

Child Care Subsidies, Quality, and Optimal Income Taxation

Author(s): Spencer Bastani, Sören Blomquist and Luca Micheletto

Source: *American Economic Journal: Economic Policy*, November 2020, Vol. 12, No. 4 (November 2020), pp. 1-37

Published by: American Economic Association

Stable URL: <https://www.jstor.org/stable/10.2307/27028629>

REFERENCES

Linked references are available on JSTOR for this article:

https://www.jstor.org/stable/10.2307/27028629?seq=1&cid=pdf-reference#references_tab_contents

You may need to log in to JSTOR to access the linked references.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

American Economic Association is collaborating with JSTOR to digitize, preserve and extend access to *American Economic Journal: Economic Policy*

Child Care Subsidies, Quality, and Optimal Income Taxation[†]

By SPENCER BASTANI, SÖREN BLOMQUIST, AND LUCA MICHELETTI*

We study child care subsidies in a Mirrleesian optimal tax framework where parents choose both the quantity and quality of child care. Child care services not only enable parents to work, but also contribute to children's human capital. We examine the conditions under which child care expenditures should be encouraged or discouraged by the tax system under different assumptions regarding the available policy instruments. Using a quantitative model calibrated to the US economy, we illustrate the possibility that child care expenditures should be taxed rather than subsidized, and we discuss the merits of public provision schemes for child care. (JEL H21, H24, H41, J13)

Several arguments have been proposed to justify subsidizing child care expenditures. One common claim is that child care subsidies are desirable since they enable both parents to work.¹ Another argument in favor of child care subsidies is that they represent a means to increase fertility (see, for example, OECD 2011). Perhaps most importantly, child care outside the home may serve the purpose of improving child outcomes, in particular for children with a poor social background.²

If one takes the benefits of child care as given and decides that child care should be subsidized, it would seem natural to argue that the gains from subsidizing child

*Bastani: Department of Economics and Statistics, Linnaeus University; Uppsala Center for Fiscal Studies, Uppsala Center for Labor Studies, CESifo (email: spencer.bastani@lnu.se); Blomquist: Uppsala Center for Fiscal Studies at the Department of Economics, Uppsala University (email: soren.blomquist@nek.uu.se); Micheletti: Department of Law, University of Milan; Dondena Center for Research on Social Dynamics, Bocconi University; Uppsala Center for Fiscal Studies; CESifo (email: luca.micheletti@unibocconi.it). Dan Silverman was coeditor for this article. This is a substantially revised and extended version of Bastani et al. (2012), which circulated under the title “Child Care Subsidies, Quality and the Optimal Income Tax.” We are grateful to Dan Anderberg, Robin Boadway, Bradley Heim, Wojciech Kopczuk, Pierre Pestieau, Panu Poutvaara, and Andrew Shephard as well as numerous seminar, workshop, and conference participants for comments on earlier drafts. Financial support from the Jan Wallander and Tom Hedelius Foundation and the Nordic Tax Research Council is gratefully acknowledged. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

[†]Go to <https://doi.org/10.1257/pol.20170648> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

¹See, for example, OECD (2006), as well as Blau and Robins (1988), Gustafsson and Stafford (1992), Ribar (1995), and Powell (2002).

²See, for example, Blau and Currie (2006), Currie (2006), and Waldfogel (2006), as well as Heckman (2006) and Heckman and Masterov (2007), who argue that high-quality child care has the potential to help with the promotion of social skills and reduce rates of crime, teenage pregnancy, high school dropout rates, adverse health conditions, and other social problems. In more recent work, Havnes and Mogstad (2011) find that subsidized child care has large positive effects on children's long-run adult outcomes, and Havnes and Mogstad (2015) report that the positive effects are particularly large for children from families below median levels of income. The benefits of subsidized child care are, however, not undisputed. For instance, Baker, Gruber, and Milligan (2008) find a negative short-run effect of child care on children's noncognitive development.

care expenditures ought to be traded off against the deadweight losses of the taxes that are needed to finance the subsidies. The optimal tax literature has, however, pushed the opposite argument—namely, that subsidies to child care have the potential to increase the overall efficiency of the tax system. The argument is related to the well-known result in the optimal tax literature that goods complementary to labor supply should be subject to a more lenient tax treatment and dates back to Corlett and Hague (1953).

In this paper, we evaluate—both theoretically and quantitatively—the desirability of child care subsidies in a model where the quantity and quality of care that children receive, both at home (in terms of informal care from parents) and outside the home (at child care facilities), affects the children’s human capital that enters into the utility function maximized by parents.

Our model economy consists of households that have the same (positive) number of children in child care ages but differ in terms of market ability and, possibly, in terms of the production function for the children’s human capital. The analysis is cast in a Mirrleesian optimal tax framework where the government pursues redistributive goals, but is constrained by asymmetric information. In particular, while the government knows the structure of households’ preferences and the distribution of household types, it cannot tell “who is who.” This implies that personalized lump-sum taxes and transfers are not feasible. The set of available tax instruments will then depend on which variables are publicly observable at the individual level and, possibly, on other constraints. In accordance with the bulk of the optimal tax literature, earned income will be assumed to be publicly observable at the individual level, so that a nonlinear earned income tax is an available policy instrument. Due to the asymmetric information problem, the government will in general need to impose distortions on agents’ behavior in order to achieve redistributive goals. These distortions represent the efficiency costs that have to be incurred in order to induce agents to truthfully reveal their types and choose the bundle intended for them by the government rather than behave as “mimickers” to reduce their tax burden. In such a setting, the desirability of supplementing income taxation with other instruments hinges crucially on the possibility to use them to make mimicking less attractive. As several previous contributions (that will be reviewed in Section I) have pointed out, child care subsidies could be justified precisely on these grounds.

Compared to previous contributions in the optimal tax literature, the main distinguishing contribution of our paper is to emphasize the key importance of the quality dimension of child care (and of the human capital formation process more generally) in determining whether or not child care subsidies are useful as an instrument to achieve redistribution at lower efficiency costs.

In the first part of the paper, we theoretically analyze the role for child care subsidies in a simplified two-type setting, focusing on various government problems that differ with respect to the economic variables that are assumed to be publicly observable at the individual level and with respect to the policy instruments that are available to the government.

We start by characterizing the properties of a constrained efficient allocation under the assumption that earned income, hours spent at a child care facility, and quality of the chosen facility are all variables that are publicly observable at the

individual level. We then analyze the properties of a constrained efficient allocation under the assumption that only earned income and child care expenditures are publicly observable at the individual level. Finally, we consider the possibility that a nonlinear income tax is supplemented with child care subsidies that are administered through an “opting-out” public provision scheme.

The reason for considering these different cases is that it illustrates that the observational assumptions that one makes, as well as the assumptions regarding the policy instruments available to the government, can be relevant to determining whether child care expenditures should be encouraged or discouraged by the tax system.

In the second part of the paper, we construct a more general model that we calibrate to empirical wage distributions and time use patterns based on US data. Assuming that the government chooses optimally a nonlinear income tax on total household income, we evaluate the welfare effects of two alternative ways of subsidizing child care expenditures.

First, we consider offering child care subsidies at rates that are allowed to depend on both the household income and the labor force participation decision of the secondary earner. This is equivalent to allowing a fraction (that is dependent both on the household income and the labor force participation decision of the secondary earner) of child care expenditures to be claimed as a refundable tax credit.

Second, we consider a public provision scheme where parents can choose between (i) getting center-based child care services of a given fixed quality (chosen by the government) for as many hours as they want for free (in which case households are said to “opt in”), and (ii) choosing their preferred quality of center-based services but bearing the full cost (in which case households are said to “opt out”).³

Our numerical results, based on data for the United States, indicate that subsidizing child care expenditures by means of refundable tax credits cannot be justified as a way to achieve redistribution at lower efficiency costs, at least not when the nonlinear income tax is optimally chosen. Intuitively, the reason is due to the fact that the tax-saving value of a refundable tax credit depends positively on the amount of child care expenditures. High-skilled mimickers (i.e., high-skilled individuals who reduce their labor supply in order to qualify for a more lenient tax treatment) can be expected to use formal child care for fewer hours than households being mimicked (i.e., actual low-skilled households). However, with quality of center-based care being a choice variable for households, high-skilled mimickers might still spend more on child care than low-skilled agents. This is indeed what happens in our simulations. In contrast, we find that administering a subsidy through an opting-out public provision scheme allows achieving redistribution at lower efficiency costs. By conditioning the subsidy on the acceptance of a quality that is set by the government, the tax-saving value of the subsidy for opting-in households only depends on the number of hours of center-based child care.

³The study of child care subsidies in the form of opting-out public provision schemes is a novelty of our analysis. With a few exceptions (such as Blomquist and Christiansen 1995), previous contributions in the optimal tax literature have either considered child care subsidies that are equivalent to refundable tax credits or subsidies administered through topping-up public provision schemes. In addition, their quantitative importance has not been assessed.

The rest of the paper is organized as follows. In Section I, we discuss in more details how our contribution relates to some recent papers analyzing the welfare consequences of child care subsidies. In Section II, we set up a simplified theoretical model that is used to evaluate the desirability of subsidizing child care under different assumption regarding which economic variables are publicly observable at the individual level and which policy instruments are available to the government. In Section III, we describe the model, calibrated on US data, that we employ to evaluate the social welfare effects of subsidizing child care; we also present the government's problem that we consider in the quantitative analysis as well as the computational approach used to solve the model. Section IV provides the quantitative results of our numerical simulations. Finally, Section V offers some concluding remarks.

I. Related Literature

While the emphasis on the importance of the quality dimension of child care (and the human capital formation process more generally) is a novelty of our analysis and suggests that child care subsidies might not necessarily be desirable, our paper relates to various previous contributions that have provided a role for child care subsidies as an efficiency-enhancing device.

First and foremost, our paper relates to the Atkinson and Stiglitz (1976) theorem on the usefulness of commodity taxes in the presence of a general (nonlinear) labor income tax. According to that theorem, if the income tax is allowed to be nonlinear and set optimally, commodity taxes are a redundant policy instrument when preferences are separable between leisure and other goods. If the separability condition is not satisfied, the theorem prescribes to use commodity taxes and subsidies to discourage the consumption of goods/services that are substitutes for labor supply and encourage the consumption of goods/services that are complements to labor supply. Viewing child care services as a primary example of services that are complements with labor supply, it would seem natural to argue that they should be subsidized or in any case be subject to a more lenient tax treatment (compared with other goods/services).⁴ This is indeed the result obtained by Blomquist, Christiansen, and Micheletto (2010) in a model where (i) the quality of center-based child care services does not enter into the utility function of agents, and (ii) parents only use child care to cover for hours spent working in the market.

In another contribution, Domeij and Klein (2013) study how child care subsidies can help achieve efficient labor wedges (both across time and across agents) in a dynamic Ramsey optimal tax problem. They recommend that child care expenditures should be made tax deductible. However, they do not consider a Mirrleesian income tax setting, and they disregard the quality dimension of child care services and assume that they are only needed when both parents work (one hour of child care is needed for every hour that both parents work).

In a more recent paper, Guner, Kaygusuz, and Ventura (2020) extend the analysis by Domeij and Klein (2013) in several directions and study the macroeconomic

⁴This is also the view expressed by Crawford, Keen, and Smith (2010) in one of the chapters contained in *The Mirrlees Review*. See Bastani, Blomquist, and Pirttilä (2015) for a recent discussion of this result.

and welfare implications of transfers to households with children, including subsidies to child care. However, similar to Domeij and Klein (2013), they do not consider a Mirrleesian optimal income tax setting, and even though the quantity of child care in their model is a choice variable (and not strictly related to hours of work) and agents face different (exogenous) child care costs, child care quality is not a choice variable of agents.⁵

Another related paper is by Koehne and Sachs (2020), who study the benefits of providing tax deductions for household services in a Mirrleesian framework. However, whereas our focus is on describing properties of particular welfare optima in the context of child care, their work is focused on Pareto-improving reforms in the context of household services more generally. Hence, we view their work as complementary to ours.

Finally, the paper most closely related to ours is Ho and Pavoni (2020), who consider a static Mirrleesian setting and provide a rich set of results on the optimal manner in which to subsidize child care. Even though their setting is similar to ours, the two main differences are that they consider a different model of household decision-making and analyze the government's problem in a different way. The most important difference with respect to the first point is that we endogenize the choice of quality of formal care. With respect to the second point, they compute the constrained efficient allocation on the basis of the informational frictions in the economy in a first step. Then, in a second step, they determine the discrepancies between marginal rates of substitution and marginal rates of transformation ("wedges") and proceed to discuss which policy instruments are needed to implement the constrained efficient allocation. In our paper, instead, we follow a complementary approach. In particular, while in the first part of the theoretical analysis we characterize the wedges prevailing at a constrained efficient allocation, in the second part of the theoretical analysis, and in our quantitative analysis, we prespecify the policy instruments and then analyze their qualitative and quantitative features. When we prespecify the policy instruments, we consider a non-linear income tax and either linear child care subsidies (linear tax credits) or a simple opting-out public provision scheme. A final difference between our paper and Ho and Pavoni (2020) is that we restrict ourselves to a two-type setting in our theoretical analysis, whereas they consider a more general theoretical framework with an arbitrary discrete number of household types.

II. A Simplified Theoretical Model

In this section, we theoretically analyze child care subsidies using a simple two-type single-parent household model. This model will then be extended in several dimensions when we perform our quantitative analysis in Section III.

We consider a subpopulation of the economy, namely families with the same (positive) number of children in child care ages, implicitly assuming that the tax system can be tagged based on the number of children in child care ages living in a given household (in accordance with what is observed in many countries, including the

⁵ See also Bick (2016), who employs a rich model of household behavior with fertility, labor force participation, and various child care choices to study the welfare effects of two child care reforms in Germany.

