OPTIMAL TAXATION OF FAMILIES:
Mirrlees meets Becker[*]

Musab Kurnaz[†]

Tepper School of Business, Carnegie Mellon University

This Version: September 23, 2016

Abstract

I study the optimal taxation of families in an environment in which (i) characteristics of a family -labor productivity and desire for children- are only observable by the family and (ii) child-rearing requires both goods and parental time. Potential parents simultaneously decide labor income and the family size. The government uses information on family income and family size to construct an optimal tax system through the combination of an income tax schedule and child tax credits. I observe that parental time and the cost of goods involved in child-rearing have distinct impacts on the shape of optimal child tax credits. As a quantitative analysis, I calibrate my model to US data and show that child-rearing costs translate into a pattern of optimal credits that is U-shaped in income. In particular, tax credits decrease during the first three-quarters of the income distribution and increase thereafter. In addition, tax credits decrease according to family size because of the economies of scale in child-rearing. For median-income families, the credit for the second child equals 64% of the credit for the first child. I find that the optimal tax schedule generates a welfare gain equal to 1.3% of aggregate consumption.

JEL-Classification: H21, H53, D82, J13

Keywords: Optimal Income Taxation, Child Tax Credits, Multidimensional Screening, Fertility

---

[†]Musab Kurnaz: mkurnaz@andrew.cmu.edu

1 Introduction

Family income and family size are two important components of family income taxation in the United States. Over many decades, the US government has gradually increased tax benefits for parents. As a result, federal spending on Earned Income Tax Credits, which are mostly obtained by parents, increased by 140% and federal spending on Child Tax Credits doubled during the first decade of 2000s. Although a large literature in labor and family economics studies the consequences of these increases, there is almost no normative work to guide governments for the design of child tax credits. My paper fills this gap. In this paper, I ask what the optimal child tax credits should be. To address this question, I study a model in which working families are heterogeneous and parents bear child-rearing costs. I find that child-rearing costs translate into U-shaped child tax credits in income. On the one hand, the cost of goods decreases family consumption and the percentage reduction is larger for lower income parents. This strengthens the government's redistribution motives for low income parents. Consequently, tax credits are decreasing in the left tail of income distribution. On the other, the opportunity cost of time devoted to childcare creates distortions in parents' labor supply, and the distortions are greater for higher income parents. To reduce the efficiency loss created by distortions, the government decreases the marginal taxes of high-income parents. As a result, tax credits are increasing in the right tail of income distribution.

I make both theoretical and quantitative contributions. On the theoretical side, I embed a Becker model into an optimal tax framework.[1] Family economists have studied the former to analyze fertility behavior of households who bear child-rearing costs. I show how child-rearing costs affect the optimal child tax credits. My work, which studies the optimal child tax credits together with the optimal income taxation, is one of the first studies in the normative tax literature. On the quantitative side, given a structural form of utility from a discrete number of children, I estimate the curvature of family preferences over the number of children where the curvature measures the marginal utility of a child. This estimation is new to the public finance literature.

I study a family problem in which potential parents have two unobservable characteristics: labor productivity and desire (taste) for children.[2] In contrast, the main stream of public finance literature focuses on individual workers who are characterized only by their privately observed labor productivities. The two unobservable characteristics of

---

[1] Becker (1960) and Becker (1965) suggest that children should appear as a consumption good in a family utility function and should be considered as an output of goods and time.

[2] The taste for children is a measure of the child desire of households, which, in turn, is affected by non-economic factors, such as religion and race (see Becker (1960)).

potential parents create a two-dimensional screening problem. Because of the technical difficulties, the public finance literature that deals with multidimensional screening problems is sparse. My paper is one of the exceptions. I handle the difficulties by assuming that the utility from children is additively separable within the family utility function. The implication is that the marginal benefit of a child is independent of parents' labor productivities and the marginal cost of a child is independent of parents' tastes for children. Consequently, for a given labor productivity level, there exist *threshold* tastes that equalize the marginal cost and benefit of a child. When the taste of potential parents is between n^{th} and $n+1^{th}$ threshold tastes, they choose to have n children, which is observable by society. This optimality condition provides information on the underlying tastes of parents. In addition, observable family incomes provide information on labor productivities. Using observable family incomes and family sizes, the government constructs an optimal tax schedule under which families cannot be better off by either changing their sizes or by generating another income.

The taxation of families consists of a combination of an income tax schedule and child tax credits. The former redistributes towards low earning families and the latter reduces the income tax liabilities of parents who are made monetarily worse off by child-rearing costs. The reduction at the bottom and at the top is high because of the costs of goods and the opportunity cost of the time devoted to childcare, respectively. The cost of goods decreases consumption more among low-income parents than among high-income parents. This implies that parents' marginal utility of consumption is higher at the bottom of the income distribution. This enhances government's redistribution motives for lowest income parents and consequently, the government increases tax credits for them. On the other hand, the opportunity cost of time devoted to childcare increases the cost of working of parents and disincentivizes them working hard. Especially, the cost of working gets larger towards higher-productive parents. Hereby, high-productive parents search for opportunities of being better by generating low income. In the meanwhile, the government wants to have a higher output to finance its expenditure and its redistributive motives. Hence, the government incentivizes high-productive parents working hard by decreasing their marginal taxes and increasing their tax credits. To sum up, the distinct effects of the cost of goods and the opportunity cost of time create optimal child tax credits that are U-shaped with respect to income.

I calibrate my model to the US economy and analyze the implications of child-rearing costs for optimal policy. First, I find estimates for child-rearing costs. I estimate the cost of goods using information obtained from the annual report of the US Department of Agriculture, and I calculate the opportunity cost of time devoted to childcare as it

is the percentage reduction in labor of families owing to children. Second, I use the Current Population Survey (CPS) to calculate the labor productivities of families from their consumption-labor margin. CPS includes detailed information about taxes and the weights of observations. Using this information, I derive the probability distribution of productivities. Third, I assume that child tastes are drawn from an exponential distribution and I use a maximum likelihood estimation to find the marginal benefit of each child in families.

The quantitative analysis provides two policy results which are distinct from the US child tax credits. First, the optimal child tax credits for each child are U-shaped according to income. In particular, the optimal tax credits decrease during the first three-quarters of the income distribution and increase thereafter. In contrast, the US tax credits are decreasing in income. This implies that the US government focuses on the impact of cost of goods. However, time devoted to childcare significantly affects the labor supply of parents and subsequently reduces the total output of the economy.[3] Therefore, including the effect of time cost into the design of credits may improve social welfare.

Second, I show that the optimal tax credits are not same for each child in a family because the marginal cost of child-rearing on family welfare is decreasing by the number of children. This result stems from two facts. First, families have convex preferences over their welfare. Second, there is an economies of scale in child-rearing costs. A two-child family annually spends 96 market labor hours for its first child while it only spends 40 hours for its second. Also, the cost of goods of raising two children is only 58% more than the cost of goods of raising one child. As a result, the decrease in family welfare owing to the first child is relatively larger than the decrease owing to the second child. Accordingly, the government's redistribution motives are affected by these relative changes and more credits are given for the first child. In particular, I find that the credit for the first child is 56% larger than the credit for the second given to the median income families. In contrast, the US child tax credits are almost same for each child in families.

The sharp differences between the optimal and the actual tax schedule raise an important question: What is the potential welfare gain from implementing the optimum? I find that the gain is 1.3% in terms of an equivalent increase in consumption for all families. This gain suggests that a U-shaped tax credit schedule improves social welfare significantly. In addition, the optimal tax schedule makes 50% of the population better off, which implies that the optimal policy can socially be quite attractive.

The remainder of the paper is organized as follows. After a brief review of the litera-

---

[3]According to Haveman and Wolfe (1995), the opportunity cost of time for childcare was 1.8% of 1992 GDP.

ture, I provide in Section 2 an institutional background for the taxation of families. Next, in Section 3, I introduce the model, within which I derive the two-dimensional optimal tax schedule. Section 4 quantitatively analyzes the model. I check the robustness of results in Section 6 and conclude in Section 7.

Related Literature: This paper links the literature on fertility theories to public finance literature. Most of the public finance literature abstracts from the child decision and the majority of fertility theories abstract from optimal taxation. My paper fills this gap.

There is a voluminous literature on fertility theories started with Becker (1960) and Becker (1965). These two seminal works suggest that children should appear in family utility as a consumption good that is produced by inputs of goods and time. I refer to these inputs as child-rearing costs. Haveman and Wolfe (1995) use Consumer Expenditure Survey (CEX) data to measure child-rearing costs. They find that child-rearing costs reared both by the US government and the US families are 14.5% of 1992 GDP. This shows the importance of child-rearing costs in the US economy. In addition, two-thirds of the costs are reared by parents which indicates the significance of child-rearing costs on family welfare.

Haveman and Wolfe (1995) find that the cost of goods is 82% of the total parental cost, which includes expenditures on food, housing, health care, and clothes. The impact of the goods' cost on the economy has been recently studied by Golosov, Jones, and Tertilt (2007) and Hosseini, Jones, and Shourideh (2013). The former work studies the efficiency of the allocations for future generations while the latter focuses on consumption inequality over the long run.

Time devoted to childcare is one the main reasons for the well-known empirical evidence that the fertility rate is negatively correlated with family labor income (see Jones, Schoonbroodt, and Tertilt (2010)).[4] Moreover, time cost increases the sensitivity of parental labor to wage changes. Blundell, Meghir, and Neves (1993) find that parents have a higher Frisch elasticity of labor supply than non-parents. In this paper, I endogenize the elasticity of parents and show that parents who have more children have a higher elasticity. This can be attributed to the fact that more children require more time and reduce the available time for labor. The Frisch elasticity, in particular, is important because it is one of the major components of an optimal tax system (see Saez (2001)).

In contrast to most contributions to the fertility literature, which considers that child-rearing requires either goods or parental time, I take both costs into account and show that their impacts on family welfare are very important for optimal tax policies.

---

[4]The opportunity cost of time devoted to childcare is higher for higher wage workers.

My paper also contributes to the public finance literature, which is founded on the trade-off between efficiency and equity (see Mirrlees (1971)). The trade-off arises because there is a friction in information. The main stream of this literature assumes that workers' productivities are observable only to the workers themselves. In my paper, working families privately observe their tastes for children as well as their own labor productivities. This implies that the friction in information is two dimensional. The literature on the multidimensional screening problem is sparse because of the technical challenges.[5] Kleven, Kreiner, and Saez (2009) and Jacquet, Lehmann, and der Linden (2013) are notable exceptions. The former work focuses on the jointness of couples' income taxation in which families privately observe primary earners' labor productivities and secondary earners' work costs. They show that marginal income tax rates of the primary earner should be smaller if his or her spouse works. The latter studies an environment in which workers privately know their labor productivities and their taste for work and make an extensive labor decision. The study provides a rationale for non-negative marginal rates. In order to study child tax credits, and in contrast to the studies just described, I focus both on the marginal rates and on the tax liabilities of families. Moreover, both of the studies just mentioned have only two categories of households. In this paper, I derive optimal taxes for an arbitrary number of family sizes.

