disadvantages of these data sources for our purposes are that the CPS does not report costs of preschool, while the CNLSY does not allow a distinction between public and private preschool.

The ECPP has collected cross-sectional data on children and their families from birth to age six. The survey has been conducted five times in 1991, 1995, 2001, 2005 and 2012. We focus on the 2001 and 2005 waves of the survey, which gives us a sample of 2344 four year-old children. We focus on these years because the policy environment is relatively stable. After 2005 we would be worried about changes in the policy environment in e.g. Florida, and before 2001 we would be worried about changes in the policy environment in e.g. Oklahoma. Despite restricting ourselves to these two waves of data, our sample size is still large enough to have precision in our model estimation.

The ECPP collects detailed information on participation in both center based care, i.e. day care, and center based pre-K programs, i.e. preschool. Preschool is clearly distinguished from day care in the data. Many children in the sample participate in multiple programs (up to four are observed), and this often results from participation in a preschool programs and a separate day care arrangement. The location of preschool programs is observed. Nearly 90% of the sample reports the location as either "A public preschool or school," "A private preschool or school," "A church or other place of worship," or "A community center." We classify church based preschools as private (and private preschools also as private). We classify preschools based in community centers as public preschools (and public preschools also as public). Several smaller categories that are difficult to classify remain, for example preschool located in "Its own building." We drop unclassifiable observations. For each program that a child is enrolled in we observe the cost to the parents. The treatment of these data is described in detail in section 3.2.

Many characteristics of children and their families are also observed in the data. We include some of these as covariates to explain the estimated variation in preschool preferences. We restrict ourselves to covariates that would not be directly affected by a policy change. Specifically, we include an indicator for whether one of the parents has a college degree, and indicator for whether the child has any type of developmental delay or learning impairment, and the number of siblings the child has.

Lastly, the data set includes the income of the household. Income must be observed for our analysis because distributions of preferences vary with parents ability to earn income. However, income is not perfectly observed because the data are reported in increments of $5,000, i.e. $0-$5,000, $5,001-$10,000, and so on. We take the mid-points of the intervals to be the household's income. One aspect of our analysis that this affects is classification of households into those eligible for Head Start. This is because for a given family size the federal poverty line always lies within one of the intervals. Our approach is to use the nearest threshold to the poverty line. For example, in 2005 the poverty line for a family of four was $19,350, and so if $15,001-$20,000 is the observed income we classify them as eligible. However, for a family of three in 2005 the poverty line was $16,090, and so based on the closest threshold rule if $15,001-$20,000 is the observed income we classify them as ineligible for Head Start. We have also estimated versions using the next lowest threshold and next highest threshold as classification rules, and there is little impact on the results.

3.2 Fee distributions

Although preschool fees are potentially continuous in the data, we treat the true underlying fees as discrete and measured with error. In particular, we split fees into $J = 12$ bins, which we index by j. The first cost bin includes observations of fees of exactly \$0. The remaining bins include observations within \$1000 increments (i.e. \$1-1000, \$1001-2000, etc.) and the final bin including all observations of fees exceeding \$10,000. There are separate distributions for public and private fees, but both distributions use the same bin cut-offs. The average fee paid within each bin, as well as the frequency of observations within each bin, are presented in Table 1. We take this means within these bins as given below, and treat fee observations as the discrete variable indicating which bin the household is observed in. As such, part of our empirical exercise will involve estimating the true probability mass function of fee offers that generates this distribution of observed fees. Alternatively, we would need to specify a continuous distribution of the true underlying fee offers, which would have the disadvantage of being in many ways less flexible than the distribution we estimate. For example, our approach allows for a large mass of zero fee observations, which most continuous distributions would not allow.

Table 1: Distribution of Discretized Fees

	Public		Private	
bin	frequency	avg. fee	frequency	avg. fee
\$0	536	\$0	65	\$0
\$1-1000	82	\$611	135	\$751
\$1001-2000	43	\$1417	207	\$1450
\$2001-3000	34	\$2438	94	\$2464
\$3001-4000	28	\$3469	63	\$3564
\$4001-5000	27	\$4523	49	\$4554
\$5001-6000	21	\$5546	41	\$5648
\$6001-7000	9	\$6695	24	\$6654
\$7001-8000	14	\$7441	18	\$7452
\$8001-9000	5	\$8473	12	\$8603
\$9001-10,000	1	\$9100	8	\$9646
\$10,001+	4	\$11,250	10	\$14,124

We denote by $\xi_{j,pu}$ and $\xi_{j,pr}$ the probabilities of a household drawing fee bin j for public and private preschools, respectively. We should emphasize that these probabilities are not equal to the observed frequencies reported in Table 1, but rather are estimated along with the other model parameters. The reason they are not equal to the observed frequencies is that the propensity to opt into a particular type of preschool decreases as the fee grows. We also use the notations $F_{pu,j}$ and $F_{pr,j}$ to represent the costs associated with public and private costs within bin j. For example, if a household draws public cost bin 3 their public fee would be $F_i(pu, y_i^*) = F_{pu,3}$.

Head Start and the Poverty Line: The federal Head Start program implies that the accessibility of free public preschool is mich higher for households below the poverty line than above it. We capture this by introducing an additional probability term ξ_{hs} that re-weights the fee distribution below the poverty line towards the zero fee possibility. This has the effect for households below the poverty line the probability of being offered cost-free public preschool is $\xi_{hs} + (1 - \xi_{hs})\xi_{pu,1}$, the probability of drawing the second cost bin is $(1 - \xi_{hs})\xi_{pu,2}$, and so on. This distinction between above and below poverty line policy environments provides quasi-experimental variation that is crucial to our identification strategy.

3.3 Preschool Preference Heterogeneity Distributions

Above we introduced θ_i and $\theta_{pr,i}$ as notation for monetized values derived from preschool choices. As we will discuss in detail below, our model is identified if we restrict ourselves to a two parameter distribution function for θ_i and a single parameter distribution function for $\theta_{pr,i}$. To this end, we assume the general preschool preference in normally distributed: $\theta_i \sim N(\bar{\theta}(y_i, X_i), \sigma_\theta)$, where the mean depends on parental income and a set of exogenous covariates X_i (these are described in the data section). Furthermore, we assume the additional utility from private preschool is exponentially distributed with a mean that depends on income and covariates: $\theta_{pr,i} \sim \mathrm{Exp}(\lambda(y_i, X_i))$. Although these distributional assumptions are restrictive, our identification does not rely specifically on them, rather the small number of parameters they depend on. For example, we have also estimated (see Appendix C.2) a version of our model with $\theta_{pr,i}$ distributed triangularly, with little effect on our results.

The structure assumed in the previous paragraph needs to be reconciled with the notation introduced in section 2.2 in two ways. First, and most important, is that in section 2.2 the distributions were conditional on unobservable ability ω_i, whereas they are now conditional on observable income y_i. When we estimate we do in fact condition on y_i – including income as a covariate that shifts mean preferences – and then later when we analyze policy we use household optimal labor supply rules to recover distributions that depend on ω_i as originally suggested by Saez (2001).[10] By following this procedure, we do not force our parameter estimates to depend on particular assumptions about labor supply elasticities, and can easily conduct robustness exercises that consider alternative ways to recover $G_\theta(\theta_i|\omega_i, X_i)$ and $G_{pr}(\theta_{pr,i}|\omega_i, X_i)$ from the estimated distributions.

[10] We use the Gouveia-Strauss specification of Guner, Kaygusuz, and Ventura (2014, table 12) for the marginal tax rates. We use the specification including local sales taxes and take the average over all individuals.

We also need an extension of the notation in order to keep distributions depending on income and ability distinct. It is natural to now use the common standard normal CDF and PDF notations $\Phi(\cdot)$ and $\phi(\cdot)$, respectively, to denote the distribution of θ_i conditional on y_i and X_i. For private preschool preferences we adjust the notation by using $G^y_{pr}(\theta_{pr,i};\lambda(y_i,X_i))$ and $g^y_{pr}(\theta_{pr,i};\lambda(y_i,X_i))$ to denote the CDF and PDF of exponential distribution with mean $\lambda(y_i,X_i)$.