United States). For illustrative purposes, and given that the same kind of analysis can be carried out for each specific tagged group, we will assume that each household has one child in child care ages. This subgroup consists of two types of single-parent households who differ in terms of market ability and, possibly, also in terms of the production function for the child's human capital. We let households of type 1 be the low-ability households and households of type 2 be the high-ability ones, and we denote by w^1 and w^2 the wage rate paid to, respectively, a high-ability agent and a low-ability agent (with $w^2 > w^1$). The total number of households is normalized to unity and the proportion of households of type 1 is denoted by π (so that $1 - \pi$ is the proportion of households of type 2).

We also assume that children always need someone to take care of them. Care can either be provided by the parent (when he/she is not working in the market or is not engaged in other activities without the child) or by means of external child care services offered by centers that differ in quality.

Denoting by Θ the time endowment of a parent, his/her time constraint is given by

$$(1) \quad L + h + \ell = \Theta,$$

where L represents the time devoted to work in the market, h denotes the time spent by parent with the child, and ℓ represents leisure time spent by the parent without the child.

Based on our assumption that children always need someone to take care of them, the time constraint for a child is

$$(2) \quad h_c + h = \Theta,$$

where h_c denotes the number of hours that a child spends in a child care facility.

To capture in a simple way the idea that a higher quality of the early childhood environment fosters the human capital development of the child and ameliorates his/her future prospects as an adult, we assume that human capital is built according to a function that depends on: (i) the amount of time spent by a parent with the child; (ii) the amount of time spent by the child at a child care facility; (iii) the nurturing ability of the parent; (iv) the quality of the child care facility chosen by the parent; (v) the innate ability of the child. Formally, we let human capital be given by a function

$$\gamma f(\omega h, q_c h_c),$$

where ω (> 0) denotes the nurturing ability of the parent, q_c (≥ 0) denotes the quality of the child care facility chosen by the parent, and γ (> 0) is a scalar reflecting the innate ability of the child.

Parents derive utility from the consumption of a composite good denoted by c (treated as the numéraire of our economy), from leisure, and from the human capital of their child. In particular, households' preferences are represented by the following utility function:

$$(3) \quad U = u(c) + \gamma f(\omega h, q_c h_c) + v(\ell),$$

where we assume that $u(\cdot)$, $f(\cdot, \cdot)$, and $v(\cdot)$ are concave functions that are increasing in each argument.

Regarding the values of the parameters γ and ω , we will assume that they are weakly higher for parents of high market ability than for parents of low market ability.⁶ Thus, we will focus on a setting where $w^2 > w^1$, $\gamma^2 \geq \gamma^1$, and $\omega^2 \geq \omega^1$.⁷

The laissez-faire hourly price of center-based child care services, denoted by p , is assumed to depend on the quality q_c of the child care facility through the isoelastic function $p(q_c) = k(q_c)^\sigma$, with $k > 0$ and where we assume $\sigma \geq 1$.

Substituting the time constraints (1) and (2) into the utility function (3), we can write the problem solved by a parent under laissez-faire as

$$\max_{L, h, q_c} u(wL - (\Theta - h)p(q_c)) + \gamma f(\omega h, (\Theta - h)q_c) + v(\Theta - L - h).$$

Using primes to denote derivatives and denoting by f'_j the derivative of the $f(\cdot, \cdot)$ -function with respect to its j th argument (with $j = 1, 2$), the first-order conditions of the maximization problem above are given by

$$(4) \quad L: \quad 1 - \frac{v'(\ell)}{wu'(c)} = 0,$$

$$(5) \quad h: \quad 1 - \frac{v'(\ell)}{p(q_c)u'(c) + (\omega f'_1 - q_c f'_2)\gamma} = 0,$$

$$(6) \quad q_c: \quad 1 - \frac{p'(q_c)u'(c)}{\gamma f'_2} = 0.$$

Below we will analyze various government problems, always maintaining the following two assumptions: (i) the government aims at redistributing toward agents of type 1 (who are worse off under laissez-faire), and (ii) the government knows the distribution of types in the population but does not know “who is who,” so that first-best personalized lump-sum taxes and transfers are not feasible. The problems that we consider will instead differ in terms of other observability assumptions and/or assumptions about the policy instrument available to the government. In particular, in the first problem (Case 1) we characterize the properties of a constrained efficient allocation under the assumption that earned income, hours spent by a child at a child care facility, and quality of the chosen child care facility are all variables that are publicly observable at the individual level. In the second problem (Case 2), we analyze the properties of a constrained efficient allocation under the assumption that only earned income and child care expenditures are publicly observable at the individual level. Finally, in the third problem (Case 3), we consider the

⁶We have previously said that the parameter γ captures the innate ability of the child in acquiring human capital. Alternatively, differences in γ across households could be interpreted as reflecting heterogeneity in preferences. Under this alternative interpretation, the human capital of the child carries a larger weight in the utility function of households characterized by a higher value for γ .

⁷The case $\gamma^2 > \gamma^1$ can be seen as reflecting the genetic intergenerational transmission of ability.

possibility that a nonlinear income tax is supplemented with child care subsidies that are administered through an “opting-out” public provision scheme.⁸

As will be clear from our analysis, one reason for considering these various cases is that it allows highlighting that, depending on the observability assumptions that one makes and on the set of policy instruments available to the government, one can get opposite conclusions about the desirability to subsidize formal child care.

A. Case 1: Earned Income, Hours of Child Care, and Quality Observable

Assume that earned income, hours spent by a child at a child care facility, and the quality of the chosen facility are all variables that are publicly observable at the individual level. Denoting earned income by y (with $y \equiv wL$), these observability assumptions imply that the tax function can be a general function $T = T(y, h_c, q_c)$ or equivalently, exploiting the time constraint (2), $T = T(y, h, q_c)$. A constrained efficient allocation can then be found as a solution to the following government’s program (problem (P1)):

$$(P1) \quad \max_{c^1, y^1, h^1, q_c^1, c^2, y^2, h^2, q_c^2} u(c^1) + \gamma^1 f(\omega^1 h^1, (\Theta - h^1)q_c^1) + v\left(\Theta - \frac{y^1}{w^1} - h^1\right)$$

subject to

$$u(c^2) + \gamma^2 f(\omega^2 h^2, (\Theta - h^2)q_c^2) + v\left(\Theta - \frac{y^2}{w^2} - h^2\right) \geq \bar{V},$$

$$\begin{aligned} & u(c^2) + \gamma^2 f(\omega^2 h^2, (\Theta - h^2)q_c^2) + v\left(\Theta - \frac{y^2}{w^2} - h^2\right) \\ & \geq u(c^1) + \gamma^2 f(\omega^2 h^1, (\Theta - h^1)q_c^1) + v\left(\Theta - \frac{y^1}{w^2} - h^1\right), \end{aligned}$$

$$\left[y^1 - c^1 - p(q_c^1)(\Theta - h^1)\right]\pi + \left[y^2 - c^2 - p(q_c^2)(\Theta - h^2)\right](1 - \pi) \geq \bar{R}.$$

In the maximization problem above, the first constraint prescribes a minimum utility requirement for the high-skilled households; the second constraint is a self-selection (incentive-compatibility) constraint requiring that high-skilled households have no incentive to mimic low-skilled households by choosing a bundle intended only for the latter. This constraint descends from the informational frictions in the economy (the fact that the government cannot directly observe the true type of an agent) and from the assumption that the socially desirable direction of redistribution is from

⁸In online Appendix A, we also analyze a fourth case in which we adopt the assumption that, while earned income can be subject to a nonlinear income tax, child care expenditures can only be subsidized or taxed at a proportional rate. This constraint can either be viewed as descending from the fact that child care expenditures are not publicly observable at the household level (in which case the constraint would be a direct consequence of the informational frictions characterizing the economy), or as capturing the idea that a proportional subsidy/tax is easier to administer and therefore allows to save on the administrative costs of taxes.

type 2 to type 1. Finally, the last constraint represents the resource constraint of the economy (with \bar{R} denoting an exogenous revenue requirement for the government).

The following proposition characterizes the properties of the solution to the government's program.

PROPOSITION 1: *No distortion should be imposed on the choices made by high-skilled households. For low-skilled households, y^1 should be downward distorted so that*

$$1 - v'\left(\Theta - \frac{y^1}{w^1} - h^1\right) / w^1 u'(c^1) > 0;$$

q_c^1 should be left undistorted if households only differ in terms of market ability ($\gamma^1 = \gamma^2 \equiv \gamma$ and $\omega^1 = \omega^2 \equiv \omega$); otherwise, if either $\omega^2 > \omega^1$ or $\gamma^2 > \gamma^1$ (or both $\omega^2 > \omega^1$ and $\gamma^2 > \gamma^1$), q_c^1 should be downward distorted so that

$$1 - p'(q_c^1) u'(c^1) / \gamma^1 f_2(\omega^1 h^1, (\Theta - h^1) q_c^1) > 0.$$

Finally, h_c^1 should be upward distorted (unless $y^1 = 0$, in which case h_c^1 should be left undistorted) if households only differ in terms of market ability;⁹ otherwise, if either $\omega^2 > \omega^1$ or $\gamma^2 > \gamma^1$ (or both $\omega^2 > \omega^1$ and $\gamma^2 > \gamma^1$), h_c^1 should be upward (downward, resp.) distorted when the following condition holds:

$$\begin{aligned} & \gamma^2 \left[\omega^2 f_1'(\omega^2 h^1, (\Theta - h^1) q_c^1) - q_c^1 f_2'(\omega^2 h^1, (\Theta - h^1) q_c^1) \right] - v'\left(\Theta - \frac{y^1}{w^2} - h^1\right) \\ & > (<) \gamma^1 \left[\omega^1 f_1'(\omega^1 h^1, (\Theta - h^1) q_c^1) - q_c^1 f_2'(\omega^1 h^1, (\Theta - h^1) q_c^1) \right] - v'\left(\Theta - \frac{y^1}{w^1} - h^1\right). \end{aligned}$$

PROOF OF PROPOSITION 1:

See online Appendix B1. ■

According to Proposition 1, all margins of choice for a high-skilled household should be left undistorted. Intuitively, since the government aims at redistributing from high- to low-skilled households, the latter have no incentive to behave as mimickers by falsely claiming to be of type 2 (i.e., high-skilled) in order to get the bundle (c^2, y^2, h^2, q_c^2) . Therefore, distorting the choices of high-skilled households would entail efficiency losses without any offsetting benefit in terms of mimicking-deterring effects.

Proposition 1 also states that incentive compatibility considerations require to distort downward the labor supplied in the market by low-skilled households, both in the case where market ability is the only dimension of heterogeneity among agents and in the case where high-skilled households are also more effective in building up human capital for their children ($\omega^2 > \omega^1$ or $\gamma^2 > \gamma^1$). The intuition for this

⁹That h_c^1 should be upward distorted is equivalent to saying that h^1 should be downward distorted. Formally, it means that the following condition is satisfied at a solution to the government's program: $1 - v'(\Theta - (y^1/w^1) - h^1) / (p(q_c^1) u'(c^1) + [\omega f_1'(\omega h^1, (\Theta - h^1) q_c^1) - q_c^1 f_2'(\omega h^1, (\Theta - h^1) q_c^1)] \gamma) > 0$.

result comes from the fact that for a given amount of time devoted to children (h), the marginal disutility (in terms of foregone leisure) of earning an additional dollar is in both cases higher for a low-skilled household than for a high-skilled household (formally, $-v'(\Theta - (y^1/w^1) - h^1)/w^1 < -v'(\Theta - (y^1/w^2) - h^1)/w^2$).

Regarding the quality of center-based care, q_c^1 should be left undistorted if households only differ in terms of market ability. For a given amount of time devoted to children (h), the marginal return of raising q_c^1 in terms of higher human capital for the child is the same for a low-skilled agent and for a high-skilled mimicker when $\gamma^1 = \gamma^2 = \gamma$ and $\omega^1 = \omega^2 = \omega$ (being equal to $\gamma f_2'(\omega h^1, (\Theta - h^1)q_c^1)$). However, if either $\omega^2 > \omega^1$ or $\gamma^2 > \gamma^1$ (or both $\omega^2 > \omega^1$ and $\gamma^2 > \gamma^1$), the marginal return of raising q_c^1 would be lower for a low-skilled household than for a high-skilled mimicker ($\gamma^1 f_2'(\omega^1 h^1, (\Theta - h^1)q_c^1) < \gamma^2 f_2'(\omega^2 h^1, (\Theta - h^1)q_c^1)$), in which case a downward distortion on q_c^1 would be optimal since it would make mimicking less attractive for high-skilled households.

Finally, according to Proposition 1, the direction of the optimal distortion on h_c^1 cannot be unambiguously determined unless households only differ in market ability. In such a case, h_c^1 should be upward distorted (or left undistorted if at an optimum $y^1 = 0$), which is equivalent to saying that h^1 should be downward distorted (since from (2) it is $h_c = \Theta - h$). The intuition for this result comes from the fact that for a given quality of the child care facility, the marginal return of raising h^1 , in terms of higher human capital for the child, is the same for a low-skilled agent and for a high-skilled mimicker when $\gamma^1 = \gamma^2 \equiv \gamma$ and $\omega^1 = \omega^2 \equiv \omega$ (being equal to $\gamma[\omega f_1'(\omega h^1, (\Theta - h^1)q_c^1) - q_c^1 f_2'(\omega h^1, (\Theta - h^1)q_c^1)]$). At the same time, however, the marginal disutility (in terms of foregone leisure) of increasing h^1 is, for a given value of $y^1 > 0$, higher for a low-skilled household than for a high-skilled household (formally, $-v'(\Theta - (y^1/w^1) - h^1) < -v'(\Theta - (y^1/w^2) - h^1)$).