2 Children in the US Income Tax Code

To motivate my analysis I document changes in federal spending in the most important US government policies. In over 100 programs, 28% of the federal budget for welfare programs is spent for children and 33% of this expense is related to tax credits and exemptions such as the Child Tax Credit (CTC), the Child and Dependent Care Tax Credit (CDCTC), Dependent Exemptions, and the Earned Income Tax Credit (EITC).[6] In Appendix A.1 I present detailed information about these programs. In this section I focus on drastic changes made in the federal budget for the CTC and the EITC: an increase of 210% and 140%, respectively (see Figure 1). The increase in spending for the CTC occurred because per child credit was gradually increased from $400 to $1,000 and the eligibility conditions to claim CTC were relaxed. The increase in EITC spending was mostly a reflection of expansions of the program during 2000s.

The change in the federal budgets for tax credits (especially changes in the EITC pro-

---

[5] In the industrial organization literature, Baron and Myerson (1982) and Rochet and Chone (1998) identify additional requirements needed to solve multidimensional screening problems.

[6] http://www.urban.org/sites/default/files/alfresco/publication-pdfs/412599-Data-Appendix-to-Kids-Share--.PDF
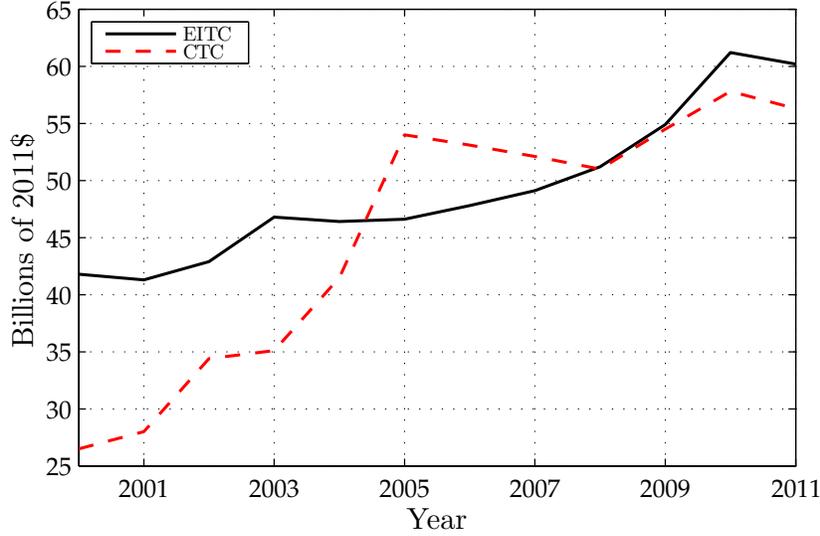
Figure 1: Government Spending for EITC and CTC

Source: A work by the Tax Policy Center, which is a joint venture of the Urban Institute and Brookings Institution, uses expenditures on EITC and CTC taken from Internal Revenue Service Statistics of Income.

gram) is in the focus of positive literature (see Hotz and Scholz (2003) and their references). However, the normative analysis is very sparse. My paper is one of the first to study the design of tax credits. In the following section, I introduce a static model wherein heterogeneous potential parents make decisions about labor choice and the number of children based on their observations of their own labor productivities and their desire for children. Raising children requires both goods and parental time. The effects of these costs on family welfare are the key determinants of the tax credit design.

3   Model

A continuum of potential parents (hereafter families), whose size is normalized to 1, has identical preferences over consumption $c \in \mathbb{R}_+$, labor income $z \in [0, \bar{z}]$, and number of children $n \in \mathcal{N} = \{0, 1, \ldots, N\}$ described by a utility function $\mathcal{U} : \mathbb{R}_+ \times [0, \bar{z}] \times \mathcal{N} \to \mathbb{R}$. $\mathcal{U}(\cdot, \cdot, n)$ is assumed to be concave, twice differentiable on the interior of its domain, with for each $z \in [0, \bar{z}]$ and for all $n \in \mathcal{N}$, $\mathcal{U}(\cdot, z, n)$ increasing and for each $c \in \mathbb{R}_+$ and for all $n \in \mathcal{N}$, $\mathcal{U}(c, \cdot, n)$ decreasing and strictly concave. First and second partial derivatives of $\mathcal{U}$ are denoted $\mathcal{U}_x(\cdot, \cdot, n)$ and $\mathcal{U}_{xy}(\cdot, \cdot, n)$ with $x, y \in \{c, z\}$. $\mathcal{U}$ satisfies the Inada conditions: for all $c \in \mathbb{R}_{++}$ and for all $n \in \mathcal{N}$, $\lim_{z \downarrow 0} \mathcal{U}_z(c, \cdot, n) = 0$ and $\lim_{z \uparrow \bar{z}} \mathcal{U}_z(c, \cdot, n) = -\infty$. In addition, $\mathcal{U}$ satisfies the Spence-Mirrlees single crossing property: $\frac{-\mathcal{U}_z(c, z/\theta, \cdot)}{\theta \mathcal{U}_c(c, z/\theta, \cdot)}$ is decreasing in θ.

7

A family is characterized by a labor productivity $\theta \in \Theta = (\underline{\theta}, \overline{\theta})$ and a taste for children: $\beta \in \mathcal{B} = (\underline{\beta}, \overline{\beta})$.[7] The family characteristics $\gamma = (\beta, \theta)$ are distributed according to a continuous density distribution over $\Gamma = \mathcal{B} \times \Theta$. Let $\Pi(\cdot, \cdot)$ be the joint distribution function over (β, γ) and let $P(\beta|\theta)$ be the cumulative distribution of β conditional on θ: $\Pi(\beta, \theta) = \int_{\underline{\theta}}^{\theta} P(\beta|\theta') f(\theta') d\theta'$ where $f(\cdot)$ is the marginal distribution over Θ.

The government cannot observe γ and constructs a nonlinear tax system according to observed family labor income and the number of children. Let $T(z, n)$ be the tax liabilities of n-child families who generate z income. I define $k(z, j) := T(z, j-1) - T(z, j)$ as the tax credit for the jth child. Note that tax credits depend both on family income and family size. For notational sake, I represent $T(z, n)$ by $T_n(z)$ and $k(z, j)$ by $k_j(z)$.

Following Becker (1965), rearing n children requires an input of e_n goods and b_n units of of parental time.[8] This assumption implies that child-rearing costs are unconditional on family income. Another implication is that, according to the government's perspective, child-rearing requires same amount inputs regardless of their parents' welfare.

A family with characteristics $\gamma = (\beta, \theta)$ solves the problem:

$$\max_{c,z,n} \; \mathcal{U}(c, z, n) \qquad \text{subject to} \qquad c \leq z - T_n(z) - e_n \qquad \text{(FP)}$$

where

$$\mathcal{U}(c, z, n) = \Psi \left(u(c) - \theta h \left(\frac{z}{\theta} + b_n \right) + m(n, \beta) \right). \qquad (1)$$

Ψ and u are weakly concave functions and h is an increasing and convex function of class $\mathcal{C}^2$ and normalized so that $h(0) = 0$ and $h'(1) = 1$ (see Kleven, Kreiner, and Saez (2009)) and $m(n, \beta)$ is concave in n. Note that the utility function, $\mathcal{U}$, is in line with Becker (1960) who suggests children should appear as a consumption good in family utilities.

The framework I use captures the effects of child-rearing costs on family welfare in simplest way. It omits intra-family bargaining and the marriage decision. Consequently, there is no conflict between partners in allocating their labor supply in my model.[9] Bar-

---

[7] I follow the main stream in public finance literature and do not specifically model the production side of the economy. I consider the labor productivities are exogenously observed by families. There are few exceptions which specifically study endogenous productivities and its consequences on taxes. See Ales, Kurnaz, and Sleet (2015) and references there.

[8] Alternatively, e_n and b_n can be interpreted as the minimum spending on n children, with additional spending incorporated into family consumption and leisure.

[9] There exists a literature on collective labor supply which studies the determinants and consequences of intra-household bargaining on family labor supply (see Chiappori (1988) and Chiappori (1992)). In addition, intra-household allocation might play important role on family structure (see Chiappori, Fortin, and Lacroix (2002)).

gaining on allocating labor supply would be useful to analyze the design of child tax credits conditional on spouses' working status. I leave this important extension for later work.

Next subsection introduces family problem solution. Since the child choice is discrete, the solution includes two steps.

3.1 Family Problem Solution

I use backward induction to solve (FP): Given n, the optimal consumption, c_n, and optimal income, z_n, satisfy the first-order condition and the budget set of families:

$$u'(c_n)\left(1 - \frac{\partial T_n(z_n)}{\partial z}\right) = \frac{\partial h\left(\frac{z_n}{\theta} + b_n\right)}{\partial z} \tag{2}$$

$$c_n = z_n - T_n(z_n) - e_n. \tag{3}$$

These equations imply c_n and z_n are independent of β. I define the indirect utility of n-child families as

$$V_n(\theta) := u(c_n) - \theta h\left(\frac{z_n}{\theta} + b_n\right). \tag{4}$$

This definition implies that the marginal cost of nth child, $V_{n-1}(\theta) - V_n(\theta)$, is independent of families' child taste. On the other hand, the marginal benefit of nth child depends on the child taste: $\mathcal{M}(n, \beta) := m(n, \beta) - m(n-1, \beta)$. A (β, θ) family has n children if and only if the marginal benefit of nth child is larger than its marginal cost while the marginal benefit of $n + 1^{th}$ child is less than its marginal cost. Formally, a (β, θ) family choose to have n children if and only $\beta \in (\beta_n(\theta), \beta_{n+1}(\theta))$, where

$$\beta_n(\theta) := \mathcal{M}^{-1}(V_{n-1}(\theta) - V_n(\theta)), \tag{5}$$

for $n = 1, 2, \ldots, N$. Note that $\beta_n(\theta)$ equalizes the marginal benefit and the cost of nth child. I define $\beta_n(\theta)$ as the threshold taste of nth child for families whose labor productivities are θ. Note that threshold tastes depend on productivities. Given a productivity level, $\{\beta_n(\theta)\}_{n \in \mathcal{N}}$ provides information on the underlying child tastes of families.[10]

Using the family problem solution, I state the definition of a tax equilibrium:

Definition 1. Let $G \in \mathbb{R}_+$ be a fixed public spending amount. Given G, a *tax equilib-*

---

[10]Families' income and size are observable to the government. Consequently, can infer the tastes for children of n-child families whose productivies are θ.

rium is a tax system $T : \mathbb{R}_+ \times \mathcal{N} \to \mathbb{R}$, and allocation $\{c(\gamma), z(\gamma), n(\gamma)\}_{\gamma \in \Gamma}$ such that $(c(\gamma), z(\gamma), n(\gamma))$ solves (FP) and $G \leq \int_\Gamma [z(\gamma) - c(\gamma) - e_{n(\gamma)}] d\Pi(\gamma)$. Let $\mathcal{T}$ be the set of tax equilibria (given G), which I take to be nonempty.