3.4 The Likelihood Function

3.4.1 Likelihood of Observing Public Preschool

When we observe a child is enrolled in public preschool we also observe their cost of public preschool $F_i(pu,y_i)$, but not their counterfactual cost of private preschool $F_i(pr,y_i)$. Therefore, private fees are treated as unobserved heterogeneity and integrated out. The likelihood of the public preschool observation is thus equal to:

$$\begin{aligned}
&\Pr\left[ps_i = pu, F_i(pu,y_i) = F_{pu,k} | y_i, X_i\right] = \\
&= \sum_{j=1}^{J} \xi_{pr,j} G^y_{pr}(F_{pr,j} - F_{pu,k}; \lambda(y_i,X_i)) \times \left(1 - \Phi\left(\frac{F_{pu,k} - \bar{\theta}(y_i,X_i)}{\sigma_\theta}\right)\right) \times \xi_{pu,k}.
\end{aligned} \tag{14}$$

This expression is a weighted average over possible counterfactual private fees of the probability the private fee is too large for i to prefer private preschool over public, multiplied by the probability the public fee is low enough that public preschool is preferred over none. Lastly, we multiply by the PMF of public fees at the observed draw.

Below the Poverty Line: Adjusting for the additional probability of free public preschool being available in the form of Head Start, we have the following likelihood of observing private preschool chosen for those below the poverty line:

$$\begin{aligned}
&\Pr\left[ps_i = pu, F_i(pu,y_i) = F_{pu,k} | y_i, X_i, y_i \le y_{pov}\right] = \\
&= \sum_{j=1}^{J} \xi_{pr,j} G^y_{pr}(F_{pr,j} - F_{pu,k}; \lambda(y_i,X_i)) \times \left(1 - \Phi\left(\frac{F_{pu,k} - \bar{\theta}(y_i,X_i)}{\sigma_\theta}\right)\right) \\
&\times \begin{cases} \xi_{hs} + (1-\xi_{hs})\xi_{pu,1} & if \quad F_i(pu,y_i) = 0 \\ (1-\xi_{hs})\xi_{pu,k} & if \quad F_i(pu,y_i) > 0 \end{cases}.
\end{aligned} \tag{15}$$

3.4.2 Likelihood of Observing Private Preschool

As was the case for public preschool, we only observe the private fee paid $F_i(pr,y_i)$ if the child is enrolled in private preschool, but not their counterfactual cost of public preschool. Again, we treat the cost of public as unobserved heterogeneity and thus can derive the likelihood of a private preschool observation:

$$\begin{aligned}
&\Pr\left[ps_i = pr, F_i(pr, y_i) = F_{pr,k} | y_i, X_i\right] = \\
&= \sum_{j=1}^{J} \xi_{pu,j} \int_{F_{pr,k}-F_{pu,j}}^{\infty} \left(1 - \Phi\left(\frac{F_{pr,k} - \theta_{pr,i} - \bar{\theta}_i(y_i, X_i)}{\sigma_\theta}\right)\right) g_{pr}^{y}(\theta_{pr,i}; \lambda(y_i, X_i)) d\theta_{pr,i} \\
&\times \xi_{pr,k}.
\end{aligned} \tag{16}$$

Note that it is understood that $g_{pr}^{y}(\theta_{pr,i}|y_i, X_i) = 0$ outside the support of $\theta_{pr,i}$. This likelihood component depends on the probability that $\theta_{pr,i}$ is large enough to overcome the difference in private versus public fees – which could be unity if private fees are smaller – and on the probability that θ_i is large enough that private is preferred over no preschool.

Below the Poverty Line: Adjusting for the additional probability of free public preschool being available in the form of Head Start, we have the following likelihood of observing no preschool chosen for those below the poverty line:

$$\begin{aligned}
&\Pr\left[ps_i = pr, F_i(pr, y_i) = F_{pr,k} | y_i, X_i, y_i \leq y_{pov}\right] = \\
&= \left[\xi_{hs} + (1 - \xi_{hs})\xi_{pu,1}\right] \left\{ \int_{F_{pr,k}}^{\infty} \left(1 - \Phi\left(\frac{F_{pr,k} - \theta_{pr,i} - \bar{\theta}_i(y_i, X_i)}{\sigma_\theta}\right)\right) \cdot \right. \\
&\left. \cdot\, g_{pr}^{y}(\theta_{pr,i}; \lambda(y_i, X_i)) d\theta_{pr,i} \right\} \times \xi_{pr,k} + \\
&+ \sum_{j=2}^{J} (1 - \xi_{hs})\xi_{pu,j} \left\{ \int_{F_{pr,k}-F_{pu,j}}^{\infty} \left(1 - \Phi\left(\frac{F_{pr,k} - \theta_{pr,i} - \bar{\theta}_i(y_i, X_i)}{\sigma_\theta}\right)\right) \cdot \right. \\
&\left. \cdot\, g_{pr}^{y}(\theta_{pr,i}; \lambda(y_i, X_i)) d\theta_{pr,i} \Big) \right\} \times \xi_{pr,k}.
\end{aligned} \tag{17}$$

3.4.3 Likelihood of Observing No Preschool

Neither $F_i(pu, y_i)$ nor $F_i(pr, y_i)$ is observed when $ps_i = no$ is chosen, thus both are treated as unobserved heterogeneity. The likelihood of a *no* observation is:

$$\begin{aligned}
&\Pr\left[ps_i = no | y_i, X_i\right] = \\
&= \sum_{j=1}^{J} \sum_{k=1}^{J} \xi_{pu,j} \xi_{pr,k} \int_{-\infty}^{F_{pu,j}} G_{pr}^{y}\left(F_{pr,k} - \theta_i; \lambda(y_i, X_i)\right) \phi\left(\frac{\theta_i - \bar{\theta}(y_i, X_i)}{\sigma_\theta}\right) d\theta_i.
\end{aligned} \tag{18}$$

The likelihood that no preschool is chosen depends on the probability that θ_i and $\theta_{pr,i}$ are both sufficiently small.

Below the Poverty Line: Adjusting for the additional probability of free public preschool being available in the form of Head Start, we have the following likelihood of observing public preschool chosen for those below the poverty line:

$$
\begin{aligned}
&\Pr\left[ps_i = no | y_i, X_i\right] = \qquad (19)\\
&= \left[\xi_{hs} + (1-\xi_{hs})\xi_{pu,1}\right] \sum_{k=1}^{J} \xi_{pr,k} \int_{-\infty}^{0} G_{pr}^{y}\left(F_{pr,k} - \theta_i; \lambda(y_i, X_i)\right) \phi\left(\frac{\theta_i - \bar{\theta}(y_i, X_i)}{\sigma_\theta}\right) d\theta_i + \\
&\sum_{j=2}^{J} (1-\xi_{hs})\xi_{pu,j} \sum_{k=1}^{J} \xi_{pr,k} \int_{-\infty}^{F_{pu,j}} G_{pr}^{y}\left(F_{pr,k} - \theta_i; \lambda(y_i, X_i)\right) \phi\left(\frac{\theta_i - \bar{\theta}(y_i, X_i)}{\sigma_\theta}\right) d\theta_i.
\end{aligned}
$$

3.5 Overall Parameter Identification

Each observation in our sample can be classified into one of the three situations for which the likelihood is given above. We take the logs of the applicable likelihood and sum across all observations, then maximize this sum. Maximization is subject to constraints $\sum_{j=1}^{J} \xi_{pu,j} = 1$ and $\sum_{j=1}^{J} \xi_{pr,j} = 1$. We explain how maximization of this likelihood identifies all of the parameters in the following section.

The easiest way to understand our identification is to first consider a case in which there are unique public and private fees, i.e. $J = 1$, and no dependence of preference distributions on covariates. Head Start is still possibly available for free for households below the poverty line. In this example there will be four parameters to identify: $\bar{\theta}$, σ_θ, λ and ξ_{hs}. There will also implicitly be four probabilities with which to identify these parameters: the probabilities of observing public and private preschool above the poverty line, and the probabilities of observing public and private preschool below the poverty line. Therefore, in this simplified example, a unique combination of the four parameters would align the model and raw data sample probabilities.

We should emphasize how important quasi-experimental variation is for identification of the parameters of our model. Without the sharp change in the policy environment at the poverty line we would only have two probabilities with which to identify three distributional parameters. The means of θ_i and $\theta_{pr,i}$ could be identified, but the variance of θ_i would not be. Obviously, without information about the extent of variation in households valuations of preschool we would have a very limited understanding of how preschool enrollment would respond to policy changes because the density of 'marginal' enrollees would be unknown. However, by leveraging the Head Start qualification threshold to learn about variation in preschool preferences we are able to understand responsiveness to policy changes. Put differently, we are using an observed policy variation to make inferences about behaviour, which allows us to then study counterfactual policy changes.