If instead either $\omega^2 > \omega^1$ or $\gamma^2 > \gamma^1$ (or both $\omega^2 > \omega^1$ and $\gamma^2 > \gamma^1$), one cannot rule out the possibility that h_c^1 should be downward distorted (equivalently, that h^1 should be upward distorted). Intuitively, the reason is that, while it is still true that the marginal disutility (in terms of foregone leisure) of increasing h^1 (conditional on y^1) is higher for a low-skilled than for a high-skilled household, it might be the case that the marginal return of raising h^1 , in terms of higher human capital for the child, is larger for a low-skilled agent than for a high-skilled mimicker—i.e., it might be that

$$(7) \quad \gamma^2 [\omega^2 f_1'(\omega^2 h^1, (\Theta - h^1)q_c^1) - q_c^1 f_2'(\omega^2 h^1, (\Theta - h^1)q_c^1)] \\ < \gamma^1 [\omega^1 f_1'(\omega^1 h^1, (\Theta - h^1)q_c^1) - q_c^1 f_2'(\omega^1 h^1, (\Theta - h^1)q_c^1)].$$

Whereas the first circumstance favors distorting h_c^1 upward, inequality (7) would call for distorting h_c^1 in the opposite direction.¹⁰

¹⁰Rewriting (7) as LHS > RHS, where LHS = $[\gamma^2 f_2'(\omega^2 h^1, (\Theta - h^1)q_c^1) - \gamma^1 f_2'(\omega^1 h^1, (\Theta - h^1)q_c^1)]q_c^1$ and RHS = $\gamma^2 \omega^2 f_1'(\omega^2 h^1, (\Theta - h^1)q_c^1) - \gamma^1 \omega^1 f_1'(\omega^1 h^1, (\Theta - h^1)q_c^1)$, we can notice that LHS is strictly positive when either $\omega^2 > \omega^1$ (assuming $f_{12}' > 0$) or $\gamma^2 > \gamma^1$. RHS, instead, is necessarily positive when $\gamma^2 > \gamma^1$ and $\omega^2 = \omega^1$, but has an ambiguous sign when $\omega^2 > \omega^1$. For instance, denoting by $\epsilon_{f_1, \omega}$ the elasticity of f_1' with respect to ω (i.e., $\epsilon_{f_1, \omega} \equiv \omega h f_{11}''/f_1'$), the sign of RHS is negative if $\gamma^2 = \gamma^1$ and $\epsilon_{f_1, \omega} < -1$.

B. Case 2: Only Earned Income and Child Care Expenditures Publicly Observable

Assume now that y is publicly observable at the individual level—albeit h_c and q_c are not publicly observable at the individual level—and that the government can observe how much each household spends on child care. Denoting child care expenditures by D (i.e., $D \equiv p(q_c)h_c = p(q_c)(\Theta - h)$), these observability assumptions imply that the tax function can be a general function of the form $T = T(y, D)$.

Since a tax function $T(y, D)$ associates to each pair (y, D) a corresponding amount of consumption c , the problem of choosing an optimal tax schedule can be equivalently stated as the problem of selecting two triplets (y^i, D^i, c^i) , one for each type of household. Since from the time constraint (2) we have that $\Theta - h = h_c$, and given that $h_c = D/p(q_c)$, it follows that $\Theta - h = D/p(q_c)$. Therefore, we can equivalently express leisure ℓ as $D/p(q_c) - y/w$. Thus, given any triplet (y, D, c) , an agent of type i solves

$$\max_{q_c^i} u(c) + \gamma^i f\left(\left(\Theta - \frac{D}{p(q_c^i)}\right)\omega^i, \frac{Dq_c^i}{p(q_c^i)}\right) + v\left(\frac{D}{p(q_c^i)} - \frac{y}{w^i}\right).$$

Denote the resulting “conditional” demand function by $Q_c^i(y, D, c)$ and notice that, due to the assumed separability between c and other arguments in the utility function, the conditional demand function is really only a function of y and D —i.e., $Q_c^i(y, D)$. Furthermore, denote the indirect utility function by

$$\begin{aligned} V^i(y, D, c) \equiv & u(c) + \gamma^i f\left(\left(\Theta - \frac{D}{p(Q_c^i(y, D))}\right)\omega^i, \frac{DQ_c^i(y, D)}{p(Q_c^i(y, D))}\right) \\ & + v\left(\frac{D}{p(Q_c^i(y, D))} - \frac{y}{w^i}\right). \end{aligned}$$

A constrained efficient allocation can then be found as a solution to the following government’s program (problem (P2)):

$$(P2) \quad \max_{y^1, D^1, c^1, y^2, D^2, c^2} V^1(y^1, D^1, c^1)$$

subject to

$$V^2(y^2, D^2, c^2) \geq \bar{V},$$

$$V^2(y^2, D^2, c^2) \geq V^2(y^1, D^1, c^1),$$

$$(y^1 - c^1 - D^1)\pi + (y^2 - c^2 - D^2)(1 - \pi) \geq \bar{R}.$$

Define $q_c^1 \equiv Q_c^1(y^1, D^1)$, $q_c^2 \equiv Q_c^2(y^2, D^2)$, and $\hat{q}_c \equiv Q_c^2(y^1, D^1)$. The following proposition characterizes the properties of the solution to the government’s program.

PROPOSITION 2: *No distortion should be imposed on the choices made by high-skilled households. For low-skilled households, y^1 should be distorted downward (i.e., $1 - v'(\ell^1)/w^1 u'(c^1) > 0$) if households only differ in terms of market ability ($\gamma^1 = \gamma^2$ and $\omega^1 = \omega^2$); otherwise, if either $\omega^2 > \omega^1$ or $\gamma^2 > \gamma^1$ (or both $\omega^2 > \omega^1$ and $\gamma^2 > \gamma^1$), y^1 should be distorted downward (upward, resp.) when*

$$\frac{v'\left(\frac{D^1}{p(q_c^1)} - \frac{y^1}{w^1}\right)}{w^1} > (<) \frac{v'\left(\frac{D^1}{p(\hat{q}_c)} - \frac{y^1}{w^2}\right)}{w^2}.$$

Finally, child care expenditures D^1 should be distorted downward (upward, resp.) if the following condition holds:

$$\frac{\gamma^2 f'_2\left(\left(\Theta - \frac{D^1}{p(\hat{q}_c)}\right)\omega^2, \frac{D^1 \hat{q}_c}{p(\hat{q}_c)}\right)}{p'(\hat{q}_c)} > (<) \frac{\gamma^1 f'_2\left(\left(\Theta - \frac{D^1}{p(q_c^1)}\right)\omega^1, \frac{D^1 q_c^1}{p(q_c^1)}\right)}{p'(q_c^1)}.$$

PROOF OF PROPOSITION 2:

See online Appendix B2. ■

Compared to the results obtained in the previous subsection, the first thing to notice is that the changes in the informational assumptions do not impair the result that the choices made by high-skilled households are left undistorted. Since low-skilled households have no incentive to behave as mimickers by falsely claiming to be of type 2 (i.e., high-skilled), nothing can be gained by distorting the triplet (y^2, D^2, c^2) intended for the high-skilled.

Instead, the changes in the informational assumptions have an impact on the results pertaining to low-skilled households.

With respect to the optimal distortion on y^1 , the result that the labor wedge ought to be positive (i.e., y^1 distorted downward) does not necessarily hold. In particular, while it holds for sure if households differ only in market ability, it might be violated if high-skilled households are also more effective in building up human capital for their children ($\omega^2 > \omega^1$ or $\gamma^2 > \gamma^1$). The reason why it is no longer possible to unambiguously establish that the labor wedge is positive stems from the fact that, under the current informational assumptions, the government has a weaker control on h^1 . The public observability of expenditures on child care by households implies that the government can control D^1 . However, households maintain freedom of choice regarding the way to achieve a given level of child care expenditures—i.e., they can freely choose among all combinations of q_c^1 and h^1 satisfying $(\Theta - h^1)p(q_c^1) = D^1$. When households only differ in market ability, one can prove that, even though a high-skilled mimicker would choose a child care facility of higher quality than the one chosen by a low-skilled household ($\hat{q}_c > q_c^1$) and would therefore spend more time with his child ($\hat{h} > h^1$), the amount of leisure enjoyed by a mimicker would be higher than for a low-skilled household. This, in turn, implies that given D^1 , the marginal disutility (in terms of foregone leisure) of earning an additional dollar is higher for a low-skilled than for a high-skilled

mimicker, and therefore y^1 should be distorted downward. However, when $\omega^2 > \omega^1$ or $\gamma^2 > \gamma^1$ (or both $\omega^2 > \omega^1$ and $\gamma^2 > \gamma^1$), one cannot rule out the possibility that the quality of the child care facility chosen by a mimicker is so much greater than the one chosen by a low-skilled household—and therefore \hat{h} so much larger than h^1 —that the amount of leisure enjoyed by a mimicker ends up being lower than for a low-skilled household.¹¹

Regarding the optimal distortion on D^1 , remember that in the previous subsection we obtained the result that when $\gamma^1 = \gamma^2$ and $\omega^1 = \omega^2$, q_c^1 should be left undistorted and h_c^1 should be upward distorted. Under the current informational assumptions, instead, the direction of the optimal distortion on D^1 cannot be in general unambiguously determined. This result indicates that the observability assumptions can play a key role in determining whether child care expenditures should be encouraged or discouraged by the tax system.

The following corollary of the second part of Proposition 2 illustrates that both a downward and an upward distortion on D^1 might be optimal when $\gamma^1 = \gamma^2$ and $\omega^1 = \omega^2$.¹²

COROLLARY 1:

- (i) Suppose $f''_{12} > 0$, $p'' = 0$, $\gamma^2 \geq \gamma^1$, and $\omega^2 \geq \omega^1$; then, a downward distortion on D^1 is desirable.
- (ii) Suppose $f''_{12} = f''_{22} = 0$, $p'' > 0$, $\gamma^1 = \gamma^2$, and $\omega^1 = \omega^2$; then, an upward distortion on D^1 is desirable.

PROOF OF COROLLARY 1:

See online Appendix B3. ■

The corollary above highlights the role played by the assumptions about the human capital production function and about the function relating the hourly price of center-based child care to its quality. The fact that both a downward and an upward distortion on D^1 might be optimal is interesting, especially when compared to the result that one would obtain in a model where agents only differ in market ability and where there is no heterogeneity in the quality of the available child care facilities. Assuming that all child care facilities offer the same quality at a common hourly price, child care expenditures only depend on h_c , and one is immediately led to the conclusion that h_c^1 should be distorted upward for mimicking-detering

¹¹For instance, this would be the case if $w^2 = w^1$, $\omega^2 = \omega^1$, $\gamma^2 - \gamma^1 > 0$, and the price function is of the form $p(q_c) = kq_c$ (with $k > 0$). Under these assumptions, equation (B31) in online Appendix B2 ensures that $\hat{q}_c > q_c^1$ and, therefore, due to the assumption $w^2 = w^1$, $\hat{\ell} = (D^1/p(\hat{q}_c)) - (y^1/w^2) < (D^1/p(q_c^1)) - (y^1/w^1) = \ell^1$ and $v'(D^1/p(q_c^1) - y^1/w^1)/w^1 < v'(D^1/p(\hat{q}_c) - y^1/w^2)/w^2$. The same would be true if, instead of $w^2 = w^1$, we had that $w^2 > w^1$ with the difference $w^2 - w^1$ sufficiently small. In fact, from equations (B29) and (B31) in online Appendix B2 we know that we would still have $\hat{q}_c > q_c^1$. Therefore, if the difference $w^2 - w^1$ were sufficiently small, it would still be that $v'(D^1/p(q_c^1) - y^1/w^1)/w^1 < v'(D^1/p(\hat{q}_c) - y^1/w^2)/w^2$.

¹²With respect to the properties of the tax function $T(y, D)$ that allow implementing the constrained efficient allocation, a downward (upward, resp.) distortion on D^1 implies that $\partial T(y^1, D^1)/\partial D^1 > 0$ ($\partial T(y^1, D^1)/\partial D^1 < 0$, resp.) at the (y^1, D^1, c^1) -bundle intended for low-skilled households.

purposes (i.e., child care expenditures should be subsidized at the margin for low-skilled households). Previous contributions in the optimal tax literature have focused on this case.

One strand of the literature (see, e.g., Blomquist, Christiansen, and Micheletto 2010; Bastani, Blomquist, and Pirttilä 2015) simply assumed that agents demand a number of hours of center-based child care that is equal to their labor supply in the market (i.e., using our notation, $h_c = y/w$ for all agents), implicitly assuming that for parents, time spent with the children is a perfect substitute for leisure time spent without the children. In this case, child care services are simply viewed as something that needs to be purchased in order to enable agents to work.

Another strand of the literature (see, e.g., Ho and Pavoni 2020) assumed instead that time spent by parents with the children is a perfect substitute for time spent working in the market; however, in this case, the longer an agent works in the market, the less time he will spend with the children. Therefore, even though it is possible that agents demand child care hours in excess of their working hours (i.e., it is possible that $h_c > y/w$), it is still the case that center-based child care hours (and therefore child care expenditures) are increasing in the hours spent working in the market (y/w).

With child care expenditures increasing in the number of hours spent working in the market, a high-skilled mimicker would always like to spend less on child care than a low-skilled agent (since for any given value for gross income y , a high-skilled mimicker needs to work fewer hours in the market). In accordance with the Atkinson-Stiglitz (1976) theorem, one would then benefit, for mimicking-deterring purposes, from supplementing a nonlinear income tax with other tax instruments aimed at encouraging the demand for child care services.¹³ This is what drives the result obtained in the previous literature that child care expenditures should be subsidized (either linearly or nonlinearly).¹⁴

However, when agents have the possibility to pay a higher hourly price to get their children at a higher-quality facility, child care expenditures depend both on h_c and q_c . Thus, even if it were the case that a low-skilled agent favors having his child at a child care facility for more hours than a high-skilled mimicker, it would not necessarily follow that mimicking-deterring considerations require to subsidize at the margin child care expenditures by low-skilled agents. Contrary to previous findings in the optimal tax literature, downward distorting the child care expenditures of low-skilled households might then become desirable. In our model, as Corollary 1 shows, key elements to assess the direction of the optimal distortion on D^1 are given

¹³ According to the AS theorem, if the income tax is allowed to be nonlinear, commodity taxes are a redundant policy instrument when preferences are separable between leisure and other goods. Instead, if the separability condition is not satisfied, one should use commodity taxes and subsidies to discourage the consumption of goods/services that are substitutes for labor supply and encourage the consumption of goods/services that are complements to labor supply.