3.2   The Government's Problem

A government attaches a Pareto weight $\xi(\gamma)$ to families whose characteristics are γ and weights are normalized to satisfy $\int_\Gamma \xi(\gamma) d\gamma = 1$. The government selects a tax equilibrium to solve:

$$\max_{\mathcal{T}} \int_\Gamma \xi(\gamma) \, \mathcal{U}((c(\gamma), z(\gamma), n(\gamma)) \, d\Pi(\gamma). \tag{GP}$$

Let T^* and $\{c^*(\gamma), z^*(\gamma), n^*(\gamma)\}_{\gamma \in \Gamma}$ denote an optimal tax equilibrium. Optimal marginal tax rate of families with n children is defined as:

$$\frac{\partial T_n^*(z_n)}{\partial z} = 1 + \frac{\mathcal{U}_z(c^*(\gamma), z^*(\gamma), n^*(\gamma))}{\mathcal{U}_c(c^*(\gamma), z^*(\gamma), n^*(\gamma))} \qquad \forall n \in \mathcal{N}.$$

I follow the conventional procedure of characterizing optimal tax equilibria by recovering their associated allocations from a mechanism design problem. Subsequently, prices and optimal taxes are determined to ensure implementation of this allocation as a part of tax equilibrium. The mechanism design problem associated with (GP) can be formulated as:

$$\max_{\{c(\gamma), z(\gamma), n(\gamma)\} \in \mathbb{R}_+ \times [0, \bar{z}], \times \mathcal{N}} \int_\Gamma \xi(\gamma) \, \mathcal{U}((c(\gamma), z(\gamma), n(\gamma)) \, d\Pi(\gamma) \tag{MDP}$$

subject to the incentive constraints

$$\max_{\gamma' \in \Gamma} \Psi \left(u(c(\gamma')) - \theta h \left(\frac{z(\gamma')}{\theta} + b_n \right) + m(n(\gamma'), \beta) \right) \leq \mathcal{U}((c(\gamma), z(\gamma), n(\gamma)) \quad \forall \gamma \in \Gamma \tag{6}$$

and the resource constraint

$$G \leq \int_\Gamma [z(\gamma) - c(\gamma) - e_{n(\gamma)}] d\Pi(\gamma).$$

In (MDP), the government chooses a report-contingent allocation of consumption, income, and number of children $\{c(\gamma), z(\gamma), n(\gamma)\}$ for all $\gamma \in \Gamma$ that induces every family to truthfully report its characteristics γ and generate $z(\gamma)$ income and have $n(\gamma)$ chil-

dren. Incentive constraints (6) ensure the optimality of truthful reporting. If a family with γ characteristics pretends to have γ' characteristics, the family must generate $z(\gamma')$ income and have $n(\gamma')$ number of children.

Families have two unobservable characteristics, therefore, families can misrepresent any of the two. However, separability assumption in the family utility (1) disentangles the effect of each characteristics on child decision. The marginal cost of a child depends only on θ and the marginal benefit of a child depends only on β. This allows to prevent misrepresentation of characteristics sequentially. Given families do not find misrepresenting their β beneficial, I use a first-order approach, which is the conventional procedure to prevent misrepresentation of θ:

$$\dot{V}_n(\theta) := \frac{\partial V(\theta)}{\partial \theta} = -h\left(\frac{z_n}{\theta} + b_n\right) + h'\left(\frac{z_n}{\theta} + b_n\right)\frac{z_n}{\theta} \geq 0 \quad \forall n \in \mathcal{N}. \tag{7}$$

Second, as the inequality in (7) holds for all $n \in \mathcal{N}$, I use the set of threshold tastes to prevent misrepresentation of β satisfying (5).

I adjust (MDP) with (4), (5), and (7). The new problem is a sophisticated version of the original mechanism design problem, and without loss of generality, I call the new problem as the "pseudo-mechanism design problem".

3.2.1 Pseudo-Mechanism Design Problem

The government solves the following problem:

$$\max_{\{\{c_n(\theta),z_n(\theta)\}\in\mathbb{R}_+\times[0,\bar{z}]\}_{n\in\mathcal{N}}} \int_\Theta \sum_{n\in\mathcal{N}} \int_{\beta_n(\theta)}^{\beta_{n+1}(\theta)} \xi_n(\theta)\Psi(V_n(\theta) + m(n,\beta))p(\beta|\theta)f(\theta)d\beta d\theta \quad \text{(PMDP)}$$

subject to (4), (5), and (7) and the resource constraint:

$$G \leq \int_\Theta \sum_{n\in\mathcal{N}} \int_{\beta_n(\theta)}^{\beta_{n+1}(\theta)} \left(z_n(\theta) - c_n(\theta) - e_n\right)p(\beta|\theta)f(\theta)d\beta d\theta \tag{8}$$

where $\beta_0(\theta) = \underline{\beta}$ and $\beta_{N+1}(\theta) = \bar{\beta}$. The adjustment in the government's problem does not alter the solution:

Lemma 1. *The solution of (MDP) equals to the solution of (PMDP).*

Proof. See Appendix A.2. □

The solution of (PMDP) gives the optimal taxation of families which is characterized by the marginal income tax rates for each family size:

11

Proposition 1. *The solution of ([PMDP](#)) satisfies the following differential equation for all $n \in \mathcal{N}$:*

$$\frac{T'_n(\theta)}{1 - T'_n(\theta)} = \frac{1}{\varepsilon_n(\theta)} \times \frac{1}{\theta H_n(\theta)} \times \mathcal{R}_n(\theta) \tag{9}$$

where

$$\varepsilon_n(\theta) = \frac{\partial \log z_n(\theta)}{\partial \log(1 - T'_n(z_n(\theta)))} = \frac{h'(z_n(\theta)/\theta + b_n)}{(z_n(\theta)/\theta)h''(z_n(\theta)/\theta + b_n)},$$

$$H_n(\theta) = f(\theta)(P(\beta_{n+1}(\theta)|\theta) - P(\beta_n(\theta)|\theta)),$$

$$\mathcal{R}_n(\theta) = \int_{\theta}^{\bar{\theta}} \left[\frac{(1 - g_n(\theta'))}{u'(c_n(\theta'))} (P(\beta_{n+1}(\theta')|\theta') - P(\beta_n(\theta')|\theta'))) \right.$$

$$\left. + \sum_{j=n}^{n+1} \Delta T_{j-1}(\theta') p(\beta_j(\theta')|\theta') \frac{\partial \beta_j(\theta')}{\partial V_n(\theta')} \right] u'(c_n(\theta)) f(\theta') d\theta'$$

where z_n, T_n is continuous in θ, and $\varepsilon_n(\theta)$ is the elasticity of income of n-child families with respect to marginal taxes, and $g_n(\theta)$ is the weight assigned by the government to θ-productive families with n children, and $\Delta T_n(\theta) := T_n(\theta) - T_{n+1}(\theta)$ where $T_n(\theta) = z_n(\theta) - c_n(\theta)$.[11]

Proof. See Appendix. ☐

To understand the economic intuition behind equations of Proposition [1](#), I provide an heuristic derivation of ([9](#)) for one-child families based on the effects of a small tax reform around the optimal tax system. For simplicity, let $u(c) = c$ and suppose that the government increases the taxes of one-child families with $\theta' \geq \theta$ productivities by dT (see Figure [2a](#)).

This change creates three effects. First, there will be a welfare loss for the society because consumption of one-child families is decreased by dT. The welfare cost of a dollar reduction in consumption of one-child families with θ labor productivity is

$$g_1(\theta) := \mathbb{E}_\beta \left[\frac{\xi_1(\theta) \Psi'(V_1(\theta) + m(1, \beta)))}{\lambda} | \beta_1(\theta) < \beta < \beta_2(\theta) \right]$$

in terms of the public goods. Note that the government collects dT amount of taxes from all $\theta' \geq \theta$ productive families. Therefore the net effect for a θ' family with one-child is

---

[11]I let $\Delta T_{-1}(\theta) = 0$ when $n = 0$ and $\Delta T_N(\theta) = 0$ when $n = N$.

(a) Increase in $T_1(\theta)$ by dT (b) Changes in Threshold Child Taste Values
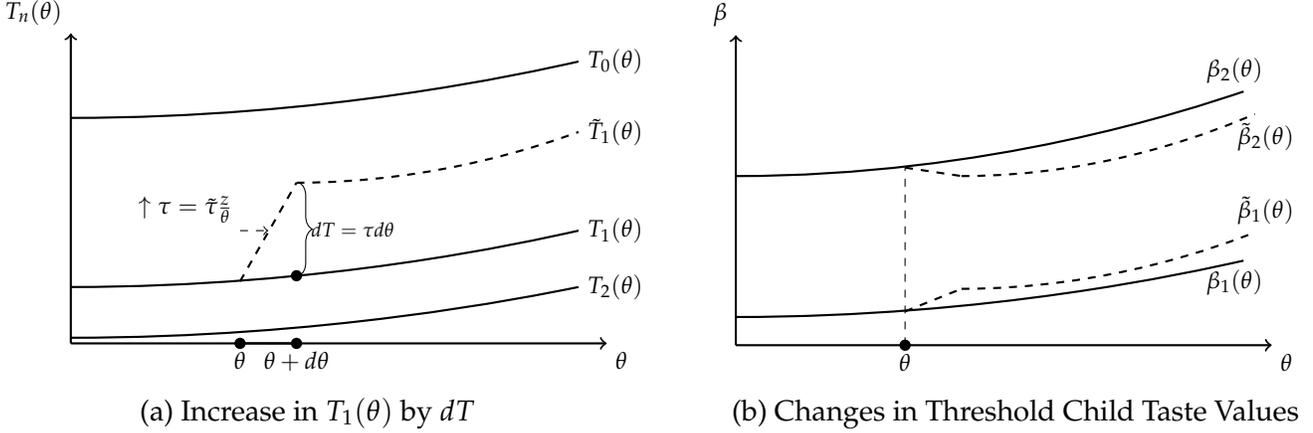
Figure 2: Heuristic Proof of Proposition 1

$dT(1 - g_1(\theta'))$. Integration of net effects gives the aggregate effect:

$$dM = dT \int_{\theta}^{\bar{\theta}} (1 - g_1(\theta')) \underbrace{[P(\beta_1(\theta')|\theta') - P(\beta_2(\theta')|\theta')]f(\theta')}_{\text{density of 1 child-families}} d\theta'.$$

dM is a mechanical effect and does not include any behavioral responses. Next, I focus on the behavioral responses of families to dT.

Second, families whose productivities are in $[\theta, \theta + d\theta]$ generate less income. To increase taxes by dT, the government should increase the marginal taxes of families whose productivities are in $[\theta, \theta + d\theta]$ by $\tau = \tilde{\tau} \frac{z_1}{\theta}$, where $\tilde{\tau}$ represents the change in the marginal tax rates on income (see Figure 2a).[12] Consequently, these families reduce their labor supply and their incomes decrease by $dz = \frac{z_1 \varepsilon_1(\theta) \tilde{\tau}}{1 - T_1'(z_1(\theta))}$, where $\varepsilon_1(\theta) := \frac{\partial \log z_1(\theta)}{\partial \log(1 - T_1'(z_1(\theta)))}$ is the elasticity of income of one-child families with respect to marginal tax rates. Combining the terms gives the first behavioral effect is:

$$dB_1 = -T_1'(\theta)dzf(\theta)d\theta = -dT \frac{T_1'(\theta)}{1 - T_1'(\theta)} \varepsilon_1(\theta)\theta \underbrace{[P(\beta_1(\theta)|\theta) - P(\beta_2(\theta)|\theta)]f(\theta)}_{\text{density of 1 child-families}}.$$

Third, families whose tastes for children are in the neighborhood of $\beta_1(\theta)$ and $\beta_2(\theta)$ alter their sizes (see Figure 2b). The one-child families whose tastes for children are in the neighborhood of $\beta_1(\theta)$ prefer to have no children after the increase in their taxes. As a result, their tax liabilities are changed by: $\Delta T_0(\theta') := T_0(\theta') - T_1(\theta')$ for all $\theta' \geq \theta$. For a particular θ', the effective change is: $\Delta T_0(\theta') \frac{\partial \beta_1(\theta')}{\partial V_1(\theta')} p(\beta_1(\theta')|\theta')f(\theta')$ where $\frac{\partial \beta_1(\theta')}{\partial V_1(\theta')}$ is the

---

[12]To change the marginal rates over productivities by τ, the marginal rates on income should increase by $\tilde{\tau}$.