Adding variation in fees, i.e. the $\xi_{pu,j}$ and $\xi_{pr,j}$ parameters, adds one empirical probability per parameter. The reason there is only one additional empirical probability per additional discrete fee level is that fees are only observed for the chosen form of preschool. That is, adding an additional level of public preschool fees adds an empirical probability by introducing a variation of the public preschool enrollment rate, but does not add an observable variation of the private preschool enrollment rate because public fees are integrated out when computing the likelihood of a privately enrolled observation.

Lastly, the influence of covariates on mean preferences for public preschool is modelled as linear, i.e. $\bar{\theta}(y_i, X_i) = \beta_{\bar{\theta}}' X_i + \beta_{\bar{\theta}}^y y_i$ is modelled as linear. The mean preference for private preschool is modelled as log-linear, i.e. $\lambda(X_i) = \exp(\beta_\lambda' X_i + \beta_\lambda^y y_i)$, in order to ensure a non-negative mean. The loadings in the vectors $\beta_{\bar{\theta}}$ and β_λ will obviously be identified by covariation between the corresponding variables and preschool enrollment propensities.

3.6 Parameter Estimates

Table 2 presents our estimates of the main model parameters. In Appendix B we provide estimates of the fee distribution PMF, i.e. the $\xi_{pu,j}$ and $\xi_{pr,j}$ parameters. Because of its importance for our empirical strategy, we present the effect of Head Start eligibility on the probability of free public preschool here in the main results.

Table 2: Maximum Likelihood Parameter Estimates

Parameter	Estimate	Standard Error
Parameters of θ_i Distribution		
$\beta_{\bar{\theta}}$ - Constant	2.1253	(0.3408)
$\beta_{\bar{\theta}}$ - College Parent	1.1430	(0.4092)
$\beta_{\bar{\theta}}$ - Number of Siblings	-0.9438	(0.1663)
$\beta_{\bar{\theta}}$ - Developmental Index	2.1325	(0.5224)
$\beta_{\bar{\theta}}^{y}$ - Income (\$1000s)	0.0399	(0.0057)
σ_θ	4.4828	(0.7037)
Parameters of $\theta_{pr,i}$ Distribution		
β_λ - Constant	-0.7488	(0.2467)
β_λ - College Parent	1.2478	(0.2109)
β_λ - Number of Siblings	-0.3363	(0.1192)
β_λ - Developmental Index	-1.2179	(0.4806)
β_λ^{y} - Income	0.0163	(0.0026)
Head Start Eligibility Effect		
ξ_{hs}	0.5295	(0.0599)

Notes: Mean preferences are $\bar{\theta} = \beta_{\bar{\theta}}' X_i$ and $\lambda = \exp(\beta_\lambda' X_i)$. Standard errors are based on the BHHH estimator for the Information Matrix. The estimated parameters of the fee distributions are presented in Appendix table 3.

Interpretation of the parameters of the θ_i distribution is straightforward because preferences are in monetary terms. The constant indicates that for a child whose parents have zero income, for whom neither parent has a college degree, who has no siblings, and who does not have a developmental problem, the value of preschool to their parents is \$2,125. When one of the parents has a college degree, their valuation of preschool increases by \$1,143, still holding income constant at zero. Each additional sibling reduces the valuation of preschool by \$944, and if the child has a developmental problem the valuation of preschool increases by \$2,133. Parents with higher incomes have greater valuations of preschool, to the extent that a household with \$100,000 of income values preschool \$4,000 more than a household with zero income. There is also wide variation in valuations of preschool as indicated by the standard deviation of θ_i being \$4,483.

Whereas θ_i reflects general preferences for preschool, $\theta_{pr,i}$ reflects differential valuations of private versus public preschool. Interpretation is somewhat more difficult because of the log-linear specification in this dimension. The constant indicates a moderate baseline valuation of private preschool of \$473. If one of the parents has a college degree the valuation of private versus public preschool increases considerably. Additional siblings or a developmental problem lowers parents valuations of private versus public preschool, and additional income moderately increases this relative preference.

When interpreting these results it is important to remember that valuations are relative to marginal utility from wealth. Thus, preschool preferences increasing with income can be due to a combination of greater ability to pay for preschool and other deeper preferences, such as altruism.

The last parameter estimate reported in Table 2 is the additional probability of being offered free public preschool if the family is below the poverty line. The point estimate is 0.53, and the standard error is quite small relative to this. More formally, we re-estimate a restricted version of the model with $\xi_{hs} = 0$, and the Likelihood-Ratio statistic is 50.8, clearly indicating an important role for this parameter. We emphasize this result because it is evidence that the policy variation we are leveraging in order to identify our model is substantial.

3.7 Comparisons to Other Data and Results

To demonstrate the suitability of our model for the type of policy analysis we intend to use it for, we simulate the effects of a previously studied policy change on the behavior of households in our model. We also discuss here the fit of our model to data around the Head Start eligibility threshold.

3.7.1 Oklahoma and Georgia Universal Preschool

In 1995 and 1998 the states of Georgia and Oklahoma, respectively, introduced universal preschool programs. The effects of these program introductions on preschool enrollment were studied in detail by Cascio and Schanzenbach (2013). Using a difference in differences framework with control states, Cascio and Schanzenbach found that introducing free public preschool increased the overall preschool enrollment rate by 12-15 percentage points. However, they also found evidence of considerable 'crowding-out' of private preschool investment.

We can easily simulate the effect of making public preschool universally free within the context of our model, and when we do so the overall preschool enrollment rate increases by 11.8 percentage points. We also find that considerable crowding out of private preschool enrollment occurs, with private preschool enrollment falling by 16.5 percentage points (as a proportion this reduction is from 48.8% of all preschool enrollment being private to only 19.8%). However, it is difficult to compare the rate simulated in our model to the findings in Cascio and Schanzenbach (2013) because the Georgia universal preschool program has a large voucher component. Public vouchers can be used to pay for private preschool in Georgia, meaning that the observed reductions in private preschool enrollment underestimate the extent to which there was crowding out in Georgia. Indeed, the extent of crowding out in our model is somewhat larger than Cascio and Shanzenbach estimate for Georgia and Oklahoma combined, but is somewhat smaller than the reduction in the average private preschool enrollment rate observed in CPS data for

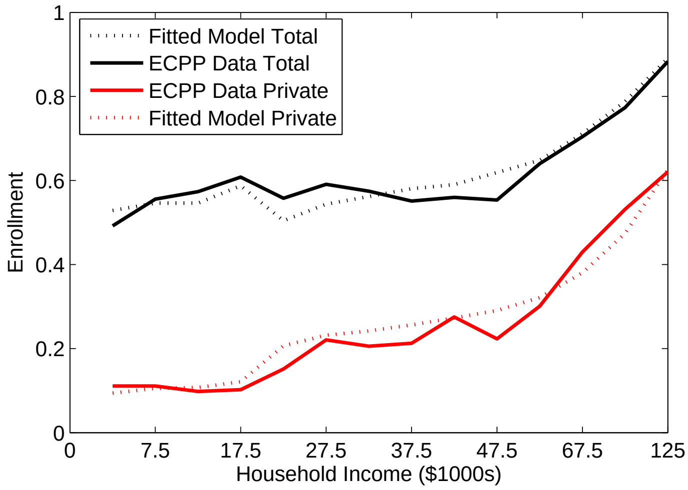


Figure 2: Overall and Private Preschool by Income.

Oklahoma alone. In our own analysis of October CPS Education Supplement data for Oklahoma we found that pre-1998 46.7% of four year-old preschool enrollment was in private preschools whereas only 17.4% of enrollment was in private preschool post-1998. These numbers do not control for trends in other states and so likely over represent the degree of crowding out in Oklahoma, but nonetheless show that crowding-out in our model is reasonable. (We discuss crowding-out again below in the context of the Head Start threshold.)