¹⁴ In Blomquist, Christiansen, and Micheletto (2010) and Bastani, Blomquist, and Pirttilä (2015), the assumption that $h_c = y/w$ for all agents leads to the conclusion that child care expenditures should be subsidized at a 100 percent rate. Ho and Pavoni (2020), instead, by relaxing the assumption $h_c = y/w$ and allowing for nonlinear subsidies, find that child care subsidies should optimally follow a sliding scale. In a model with many consumption goods, one does not necessarily need to subsidize child care expenditures in order to encourage them; it would be enough to subject them to a more lenient tax treatment compared with other goods.

by the shape of the price function $p(q_c)$ and the properties of the human capital production function.

C. Case 3: An Opting-Out Public Provision Scheme

The analysis carried out in the previous subsections shows that the case for subsidizing center-based child care expenditures is significantly weakened once one takes into account that the market provides households with an array of choices regarding the quality of child care services. Ideally, one would like to have at disposal tax instruments that can distinguish between the share of child care expenditure that is due to a specific quality of the chosen facility and the share that is due to the number of hours that a child spends at a facility. If a subsidy is “blind” to this difference, then a tax on child care might well be optimal, which is in contrast to previous findings in the literature.

However, even when hours spent by a child at a child care facility and the quality of the chosen facility are not publicly observable at the household level, there might be a way to come close to replicating the constrained efficient allocation characterized in Proposition 1. In particular, suppose that a nonlinear income tax $T(y)$ is supplemented with an opting-out public provision scheme that works as follows. The government provides center-based child care services of quality \bar{q}_c at a given hourly user charge. Households can decide to “opt in” and use the publicly provided services for as many hours as they need. Alternatively, if households are not happy with the quality provided by the government, they can “opt out” and choose their preferred quality of center-based services paying the full market price.

To illustrate the merits of such a public provision scheme, we will here focus on the case where $p'' = 0$, $w^2 > w^1$, $\gamma^2 = \gamma^1 \equiv \gamma$, and $\omega^2 = \omega^1 \equiv \omega$. Notice that under these assumptions, a downward distortion on D^1 is optimal when the tax function is of the form $T = T(y, D)$ (see the first part of Corollary 1).¹⁵

When a nonlinear income tax is the only available policy instrument, the government’s problem can be equivalently formulated as the problem of selecting two bundles in the (b, y) -space, where b denotes disposable income (i.e., after-tax income to be allocated between child care expenditures and consumption of the composite good c). Clearly, to be admissible, the two bundles (b^i, y^i) for $i = 1, 2$ should satisfy public-budget balance and incentive compatibility, the latter meaning that each household must (weakly) prefer, and therefore select, the bundle intended for his own type rather than behave as a “mimicker” by choosing the bundle intended for a different type. Now suppose that the government can supplement the nonlinear income tax with an income-dependent opting-out public provision system such that households earning y^1 have the possibility to choose between (i) getting publicly provided child care services of a given quality \bar{q}_c (chosen by the government)

¹⁵These assumptions also imply that when a proportional tax/subsidy on child care expenditures is used alongside an optimal nonlinear income tax, child care expenditures should be taxed (see Case 4 in online Appendix A).

at an hourly user charge $(1 - \beta)p(\bar{q}_c)$, and (ii) choosing their preferred quality of center-based child care services but bearing the full cost.¹⁶

Formally, the problem solved by a household of type i who chooses to opt in is

$$\max_{h^i} u\left(b^1 - (1 - \beta)(\Theta - h^i)p(\bar{q}_c)\right) + \gamma f(\omega h^i, (\Theta - h^i)\bar{q}_c) + v\left(\Theta - \frac{y^1}{w^i} - h^i\right),$$

with associated first-order condition

$$(8) \quad (1 - \beta)p(\bar{q}_c)u' + \gamma(\omega f'_1 - \bar{q}_c f'_2) = v'.$$

Denote by $H^i(\bar{q}_c, \beta, b^1, y^1)$ the “conditional” demand function for a household of type i who opts in and denote the corresponding indirect utility function by

$$\begin{aligned} V^i(\bar{q}_c, \beta, b^1, y^1) &\equiv u\left(b^1 - (1 - \beta)(\Theta - H^i(\bar{q}_c, \beta, b^1, y^1))p(\bar{q}_c)\right) \\ &\quad + \gamma f\left(\omega H^i(\bar{q}_c, \beta, b^1, y^1), (\Theta - H^i(\bar{q}_c, \beta, b^1, y^1))\bar{q}_c\right) \\ &\quad + v\left(\Theta - \frac{y^1}{w^i} - H^i(\bar{q}_c, \beta, b^1, y^1)\right). \end{aligned}$$

For a household of type i choosing a bundle (b, y) and opting out, the optimization problem is instead the following:

$$\max_{h^i, q_c^i} u\left(b - (\Theta - h^i)p(q_c^i)\right) + \gamma f(\omega h^i, (\Theta - h^i)q_c^i) + v\left(\Theta - \frac{y}{w^i} - h^i\right),$$

with associated first-order conditions

$$p(q_c^i)u' + \gamma(\omega f'_1 - q_c^i f'_2) = v',$$

$$p'(q_c^i)u' = \gamma f'_2.$$

Denote by $H^i(b, y)$ and $Q_c^i(b, y)$ the “conditional” demand functions for a household of type i who chooses the bundle (b, y) and opts out. The corresponding indirect utility function is given by

$$\begin{aligned} V^i(b, y) &\equiv u\left(b - (\Theta - H^i(b, y))p(Q_c^i(b, y))\right) \\ &\quad + \gamma f\left(\omega H^i(b, y), (\Theta - H^i(b, y))Q_c^i(b, y)\right) + v\left(\Theta - \frac{y}{w^i} - H^i(b, y)\right). \end{aligned}$$

Finally, define $h^{1, in} \equiv H^1(\bar{q}_c, \beta, b^1, y^1)$, $h^2 \equiv H^2(b^2, y^2)$, $q_c^2 \equiv Q_c^2(b^2, y^2)$, $\hat{h}^{in} \equiv H^2(\bar{q}_c, \beta, b^1, y^1)$, $\hat{h} \equiv H^2(b^1, y^1)$, and $\hat{q}_c \equiv Q_c^2(b^1, y^1)$. The government’s problem can be formally stated as (problem (P3)):

$$(P3) \quad \max_{y^1, b^1, y^2, b^2, \bar{q}_c, \beta} V^1(\bar{q}_c, \beta, b^1, y^1)$$

¹⁶The assumption that only households earning a given amount of income are eligible to opt in is made to keep the analysis as simple as possible. We will come back to this later.

subject to

$$\begin{aligned} V^2(b^2, y^2) &\geq \bar{V}, \\ V^2(b^2, y^2) &\geq V^2(b^1, y^1), \\ V^2(b^2, y^2) &\geq V^2(\bar{q}_c, \beta, b^1, y^1), \\ [y^1 - b^1 - \beta p(\bar{q}_c)(\Theta - h^{1,in})]\pi + (y^2 - b^2)(1 - \pi) &\geq \bar{R}. \end{aligned}$$

In the problem above, there are two self-selection constraints since there are two alternative mimicking strategies available to high-skilled households. One mimicking strategy is for them to choose the bundle (b^1, y^1) and opt out; the other available strategy is for them to choose the bundle (b^1, y^1) and opt in. The second and the third constraint in the government's problem above are the two self-selection constraints that jointly ensure incentive compatibility.¹⁷

For the purpose of shedding light on the welfare-enhancing power of an opting-out public provision scheme, rather than solving the problem above to characterize the optimal values for β and \bar{q}_c , we will rely on a policy reform approach and assume that the initial equilibrium is a pure income tax optimum (i.e., an optimum where only a nonlinear income tax is used). Denote by (b^{1*}, y^{1*}) and (b^{2*}, y^{2*}) the two optimal (b, y) -bundles characterizing a pure income tax optimum, and denote by q_c^{1*} the quality optimally chosen by low-skilled households at this equilibrium. Notice that by setting $\bar{q}_c = q_c^{1*}$ and $\beta = 0$ in the government's problem (P3), one replicates the pure income tax optimum.¹⁸

Even though the public provision scheme with $\bar{q}_c = q_c^{1*}$ and $\beta = 0$ is welfare-neutral, and therefore is really a redundant instrument, it is a useful starting point for our policy reform analysis. This will consist of the following steps. We will construct a policy reform that is welfare-neutral for low-skilled households and that at the same time replicates the pattern of distortions for h_c^1 and q_c^1 that are deemed desirable according to the constrained optimum characterized in Proposition 1. We will then evaluate the effects of the envisaged reform on the constraints faced by the government in maximizing the utility of low-skilled households. To keep the presentation as simple as possible, below we will characterize the structure of our policy reform for the case in which utility is quasilinear in consumption so that $u''(c) = 0$.¹⁹

¹⁷Notice that, due to our assumption that the government aims at redistributing toward low-skilled households, we disregard the self-selection constraint requiring that low-skilled households are not tempted to choose the (b, y) -bundle intended for high-skilled households.

Relaxing the assumption that eligibility for public provision is conditional on income would require to also consider in the government's problem how a policy change affects the opting-in/opting-out choice for high-skilled households not behaving as mimickers. This, however, would not affect the main qualitative insights of our analysis.

¹⁸One would obtain the same values for y^1, b^1, y^2 , and b^2 that prevail at a pure income tax optimum. Low-skilled households would be indifferent between (i) choosing (b^1, y^1) and opt in, and (ii) choosing (b^1, y^1) and opt out. The maximum utility that high-skilled households could obtain by behaving as mimickers is the same as under a pure income tax optimum given that they can always choose (b^1, y^1) and opt out.

¹⁹We have performed the same analysis for the case when $u'' < 0$. The qualitative results remain unscathed but the expression becomes much more complicated.

Remember that according to the results stated in Proposition 1, when agents only differ in market ability ($\gamma^2 = \gamma^1 \equiv \gamma$ and $\omega^2 = \omega^1 \equiv \omega$), a constrained efficient allocation requires to impose an upward distortion on h_c^1 and to leave q_c^1 undistorted.

Thus, starting from our initial equilibrium, where $\bar{q}_c = q_c^{1*}$ and $\beta = 0$, we will consider a reform that marginally raises β by $d\beta > 0$, while at the same time adjusts \bar{q}_c and b^1 in such a way that the following conditions are satisfied:

$$(9) \quad dV^1(\bar{q}_c, \beta, b^1, y^1) = (\Theta - h^{1,in})p(\bar{q}_c)u'd\beta + u'db^1 \\ + [\gamma f'_2 - (1 - \beta)p'(\bar{q}_c)u'](\Theta - h^{1,in})d\bar{q}_c = 0,$$

$$(10) \quad \gamma(\omega f''_{12} - \bar{q}_c f''_{22})\left(\frac{dh^{1,in}}{d\beta} + \frac{dh^{1,in}}{d\bar{q}_c}d\bar{q}_c + \frac{dh^{1,in}}{db^1}db^1\right) \\ + (\Theta - h^{1,in})\gamma f''_{22}d\bar{q}_c = 0,$$

where all the various derivatives of the f -function are evaluated at $(\omega h^{1,in}, (\Theta - h^{1,in})\bar{q}_c)$.

Condition (9) postulates that the reform is welfare-neutral for low-skilled households; condition (10) postulates that the overall effect of the reform is to leave the quality of child care services at an undistorted value for low-skilled households who opt in.²⁰

Exploiting the fact that at the initial equilibrium, $\beta = 0$ and $\bar{q} = q_c^{1*}$, so that $\gamma f'_2 - (1 - \beta)p'(\bar{q}_c)u' = 0$, we show in online Appendix B4 that we must have

$$(11) \quad d\bar{q}_c = \frac{(\omega f''_{12} - \bar{q}_c f''_{22})p(\bar{q}_c)u'}{\gamma(\Theta - h^{1,in})(\omega)^2[(f''_{12})^2 - f''_{11}f''_{22}] - (\Theta - h^{1,in})v''f''_{22}}d\beta < 0,$$

$$(12) \quad db^1 = -(\Theta - h^{1,in})p(\bar{q}_c)d\beta < 0.$$

Notice that since low-skilled households were indifferent between opting in and opting out at the pre-reform equilibrium, the proposed reform, while keeping their utility unchanged, makes them strictly prefer to opt in. Notice also that since low-skilled households prefer to opt in, they will choose $h^{1,in}$ in accordance to (8), which implies (given that $\beta > 0$ at the post-reform equilibrium) that their choice for hours of center-based child care is upward distorted.

Consider now the effects of the proposed reform, which is welfare-neutral by construction for low-skilled households, on the constraints faced by the government.

²⁰Condition (10) is obtained by totally differentiating the no-distortion condition

$$\gamma f'_2(\omega h^{1,in}, (\Theta - h^{1,in})\bar{q}_c) - p'(\bar{q}_c)u'(b^1 - (1 - \beta)(\Theta - h^{1,in})p(\bar{q}_c)) = 0,$$

and taking into account that we are assuming $u'' = 0$. Notice that at the initial equilibrium, where $\bar{q}_c = q_c^{1*}$ and $\beta = 0$, the condition above is satisfied. This is obvious since the initial equilibrium replicates the pure income tax optimum. At a pure income tax optimum, the individual choices for h_c and q_c satisfy the no-distortion conditions (5) and (6).

Since the reform did not change y^2 and b^2 , the reform is clearly also welfare-neutral for high-skilled households who do not behave as mimickers. Regarding the effect on the government's budget constraint, denoting by μ the Lagrange multiplier attached to this constraint, we have that the effect is given by

$$\mu\pi\left\{-db^1 - p(\bar{q}_c)(\Theta - h^{1,in})d\beta - \beta p'(\bar{q}_c)(\Theta - h^{1,in})d\bar{q}_c + \beta p(\bar{q}_c)dh^{1,in}\right\}.$$

However, taking into account that β was equal to zero at the pre-reform equilibrium and substituting $-(\Theta - h^{1,in})p(\bar{q}_c)d\beta$ for db^1 , the expression above is equal to zero. Thus, the proposed reform is also budget-neutral for the government.