13

mechanical effect of $V_1(\theta')$ on $\beta_1(\theta')$ and $p(\beta_1(\theta')|\theta')f(\theta')$ is the density of these families. Similarly, one-child families in the neighborhood of $\beta_2(\theta)$ prefers to have two children. For this case, the effective change is: $\Delta T_1(\theta')\frac{\partial\beta_2(\theta')}{\partial V_1(\theta')}p(\beta_2(\theta')|\theta')f(\theta').$[13] The aggregate effect is represented by:

$$ dB_2 = \int_\theta^{\bar{\theta}} \left(\Delta T_0(\theta')p(\beta_1(\theta')|\theta')\frac{\partial\beta_1(\theta')}{\partial V_1(\theta')} + \Delta T_1(\theta')p(\beta_2(\theta')|\theta')\frac{\partial\beta_2(\theta')}{\partial V_1(\theta')} \right) f(\theta')d\theta'. $$

To sum up, the total effect of any small changes in the tax system should be zero, if the tax system is optimal: $dM + dB_1 + dB_2 = 0$. (9) directly emerges by the previous equality for $u(c) = c$ and $n = 1$. Note that, this method can be processed for any $n \in \mathcal{N}$ to get the full schedule.

Next, I provide an interpretation of the terms of the tax code in Proposition 1.

Interpretation of the terms: I provide interpretation on the effect of each term on the marginal taxes (9). First, the elasticity, $\varepsilon_n(\theta)$, is reciprocally correlated with the marginal taxes. The marginal taxes create distortions on labor supply and hence family income. The distortions are greater for families with higher elasticity of income and create a deadweight loss for the economy. Hence, the government reduces the marginal taxes of those with higher income elasticity.

Second, the density of family sizes, $H_n(\theta)$, decreases the marginals. Intuitively, if the density is large, the impact of the distortions created by the marginal taxes will be large and reduce efficiency. Therefore, the optimal marginal rates are negatively correlated with family sizes.

Third, $\mathcal{G}_n(\theta)$ measures the redistribution tastes of the government. If the government has a high redistribution motives for a particular family, the government wants that family to have more consumption and more leisure. This is satisfied by increasing their marginal tax rates and reducing their tax liabilities.

Next, I state how the tax code in Proposition 1 differ from tax codes in the literature.

Novelty of the tax code: One the most important novelties is that the tax code depends on family income and size together. As a result, I can shed light on child tax credits. In addition, there are three more novelties owing to the number of children dimension.

First, the elasticity of income with respect to taxes, ε_n, is *endogenous*. The endogeneity arises because time is perfectly substitutable between childcare and market labor. Consequently, time devoted to childcare reduces market labor and makes labor supply (income) more sensitive to tax changes. In the following lemma, I prove this for a particular case:

---

[13]The new threshold tastes are represented by $\tilde{\beta}_1(\theta)$ and $\tilde{\beta}_2(\theta)$ in Figure 2b.

Lemma 2. *Let $u(c) = c$ and $h(x) = \frac{x^{1+\frac{1}{\varepsilon}}}{1+\frac{1}{\varepsilon}}$. Then: $\varepsilon_n(\theta) = \varepsilon(1 + \frac{b_n}{z_n/\theta})$.*

Proof. Note that $\varepsilon_n := \frac{\log \partial z_n}{\log \partial(1-T'_n)} = \frac{1-T'_n}{z_n} \frac{\partial z_n}{\partial(1-T'_n)}$. Equation (2) implies: $(1-T'_n) = h'\left(\frac{z_n}{\theta} + b_n\right)$. I take derivative with respect to $(1-T'_n)$ and get: $\varepsilon_n = \frac{h'\left(\frac{z_n}{\theta} + b_n\right)}{h''\left(\frac{z_n}{\theta} + b_n\right)\frac{z_n}{\theta n}} = \varepsilon\left(1 + \frac{b_n}{z_n/\theta}\right)$. $\square$

It is straightforward to see that $\varepsilon_n(\theta)$ depends on z_n, and hence the elasticity of parental income is endogenous. Moreover, parents' labor supply is more elastic than the non-parents', $\varepsilon_n(\theta) > \varepsilon_0(\theta) = \varepsilon$, which is in line with Blundell, Meghir, and Neves (1993).

Second, $H_n(\theta)$, the endogenous density of family size, appears in the code. The term provides information about the underlying tastes for children. The government knows that families whose tastes for children are in $(\beta_n(\theta), \beta_{n+1}(\theta))$ generate same income and have same number of children if they face same marginal tax rates.

Third, $\Delta T_n(\theta)$, the tax difference owing to number of children, shows up in the code. This is mainly because $\Delta T_n(\theta)$ affects the marginal utility of family consumption which shapes the government's redistribution motives.

3.3 Understanding the Shape of Credits

In this subsection, I provide an example to show the main forces behind the shape of tax credits. For simplicity, I let $u(c) = c$ and $\mathcal{N} = \{0, 1\}$. Moreover, $\mathcal{B} := \{\underline{\beta}, \overline{\beta}\}$ with $\underline{\beta} = 0$ and $\overline{\beta} = \infty$. This implies some families have 0 children and some have 1 child. Also, let $\Theta := \{\theta_L, \theta_M, \theta_H\}$ with $\theta_L < \theta_M < \theta_H$. Three productivity levels ensure to compare tax credits on income base. Families who generate z income pay $T(z)$ of income taxes and those who have a child get a tax credit of $k(z)$. Next, I consider environments in which parents rear only one type of child-rearing costs.

3.3.1 Only Goods Cost

Suppose that $b_1 = 0$. Without government intervention, families generate $z_n(\theta_j) = \theta_j$ for all $n \in \mathcal{N}$ and $j = L, M, H$. The percentage change in potential consumption of parents owing to the cost of goods is much higher for low productivities (incomes). Therefore, the change in the marginal utility of consumption is much higher for lower income families. Consequently, the government redistribution motives are stronger at the bottom and hence the lowest income parents receive the highest tax credits.

3.3.2 Only Time Cost

Consider $e_1 = 0$. I focus on the child tax credits under an optimal tax schedule. To understand the effect of time cost of childcare on optimal child tax credits, I study the incentive constraints of zero-child families whose labor productivities are θ_M, (IC-0M), and the incentive constraints of one-child families whose labor productivities are θ_H, (IC-1H). [14]

$$z_0(\theta_M) - h\left(\frac{z_0(\theta_M)}{\theta_M}\right)\theta_M - T(z_0(\theta_M)) \geq z_0(\theta_L) - h\left(\frac{z_0(\theta_L)}{\theta_M}\right)\theta_M - T(z_0(\theta_L)) \quad \text{(IC-0M)}$$

$$z_1(\theta_H) - h\left(\frac{z_1(\theta_H)}{\theta_H} + b_1\right)\theta_H - T(z_1(\theta_H)) + k(z_1(\theta_H)) \geq$$

$$z_1(\theta_M) - h\left(\frac{z_1(\theta_M)}{\theta_H} + b_1\right)\theta_H - T(z_1(\theta_M)) + k(z_1(\theta_M)). \quad \text{(IC-1H)}$$

To compare child tax credits on income base, I assume $z_0(\theta_M) = z_1(\theta_H)$ and $z_0(\theta_L) = z_1(\theta_M)$.[15] Next, I define

$$K(z) := \left[h\left(\frac{z}{\theta_M}\right)\theta_M - h\left(\frac{z}{\theta_H} + b\right)\theta_H \right]. \tag{10}$$

The implication of (10), (IC-0M), and (IC-1H) is:

$$k(z_1(\theta_H)) - k(z_1(\theta_M)) = K(z_1(\theta_M)) - K(z_1(\theta_H)).$$

Note that $K'(z) < 0$ if and only if $\frac{z}{\theta_M} < \frac{z}{\theta_H} + b$ which is satisfied for all $z < z_1(\theta_H) = z_0(\theta_M)$.[16] Therefore

$$k(z_1(\theta_H)) - k(z_1(\theta_M)) > 0. \tag{11}$$

The intuition behind (11) is the following. To generate a particular income level, parents have to have higher labor productivities than non-parents owing to the opportunity cost of time devoted to childcare. Consequently, the incentive constraints of one-child families are tighter comparing to those of zero child families, because producing low income is much more easier (tempting) for parents owing to the convexity of h. The tight-

---

[14](IC-0M) and (IC-1H) are the downward (binding) incentive constraints which is assumed to be binding by following the main stream of literature.

[15]Consider a set of exogenous variables that satisfies this assumption. If one of the equalities is not satisfied, I cannot compare tax credits on income base.

[16]$z_1(\theta_H) = \theta_H(1-b) = z_0(\theta_M) = \theta_M(1 - T'(z_0(\theta_M)))$, where equalities come from first order conditions of families' problem for $z_1(\theta_H)$ and $z_0(\theta_M)$, respectively.

ness gets stronger towards highest income families. Hence, the government reduces the marginal taxes of high-income parents and increases their tax credits to relax their incentive constraints.

The cost of goods and the opportunity cost of time devoted to childcare distinctly affects optimal tax credits. To see the overall effect of child-rearing costs on an optimal policy, I bring my model to US data.

4   Quantitative Analysis

In this section, I estimate the child-rearing costs, the marginal utility of each child in families, and the productivity distribution to examine the government's problem (PMDP) quantitatively. First, I focus on the family utility function (1). Second, I state which data set is the most suitable for my model and I create criteria for the sample I use. Third, I find estimates for child-rearing costs. Fourth, I derive the productivity distribution and estimate the parameter that determines marginal benefit of a child given a structural child taste distribution.

4.1   Family Preferences

This subsection introduces family preferences in the benchmark. Section 6 provides robustness analysis on the preferences.

I specify functions stated in (1). First, I assume $u(c) = c$ based on the fact that most of empirical studies find small income effect comparing to the substitution effect (see Blundell and Macurdy (1999)). In addition, it is natural to eliminate the non-labor income effect on labor to understand the relationship between labor income and fertility.

Second, I set $h\left(\frac{z}{\theta} + b_n\right) = \frac{\left(\frac{z}{\theta} + b_n\right)^{1+\frac{1}{\varepsilon}}}{1+\frac{1}{\varepsilon}}$ which implies childless families ($b_n = 0$) have a constant elasticity of income with respect to marginal rates: $\varepsilon := \frac{\partial \log z}{\partial \log(1-T')}$. I set $\varepsilon = 0.56$ by using individual level of elasticities in the literature (see Appendix A.3). Note that this number is quite close to the elasticity estimates in Chetty (2012).[17]

Third, I assume that $m(n, \beta)$ is a weakly super-modular function, $m(n, \beta) := -(N - n)^\rho(\beta - \underline{\beta})$. I estimate ρ in Section 4.4.2.