3.7.2 Variation Around the Head Start Eligibility Threshold

A big component of our model identification involves using policy variation implied by the Head Start eligibility threshold at the poverty line. Here we analyze how our model performs in that dimension be examining how well it fits the changes of preschool enrollment at the threshold. Because of variation in household size, there is variation in the income levels at which Head Start eligibility ends, but for most households the cutoff is between the income ticks at $17,500 and $27,500 on Figure 2. As the figure shows, in the ECPP data private preschool increases sharply over this income range, but overall preschool enrollment is stable or possibly falling slightly. The model replicates this pattern very well, and this is entirely because of the substantial shift in the policy environment implied by ξ_{hs} (see section 3.6 for an LR-test of the importance of this policy variation). Overall, we conclude from this that our model replicates changes in preschool enrollment behaviour around the poverty line well, and emphasize that this is exactly how we identify the responsiveness of behavior to the counterfactual policies we will analyze in the remainder of this paper.

3.7.3 Head Start Impact Study

The Head Start Impact Study (HSIS) was a randomized experiment in which a nationally representative sample of three- and four-year old Head Start applicants were randomly accepted or declined admission into over-subscribed Head Start centers. Those who were declined admission were allowed to pursue other substitute preschool options. The aspect of substitute programs was analyzed by Kline and Walters (2015), who found that total preschool enrollment (including enrollment in non-Head Start alternatives) was 39.3% lower among those randomly declined Head Start admission.

In our model the probability that a child below the poverty line has access to free public preschool is $\xi_{hs} + (1 - \xi_{hs})\xi_{pu,1}$. While it is tempting to interpret ξ_{hs} as the component of this due to Head Start access, this would not be correct. Our estimation procedure does not separately identify the probabilities Head Start and non-Head Start access below the poverty line, rather the joint probability of access to either. If Head Start crowds-out other free public preschool programs that are available for children above the poverty line, then the probability of Head Start access would be larger than ξ_{hs}.[11] This implies that we should not simulate the HSIS by reducing the probability of free public preschool by ξ_{hs}, but rather some larger number. Because we do not know what that larger number is we bound the effect of the HSIS by first just setting $\xi_{hs} = 0$, and second by setting the entire probability $\xi_{hs} + (1 - \xi_{hs})\xi_{pu,1} = 0$ (and adjusting the remaining public fee probabilities proportionally).[12] In the first case overall preschool enrollment falls by 25.6%, and in the second case overall preschool enrollment falls by 47.1%. Because the actual result of the HSIS was roughly in the middle of these two numbers we conclude that our model captures this dimension of the policy environment well.

4 Optimal Preschool Policies in the U.S.

4.1 Benchmark Case

Figure 3 illustrates the optimal fee schedule as a function of family income. It increases concavely with an average marginal preschool fee of 2%. As opposed to the adjusted benchmark, not surprisingly, the schedule is much smoother. Parents below the poverty line would pay higher fees whereas families with income between $20,000 and $60,000 would pay lower fees. The fee increases up to an income of around $200,000 and then stays constant.

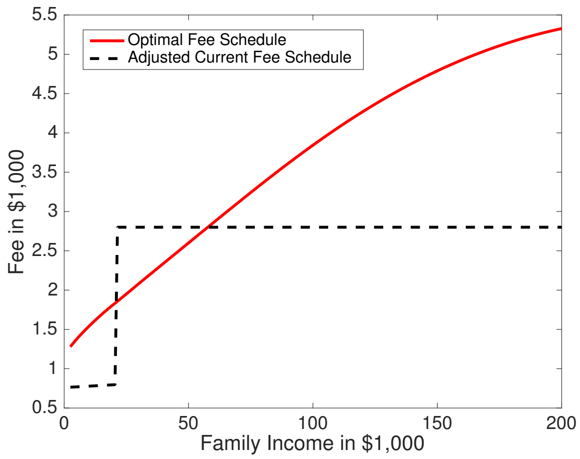


Figure 3: Optimal Fee Schedule

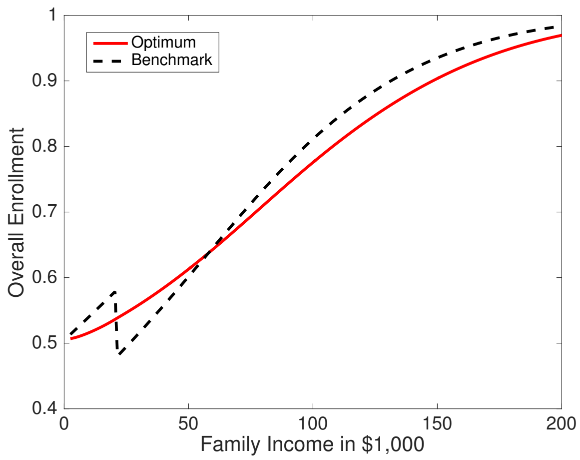
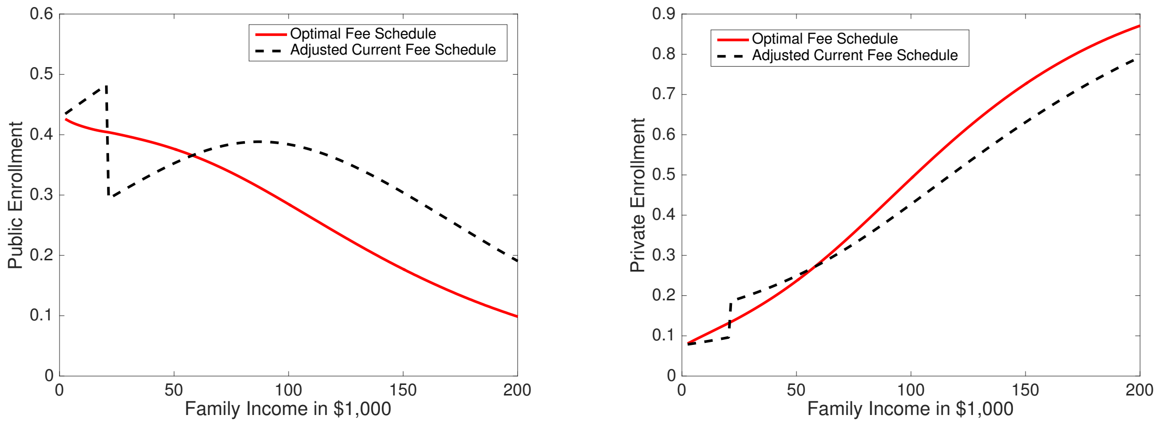


Figure 4: Overall Enrollment as a Function of Family Income



(a) Public Preschool Enrollment

(b) Private Preschool Enrollment

Figure 5: Decomposition of Enrollment

The resulting enrollment is illustrated in Figure 4 and mirrors the fee schedule: enrollment increases for families between $20,000 and $60,000 and decreases for the other income levels. By construction the overall level of enrollment is fixed. So the questions are how much better the optimal schedule performs in terms of costs and how it affects private and public enrollment. We first look at the latter question. Figure 5 illustrates public and private enrollment as a function of family income. The left panel shows public enrollment and again mirrors the fee schedule from Figure 3. Public enrollment is higher (lower) for income levels where public fees are lower (higher) than in the adjusted benchmark. Importantly, an increase (or a decrease) in public enrollment at a certain income level can be decomposed into a decrease (or an increase respectively) in private enrollment and the share of children who do not visit any preschool. As a consquence, the difference between private enrollment in the optimum and the adjusted benchmark is not as high as for public enrollment. What does all this imply for cost savings?

We find that public expenditure for public preschool decreases by 5.33% through better targeting of subsidies. This change in costs can be decomposed into three parts. This is illustrated in Figure 6. First of all, holding public enrollment constant, parents pay different fees in the optimum than in the current system. In particular, many richer parents that were inframarginal in their decision pay higher fees in the optimum. This amounts to 6.4% less of public spending. At the same time, some parents were at the margin of the decision whether to send their child to private or public preschool and send their child now to private preschool given the higher fee. This effect amounts to approximately 1.2% of less public spending. Finally, the

[11] This would be true if, for example, there were a positive probability that a community group would organize a replacement in the event that a Head Start center closed-down.

[12] In the estimated model there is a 67.1% chance of access to free public preschool below the poverty line. The first version of the HSIS simulation reduces this to 30.6%, and the second version further reduces this to 0%.