What is left to assess is the impact of the reform on high-skilled households who were to behave as mimickers. In online Appendix B5, we show that the proposed reform is mimicking-deterring in the sense that it mitigates both the self-selection constraints that appear in the government's problem.

In summary, we have shown that, starting from a pure income-tax optimum, it is possible to implement an opting-out public provision scheme that (i) replicates the pattern of distortions for h_c^1 and q_c^1 that are deemed desirable according to a constrained optimum; (ii) is welfare-neutral for all households not behaving as mimickers; (iii) is budget-neutral for the government; and (iv) lowers the utility of high-skilled mimickers under both available mimicking strategies. Hence, a reform along the suggested lines opens the possibility for the government to achieve a Pareto improvement upon the initial equilibrium.

III. Quantitative Model

We now proceed with the quantitative analysis. We first describe the quantitative model and our calibration that is carried out taking into account the current US tax system and existing child care subsidies. In Section IIIG, we turn to our optimal tax analysis where we let the government optimize the nonlinear income tax schedule and the child care subsidy scheme.

A. General Setting

The quantitative model is an extension of the simplified model considered in Section II. Most importantly, we extend the setting to more than two types and consider two-earner couples consisting of mothers and fathers. We make the assumption that market ability and nurturing ability are positively correlated and that there is assortative mating. This allows us to summarize the ability type of a given household by means of a single parameter $i \in \{1, \dots, N\}$, which we refer to as the *household skill type* (where a higher index corresponds to a higher ability).²¹ Moreover, in line with previous literature (e.g., Cogan 1981, Blundell and Shephard 2012), we also assume that in each household, there is a fixed cost associated with the mother's

²¹ To completely represent all the different types of couples in the economy, we would need to consider a matrix of household types where each element ij corresponds to the couple where the father has skill type i and the mother has skill type j . Allowing for such a rich type structure would, however, be computationally intractable.

labor force participation.²² More specifically, we assume that mothers in households of type i differ in their fixed cost type $j \in \{1, \dots, \Psi\}$ and incur a fixed cost of χ_{ij} when entering the labor force. Thus, the type space is fully characterized by the tuple $(i, j) \in \{1, \dots, N\} \times \{1, \dots, \Psi\} \equiv \Theta$. This means that, even though we have assumed that household wage rates can be summarized into a unidimensional parameter, the type-space is still bidimensional by virtue of the heterogeneity in the fixed costs of work.

Since our emphasis in the quantitative analysis is on the labor market decisions made by mothers, we have chosen to let the skill type of the household correspond to the skill category of the mother. Thus, the wage pair (w_m^i, w_f^i) , where the subscript m refers to “mother” and the subscript f refers to “father,” represents households where the mother belongs to the i th skill category in the wage distribution for mothers and w_f^i is the average wage of all fathers matched with type- i mothers. This implies that the skill level of the household is described by the mother’s relative position (rank) in the wage distribution of mothers.²³ The procedure to compute the wages based on actual data is described in detail in Section IIIC below. In our simulations we consider $N = 5$ skill types and a continuous distribution of cost types (approximated by $\Psi = 1,000$).²⁴

B. Household Decision Problem

Let L_k , h_k , and ℓ_k denote, respectively, the labor supply, domestic care, and leisure of spouse j (for $k = m, f$) and let w_k and ω_k denote, respectively, the wage rate and nurturing ability of spouse k . Moreover, denote by h_c and q_c the quantity and quality of formal care, and denote by q the human capital of the child, with q given by

$$q = \gamma^i f(\omega_m^i h_m, \omega_f^i h_f, q_c h_c).$$

The general formulation of the problem solved by each household $(i, j) \in \Theta$ in the presence of existing taxes and child care subsidies can be described as follows:

$$\max_{c, L_m, L_f, h_m, h_f, \ell_m, \ell_f, h_c, q_c} \{u(c) + g(q) + v(\ell_m) + v(\ell_f) - \mathbf{1}[L_m > 0] \cdot \chi^{ij}\}$$

subject to the household budget constraint

$$(13) \quad c = w_f^i L_f + w_m^i L_m - T^{US}(w_f^i L_f + w_m^i L_m) - CE(D, w_f^i L_f, w_m^i L_m),$$

²² These costs enable the model to replicate empirically relevant shares of nonworking households and can be viewed, for example, as representing the psychological costs associated with leaving a child in the care of a nonparent.

²³ In principle, the skill rate of the household could be constructed based on the mother’s skill, the father’s skill, or any combination of the two. We have chosen to let the mother dictate the skill level of the household since we focus on their labor force participation decisions.

²⁴ Computational considerations prevent us from expanding the model beyond five skill types. However, taking into account that for each skill type we have both two-earner households and one-earner households, the model features ten household types.

and the following time constraints (for, respectively, the mother, the father, and the child):

$$(14) \quad L_m + h_m + \ell_m = 1,$$

$$(15) \quad L_f + h_f + \ell_f = 1,$$

$$(16) \quad h_c + h_m + h_f = 1.$$

In the above formulation, we have normalized to 1 the time endowment for each member of the household. Moreover, $\mathbf{1}[L_m > 0]$ is an indicator function that takes the value 1 when both spouses are working, T^{US} is the tax function and CE is the *net* (of subsidy) child care expenditure of the household as a function of gross child care expenditure $D \equiv p(q_c)h_c$ and the income of both spouses, $w_f^i L_f$ and $w_m^i L_m$.

The functions T^{US} and CE are chosen to approximate the rules governing taxes and child care subsidization in the United States. In specifying T^{US} , we follow Heathcote, Storesletten, and Violante (2014) and assume the following parametric form for T^{US} :

$$T^{US}(y) = y - \lambda y^{1-\tau},$$

which implies that the relationship between post-tax income \tilde{y} and pre-tax income y is given by $\tilde{y} = \lambda y^{1-\tau}$ or, equivalently, $\log(\tilde{y}) = \lambda + (1 - \tau)\log(y)$, which we estimate by OLS using information on the relationship between \tilde{y} and y provided by NBER TAXSIM. Using a sample of households filing jointly with small children (below the age of 6) and assuming *zero* child care expenditure, we find $\tau = 0.164$ and $\lambda = 1.31$.

In constructing the functional form for the CE function, we take into account the CTC (Child Tax Credit) and the CDCTC (Child and Dependent Care Tax Credit), which are federal tax credits.²⁵ In addition, we model the state tax credit that applies in California (which is a function of the CDCTC) and the subsidies that are offered through the CCDF (Child Care and Development Fund).²⁶ Since the actual rules governing child care involve various kinks, we calculate the CE function using a smooth approximation to facilitate incorporation in our computational model.

The purpose of our quantitative exercises will later be to consider various reforms where we replace the T^{US} function with an optimally chosen nonlinear income tax schedule and where we replace the CE function with optimized child care subsidies.

²⁵The CTC is not, strictly speaking, a child care subsidy since there is no requirement that paid child care be used. Thus, it is actually a child subsidy. Eligibility for the CTC is not conditional on being employed and its amount is (weakly) decreasing in the household AGI (adjusted gross income). The CDCTC is a nonrefundable tax credit available to families with children aged under 13 and covers part of child care expenses. Being employed is a requirement that must be fulfilled to be eligible for the CDCTC.

²⁶US states are offered a block grant from the federal government in the form of the CCDF. The purpose of the CCDF is to increase the availability, affordability, and quality of child care services. In online Appendix C, we describe in more detail the rules governing the various subsidy programs that we model in our analysis.

C. Data on Market Wage Rates, Participation, and Hours of Work

We use the Current Population Survey (CPS) Labor Extracts 2003–2006 (NBER 2007) as our main data source. We compute the average wage, labor market participation rate, and work hours for husbands and wives by household skill level (as defined above) for each of the years 2003–2006 and then take averages to make our calculations less sensitive to year-specific shocks. Wage rates are obtained by dividing weekly earnings by weekly hours of work. Our sample contains all married couples between age 20 and 65 who were not self-employed and who had at least one child below the age of 6. All wages are expressed in terms of 2006 US dollars.

To obtain a wage rate for mothers who lack a wage observation we follow an imputation procedure. We regress log wages on a set of covariates, including flexible controls for age and education. We also include the education level and age of the husband in this regression. The regression generates a set of predicted values for mothers who lack a wage observation. However, all these predictions lie on the regression surface. To obtain correct moments of the distribution of female wages, we draw a large number of samples from the empirical distribution of the residuals in the prediction regression and add these to the predicted wages. The final measure of the wage rate for mothers is equal to the actual wage, whenever it exists, and equal to the predicted wage otherwise.²⁷ The wage distributions for mothers and fathers are approximated using the deciles of the predicted wage distributions.

We focus on hours worked per week, measured in terms of the “usual weekly working hours” during a typical work week. This might be missing some variation that stems from the fact that some workers can have more than one job.²⁸ In addition, some variation in annual hours of work stems from the number of weeks worked during a year. For simplicity, we multiply weekly earnings by 48 to get a measure of annual earnings. Since we do not want to overfit our model to potentially noisy wage data, our calibration procedure targets average working hours for mothers and fathers.

The spousal wage rates, hours of work, and the labor market participation of mothers and fathers associated with the different household types $i = 1, \dots, 5$ are displayed in Table 1.

D. Functional Forms

We assume that the utility function of households takes the following form:

$$(17) \quad U = c^\alpha + q^\beta - \frac{1}{\ell_m} - \frac{1}{\ell_f} - \mathbf{1}_{[L_m > 0]} \cdot \chi^{ij}.$$

²⁷This procedure neglects the fact that workers and nonworkers might be different along unobservable dimensions, resulting in selection. This is a standard issue in the literature and is usually addressed by adding a selection term to the prediction equation. However, the credibility of such corrections is severely hampered by functional form assumptions and lack of suitable instruments. For robustness, we have performed a selection correction using county as instrument. This turned out to have a very minor impact on the discrete wage distributions that we use in our simulations.

²⁸There is also a variable in the CPS called “hours worked last week” that potentially could capture the labor supply associated with multiple jobs. However, we did not use this variable since it is plagued by measurement error (e.g., some workers report that last week they worked 0 whereas in a *usual* working week they would work 40 hours).

TABLE 1—HOURLY WAGE RATES (2006 US\$), WEEKLY HOURS OF WORK, AND LABOR FORCE PARTICIPATION RATES (LFP) FOR MOTHERS AND FATHERS IN OUR DATA

Type	w_m	w_f	$Hours_m$	$Hours_f$	LFP_m	LFP_f
1	7.00	14.01	31.99	43.03	0.53	0.95
2	10.62	16.07	34.09	43.74	0.61	0.96
3	14.25	18.49	35.61	43.96	0.63	0.96
4	19.40	21.53	35.84	44.40	0.65	0.97
5	30.47	27.39	34.38	44.74	0.72	0.97

In the above specification, we have employed functional forms that are suitable for numerical computation. The household derives utility from private consumption and the child’s human capital through the concave functions c^α and q^β where $0 < \alpha < 1$ and $0 < \beta < 1$. In addition, the household derives utility from spousal leisure ℓ_m and ℓ_f through the increasing and concave function $v(x) = -1/x$ in order to produce desirable labor supply behavior.²⁹ The human capital of the child is assumed to depend on the overall quality of the child care arrangement, q , that is given by

(18)
$$q = \gamma^i f(\omega_m^i h_m, \omega_f^i h_f, q_c h_c) = \gamma^i [\omega_m^i h_m^\rho + \omega_f^i h_f^\rho + q_c h_c^\rho]^\frac{1}{\rho}.$$

Equation (18) is a three-input CES-production function where the relative importance of the production factors is endogenous due to the fact that q_c is a choice variable. There are few estimates available on the elasticity of substitution between different modes of child care, and this is clearly an area where more research is needed. Since our model is already quite rich, we adopt the normalization $\rho = 0.5$. We furthermore assume that home care productivities are related to market productivity by assuming that $\omega_m^i = \omega_f^i = \nu((w_m^i + w_f^i)/2)$, where $\nu > 0$ is a type independent scale parameter used to ensure that the scale of ω and q_c are comparable inside the production function. Notice that since mothers and fathers in a given household have the same ω , mothers generally have a comparative advantage in household work due to their lower market wage rate (except in the highest-skilled household where the market wage rate is higher for mothers). Finally, the production function in (18) features the parameter γ^i , which is set equal to w_m^i/\bar{w}_m (where \bar{w}_m is the average wage rate of mothers). This parameter captures differences in the innate abilities of children, due to, for instance, some degree of genetic transmission of ability within households.³⁰

The hourly price of child care is a function of the quality chosen by agents and is described by the isoelastic function $p(q_c) = kq_c^\sigma$. This is equivalent to the log-log relationship between the hourly price of child care and child care quality used by Blau

²⁹ The form of the leisure term in the utility function has previously been employed in the optimal tax context by, for example, Tuomala (2010).

³⁰ The particular specification we use is motivated by the fact that we let the skill type of the household be determined by the mother’s wage rate. In online Appendix E, we provide a robustness check where we let the parameter be defined based on the average household wage rate.

and Mocan (2002). We further assume that $k > 0$ and $\sigma > 1$, implying a convex price function in line with the relationship described by Havnes and Mogstad (2015).

E. Calibration

The parameters of the model are jointly optimized in order to minimize a loss function with the purpose of matching the following empirical targets: (i) average hours of work for mothers and fathers, (ii) female labor force participation rates, (iii) the profile of child care expenditure as a fraction of total household income across the household skill distribution, and (iv) the profile of the average hourly cost of child care across the household skill distribution. We describe how we match each of these targets in the paragraphs below.

Hours of Work.—As already mentioned, each adult household member is endowed with one unit of time that can be allocated to hours at the job, hours in maternal/paternal care, and leisure. We interpret the unitary time endowment as representing the time available during a year after having deducted the time needed for sleep. Thus, the unitary time endowment corresponds to 5,840 hours. Since children aged 0–6 sleep more than adults, the time endowment for the child (i.e., the time during which the child's human capital can be affected) is set to 80 percent of the adult time endowment. Our dataset contains hours worked during a usual working week. Given our assumption that each agent works 48 weeks during a year, an agent that works 40 hours per week spends a fraction $(48 \times 40)/5,840 \approx 0.33$ of his/her time endowment on the job. Thus, a father who works 40 hours a week will have $L_f = 0.33$. Consistent with the data reported in Table 1, we calibrate the average labor supply of mothers in the model to approximately be equal to $(34.4 \times 48)/5,840 \approx 0.28$ and $(44 \times 48)/5,840 \approx 0.36$ for fathers.