Finally, I choose a constant relative risk averse function over the family welfare: $\Psi(\mathcal{U}) = \frac{\mathcal{U}^{1-\sigma}}{1-\sigma}$. Chetty (2006) suggests that $\sigma := -\frac{\mathcal{U}_{cc}}{\mathcal{U}_c}c$ is an upper bound of the curvature of utility

---

[17]Chetty (2012) construsts a common confidence interval for the elasticities of many different studies.

over the family wealth and should be less than two to have positive labor supply response to positive wage changes. Jacquet, Lehmann, and der Linden (2013) set $\sigma = 0.8$ for their numerical analysis using the same data source I use. Hence, I follow them and set $\sigma = 0.8$.[18]

4.2   Sample Selection

Data source is the March release of the CPS administered by the US Census Bureau and the US Bureau of Labor Statistics.[19] I use 2005-2014 years' sample because families report both their federal taxes, child tax credits, and their marginal tax rates in this period. In addition, I restrict this sample based on four main criteria: number of spouses, employment status, age of spouses, and family income.

First, the sample has families with same number of spouses. This restriction eliminates potential time difference between one-spouse and two-spouse families. In the benchmark, the sample consists of two-spouse families. I study with the sample of one-spouse families in Section 6.

Second, spouses in families are employed which rules out the extensive margin decision and helps to capture a fine estimate for ε, b_n, and e_n.

Third, the age of each spouse in families are between 35-45. This restriction is used by many works in family economics literature that study the relationship between fertility and family income (see Docquier (2004), Jones and Tertilt (2008), and Jones, Schoonbroodt, and Tertilt (2010)). The aim of this restriction is to eliminate the age effect on income stream of spouses as well as parenthood decision. According to Bureau of Labor Statistics, (median) earnings of households increase in the early ages (16-35) and become stabilized after the age of 35.[20] Moreover, early age households may postpone parenthood decision owing to socioeconomic factors and the age restriction rules out this postponement. Another aim is to minimize the possibility that some children have grown up and left the family. The restriction implies parents of the sample bear child-rearing costs.

Fourth, families' total labor income is between 80% and 120% of their total income. The former restriction guarantees that labor is the main source of family income. The latter sets a boundary on a negative business or other income. In addition, I make same restriction on individual labor income of each spouse. In total, income restriction min-

---

[18]Kleven, Kreiner, and Saez (2009) focus on the UK data and set $\sigma = 1$ for their quantitative exercise.

[19]Data is taken from: Sarah Flood, Miriam King, Steven Ruggles, and J. Robert Warren. Integrated Public Use Microdata Series, Current Population Survey: Version 4.0. [Machine-readable database]. Minneapolis: University of Minnesota, 2015.

[20]See http://www.bls.gov/news.release/wkyeng.t03.htm

imizes the non-labor income effect on fertility decision.[21] Finally, families who can get other mean-tested transfers and who do not benefit from tax credit programs are extracted from sample. Therefore, families' labor incomes are between $20,000 and $200,000.[22] The sample consists of 28,303 families within these restrictions.

   With these restrictions on the sample, fertility rate is negatively correlated with family labor income, which is an well-known empirical evidence (see Figure 3).
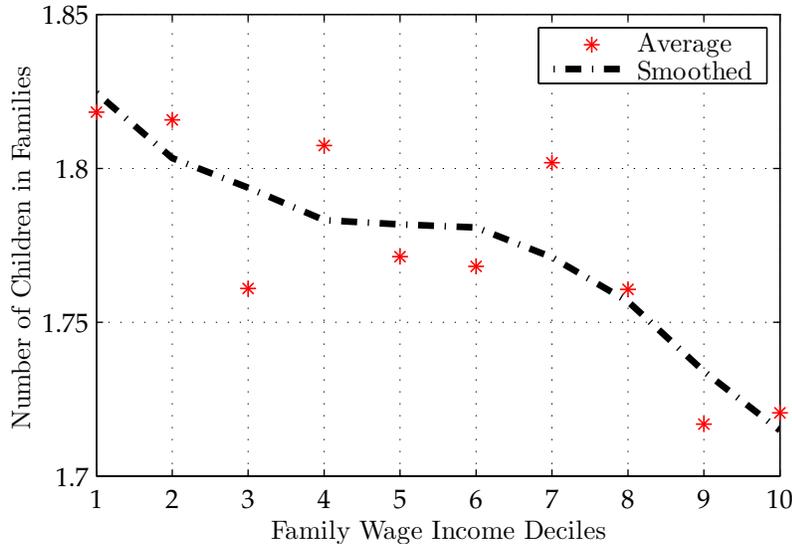

Figure 3: Income-Fertility Relation

Data: CPS 2005-2014. The sample is restricted to married (both spouses are present) households whose main source of income is labor. Total family wage income is greater than $20,000 and less than $200,000 and converted to 2011$ using CPI deflator. Spouses are aged from 35 to 45. The sample size is 28,303.

4.3   Child-rearing Costs

I find estimates of child-rearing costs in this subsection. First, I analyze the cost of goods. Next, I focus on the opportunity cost of time devoted to childcare and find estimates using CPS. I set $N = 2$ because the marginal cost of rearing the third child is relatively low (See Table 1 and See Table 2) and families with three or more children received EITC of two-child families until 2009. A robustness analysis on N is studied in Appendix A.4.

---

[21]Non-labor income is positively correlated with fertility (see Jones, Schoonbroodt, and Tertilt (2010)).

[22] The former level is around 130% of federal poverty level for a two-people family in 2011. Families below threshold may take benefits from other mean-tested programs which are beyond this paper. The latter level is in the beginning of the phase-in region of current tax credit programs.

4.3.1 Goods Cost: e_n

Haveman and Wolfe (1995) find that the annual cost of goods for child-rearing is $12,151 per child (in terms of 2011$) based on Consumer Expenditure Survey (CEX). Examples of such costs include expenditures on food, housing, transportation, clothing, and health care. More recently, Lino (2012) analyzes the cost of child-rearing for families with different wealth and different size. This work particularly provides information on expenditures for children with different age. Using this information, I create a range of expenditures for two-spouse families in Table 1. Table 1 shows there is economies of scale in expenses and expenditures are increasing in family income.

Families with	Average Income	1 child	2 children	3 children
Low Income	38,000	11,313-12,463	18,100-19,940	21,177-23,330
Middle Income	79,940	15,463-17,900	24,740-26,690	28,946-33,509
High Income	180,040	25,575-30,638	40,920- 49,020	47,876-57,353

Table 1: Expenditures on Child-rearing

The ranges are constructed by the Table-1 of Lino (2012). The first column categorizes families according to their income. The second column presents the average income for each category. The last three columns represent the range of expenditures on child-rearing. The expenditures are in 2011$.

I consider the minimum requirement of goods is a bundle of food, clothing, health care, and childcare and education expenditures. Figure 2 of Lino (2012) implies the share of these expenses is around 48% of total expenses. Hence, I set the cost of goods as 48% of expenses of low income families: $e_1 = \$5,500$ and $e_2 = \$9,500$. The remaining expenditures are considered as a part of family consumption.

4.3.2 Parental Time: b_n

The assumption on h normalizes the cost of working (see Kleven, Kreiner, and Saez (2009)). In laissez faire, (2) implies $z_n(\theta) = \theta(1 - b_n)$, hence θb_n is forgone family earnings owing to childcare. Accordingly, I interpret b_n as the opportunity cost of time devoted to childcare. Knowles (1999) and de la Croix and Doepke (2003) use the estimate of Haveman and Wolfe (1995) and set opportunity cost as 5-10% and 15% respectively. Haveman and Wolfe (1995) calculate opportunity cost as forgone income of full time working females owing to having children.[23] I follow a similar approach and calculate b_n as the

---

[23]Knowles (1999) considers the opportunity cost calculated by Haveman and Wolfe (1995) an upper bound because some of childcare can be considered as leisure (see Godbey and Robinson (1999)).

fraction of forgone labor work (hence income) of a family owing to children. CPS has information on weekly labor hours and number of weeks worked in a year. Therefore, I calculate the average of annual labor hours of families with different sizes (see Table 2). The opportunity cost of time is lower comparing to Haveman and Wolfe (1995) because males in families are also in the sample. However, when I particularly compare the opportunity cost of time of females, the measures are in line with Knowles (1999).

Families with	0 children	1 child	2 children	2+ children
total labor (l_n)	4258	4162	4122	4098
b_n	0	0.022	0.032	0.037
sample size	2,940	5,818	13,013	19,545
female labor (l_n)	2044	1937	1863	1834
b_n (for females)	0	0.052	0.088	0.103
sample size	1,088	2,299	5,095	7,731

Table 2: Opportunity Cost of Childcare

Data: CPS 2005-2014. l_n represent the weighted average hours per year devoted to the earnings by a family of two spouses. The sample is restricted to married, 35-45 years aged, and working households. Opportunity cost of n children is calculated as $b_n := \frac{l_0 - l_n}{l_0}$. Since $b_{2+} \simeq b_2$, I work with b_{2+}.

One might consider that the opportunity cost of time of high-income families are lower. However, the opportunity cost is, interestingly, very high for high-income families (see Table 3). This result can be attributed to positive correlation between childcare and education level of high-income earners (see Ramey and Ramey (2009)).

Quintiles	1	2	3	4	5
Average Income:	$45,726	$69,416	$89,452	$113,703	$157,960
b_n:	0.04	0.03	0.01	0.03	0.06

Table 3: Opportunity Cost across Income Groups

4.4 Estimations of Productivity Distribution and ρ

This subsection finds estimates related to families' non-observable characteristics. First, I estimate the probability distribution of productivities. Second, I estimate the marginal utility of each child in families.

4.4.1 Estimation of the Distribution of Productivities: $f(\theta)$

Assuming the data was generated by a suboptimal tax equilibrium, the optimality condition (2) of the quasi-linear preference structure allows me to find productivities of families:

$$\theta = \frac{z_n}{(1 - T'_n(z_n))^{\varepsilon} - b_n}. \tag{12}$$

Note that CPS has information about family structure and detailed sources of family income and taxes. I calculate $T'_n(z_n)$ by adding federal earned income tax rates to the reported federal marginal taxes to find effective marginal rates.[24] Using b_n values from Table 2 and the weights assigned to families by the data, I show the distribution of productivities in Figure 4.
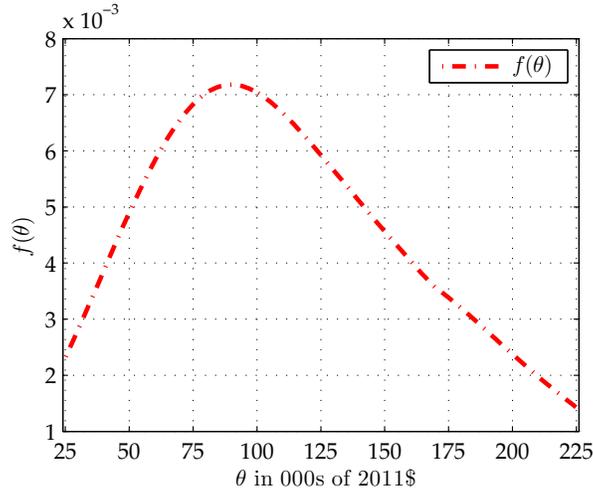
Figure 4: Productivity Distribution: $f(\theta)$

4.4.2 Estimation of the Marginal Utility Parameter: ρ

An important contribution of this paper is the estimation of family preferences over a number of children. First, I assume that $\beta \overset{i.i.d}{\sim} [1, \infty)$ is distributed according to an exponential function: $P(\beta) = 1 - \beta^{-\lambda}$ where $\lambda := -\frac{\partial \log(1 - P(\beta))}{\partial \log \beta}$ is the negative of the elasticity of having N children with respect to the marginal cost of Nth children.[25] Unfortunately, the literature on extensive elasticity of fertility with respect to marginal cost of a child is

---

[24]Earned income credits are reported in CPS, however, their marginal effects are not. I use the information on EITC for years 2005-2014 and impose its marginal effects.