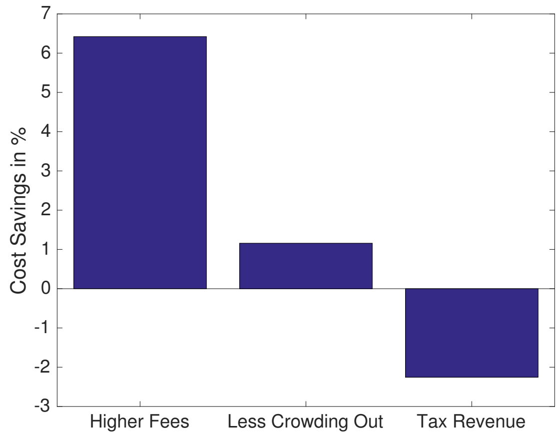


Figure 6: Decomposition of Cost Savings

steep fee schedule implies a lowering of income for all parents that do send their child to public preschool. This change in earnings does not only lower the fee they pay (which is accounted for by (i)) but also the amount of taxes they pay. This implies a decrease in tax revenue that is 2.3% of the public spending.

4.2 Labor Supply Incentives

The previous section has revealed that labor supply incentives do matter. For example, we have seen that the optimal fee schedule implies a reduction in tax revenue that is about 2% of the preschool budget. We now elaborate the role of labor supply further by exploring to what extent the shape of the fee schedule is determined by the endogeneity of income. We therefore simulate the optimal policy under the assumption that income is exogenous. The optimal fee schedule for this case is given by the black dashed-dotted line in Figure 7; the optimal fee taking into account labor supply incentives is still given by the red bold line. One can see that the endogeneity of income implies that the optimal fee schedule is less progressive. The change in the average marginal preschool fee is about 0.7 percentage points.

4.3 The Interaction with other Transfer Programs

The importance of endogenous labor supply depends on the preexisting distortions through the tax-transfer systems. The implied decrease in income because of a steep fee schedule has more severe budget effect if the baseline tax-transfer system is characterized by high marginal tax rates. In our analysis so far, we only incorporated effective marginal tax rates coming from taxes. However, low-income household typically face distortions that are much higher: if they

Figure 7: Importance of Elastic Labor Supply for Optimal Policies

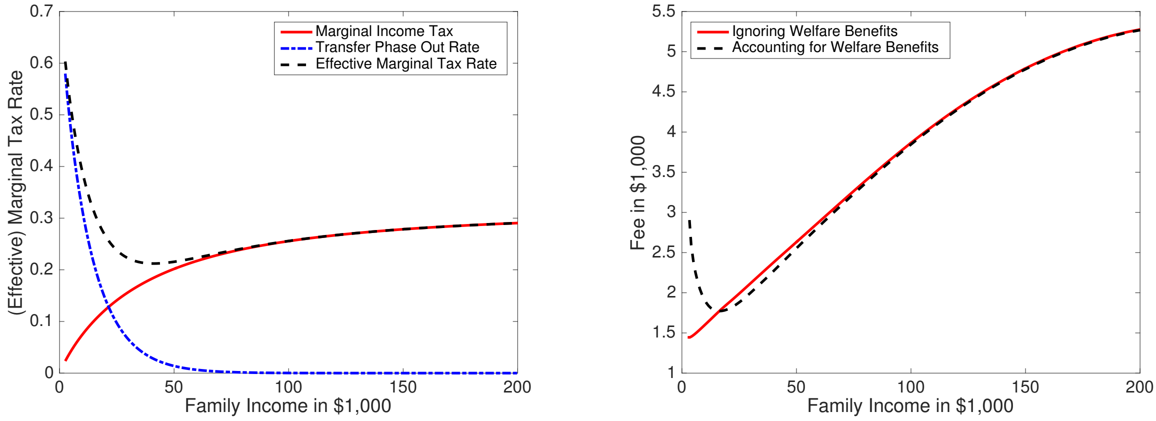
increase their earnings, they receive less transfers in the form of food stamps or temporary assistance for needy families (TANF). This transfer phase-out increases effective marginal tax rates. We take the estimates from Guner, Rauh, and Ventura (2017) on how transfer receipt varies with income. The transfer phase out rates are illustrated by the blue dashed-dotted line in Figure 8(a) and they are as high as 60% which implies that one more dollar of earned income decreases transfer receipt by 60%. Summing these phase-out rates and the marginal income tax rates (red bold line) yields the effective marginal tax rates (black dashed-dotted line).

Figure 8(b) illustrates the optimal fee schedule if the transfer receipt is taken into account (black dashed line). For low incomes, the shape differs drastically from the case where transfer receipt is not accounted for (red bold line). Preschool fees are decreasing in income up to an income of about $20,000. Afterwards the shape is very similar. Figure 10 in Appendix C.1 illustrates the cost savings due to optimal fees. The overall cost savings are 5.05% and therefore a bit lower than in the case where the distortions of the welfare benefits were not taken into account (where the number was 5,33%). The decomposition is also very similar.

4.4 Comparison of Universal Preschool

The idea of a universal preschool which is of free access to everybody is often suggested in the policy debate. How do optimal policies compare to this policy in terms of cost-efficiency? First of all, we need to establish how a universal preschool would perform in our model. As already anticipated in Section 3, we find that setting the fees for the public preschool to zero for everybody yields an increase in enrollment by 13.26 percentage points. Which fee schedule yields the same level of enrollment at minimal costs?

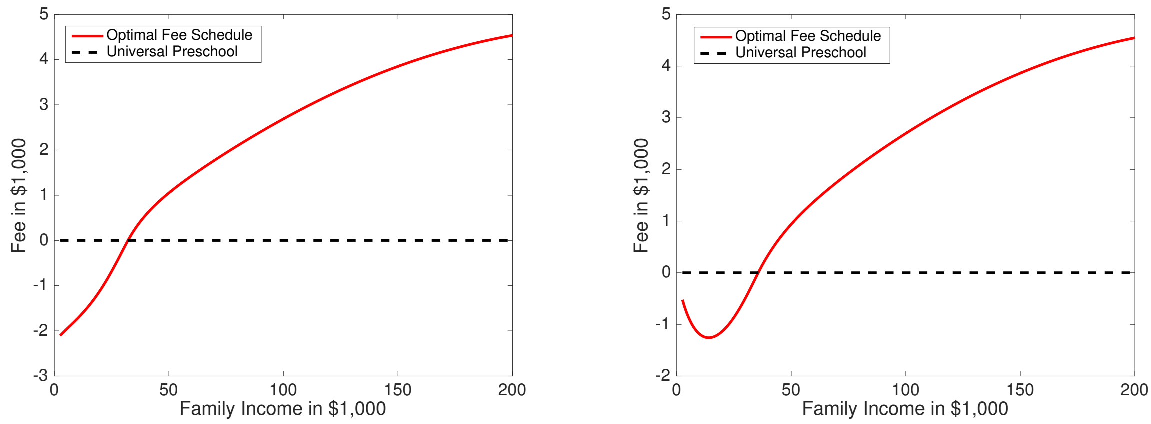
Figure 9 illustrates optimal fees that implement this level of enrollment for the case where transfer-phase out is not taken into account (left panel) and also the case where it is taken



(a) Effective Marginal Tax Rates

(b) Importance of other distortions for optimal fee schedule

Figure 8: Accounting for Distortions from Transfer Programs



(a) Ignoring Transfer Programs

(b) Accounting for Transfer Programs

Figure 9: Fee schedule that implements same level of enrollment as universal preschool

into account (right panel). In both cases, fees are negative for some income levels. This is not surprising: given that the same level of enrollment than under a universal preschool has to be implemented, the fee has to be negative for some income levels if it is positive for other income levels. Poor parents should receive up to $2,000 of transfers per year if they send their child to public preschool (for which they also do not have to pay). This resembles the concept of conditional cash transfers.

The implied cost savings of optimal policies are higher for such high target levels of enrollment. The cost saving is 10.11% and 9.44% for the case without and with welfare benefits respectively. Figure 11 in Appendix C.1 shows the decomposition of the cost savings. Interestingly, the crowding-out component now plays a much bigger role and accounts for more than 6 percentage points in costs savings in both cases.

5 Conclusion

We have taken a mechanism-design perspective on the policy debate about public preschool provision. We have developed a theory of how a public preschool fee schedule should vary with family income. We thereby did not impose any restrictions on the potential shape of such a schedule. Our formulas transparently highlight the trade-off between increasing public preschool enrollment, crowding-out private preschool enrollment and labor supply incentives.