Hours of Domestic Care.—Table 2 describes how married women with children in ages 0–6 divide their time. As already mentioned above, we assume that agents who are employed work 48 weeks per year. Thus, for employed mothers, there are $48 \times 5 = 240$ weekdays and $4 \times 7 + 48 \times 2 = 124$ holidays/weekend days. Using the value for nonemployed mothers to apply during weekends, we use $((1.9 \times 240)/5,840) + ((4.2 \times 124)/5,840) \approx 0.17$ as the target for the fraction of the time endowment full-time employed mothers spend in domestic care in our calibration.³¹

Labor Force Participation.—The fixed costs $\{\chi^{ij}\}_{(i,j) \in \Theta}$ associated with the secondary earner's labor market participation are calibrated to match the empirical skill-specific labor force participation rates. More specifically, the distributions of

³¹ For part-time employed mothers the corresponding number is $((2.7 \times 240)/5,840) + ((4.2 \times 124)/5,840) \approx 0.20$, and for nonworking mothers, the target is $((4.2 \times 52 \times 7)/5,840) \approx 0.26$. Our calibration is, by and large, consistent with these numbers as well. The labor supply for a full-time working mother in the ATUS is, according to Table 2, equal to 6.7 hours per working day. This amounts to $6.7 \times 5 = 33.5$ hours per week, which is in line with the average labor supply reported in the CPS (see Table 1).

TABLE 2—WEEKDAY TIME USE OF MARRIED WOMEN LIVING WITH YOUNG CHILDREN,
BY EMPLOYMENT STATUS (AVERAGE HOURS PER DAY)

	Not employed	Employed part-time	Employed full-time
Sleeping	8.5	8.5	8.2
Household activities	3.8	2.2	1.6
Caring for household children	4.2	2.7	1.9
Working and related activities	≈0	3.7	6.7
Leisure and sports	3.5	3.3	2.4

Notes: Data include all married women, ages 25 to 54, with a child under age 6 present in the household. Data include nonholiday weekdays and are annual averages for 2015.

Sources: Bureau of Labor Statistics (2015), American Time Use Survey 2015 (<https://www.bls.gov/tus/charts/chart2.txt>)

fixed costs associated with mothers’ labor force participation are chosen so that the model, under the benchmark US tax system, matches the household-specific motherly employment rates in Table 1.³² For this purpose, we have proceeded in the following way. For each skill type i , we compute the fixed cost that would make a mother of type i indifferent between working and not working in the calibrated benchmark economy. Denote this fixed cost threshold $\bar{\chi}^i$. Notice that mothers with $\chi^{ij} \leq \bar{\chi}^i$ will work and mothers with $\chi^{ij} > \bar{\chi}^i$ will stay out of the labor force. We further assume that the lower bound of the fixed cost is 0. If the fraction of working mothers of type i in the data is z^i , we want to assign a fixed cost of less than $\bar{\chi}^i$ to a fraction z^i of the mothers of type i , and a higher fixed cost for the remaining part of the population. We achieve this by assuming that the fixed costs are given by the power function³³

$$\chi_j = a_0(j - 1)^{a_1} + a_2, \quad j = 1, \dots, \Psi,$$

where we estimate the parameters a_0, a_1 , and a_2 using nonlinear least squares on the data points $\{0, 0\} \cup \{\Psi z^i, \bar{\chi}^i\}_{i=1}^N$.³⁴ Intuitively, the relationship is monotonically increasing.

Pattern of Child Care Expenditure.—We want our model to produce realistic patterns of child care expenditure as a fraction of family income across the household skill distribution. For this purpose, we have examined two sources of data. The Annual Social and Economic Supplement of the Current Population Survey (CPS ASEC) as well as the Survey of Income and Program Participation (SIPP).

³²Notice that by setting the fixed cost distributions appropriately, it is always possible to match any particular pattern of empirical participation rates. For example, if the fraction of mothers who work in household of type 3 is 52 percent and the number of cost types Ψ is equal to 100, we can always set $\chi_{3j} = -\infty, j = 1, \dots, 52$ and $\chi_{3j} = +\infty, j = 53, \dots, 100$. However, this would make the labor force participation of type 3 mothers completely inelastic.

³³A power distribution for the fixed costs of work has previously been used by Kleven, Kreiner, and Saez (2009).

³⁴The estimated parameters are $a_0 = 4.109\text{e-}11$, $a_1 = 3.518$ and $a_2 = 0.005028$. Notice that the slope of the fixed cost distribution is related to the concept of “participation elasticity” emphasized in the public finance literature. Our approach implies heterogeneous participation elasticities depending on the skill-specific employment level, where the relationship is established through structural assumptions and a calibration procedure.

The important empirical feature of both these data sources is that the ratio of child care expenditure to family income is higher for low-income families as compared to high-income families in households where the mother works. In our calibration, we target the numbers provided by Laughlin (2013), who finds, using the 2008 SIPP, that child care expenditure as a fraction of family income ranges between around 25 percent at the bottom to around 15 percent at the top of the income distribution.³⁵

Hourly Price of Child Care.—To pin down the parameters of our price function, we target empirical patterns of the hourly price of child care, $p(q_c)$, across the household skill distribution for two-earner couples. We use the figures reported by Whitehurst (2018, 9), who finds that the hourly price of child care ranges between around US\$3 to around US\$9 across the household income distribution.

F. The Calibrated Economy

A summary of the calibrated parameters in the model and their calibrated values is presented in Table 3. This table also contains a summary of the empirical targets that are used in our calibration as well as the corresponding variables in our model output (see also Table 4 below).³⁶

In Table 4, we show the allocation for the benchmark calibrated economy where households face the current US tax system and child care subsidies. The top panel in this table describes the time allocation for mothers and fathers in families where the mother works, and the bottom panel describes the time allocation for parents in families where the mother does not work.

The column 1 – CE/D in Table 4 shows the fraction of child care expenditure that is paid by the government. Since households where one spouse does not work are ineligible for the subsidies that we consider in the calibrated model, this fraction is equal to zero for all one-earner households. For two-earner households, the effective subsidy ranges between 51 percent and 10 percent, and is monotonically decreasing in the skill type of households.

The column T/y reports the average income tax rate paid by the various households, which ranges between 5 percent (for one-earner households of type 1) and 18 percent (for two-earner households of type 5). The column q shows that the human capital of children is increasing in the skill type of an household. It is lowest in households of type 1 where the mother does not work and highest in households of type 5 where the mother does not work. The same pattern characterizes the overall quality of the child care arrangement.³⁷

³⁵ Calculated based on Laughlin (2013, Table 6). Another calculation is Herbst (2018), who finds that the corresponding figures range between around 17 percent at the bottom and 8 percent at the top (Herbst 2018, Table 7).

³⁶ Notice that the parameters are jointly chosen to match the empirical targets using a model-fitting procedure.

³⁷ According to (18) the human capital q depends also on γ^i , which captures the effect of the innate ability of a child and is increasing in the skill type of an household. Interpreting q/γ^i as a measure of the overall quality of the child care arrangement chosen by a household, its profile is consistent with the profile of q : both among one-earner couples and two-earner couples it is increasing in the skill type of a household; moreover, it is lowest in households of type 1 where the mother does not work and highest in households of type 5 where the mother does not work.

TABLE 3—SUMMARY OF PARAMETER VALUES AND CALIBRATION TARGETS

Parameter	Value	Interpretation	
<i>Parameter values</i>			
α	0.52	Curvature utility private consumption	
β	0.49	Curvature utility human capital	
k	0.60	Scale parameter price function	
σ	4.10	Curvature parameter price function	
ν	0.10	Scale parameter household production	
Variable	Target	Model	Description
<i>Calibration targets</i>			
L_m	0.28	0.27	Average hours of work (mothers)
L_f	0.36	0.38	Average hours of work (fathers)
h_m	0.17	0.15	Average hours of maternal care
D/y	0.25–0.15	0.25–0.15	Expenditure share profile
$p(q_c)$	3–9 US\$	2.7–8.9	Hourly cost profile

TABLE 4—BENCHMARK ALLOCATION (CALIBRATED ECONOMY)

i	y	c	L_m	L_f	h_m	$\frac{D}{y}$	q	q_c
<i>Panel A. Allocation in households where the mother works</i>								
1	44.50	36.28	0.18	0.45	0.08	0.25	0.81	1.45
2	55.46	42.83	0.25	0.42	0.11	0.22	1.48	1.51
3	66.00	49.07	0.28	0.40	0.14	0.20	2.42	1.58
4	77.88	55.91	0.30	0.35	0.18	0.18	4.16	1.68
5	98.04	67.69	0.32	0.26	0.24	0.15	9.66	1.93
i	$p(q_c)$	$\frac{T}{y}$	$T'(y)$	$1 - \frac{CE}{D}$	U			
1	2.72	0.06	0.22	0.51	1.50			
2	3.29	0.10	0.25	0.40	2.34			
3	3.94	0.12	0.27	0.31	3.17			
4	5.09	0.15	0.29	0.22	4.24			
5	8.89	0.18	0.31	0.10	6.43			
i	y	c	L_m	L_f	h_m	$\frac{D}{y}$	q	q_c
<i>Panel B. Allocation in households where the mother does not work</i>								
1	41.24	34.74	0.00	0.50	0.14	0.11	0.76	1.18
2	47.09	37.93	0.00	0.50	0.21	0.12	1.50	1.31
3	53.45	41.58	0.00	0.49	0.27	0.13	2.53	1.43
4	60.13	45.32	0.00	0.48	0.34	0.14	4.44	1.58
5	70.13	51.08	0.00	0.44	0.43	0.14	10.43	1.89
i	$p(q_c)$	$\frac{T}{y}$	$T'(y)$	$1 - \frac{CE}{D}$	U			
1	1.18	0.05	0.21	0	1.36			
2	1.81	0.07	0.23	0	2.10			
3	2.59	0.09	0.24	0	2.86			
4	3.93	0.11	0.26	0	3.87			
5	8.21	0.13	0.28	0	5.98			

Notes: In the table, y and c denote annual household income and consumption, respectively, expressed in thousands of 2006 US dollars. Moreover, L_j denotes labor supply, CE net child care expenditure, D gross child care expenditure, T income tax liability, q overall quality of child care arrangement, q_c quality level of paid care arrangement, and U household utility. Finally, $1 - CE/D$ is the implicit child care subsidy rate in the current US tax system.

G. Optimal Tax Systems

The problem of finding the optimal tax and child care policy represents a bilevel programming problem. To evaluate the social welfare level associated with a particular policy set by the government it is necessary to compute how agents optimally respond to this policy. Thus, there is an upper-level (government) optimization problem and a lower-level optimization problem that is solved by each type of household in the economy. We describe the computational challenges and our computational approach in online Appendix D.

To achieve tractability, and reduce the type-space, we assume that the fixed cost of work is a utility cost entering additively in the utility function. This implies that among equally skilled households, all households will make the same choices regarding the individual decision variables provided that the mother has the same labor force participation status. Moreover, we know that among equally skilled mothers, those with a higher fixed cost will always be less likely to participate in the labor force.³⁸ This allows us to identify, for each skill group, a unique marginal worker that is indifferent between working and not working. Mothers with a fixed cost greater than the marginal worker will always stay out of the labor force, and mothers with a lower fixed cost than the marginal worker will be working. This means that at each skill level, we only need to compute the optimal individual decisions for a representative two-earner household and for a representative one-earner household, rather than computing these decisions for each possible fixed cost type.³⁹ It also implies that the government only needs to design two set of bundles for each type i . One pre-tax/post-tax income point for two-earner households of type i and one pre-tax/post-tax income point for one-earner households of type i . This drastically reduces the number of incentive constraints that need to be incorporated into the government's problem and also allows us to employ a large number of discrete cost types.

The labor force participation decision of mothers is represented by a binary matrix \mathbf{L} where $\mathbf{L}_{ij} = 1$ if the mother of type (i, j) is working and zero otherwise. Since the fixed cost of work χ_{ij} is assumed to be nondecreasing in j , the rows of \mathbf{L} will be nonincreasing when moving from the left to the right. This allows us to introduce the vector \mathbf{P} where \mathbf{P}_i is the number of leading ones along row i . Notice that \mathbf{P}_i is also equal to the fixed cost type of the worker who is, at the margin, indifferent between working and not working. The fixed cost of the marginal worker among households of type i can be computed as $x^i = F_{\chi}^{-1}(\mathbf{P}_i)$ if $F_{\chi}(x) \in [0, 1]$ is the CDF of the fixed cost distribution.

We assume that the government maximizes the sum of individuals' utilities, subject to a concave transformation $W(\cdot)$ (reflecting society's taste for redistribution). Thus, the welfare gain of subsidizing child care will be measured in terms of the effectiveness of raising social welfare. Following Brewer et al. (2010) we

³⁸In contrast, if the fixed cost of work was modeled as a monetary cost, there would be a countervailing income effect.