[25]Given a θ, $1 - P(\beta_N(\theta))$ is the fraction of N child families and the marginal cost of Nth child is $\beta_N(\theta)$.

very sparse. In an interesting work, Cohen, Dehejia, and Romanov (2013) use an Israeli data and estimate a price elasticity of -0.54 for a child and in particular the estimate for the third child is -0.42. The elasticity of fertility with respect to the benefit of a child, however, has been studied by many works (see Table 4).

Work	Region of Data	Elasticity Estimates
Cohen, Dehejia, and Romanov (2013)	Israeli	0.19
Milligan (2005)	Canadian	0.11
Laroque and Salanie (2008)	French	0.20
Whittington, Alm, and Peters (1990)	US	0.12-0.23

Table 4: Benefit Elasticity of Having Children

Table 4 suggests that the benefit elasticities are similar across countries. Therefore, I follow Cohen, Dehejia, and Romanov (2013) and set $\lambda = 0.4$.[26]

I use a Bernoulli maximum likelihood estimation to find the estimate of ρ. First, I discretize Θ to its percentiles and calculate $V_n(\theta_j)$ for each jth percentile. For all θ_j, (5) implies that the theoretical probability of having n children is:

$$P_n(\theta_j) := P(\beta_{n+1}(\theta_j)) - P(\beta_n(\theta_j)) \quad \forall n \in \mathcal{N}$$

where $\beta_0(\theta_j) = 1$ and $\beta_{N+1}(\theta_j) = \infty$. Second, I can calculate the fraction of n-child families from data: $\pi_n(\theta_j)$. Finally, I derive the Bernoulli maximum likelihood function:

$$\rho \in \arg\max \mathcal{L} = \prod_{n=0}^{2} \prod_{j=0}^{100} P_n(\theta_j)^{\pi_n(\theta_j)}. \tag{ML}$$

The solution of (ML) is $\hat{\rho} = 3.95$ where $\hat{\sigma}_\rho = 1.64$.

5   Numerical Analysis on the Optimal Tax System

In this section, I compute the optimal policy using the estimates derived in Section 4. To compare statutory and the optimal tax schedule, I set government spending as $G = \$11,383$, which is the statutory tax amount collected by the government from the sample selection.[27],[28] I numerically solve the government's problem (PMDP), which is an optimal

---

[26]I do robustness analysis for λ in Appendix A.4.

[27]Statutory taxes are calculated by TAXSIM 9.2 version where TAXSIM is the NBER's program which calculates US Federal income tax liabilities across years from individual data.

[28]Per capita taxes reported in CPS is $10,384.

control problem and its Hamiltonian is stated in Appendix, at my selected and estimated parameters using the numerical solver GPOPS-II software.[29]

Figure 5 shows that the optimal income taxes are slightly higher for incomes between $70,000 - $130,000. On the other hand, lowest income households receive around $4,000 more under the optimal tax schedule.
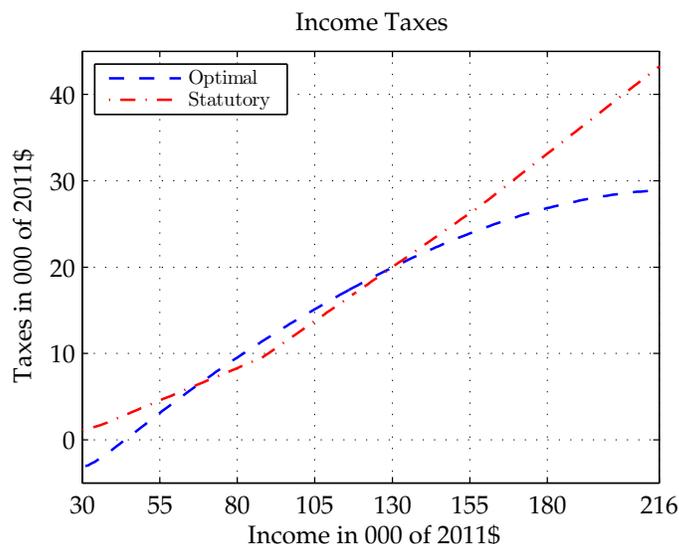
Income Taxes

Figure 5: Statutory versus Optimal Income Tax Schedules

Figure 6a and 6b show how optimal and statutory tax credits change across income. While the US tax credits are decreasing in income, the optimal credits are U-shaped.[30] The main force behind the difference is the opportunity cost of time devoted to child-rearing. Increasing credits for high-income parents do not only rise their consumption but also decreases their marginal tax rates which reduces the distortion on their labor margin, and hence incentivizes them to work harder. The shape of statutory credits implies that the US government focuses only on the cost of goods and does not deal with the time cost.

Second, the US credits are (almost) same for each child for a large range of income. In contrast, the optimal credits are decreasing in the number of children in family. There are two main sources behind this result. First, there is an economies of scale in child-rearing costs. Second, families have convex preferences over their wealth. Consequently, the welfare loss owing to the cost of rearing second child is less the loss owing to the cost of rearing the first. The optimal policy responses to these facts is to decrease credits by

---

[29]GPOPS-II is a flexible software program for solving optimal control problems, see Patterson and Rao (2014).

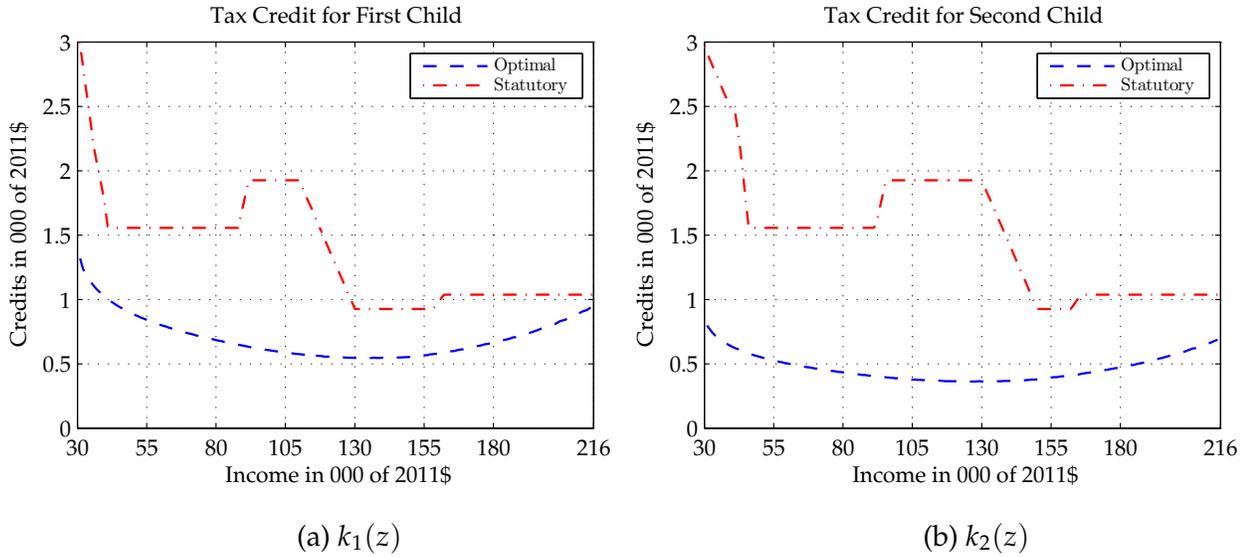[30]Small jumps in the statutory credits are owing to the effect of personal exemption when marginals change.

Figure 6: Statutory versus Optimal Tax Credits

the number of children. However, the US government disregards the difference in the welfare losses and provides same credits for each child in the family. On other hand, the UK government has recently proposed to cut the third child benefits which is in line with the optimal tax credit schedule.

These results suggest that the optimal tax schedule is strictly different than the current US tax system. Designing the tax schedule with U-shaped tax credits may improve social welfare. I find that the welfare gain from implementing the optimum is 1.3% in terms of equivalent increase in consumption for all families. In addition, 50% of families are better of with the optimal tax schedule. This suggests that the optimal tax schedule does not only increase social welfare but also is very attractive to the society.

6 Robustness Analysis

In this section, I analyze robustness of U-shaped tax credits. Among many, I show four different robustness analysis here.[31] First analysis is about the government's objective. Figure 7 shows the tax credit schedules for the Utilitarian and the Rawlasian government.[32] Both schedules are U-shaped according to income. Yet, Rawlasian government provides more credits, especially to the lowest income parents. This is because the Rawlasian government has stronger redistribution motives at the bottom of income distribution.

---

[31]The other robustness analysis can be found in Appendix A.4.

[32]The Rawlasian government's objective is to maximize lowest utility in the economy.

25

Second analysis is on the number of children. Let $n(\theta) := \mathbb{E}_\beta[n(\beta, \theta)]$ represent the average number of children of families whose labor productivities are θ. Under the optimal tax schedule in the benchmark, the total number of children in the economy is 9.7% less than the data. In addition, the solution of (PMDP) implies $n(\theta)$ is decreasing in θ because of opportunity cost of time (see Figure 8a). On the other hand, $n(\theta)$ of data circles around its average.[33] This raises the question of what are the optimal child tax credits if the

(a) Utilitarian Government   (b) Rawlasian Government
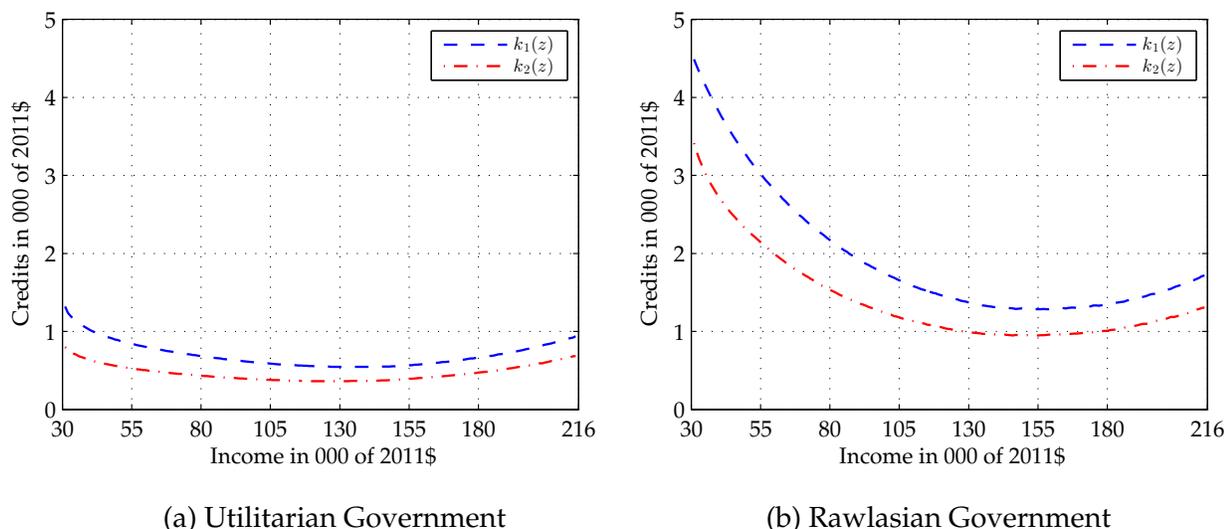
Figure 7: Government Preference Analysis

government solution should provide the same average number of children over productivities? To answer this question, I impose the constraint of $n(\theta) = 1.49$ for all $\theta \in \Theta$ into (PMDP). The optimal tax credits are stated in Figure 8b. Tax credits are much higher comparing to the benchmark and increasing over income. The reason is that the government has to provide more incentives for families to have more children. Also, the government should enhance incentives towards higher income families because the opportunity cost of time for childcare becomes heavier towards them.