We have estimated our model using microdata from the Early Childhood Program Participation Survey. Identification is achieved through a price discontinuity implied by Head Start. We find that optimal public fees can implement the same amount of enrollment as current policies with 10% less public expenditure. Taking into account the distortions from welfare programs like food stamps etc. is very important for the design of the fee schedule at very low income levels. The optimal schedule should be decreasing in income for very small incomes.

We then turn to the question how higher levels of enrollment can be implemented. The target level we choose is the one that could be implemented by a universal preschool which would imply zero fees for all families that send their children to public preschool. The cost saving is even larger here and 14%. Further, the optimal fee schedule implies negative fees for poor families which resembles conditional cash transfer programs that are often implemented in poorer countries.

In future work it would be interesting to model the private sector with more sophistication. Private preschool providers could for example respond to public providers by increasing quality. Further, we neglected the question of the optimal level of quality. To address such questions, more empirical evidence is needed about returns to the intensive margin in terms of money spend per child at preschool. Lastly, we only asked how the government should target a given level of public funds to children along the parental income distribution, but did not ask how much subsidies the government should pay overall. It would also be interesting to look to what extent higher subsidies to preschool pay for themselves through higher tax revenue in the future in the spirit of Findeisen and Sachs (2016), who look at college education subsidies.

References

Attanasio, O. P., C. Meghir, and A. Santiago (2012): "Education choices in Mexico: using a structural model and a randomized experiment to evaluate Progresa," *The Review of Economic Studies*, 79(1), 37–66.

Barnett, W. S., K. Brown, and R. Shore (2004): "The Universal vs. Targeted Debate: Should the United States Have Preschool for All?," Discussion paper, National Institute of Early Education Research.

Bassok, D., L. C. Miller, and E. Galdo (2016): "The effects of universal state pre-kindergarten on the child care sector: The case of Florida's voluntary pre-kindergarten program," *Economics of Education Review*, 53, 87–98.

Bernal, R. (2008): "The effect of maternal employment and child care on children's cognitive development," *International Economic Review*, 49(4), 1173–1209.

Bick, A. (2015): "The quantitative role of child care for female labor force participation and fertility," *Journal of the European Economic Association*.

Blundell, R., and A. Shephard (2012): "Employment, Hours of Work and the Optimal Taxation of Low-Income Families," *The Review of Economic Studies*, 79(2), 481–510.

Carneiro, P., and R. Ginja (2014): "Long-term impacts of compensatory preschool on health and behavior: Evidence from Head Start," *American Economic Journal: Economic Policy*, 6(4), 135–173.

Cascio, E. U., and D. W. Schanzenbach (2013): "The Impacts of Expanding Access to High-Quality Preschool Education," Discussion paper, National Bureau of Economic Research.

Choné, P., and G. Laroque (2011): "Optimal Taxation in the Extensive Model," *Journal of Economic Theory*, 146(2), 425–453.

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

Currie, J., and D. Almond (2011): "Human capital development before age five," *Handbook of labor economics*, 4, 1315–1486.

Diamond, P. (1998): "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates," *The American Economic Review*, 88(1), 83–95.

Domeij, D., and P. Klein (2013): “Should day care be subsidized?,” *The Review of Economic Studies*, 80(2), 568–595.

Duncan, G. J., and K. Magnuson (2013): “Investing in preschool programs,” *The Journal of Economic Perspectives*, 27(2), 109–132.

Elango, S., J. L. García, J. J. Heckman, and A. Hojman (2015): “Early Childhood Education,” Discussion paper, National Bureau of Economic Research.

Findeisen, S., and D. Sachs (2015): “Education and Optimal Dynamic Taxation: The Role of Income-Contingent Student Loans,” *Working Paper*.

——— (2016): “Optimal Need-Based Financial Aid,” *Working Paper*.

Fitzpatrick, M. D. (2008): “Starting school at four: The effect of universal pre-kindergarten on children’s academic achievement,” *The BE Journal of Economic Analysis & Policy*, 8(1).

Fitzpatrick, M. D. (2010): “Preschoolers Enrolled and Mothers at Work? The Effects of Universal Prekindergarten,” *Journal of Labor Economics*, 28(1), 51–85.

Guner, N., R. Kaygusuz, and G. Ventura (2014): “Income taxation of US households: Facts and parametric estimates,” *Review of Economic Dynamics*, 17(4), 559–581.

Guner, N., C. Rauh, and G. Ventura (2017): “Means-Tested Transfers in the US: Facts and Parametric Estimates,” *Mimeo*.

Heckman, J. J. (2013): *Giving kids a fair chance*. Mit Press.

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

Heckman, J. J., and L. K. Raut (2016): “Intergenerational long-term effects of preschool-structural estimates from a discrete dynamic programming model,” *Journal of econometrics*, 191(1), 164–175.

Jacquet, L., E. Lehmann, and B. Van der Linden (2013): “Optimal redistributive taxation with both extensive and intensive responses,” *Journal of Economic Theory*, 148(5), 1770–1805.

Kanbur, R., M. Keen, and M. Tuomala (1994): “Optimal non-linear income taxation for the alleviation of income-poverty,” *European Economic Review*, 38(8), 1613–1632.

Kanbur, R., J. Pirttilä, and M. Tuomala (2006): “Non-welfarist optimal taxation and behavioural public economics,” *Journal of Economic Surveys*, 20(5), 849–868.

Kline, P., and C. Walters (2015): "Evaluating public programs with close substitutes: The case of Head start," Discussion paper, Mimeo.

Koeniger, W., and J. Prat (2015): "Human Capital and Optimal Redistribution," *Working Paper*.

Lehmann, E., L. Simula, and A. Trannoy (2014): "Tax Me if You Can: Optimal Income Tax Competition between Competing Governments," *Quarterly Journal of Economics (forthcoming)*.

Mirrlees, J. A. (1971): "An exploration in the theory of optimum income taxation," *The Review of Economic Studies*, pp. 175–208.

Piketty, T. (1997): "La redistribution fiscale face au chomage," *Revue française d'économie*, 12(1), 157–201.

Rochet, J.-C., and L. A. Stole (2002): "Nonlinear pricing with random participation," *Review of economic Studies*, pp. 277–311.

Saez, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68(1), 205–229.

Saez, E. (2002): "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 117(3), 1039–1073.

Scheuer, F. (2014): "Entrepreneurial Taxation with Endogenous Entry," *American Economic Journal: Economic Policy*, 6(2), 126–163.

Stantcheva, S. (2015): "Optimal taxation and human capital policies over the life cycle," Discussion paper, National Bureau of Economic Research.

A Derivation of Fee Schedule with Mechanism-Design Approach

Applying a variant of the revelation principle, we know that instead of optimally choosing the function $F_{pu}(\cdot)$ (which would be an indirect mechanism), the social planner can also choose directly the respective allocation variables $\{y_{pu}(\omega), c_{pu}(\omega)\}_{\omega\in\Omega}$. This mechanism-design approach is non-standard because $\{y_{pr}(\omega), c_{pr}(\omega)\}_{\omega\in\Omega}$ and $\{y_{no}(\omega), c_{no}(\omega)\}_{\omega\in\Omega}$ are exogenously given. This is the 'direct-mechanism equivalent' of taking the income tax schedule $T(\cdot)$ as given – note that the income of parents that do not send their children to public preschools only depends on the tax schedule.[13] When optimizing over $\{y_{pu}(\omega), c_{pu}(\omega)\}_{\omega\in\Omega}$, the planner has to obey the same enrollment condition (8). In the cost equation, one has to substitute $F_{pu}(y_{pu}(\omega)) + T(y_{pu}(\omega))$ with $y_{pu}(\omega) - c_{pu}(\omega)$.