³⁹Notice that without the assumption that the utility cost is additive, there would be a huge increase in the number of individual decision problems that need to be computed, making the problem computationally intractable.

focus on a logarithmic transformation of individual utility.⁴⁰ We further assume that two distinct nonlinear income tax schedules apply to one-earner and two-earner households.⁴¹

In the absence of any kind of subsidies to child care expenditures (i.e., in the case of a pure income tax optimum), the government's problem can be described as follows:

$$(19) \quad \max_{\mathbf{P}} \Omega(\mathbf{P}),$$

$$(20) \quad \Omega(\mathbf{P}) = \max_{\{(y_1^i, b_1^i), (y_0^i, b_0^i)\}_{i=1}^N} \sum_{i=1}^N \left(\sum_{j > \mathbf{P}_i} \pi^{ij} W(V_0^i(y_0^i, b_0^i)) + \sum_{j \leq \mathbf{P}_i} \pi^{ij} W(V_1^i(y_1^i, b_1^i) - \chi_{ij}) \right),$$

subject to

$$(21) \quad V^{ij}(y_0^i, b_0^i, y_1^i, b_1^i) \geq \tilde{V}^{ij}(y_0^{i-1}, b_0^{i-1}, y_1^{i-1}, b_1^{i-1}), \quad \forall i \in \{2, \dots, N\}, \forall j,$$

$$(22) \quad V^{ij}(y_0^i, b_0^i, y_1^i, b_1^i) = \begin{cases} V_0^i(y_0^i, b_0^i) & \text{if } j > \mathbf{P}_i \\ V_1^i(y_1^i, b_1^i) - \chi^{ij} & \text{if } j \leq \mathbf{P}_i, \end{cases}$$

$$(23) \quad \tilde{V}^{ij}(y_0^{i-1}, b_0^{i-1}, y_1^{i-1}, b_1^{i-1}) = \max \{ V_0^i(y_0^{i-1}, b_0^{i-1}), V_1^i(y_1^{i-1}, b_1^{i-1}) - \chi^{ij} \},$$

$$(24) \quad V_0^i(y_0^i, b_0^i) > V_1^i(y_1^i, b_1^i) - \chi^{ij}, \quad \forall i, \forall j > \mathbf{P}_i,$$

$$(25) \quad V_0^i(y_0^i, b_0^i) < V_1^i(y_1^i, b_1^i) - \chi^{ij}, \quad \forall i, \forall j \leq \mathbf{P}_i,$$

$$(26) \quad \sum_i \left(\sum_{j \leq \mathbf{P}_i} \pi^{ij} (y_0^i - b_0^i) + \sum_{j > \mathbf{P}_i} \pi^{ij} (y_1^i - b_1^i) \right) \geq \bar{R},$$

$$(27) \quad V_0^i(y, b) = \max_{q_c, h_c, h_m} u^i \left(b - p(q_c)h_c, h_m, 0, \frac{y}{w_f^i}, h_c, q_c \right), \quad \forall i,$$

$$(28) \quad V_1^i(y, b) = \max_{q_c, h_c, h_m, L_m} u^i \left(b - p(q_c)h_c, h_m, L_m, \frac{y - w_m^i L_m}{w_f^i}, h_c, q_c \right), \quad \forall i.$$

The first thing to notice is that the government's problem features three levels of optimization. Equation (19) defines the upper-level optimization in which the government chooses the participation rate at each skill level to maximize $\Omega(\mathbf{P})$. The function $\Omega(\mathbf{P})$ is, in turn, the value function associated with the middle or "main" layer of optimization where the government strives to find the income tax schedule (defined in terms of the pre-tax/post-tax income points) that maximizes a social welfare

⁴⁰ In a previous version of the paper, we have analyzed other social welfare functions, obtaining similar qualitative results as in the present paper.

⁴¹ This only requires that the labor force participation decision is observable by the government.

function. Notice that in the main optimization problem, the parameters \mathbf{P}_i are treated as exogenous.

Turning now to the constraints of the main optimization problem, the set of incentive constraints appear in (21). These constraints ensure that each household prefers the bundle assigned to it rather than the bundle intended for the adjacent lower-skilled household. Equations (22) and (23) define the left-hand side and right-hand side of the incentive constraints where the parameters $\{\mathbf{P}_{ij}\}_{i=1}^N$ determine whether the relevant utility for an agent of type (i, j) is that which arises if the mother is not working (V_0^i) or that which arises if the mother is working ($V_1^i - \chi^{ij}$). Notice that equation (23) implies that if a type i household decides to mimic a household of type $i - 1$, it must replicate the labor force participation decision of type $i - 1$.⁴² Inequalities (24) and (25) are individual rationality constraints that ensure that the labor force participation decisions prescribed in the \mathbf{P} vector are actually the ones maximizing household utility. Constraint (26) is the government budget constraint stating that the sum of tax revenue from one-earner and two-earner households should sum up to the exogenous revenue requirement \bar{R} .⁴³ The last two equations define the indirect utilities for households where the mother does not work (equation (27)) and the indirect utilities (gross of the fixed cost of work) for households where the mother works (equation (28)). The computation of these two indirect utilities for each type- i household represents the lower-level optimization problem.⁴⁴

For the upper layer that is responsible for finding the optimal participation vector \mathbf{P} , we use a global optimization heuristic that relies on a combination of coarse searches over the full parameter space and local searches around the best coarse point.⁴⁵ For the middle and lower layers (i.e., the bilevel optimization problem), we rely on an efficient implementation in C++, interfacing the latest version of the state-of-the-art solver for nonlinear constrained optimization problems—KNITRO.

IV. Quantitative Results

In our quantitative analysis, we consider three cases. In each of these cases, the government chooses optimally a nonlinear tax on household income. In the first case, the nonlinear income tax is the only policy instrument. In the second case, the government can also subsidize child care expenditures at a proportional rate, which is allowed to depend on both the mother's employment status and household income. Finally, in the third case, we allow for a simple opting-out public provision scheme where the quality of the publicly provided care is optimally chosen by the

⁴²This is a weak simplifying assumption. The assumption that it is only possible to mimic adjacent types is potentially stronger, as letting the mother drop out of the labor force could be a way for a high-skilled household to replicate the taxable income of a much more low-skilled, two-earner household.

⁴³In our numerical simulations, the revenue requirement \bar{R} is always set equal to the fiscal surplus that arises in our US benchmark economy (described in Section IIIF).

⁴⁴Notice that these utilities must be evaluated both when a household acts truthfully and when the household behaves as a mimicker.

⁴⁵For computational tractability, we limit the precision of the search to steps of 5 percentage points in each dimension of P . In addition, we impose that the labor force participation is monotonically increasing in the household skill level, i.e., $P_i \geq P_j$ for all $i \geq j$, $i, j \in \{1, \dots, N\}$. Finally, we impose that the maximum employment rate at any skill level is 95 percent, reflecting the realistic assumption that there is a certain fraction of the population with very high fixed costs of work who would not be willing or able to work regardless of the financial incentives.

TABLE 5—OPTIMAL NONLINEAR INCOME TAX

<i>i</i>	<i>y</i>	<i>c</i>	<i>L_m</i>	<i>L_f</i>	<i>h_m</i>	$\frac{D}{\bar{y}}$	<i>q</i>	<i>q_c</i>
<i>Panel A. Allocation in households where the mother works</i>								
1	47.22	45.09	0.21	0.47	0.08	0.12	0.93	1.22
2	58.04	50.24	0.27	0.44	0.11	0.13	2.14	1.34
3	79.85	62.84	0.36	0.46	0.11	0.13	4	1.46
4	94.8	71.35	0.37	0.42	0.15	0.14	8.58	1.61
5	141.02	100.76	0.44	0.39	0.18	0.14	26.5	1.9
<i>i</i>	<i>p(q_c)</i>	$\frac{T}{\bar{y}}$	<i>T'(y)</i>	β	<i>U</i>			
1	1.34	0.05	0.19	—	2.07			
2	1.99	0.13	0.26	—	3.17			
3	2.87	0.21	0.14	—	4.43			
4	4.22	0.25	0.17	—	6.41			
5	8.31	0.29	0	—	11.13			
<i>i</i>	<i>y</i>	<i>c</i>	<i>L_m</i>	<i>L_f</i>	<i>h_m</i>	$\frac{D}{\bar{y}}$	<i>q</i>	<i>q_c</i>
<i>Panel B. Allocation in households where the mother does not work</i>								
1	38.05	37.77	0	0.47	0.14	0.12	0.98	1.2
2	45.17	40.98	0	0.48	0.21	0.14	2.27	1.33
3	48.85	42.38	0	0.45	0.27	0.14	4.49	1.44
4	61.53	49.14	0	0.49	0.34	0.14	9.2	1.6
5	83.03	62.95	0	0.52	0.43	0.14	27.98	1.92
<i>i</i>	<i>p(q_c)</i>	$\frac{T}{\bar{y}}$	<i>T'(y)</i>	β	<i>U</i>			
1	1.26	0.01	0.28	—	1.93			
2	1.91	0.09	0.25	—	2.92			
3	2.67	0.13	0.33	—	4.1			
4	4.11	0.2	0.2	—	5.87			
5	8.62	0.24	0	—	10.29			

Note: Household taxable income *y* and consumption *c* expressed in thousands of 2006 US dollars.

government and free of charge for opting-in households (i.e., opting-in households can get free of charge as many hours of formal care as they want).⁴⁶

The results for the case where the government optimizes a nonlinear income tax and where there are no subsidies to child care are shown in Table 5.

In Table 6, we show the results for the case where the government employs income-dependent child care subsidies on top of an optimal nonlinear income tax. Glancing at the column labeled by β , it can immediately be seen that child care subsidies would be suboptimal. Instead, child care expenditure should be taxed ($\beta < 0$). The implied taxes on child care appear to be quite substantial (and, conditional on the mother’s employment status, decreasing in household income), even though the implied welfare gain of these child care taxes is close to zero.⁴⁷ The result that there should be no subsidies, at least when levied as income-dependent proportional

⁴⁶In these three cases, the optimal labor force participation rates for mothers in households of type 1 through 5 are 50 percent, 60 percent, 65 percent, 75 percent, and 85 percent, respectively.

⁴⁷The welfare gain is calculated by computing the minimum amount of extra revenue that needs to be injected into pure optimal nonlinear income tax solution in order to reach the social welfare of the optimal income tax solution with child care subsidies. We then divide this amount of extra revenue by the aggregate pre-tax income in the pure optimal income tax solution to get a welfare measure expressed as a fraction of aggregate output.

TABLE 6—INCOME-DEPENDENT SUBSIDY

<i>i</i>	<i>y</i>	<i>c</i>	<i>L_m</i>	<i>L_f</i>	<i>h_m</i>	$\frac{D}{\bar{y}}$	<i>q</i>	<i>q_c</i>
<i>Panel A. Allocation in households where the mother works</i>								
1	46.79	46.1	0.2	0.47	0.08	0.1	0.87	1.16
2	58.2	51.74	0.27	0.44	0.11	0.11	2.04	1.28
3	79.65	63.29	0.36	0.46	0.11	0.12	3.94	1.44
4	94.89	71.71	0.37	0.42	0.15	0.13	8.51	1.59
5	141.12	100.61	0.44	0.39	0.18	0.14	26.48	1.9
<i>i</i>	<i>p(q_c)</i>	$\frac{T}{\bar{y}}$	<i>T'(y)</i>	β	<i>U</i>			
1	1.1	0.01	0.19	−0.18	2.08			
2	1.65	0.11	0.25	−0.17	3.17			
3	2.67	0.21	0.14	−0.06	4.42			
4	4.06	0.24	0.17	−0.03	6.4			
5	8.3	0.29	0	0	11.12			
<i>i</i>	<i>y</i>	<i>c</i>	<i>L_m</i>	<i>L_f</i>	<i>h_m</i>	$\frac{D}{\bar{y}}$	<i>q</i>	<i>q_c</i>
<i>Panel B. Allocation in households where the mother does not work</i>								
1	38.28	39.96	0	0.47	0.14	0.08	0.88	1.09
2	45.18	42.08	0	0.48	0.21	0.11	2.18	1.27
3	49.14	44.16	0	0.46	0.28	0.11	4.3	1.36
4	61.65	49.54	0	0.49	0.34	0.14	9.12	1.58
5	83.07	62.8	0	0.52	0.43	0.14	27.98	1.91
<i>i</i>	<i>p(q_c)</i>	$\frac{T}{\bar{y}}$	<i>T'(y)</i>	β	<i>U</i>			
1	0.86	−0.04	0.25	−0.38	1.95			
2	1.61	0.07	0.24	−0.16	2.92			
3	2.12	0.1	0.3	−0.22	4.09			
4	3.9	0.2	0.19	−0.05	5.86			
5	8.61	0.24	0	0	10.28			

Note: Household taxable income *y* and consumption *c* expressed in thousands of 2006 US dollars.

subsidies, challenges previous findings in the optimal tax literature advocating child care subsidies as an instrument to relax the incentive constraints faced by the government in designing an optimal nonlinear income tax.

We get an opposite result due to the fact that, while previous contributions assumed the hourly price of child care as fixed, we allow child care expenditures to depend both on the number of hours spent by a kid at a child care center and on the quality of the facility chosen by parents (which affects the hourly price of child care services). Thus, while in a model with a fixed hourly price of child care services a low-skilled agent is likely to spend more on child care services than a high-skilled mimicker (since a high-skilled mimicker needs to work fewer hours than a low-skilled agent and therefore needs fewer hours of child care for the kids), this is no longer necessarily true in our setting where the quality (and therefore the hourly price) of child care services is a choice variable for households.⁴⁸

⁴⁸ Choosing the model’s parameters in order to obtain a realistic calibration, we find that a household behaving as a mimicker (i.e., choosing the income point intended for a lower skill type) would spend more on child care than the household being mimicked, despite the fact that a mimicker demands fewer hours of center-based child care.