---

[33]The average of number of children of the sample is 1.77. Families with more than two children are considered as they have two children. Therefore, the average is 1.49.
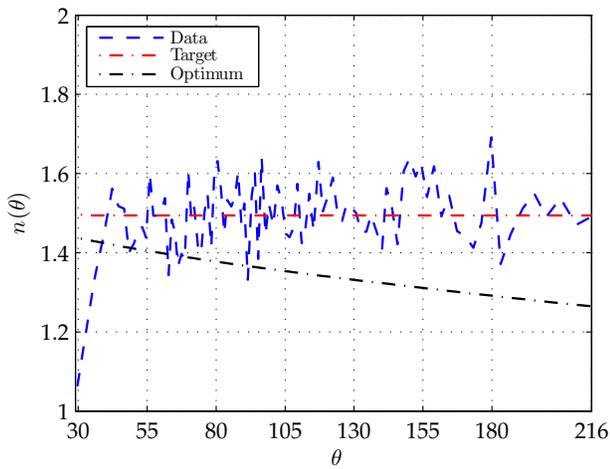
(a) Number of Children
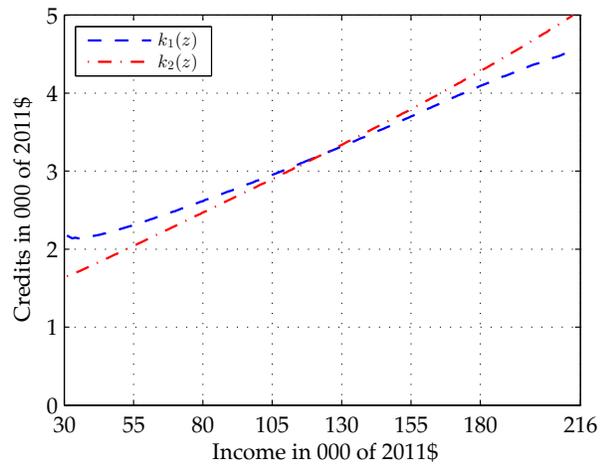
(b) Tax Credits

Figure 8: Child Quantity Criterion

Third analysis focuses on households' preferences. I consider families' preferences are represented by $\mathcal{U} = \frac{\left(c - h(\frac{z}{\theta} + b_n)\theta\right)^{1-\sigma}}{1-\sigma} + m(n, \beta)$. I let $\sigma = 0.5$. Figure 9a and 9b show the tax credits for Utilitarian and Rawlasian government, respectively. Both of the credit schedules are U-shaped in income and very similar with Figure 7.
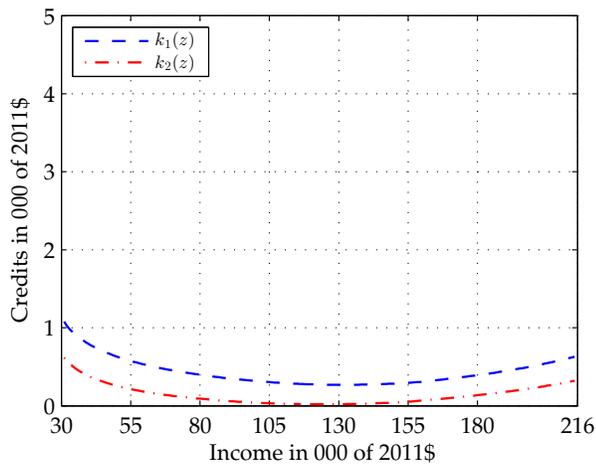
(a) Utilitarian Government
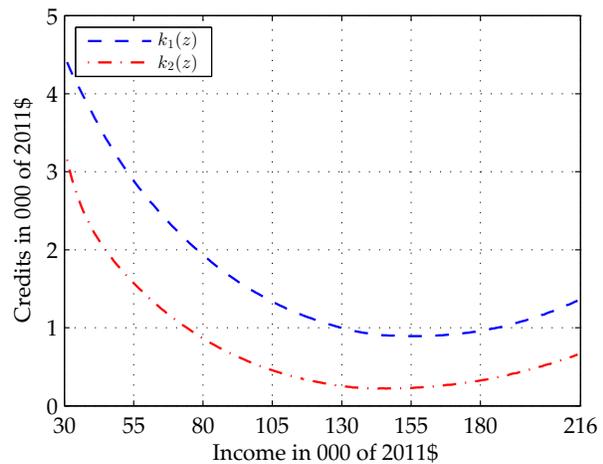
(b) Rawlasian Government

Figure 9: Households' Preference Analysis

In the fourth analysis, I study tax credits for single mothers. I adjust estimation parameters: $\varepsilon = 0.8$ (see Blundell, Pistaferri, and Saporta-Eksten (2012)), $e_1 = \$5,000$ and

27

$e_2 = \$8,500$, and $b_1 = 0.035$ and $b_2 = 0.039$.[34],[35] Figure 10a shows the productivity distribution for single females. The solution of (PMDP) with the adjusted parameters implies that the optimal credits are U-shaped in income (see Figure 10b).
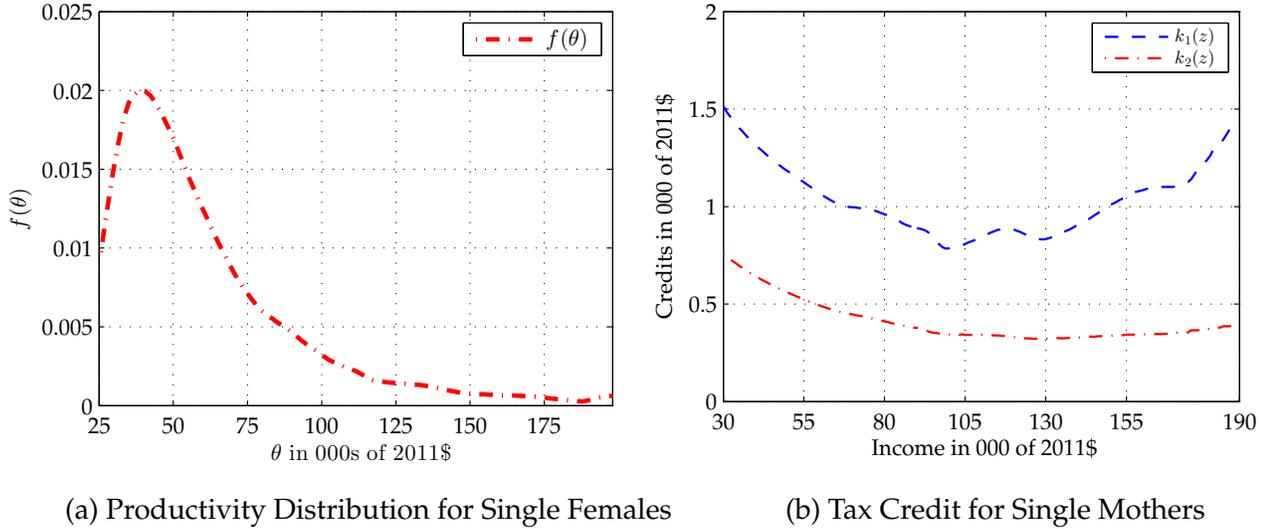

(a) Productivity Distribution for Single Females

(b) Tax Credit for Single Mothers

Figure 10: Marital Status Analysis

These four analysis suggest that U-shaped tax credit schedule is very robust. In addition to these analysis, I state more robustness analysis in Appendix A.4. I analyze the effect of λ, e_n, b_n, N, and age restriction of spouses in the sample. The optimal tax credits are U-shaped in income.

7 Conclusion

This paper studies optimal income taxation and child tax credits in a static Mirrlees model in which potential parents privately observe their characteristics-child tastes and labor productivities. Households decide how much income to generate and how many children to have by considering child-rearing costs. A Utilitarian government maximizes social welfare and determines the equity-efficiency trade-off owing to the informational friction. An optimal tax mechanism is founded on this trade-off and combines income taxation for all families and child tax credits for parents. The sufficient statistics for labor wedges and their relationship with child tax credits are derived.

---

[34] According to Table 7 of Lino (2012), low income single mothers spend $10,000 for one child and $17,000 for two children.

[35] A single woman without children works 2196 hours per year on average. A single mother with one child and two children works 2118 and 2111 hours, respectively.

Income taxes are designed to redistribute from high to low income families and child tax credits decrease tax liabilities of parents who bear child-rearing costs. The child-rearing costs are crucial inputs on the shape of the child tax credits. The cost of goods mostly affects the low-income parents and drives the government's redistribution motives towards them. On the other hand, time cost disincentives high-productive parents working hard. The government uses tax credits as incentive tools and increases provisions for the wealthier parents. As a result, the credits are U-shaped in income. Quantitatively, I find that the optimal tax credits are decreasing in the first three quarters of income distribution and are increasing afterwards. In addition, the tax credits are decreasing by family size because there is an economies of scale in child-rearing costs and families have convex preferences over their wealth.

This paper sheds light on the optimal income taxation including the child benefits for families who have multidimensional privately observed characteristics. I conclude by describing three extensions that I leave for future research. First, the paper abstracts from a dynamic setting. Such a setting can explain how the child benefit should be characterized by the age of the children. Two heterogeneous characteristics-a labor productivity and a taste for children- can be linked with the age of the parents and, consequently, the effect of optimal taxes on the time of fertility can be studied. Second, the paper abstracts from the child quality decision, which is positively correlated with parental time according to Boca, Flinn, and Wiswall (2013). Such a decision can explain why high-income parents spend more time with their children (see Guryan, Hurst, and Kearney (2008)). Third, the costs of child-rearing can be endogenous. This endogeneity can help policy makers on designing the optimal provisions via costs. For example, policies that provide a high-quality childcare in return of goods might be tempting for high-income families. This extension can also examine the current debate in the US on universal childcare provisions for working parents.

References

Ales, L., M. Kurnaz, and C. Sleet (2015). Technical Change, Wage Inequality, and Taxes. *American Economic Review 105*(10), 3061 – 3101.