Incentive Compatibility When choosing $\{y_{pu}(\omega), c_{pu}(\omega)\}_{\omega\in\Omega}$, the government has to take into account (i) that those parents who send their children to public preschool truthfully reveal their type ω and (ii) how $\tilde{\xi}(\omega, \phi)$ and $\tilde{\phi}(\omega)$ depend on $\{y_{pu}(\omega), c_{pu}(\omega)\}_{\omega\in\Omega}$. The former is equivalent to

$$\forall \omega, \omega' \text{ with } \omega \neq \omega' : U\left(c_{pu}(\omega) - \frac{y_{pu}(\omega)}{\omega}\right) \geq U\left(c_{pu}(\omega') - \frac{y_{pu}(\omega')}{\omega}\right). \tag{20}$$

Following the theory of optimal income taxation in the tradition of Mirrlees (1971), one can show that the incentive compatibility constraints (20) can be summarized by:

$$\forall \omega : \; X'(\omega) = v'\left(\frac{y_{pu}(\omega)}{\omega}\right)\frac{y_{pu}(\omega)}{\omega^2}$$

where $X(\omega) = c_{pu}(\omega) - v\left(\frac{y_{pu}(\omega)}{\omega}\right)$ and a monotonicity constraint $y'_{pu}(\omega) \geq 0$. Following common practice, we ignore the monotonicity constraint in the analytical part and numerically check ex-post, whether it is fulfilled – and it always is.

The Lagrangian associated with the mechanism-design problem then reads as (where we use the negative of the objective function to write at as a maximization problem):

[13] With income effects, the labor supply of the private preschool parents would also depend on F_{pr}. It would be straightforward to take that into account.

$$\max_{\{X(\omega),y_{pu}(\omega)\}_{\omega\in\Omega}} \mathcal{L} = \int_\Omega s_{pu}(\omega)\left[y_{pu}(\omega) - c_{pu}(\omega) - T(y^*(\omega)) - \mathcal{C}_{pu}\right] dF(\omega) \tag{21}$$

$$+\Lambda \int_\Omega (s_{pu}(\omega) + s_{pr}(\omega))\, dF(\omega) + \int_{\underline{\omega}}^{\overline{\omega}} \eta(\omega)\left(X'(\omega) - v'\left(\frac{y(\omega)}{\omega}\right)\frac{y(\omega)}{\omega^2}\right) \tag{22}$$

After applying integration by parts $\int_{\underline{\omega}}^{\overline{\omega}} \eta(\omega)X'(\omega) = \eta(\overline{\omega})X(\overline{\omega}) - \eta(\underline{\omega})X(\underline{\omega}) - \int_{\underline{\omega}}^{\overline{\omega}} \eta'(\omega)X(\omega)$, we obtain the following necessary conditions:

$$\frac{\partial \mathcal{L}}{\partial X(\omega)} = \frac{\partial s_{pu}(\omega)}{\partial X(\omega)}\left[y_{pu}(\omega) - c_{pu}(\omega) - T(y^*(\omega)) - \mathcal{C}_{pu}\right] f(\omega) - s_{pu}(\omega)f(\omega)$$
$$+\Lambda\left(\frac{\partial s_{pu}(\omega)}{\partial X(\omega)} + \frac{\partial s_{pr}(\omega)}{\partial X(\omega)}\right) - \eta'(\omega) = 0. \tag{23}$$

and

$$\frac{\partial \mathcal{L}}{\partial y_{pu}(\omega)} = \left(1 - v'\left(\frac{y_{pu}(\omega)}{\omega}\right)\frac{1}{\omega}\right) f(\omega)s_{pu}(\omega)$$
$$-\eta(\omega)\left(v'\left(\frac{y(\omega)}{\omega}\right)\frac{1}{\omega^2} + v''\left(\frac{y(\omega)}{\omega}\right)\frac{y(\omega)}{\omega^3}\right) = 0. \tag{24}$$

as well as the transversality conditions

$$\eta(\overline{\omega}) = \eta(\underline{\omega}) = 0.$$

One can show that (24) can be written as:[14]

$$\frac{\tau_{pu}(\omega)}{1-\tau_{pu}(\omega)} = \left(1 + \frac{1}{\varepsilon_{y,1-\tau}(\omega)}\right)\frac{\eta(\omega)}{\omega f(\omega)s_{pu}(\omega)} \tag{25}$$

Then, integrating (23), solving for $\eta(\omega)$ and inserting into (25) yields the result in Proposition 1. The optimal value for Λ follows from the transversality conditions.

[14] First we use just the definition of the wedge $\omega(1-\tau_{pu}(\omega)) = v'\left(\frac{y_{pu}(\omega)}{\omega}\right)$. Second – using the implicit function theorem – one can show that the elasticity satisfies $\varepsilon_{y,1-\tau}(\omega) = \frac{\omega(1-\tau_{pu}(\omega))}{\frac{y_{pu}(\omega)}{\omega}v''\left(\frac{y_{pu}(\omega)}{\omega}\right)}$. Together this implies $v'\left(\frac{y_{pu}(\omega)}{\omega}\right) + v''\left(\frac{y_{pu}(\omega)}{\omega}\right)\frac{y_{pu}(\omega)}{\omega} = \omega(1-\tau_{pu}(\omega))\left(1 + \frac{1}{\varepsilon_{pu}(\omega)}\right)$.

B Estimated Fee Distributions

Table 3: Maximum Likelihood Estimates of Fee Distributions

Parameter	Estimate	Standard Error
Public Fee Probabilities		
$\xi_{pu,1}$	0.3064	(0.0367)
$\xi_{pu,2}$	0.0759	(0.0121)
$\xi_{pu,3}$	0.0507	(0.0099)
$\xi_{pu,4}$	0.0513	(0.0106)
$\xi_{pu,5}$	0.0554	(0.0120)
$\xi_{pu,6}$	0.0732	(0.0158)
$\xi_{pu,7}$	0.0785	(0.0185)
$\xi_{pu,8}$	0.0495	(0.0168)
$\xi_{pu,9}$	0.1029	(0.0272)
$\xi_{pu,10}$	0.0503	(0.0227)
$\xi_{pu,11}$	0.0121	(0.0120)
$\xi_{pu,12}$	0.0939	(0.0438)
Private Fee Probabilities		
$\xi_{pr,1}$	0.0342	(0.0052)
$\xi_{pr,2}$	0.0978	(0.0128)
$\xi_{pr,3}$	0.1730	(0.0202)
$\xi_{pr,4}$	0.0930	(0.0129)
$\xi_{pr,5}$	0.0746	(0.0116)
$\xi_{pr,6}$	0.0692	(0.0116)
$\xi_{pr,7}$	0.0732	(0.0131)
$\xi_{pr,8}$	0.0532	(0.0117)
$\xi_{pr,9}$	0.0481	(0.0120)
$\xi_{pr,10}$	0.0435	(0.0138)
$\xi_{pr,11}$	0.0387	(0.0140)
$\xi_{pr,12}$	0.2014	(0.0389)

Standard errors are based on the BHHH estimator for the Information Matrix.

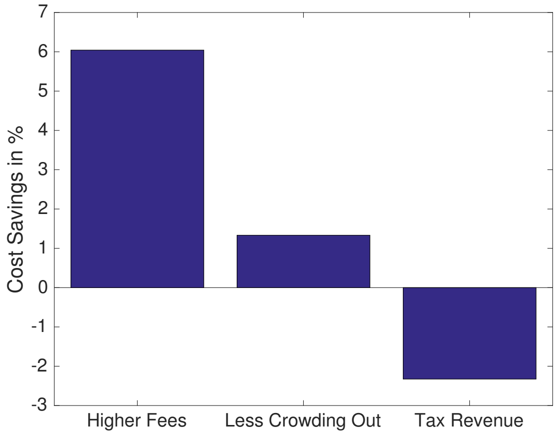
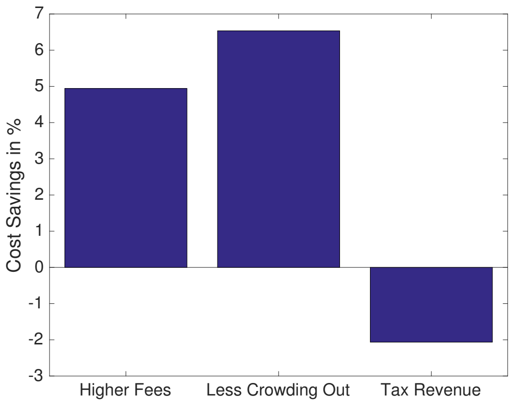
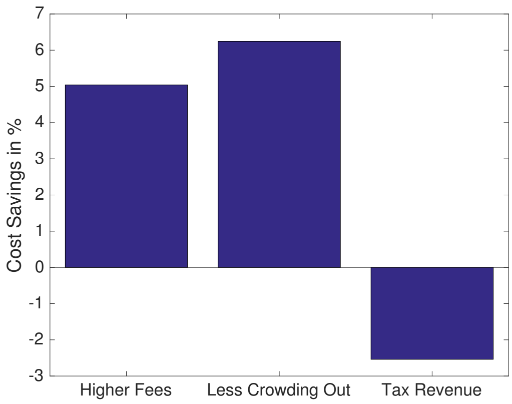