TABLE 7—PUBLIC PROVISION

<i>i</i>	<i>y</i>	<i>c</i>	<i>L_m</i>	<i>L_f</i>	<i>h_m</i>	$\frac{D}{\bar{y}}$	<i>q</i>	<i>q_c</i>
<i>Panel A. Allocation in households where the mother works</i>								
1	47.3	44.16	0.22	0.47	0.05	0.23	1.1	1.42
2	58.15	50.85	0.28	0.43	0.08	0.17	2.16	1.42
3	80.23	65.63	0.37	0.46	0.09	0.12	3.65	1.42
4	96.91	77.55	0.38	0.43	0.13	0.09	7.23	1.42
5	144.79	95.47	0.45	0.4	0.17	0.13	25.97	1.87
<i>i</i>	<i>p(q_c)</i>	$\frac{T}{\bar{y}}$	<i>T'(y)</i>	Opt in	<i>U</i>			
1	2.54	0.07	0.23	Yes	2.24			
2	2.54	0.13	0.3	Yes	3.32			
3	2.54	0.18	0.17	Yes	4.5			
4	2.54	0.2	0.19	Yes	6.35			
5	7.87	0.21	0	No	10.75			
<i>i</i>	<i>y</i>	<i>c</i>	<i>L_m</i>	<i>L_f</i>	<i>h_m</i>	$\frac{D}{\bar{y}}$	<i>q</i>	<i>q_c</i>
<i>Panel B. Allocation in households where the mother does not work</i>								
1	38.4	36.84	0	0.47	0.09	0.27	1.15	1.42
2	45	40.99	0	0.48	0.14	0.2	2.29	1.42
3	48.13	42.86	0	0.45	0.2	0.16	4.18	1.42
4	59.35	49.74	0	0.47	0.27	0.11	8.11	1.42
5	85.38	58.38	0	0.53	0.44	0.13	27.58	1.89
<i>i</i>	<i>p(q_c)</i>	$\frac{T}{\bar{y}}$	<i>T'(y)</i>	Opt in	<i>U</i>			
1	2.54	0.04	0.29	Yes	2.11			
2	2.54	0.09	0.28	Yes	3.07			
3	2.54	0.11	0.37	Yes	4.17			
4	2.54	0.16	0.28	Yes	5.81			
5	8.11	0.18	−0.01	No	9.91			

Note: Household taxable income *y* and consumption *c* expressed in thousands of 2006 US dollars.

The discussion above suggests that in order for there to be a role for child care subsidies as an instrument to achieve redistribution at a lower efficiency cost, one needs a policy instrument that allows to control the quality of the child care services that are being subsidized. For instance, this would be the case with an opting-out public provision scheme. Under an opting-out public provision scheme, given that the quality of the publicly provided care is set by the government and is the same for all households who opt in, a subsidy yields a smaller benefit to a low-skilled household than to a mimicker only if the mimicker opts in and at the same time demands more hours of center-based care than a true low-skilled agent (which is unlikely to happen given that a high-skilled mimicker needs to work fewer hours than a low-skilled agent).

The quantitative results for this case are displayed in Table 7. In the table, it can be seen that the government selects a quality level for the publicly provided child care such that all but the highest-skilled households opt in. The quality is set at 1.42, which is a quality level close to the one chosen by median households in the absence of child care subsidies. Moreover, it is worth noticing that the bottom three household types are better off under the public provision scheme, whereas the top two households are worse off. The welfare gain of adopting a public provision scheme for child

care alongside an optimal nonlinear income tax is equal 1.43 percent of GDP. Thus, the public provision scheme appears to significantly improve upon the pure optimal nonlinear income tax solution.

V. Summary and Conclusion

In this paper, we have evaluated the desirability of child care subsidies in a model where the quantity and quality of care that children receive, both at home (in terms of informal care from parents) and outside the home (at child care facilities), affects the children's human capital that enters into the utility function maximized by parents. Compared to previous contributions in the optimal tax literature, the main distinguishing contribution of our paper has been to emphasize the key importance of the quality dimension of child care (and of the human capital formation process more generally) in determining whether or not child care subsidies are useful as an instrument to achieve redistribution at lower efficiency costs.

We have assessed the case for subsidizing child care by first theoretically analyzing the welfare effects of child care subsidies under different assumptions about the economic variables that are publicly observable at the individual level and about the policy instruments that are available to the government. As shown in the first part of the paper, these assumptions can play an important role in determining whether child care expenditures should be encouraged or discouraged by the tax system.

We have also assessed the desirability of child care subsidies by means of a calibrated model, incorporating important aspects of the US economy. We have considered a joint system of taxation and employed wage distributions calibrated to fit the empirical wage distributions of mothers and fathers with kids in child care age, and we have disciplined our parameters using data from the Current Population Survey (CPS) and the American Time Use Survey (ATUS).

In contrast to what has been obtained in previous optimal taxation studies, our theoretical results highlight that it is by no means *ex ante* obvious that child care should be subsidized. Moreover, we have provided an empirical calibration where there is no scope for using child care subsidies to reduce the inefficiencies associated with income redistribution, at least when the subsidies are based on the child care expenditures incurred by households, as is the case with tax credits (rather than on the number of child care hours, as under a public provision scheme). The main reason for our results is that parents choose both the quantity and the quality of center-based child care services, whereas previous studies in the optimal tax literature have analyzed the role of child care subsidies in models where the quality of center-based child care was exogenously given. Moreover, we have allowed both the time spent by kids at a child care facility and the time spent by parents with their offspring to affect the overall quality of the child care arrangement and therefore contribute to the human capital development of children.

In a model where the quality of child care is fixed, the variation in child care expenditure is largely driven by variation in child care hours, which is strongly correlated with hours of work. This implies that if a high-skilled household were to mimic a low-skilled one, the expenditure on child care services would be higher for the low-skilled household than for the high-skilled mimicker. In our

model, instead, a high expenditure can be the result of either a high quality of the chosen child care facility or a high number of child care hours. If a high-skilled mimicker chooses a higher quality of child care services than a low-skilled household, it is then conceivable that child care expenditures are larger for the former. This undermines the role for child care subsidies as a mimicking-deterring device.

Subsidies delivered through an opting-out public provision scheme remain, however, a useful instrument for mimicking-deterring purposes. Under an opting-out public provision scheme, the quality of the center-based care is set by the government and is no longer a choice variable for the households who decide to opt in. Granting subsidies only to households who opt in implies then that the private value of a subsidy is only a function of the number of child care hours that are used. This, in turn, implies that when a low-skilled household opts in, a subsidy to child care expenditures yields a larger benefit to a low-skilled household than to a mimicker unless the latter also opts in and at the same time demands more hours of center-based care than a true low-skilled household (which is unlikely to happen given that a high-skilled mimicker needs to work fewer hours than a low-skilled household).

We have focused on the role played by child care subsidies as an instrument for the government to achieve the desired redistributive goals at lower efficiency costs. Another potential argument for subsidizing child care is that there could be externalities associated with the choices of child care arrangements that parents make for their offspring. In Bastani et al. (2017), we considered such aspects by including an externality term in the social welfare function.⁴⁹ In such a framework, child care subsidies should be positive provided the externality term carries a sufficiently large weight in the social welfare function. For countries who engage heavily in income redistribution, it might be more important to mitigate the distortionary effects associated with income taxation. In other countries that engage less in income redistribution, the externality argument in favor of subsidized child care might carry more weight.

To conclude, we would like to mention a few potentially broader implications of our results. First, the case for subsidizing child care could be greater in economies that offer a narrower or a more homogeneous selection of child care services in the private market, thereby creating a stronger link between child care expenditures and labor supply. Second, one of the most frequently occurring results in applied tax policy discussion is the recommendation that goods complementary to labor should be subsidized or taxed at lower rates than goods that are complements with leisure to mitigate the inefficiencies associated with income taxation. Our analysis highlights that such results need to be qualified to take into account the quality dimension of the goods in question.

⁴⁹The externality term can be motivated on the basis that children's human capital formation is an important determinant of the productivity and future earnings of individuals. This implies that increased human capital investments have the potential to increase the future tax base and, for a given size of public expenditures, result in lower taxes for future generations.

REFERENCES

- Atkinson, Anthony B., and Joseph E. Stiglitz. 1976. "The Design of Tax Structure: Direct versus Indirect Taxation." *Journal of Public Economics* 6 (1–2): 55–75.
- Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2008. "Universal Child Care, Maternal Labor Supply, and Family Well-Being." *Journal of Political Economy* 116 (4): 709–45.
- Bastani, Spencer, Sören Blomquist, and Luca Micheletto. 2012. "Child Care Subsidies, Quality, and the Optimal Income Tax." In *Essays on the Economics of Income Taxation*, Economic Studies 129, Department of Economics, Uppsala University.
- Bastani, Spencer, Sören Blomquist, and Luca Micheletto. 2017. "Child Care Subsidies, Quality, and Optimal Income Taxation." CESifo Working Paper 6533.
- Bastani, Spencer, Sören Blomquist, and Luca Micheletto. 2020. "Replication Data for: Child Care Subsidies, Quality, and Optimal Income Taxation." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/E116562V1>.
- Bastani, Spencer, Sören Blomquist, and Jukka Pirttilä. 2015. "How Should Commodities Be Taxed? A Counter-Argument to the Recommendation in the Mirrlees Review." *Oxford Economic Papers* 67 (2): 455–78.
- Bick, Alexander. 2016. "The Quantitative Role of Child Care for Female Labor Force Participation and Fertility." *Journal of the European Economic Association* 14 (3): 639–68.
- Blau, David, and Janet Currie. 2006. "Pre-school, Child Care, and After-School Care: Who's Minding the Kids?" In *Handbook of the Economics of Education*, Vol. 2, edited by Eric A. Hanushek and Finis Welch, 1163–278. North Holland: Elsevier.
- Blau, David M., and H. Naci Mocan. 2002. "The Supply of Quality in Child Care Centers." *Review of Economics and Statistics* 84 (3): 483–96.
- Blau, David M., and Philip K. Robins. 1988. "Child-Care Costs and Family Labor Supply." *Review of Economics and Statistics* 70 (3): 374–81.
- Blomquist, Sören, and Vidar Christiansen. 1995. "Public Provision of Private Goods as a Redistributive Device in an Optimum Income Tax Model." *Scandinavian Journal of Economics* 97 (4): 547–67.
- Blomquist, Sören, Vidar Christiansen, and Luca Micheletto. 2010. "Public Provision of Private Goods and Nondistortionary Marginal Tax Rates." *American Economic Journal: Economic Policy* 2 (2): 1–27.
- Blundell, Richard, and Andrew Shephard. 2012. "Employment, Hours of Work, and the Optimal Taxation of Low-Income Families." *Review of Economic Studies* 79 (2): 481–510.
- Brewer, Mike, Emmanuel Saez, and Andrew Shephard. 2010. "Means-Testing and Tax Rates on Earnings." In *Dimensions of Tax Design: The Mirrlees Review*, edited by James Mirrlees et al., 90–173. Oxford, UK: Oxford University Press.
- Cogan, John F. 1981. "Fixed Costs and Labor Supply." *Econometrica* 49 (4): 945–63.
- Corlett, W.J., and D.C. Hague. 1953. "Complementarity and the Excess Burden of Taxation." *Review of Economic Studies* 21 (1): 21–30.
- Crawford, Ian, Michael Keen, and Stephen Smith. 2010. "Value Added Taxes and Excises." In *Dimensions of Tax Design: The Mirrlees Review*, edited by James Mirrlees, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba, 275–422. Oxford, UK: Oxford University Press.
- Currie, Janet. 2006. *The Invisible Safety Net: Protecting the Nation's Poor Children and Families*. Princeton, NJ: Princeton University Press.
- Domeij, David, and Paul Klein. 2013. "Should Day Care be Subsidized?" *Review of Economic Studies* 80 (2): 568–95.
- Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura. 2020. "Child-Related Transfers, Household Labour Supply, and Welfare." *The Review of Economic Studies*. <https://doi.org/10.1093/restud/rdaa011>.
- Gustafsson, Siv, and Frank Stafford. 1992. "Child Care Subsidies and Labor Supply in Sweden." *Journal of Human Resources* 27 (1): 204–30.
- Havnes, Tarjei, and Magne Mogstad. 2011. "No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes." *American Economic Journal: Economic Policy* 3 (2): 97–129.
- Havnes, Tarjei, and Magne Mogstad. 2015. "Is Universal Child Care Leveling the Playing Field?" *Journal of Public Economics* 127: 100–14.

- Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante.** 2014. "Consumption and Labor Supply with Partial Insurance: An Analytical Framework." *American Economic Review* 104 (7): 2075–126.
- Heckman, James J.** 2006. "Skill Formation and the Economics of Investing in Disadvantaged Children." *Science* 312 (5782): 1900–2.
- Heckman, James J., and Dimitriy V. Masterov.** 2007. "The Productivity Argument for Investing in Young Children." *Review of Agricultural Economics* 29 (3): 446–93.
- Herbst, Chris.** 2018. "The Rising Cost of Child Care in the United States: A Reassessment of the Evidence." *Economics of Education Review* 64 (C): 13–30.
- Ho, Christine, and Nicola Pavoni.** 2020. "Efficient Child Care Subsidies." *American Economic Review* 110 (1): 162–99.
- Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez.** 2009. "The Optimal Income Taxation of Couples." *Econometrica* 77 (2): 537–60.
- Koehne, Sebastian, and Dominik Sachs.** 2020. "Pareto-Improving Reforms of Tax Deductions." CESifo Working Paper 8408.
- Laughlin, Lynda.** 2013. *Who's Minding the Kids? Child Care Arrangements: Spring 2011*. Washington, DC: US Department of Commerce.
- Mirrlees, James, Stuart Adam, Timothy Besley, Richard Blundell, Stephen Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba, eds.** 2010. *Dimensions of Tax Design: The Mirrlees Review*. Oxford, UK: Oxford University Press.
- Organization for Economic Co-operation and Development (OECD).** 2011. *Doing Better for Families*. Paris: OECD.
- Organization for Economic Co-operation and Development (OECD).** 2006. *Starting Strong II: Early Childhood Education and Care*. Paris: OECD.
- Powell, Lisa M.** 2002. "Joint Labor Supply and Childcare Choice Decisions of Married Mothers." *Journal of Human Resources* 37 (1): 106–28.
- Ribar, David C.** 1995. "A Structural Model of Child Care and the Labor Supply of Married Women." *Journal of Labor Economics* 13 (3): 558–97.
- Rosen, Sherwin.** 1996. "Public Employment and the Welfare State in Sweden." *Journal of Economic Literature* 34 (2): 729–40.
- Tuomala, Matti.** 2010. "On Optimal Non-linear Income Taxation: Numerical Results Revisited." *International Tax and Public Finance* 17 (3): 259–70.
- Waldfogel, Jane.** 2006. *What Children Need*. Cambridge, MA: Harvard University Press.
- Whitehurst, Grover.** 2018. "What is the Market Price of Daycare and Preschool?" In *Evidence Speaks Reports*, Vol. 2, 48. Washington, DC: The Brookings Institution.