Figure 10: Decomposition of Cost Savings

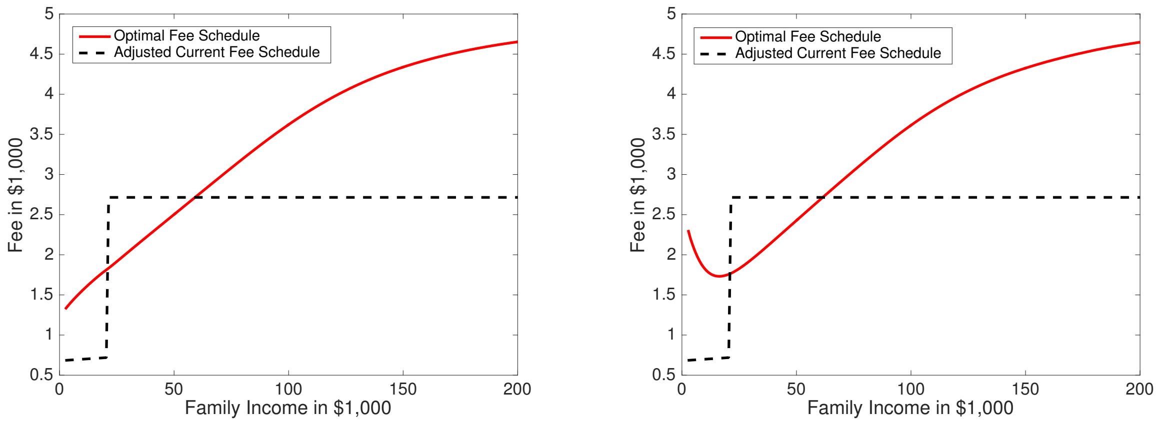


(a) Ignoring Welfare Benefits



(b) Accounting for Welfare Benefits

Figure 11: Cost Savings Decomposition



(a) Public Preschool Enrollment (b) Private Preschool Enrollment

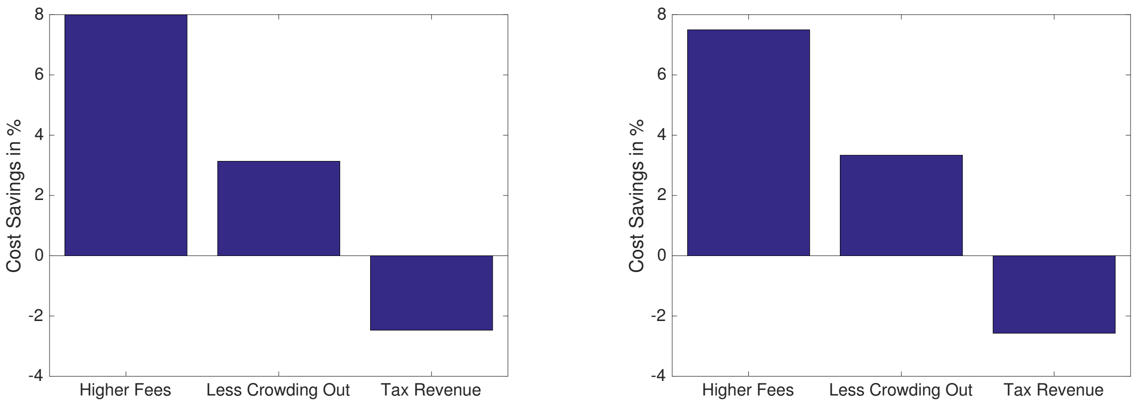
Figure 12: Fee schedule that implements same level of enrollment as universal preschool

C Appendix for Section 4

C.1 Additional Graphs

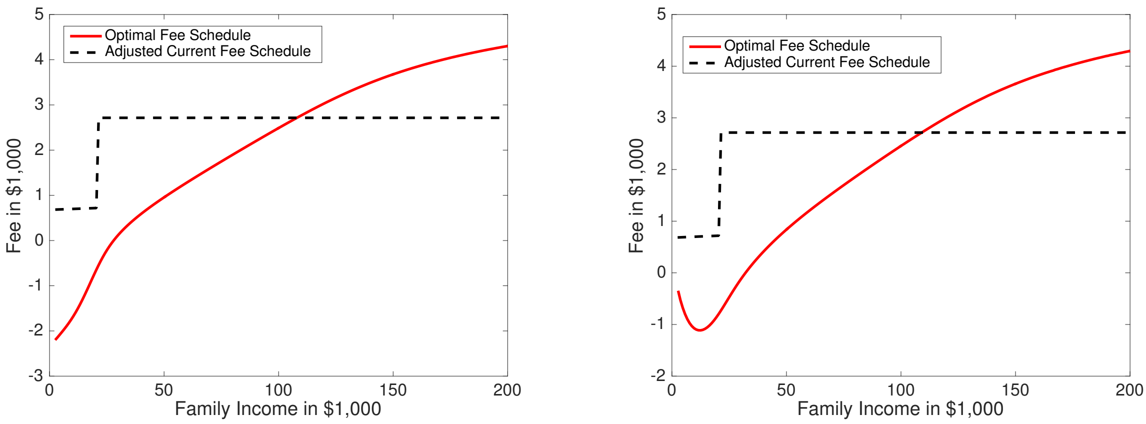
C.2 Robustness: Triangular Distribution

We now present results for optimal policies for the model version, where $\theta_{i,pr}$ follows a triangular distribution with minimal value zero. Figure 12 shows the optimal fee schedules for this case and they are very similar to the ones obtained for an exponential distribution for $\theta_{i,pr}$ as discussed in the mainbody of the text. The amount of costs that can be saved is a bit higher in this case: 8.66 and 13.4 percent respectively for the case where transfers are not and where they are taken into account. The decompositions of the cost saving is illustrated in Figure 13, again they are very similar as for the exponential distribution. Finally, we illustrate optimal fee schedules that implement the same level of enrollment as a universal preschool in Figure 14. Again, the main implications are not altered as compared to the benchmark case that we discuss in the mainbody of the paper. The costs savings are 11.54% and 10.44% in these cases. The decomposition is illustrated in Figure 15 and also resembles the one in the benchmark case.



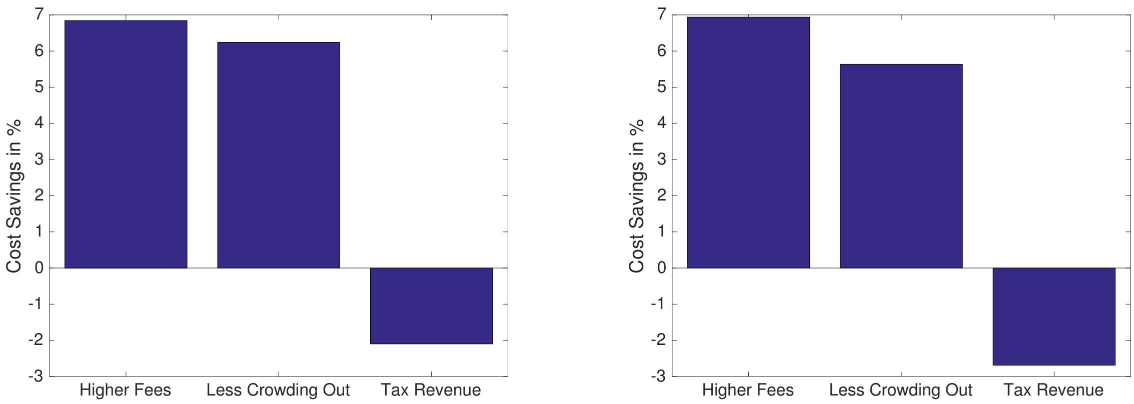
(a) Ignoring Welfare Benefits (b) Accounting for Welfare Benefits

Figure 13: Cost Savings Decomposition



(a) Public Preschool Enrollment (b) Private Preschool Enrollment

Figure 14: Fee schedule that implements same level of enrollment as universal preschool



(a) Ignoring Welfare Benefits (b) Accounting for Welfare Benefits

Figure 15: Cost Savings Decomposition