Baron, D. P. and R. B. Myerson (1982). Regulating a Monopolist with Unknown Costs. *Econometrica 50*(4), 911 – 930.

Becker, G. S. (1960). *An Economic Analysis of Fertility*, pp. 209–240. Columbia University Press.

Becker, G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal 75*(299), 493–517.

Blau, D. (2003). Child Care Subsidy Programs. In *Means-Tested Transfer Programs in the United States*, pp. 443 – 516. University of Chicago Press.

Blundell, R. and T. Macurdy (1999). Labor supply: A Review of Alternative Approaches. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3 of *Handbook of Labor Economics*, Chapter 27, pp. 1559 – 1695. Elsevier.

Blundell, R., C. Meghir, and P. Neves (1993). Labour Supply and Intertemporal Substitution. *Journal of Econometrics 59*(1), 137 – 160.

Blundell, R., L. Pistaferri, and I. Saporta-Eksten (2012, October). Consumption Inequality and Family Labor Supply. Working Paper 18445, National Bureau of Economic Research.

Boca, D. D., C. Flinn, and M. Wiswall (2013). Household Choice and Child Development. *The Review of Economic Studies 81*(1), 137 – 185.

Chetty, R. (2006). A New Method of Estimating Risk Aversion. *American Economic Review 96*(5), 1821–1834.

Chetty, R. (2012). Bounds on Elasticities With Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply. *Econometrica 80*(3), 969 – 1018.

Chiappori, P.-A. (1988). Rational household labor supply. *Econometrica 56*(1), 63–90.

Chiappori, P.-A. (1992). Collective labor supply and welfare. *Journal of political Economy 100*(3), 437–467.

Chiappori, P.-A., B. Fortin, and G. Lacroix (2002). Marriage market, divorce legislation, and household labor supply. *Journal of Political Economy 110*(1), 37–72.

Cohen, A., R. Dehejia, and D. Romanov (2013). Financial Incentives and Fertility. *The Review of Economics and Statistics 95*(1), 1–20.

de la Croix, D. and M. Doepke (2003). Inequality and Growth: Why Differential Fertility Matters. *American Economic Review 93*(4).

Docquier, F. (2004). Income Distribution, Non-convexities and the Fertility - Income Relationship. *Economica 71*(282), 261 – 273.

Falk, G. and M. L. Crandall-Hollick (2016). The Earned Income Tax Credit (EITC): An Overview.

French, E. (2005). The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour. *The Review of Economic Studies 72*(2), 395 – 427.

Godbey, G. and J. P. Robinson (1999). *Time for Life: The Surprising Ways Americans Use their Time*. The Pennsylvania State University Press.

Golosov, M., L. E. Jones, and M. Tertilt (2007). Efficiency with Endogenous Population Growth. *Econometrica 75*(4), 1039 – 1071.

Guryan, J., E. Hurst, and M. Kearney (2008). Parental Education and Parental Time with Children. *Journal of Economic Perspectives 22*(3), 23–46.

Haveman, R. and B. Wolfe (1995). The Determinants of Children's Attainments: A Review of Methods and Findings. *Journal of Economic Literature 33*(4), 1829 – 1878.

Hosseini, R., L. E. Jones, and A. Shourideh (2013). Optimal Contracting with Dynastic Altruism: Consumption and Family Size. *Journal of Economic Theory 148*(5), 1806 – 1840.

Hotz, V. J. and J. K. Scholz (2003). The Earned Income Tax Credit. In *Means-Tested Transfer Programs in the United States*, pp. 141 – 198. University of Chicago Press.

Jacquet, L., E. Lehmann, and B. V. der Linden (2013). The Optimal Marginal Tax Rates with both Extensive and Intensive Responses. *Journal of Economic Theory 148*(5), 1770 – 1805.

Jones, L. E., A. Schoonbroodt, and M. Tertilt (2010). Fertility Theories: Can They Explain the Negative Fertility-Income Relationship? In J. B. Shoven (Ed.), *Demography and the Economy*, pp. 43–100. University of Chicago Press.

Jones, L. E. and M. Tertilt (2008). An Economic History of Fertility in the U.S.: 1826 - 1960. In P. Rupert (Ed.), *Frontiers of Family Economics*, pp. 165 – 230. Bingley, UK: Emerald Press.

Keane, M. P. (2011). Labor Supply and Taxes: A Survey. *Journal of Economic Literature 49*(4), 961 – 1075.

Kleven, H. J., C. T. Kreiner, and E. Saez (2009). The Optimal Income Taxation of Couples. *Econometrica 77*(2), 537 – 560.

Knowles, J. (1999). Can Parental Decisions Explain U.S. Income Inequality? Mimeo, University of Pennsylvania, 1999.

Laroque, G. and B. Salanie (2008). Does Fertility Respond to Financial Incentives? Columbia University, Department of Economics, Discussion Paper Number 0708-15.

Lino, M. (2012). Expenditures on Children by Families, 2011. U.S. Department of Agriculture, Center for Nutrition Policy and Promotion. Miscellaneous Publication No. 1528-2011.

Milligan, K. (2005). Subsidizing the Stork: New Evidence on Tax Incentives and Fertility . *The Review of Economics and Statistics 87*(3).

Mirrlees, J. (1971). An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies 38*(2), 175 – 208.

Patterson, M. A. and A. V. Rao (2014, October). Gpops-ii: A matlab software for solving multiple-phase optimal control problems using hp-adaptive gaussian quadrature collocation methods and sparse nonlinear programming. *ACM Trans. Math. Softw. 41*(1), 1:1–1:37.

Pencavel, J. (1986). Labor supply of Men: A Survey. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics*, Volume 1 of *Handbook of Labor Economics*, Chapter 1, pp. 3 – 102. Elsevier.

Ramey, G. and V. A. Ramey (2009, August). The rug rat race. Working Paper 15284, National Bureau of Economic Research.

Rochet, J.-C. and P. Chone (1998). Ironing, Sweeping, and Multidimensional Screening. *Econometrica 66*(4), 783 – 826.

Ross, I. M. and F. Fahroo (2003). *New Trends in Nonlinear Dynamics and Control and their Applications*, Chapter Legendre Pseudospectral Approximations of Optimal Control Problems, pp. 327–342. Berlin, Heidelberg: Springer Berlin Heidelberg.

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

Whittington, L. A., J. Alm, and H. E. Peters (1990). Fertility and the Personal Exemption: Implicit Pronatalist Policy in the United States. *The American Economic Review 80*(3), 545–556.

Ziliak, J. P. and T. J. Kniesner (2005). The Effect of Income Taxation on Consumption and Labor Supply. *Journal of Labor Economics 23*(4), 769 – 796.

Appendix

Proof of Proposition 1. The Hamiltonian of the problem is:

$$
\mathcal{H} = \sum_{n=0}^{n=N} \int_{\beta_n(\theta)}^{\beta_{n+1}(\theta)} \left(\xi_n(\theta) \Psi(V_n(\theta) + m(n, \beta)) + \lambda[z_n(\theta) - c_n(\theta)] \right) p(\beta|\theta) f(\theta) d\beta
$$

$$
+ \sum_{n=0}^{n=N} \mu_n(\theta) \left(-h \left(\frac{z_n}{\theta} + b_n \right) + h' \left(\frac{z_n}{\theta} + b_n \right) \frac{z_n}{\theta} \right) \qquad \text{(Hamiltonian)}
$$

where $c_n(\theta) := u^{-1} \left(V_n(\theta) + h \left(\frac{z_n}{\theta} + b_n \right) \theta \right)$ and $\mu_n(\underline{\theta}) = \mu_n(\bar{\theta}) = 0.$[36] The first-order conditions are:

$$
\lambda \left(1 - \frac{\partial c_n(\theta)}{\partial z_n(\theta)} \right) (P(\beta_{n+1}|\theta) - P(\beta_n|\theta)) f(\theta) = -\frac{\mu_n(\theta)}{\theta} h'' \left(\frac{z_n(\theta)}{\theta} + b_n \right) \frac{z_n(\theta)}{\theta} \quad \forall n \in \mathcal{N}.
$$

Also, the co-states are:

$$
\frac{-\dot{\mu}_n(\theta)}{\lambda f(\theta)} = \int_{\beta_n(\theta)}^{\beta_{n+1}(\theta)} \left(\frac{\xi_n(\theta) \Psi'(V_n(\theta) + m(n, \beta))}{\lambda} - \frac{\partial c_n(\theta)}{\partial V_n(\theta)} \right) p(\beta|\theta) d\beta
$$

$$
+ \sum_{j=n}^{n+1} \Delta T_{j-1}(\theta) p(\beta_j(\theta)|\theta) \frac{\partial \beta_j(\theta)}{\partial V_n(\theta)} \quad \forall n \in \mathcal{N}.
$$

where $T_n(\theta) = z_n(\theta) - c_n(\theta)$ and $\Delta T_n(\theta) = T_n(\theta) - T_{n+1}(\theta).$[37] Boundary conditions imply:

$$
-\frac{\mu_n(\theta)}{\lambda} = \int_{\theta}^{\bar{\theta}} \left[\frac{(1 - g_n(\theta'))}{u'(c_n(\theta'))} (P(\beta_{n+1}(\theta')|\theta') - P(\beta_n(\theta')|\theta')) \right.
$$

$$
\left. + \sum_{j=n}^{n+1} \Delta T_{j-1}(\theta') p(\beta_j(\theta')|\theta') \frac{\partial \beta_j(\theta')}{\partial V_n(\theta')} \right] f(\theta') d\theta'
$$

where

$$
g_n(\theta) = \mathbb{E}_\beta \left[\frac{\xi_n(\theta) \Psi'(V_n(\theta) + m(n, \beta)) u'(c_n(\theta))}{\lambda} |\beta_n(\theta) < \beta < \beta_{n+1}(\theta) \right]
$$

$$
= \frac{\int_{\beta_n(\theta)}^{\beta_{n+1}(\theta)} \xi_n(\theta) \Psi'(V_n(\theta) + m(n, \beta)) u'(c_n(\theta)) p(\beta|\theta) f(\theta) d\beta}{\lambda (P(\beta_{n+1}(\theta)|\theta) - P(\beta_n(\theta)|\theta)) f(\theta)} \qquad (13)
$$

---

[36] Let $\beta_0 = \underline{\beta}$, and $\beta_{N+1} = \bar{\beta}$.

[37] $T_n(\theta)$ is the taxes paid by $\theta-$productivity families with n children. Let $\Delta T_{-1}(\theta) = 0$ and $\Delta T_N(\theta) = 0$.

is the marginal weight associated by the government to the family θ with n children, which is the cost of giving an extra dollar of consumption to the family in terms of public goods.

Combining the of previous terms shows that the optimal tax function should satisfy:

$$
\frac{T'_n(\theta)}{1 - T'_n(\theta)} = \frac{1}{\varepsilon_n(\theta)} \times \frac{1}{\theta f(\theta)(P(\beta_{n+1}(\theta)|\theta) - P(\beta_n(\theta)|\theta))} \times
$$
$$
\int_\theta^{\bar\theta} \left[\frac{(1 - g_n(\theta'))}{u'(c_n(\theta'))} (P(\beta_{n+1}(\theta')|\theta') - P(\beta_n(\theta')|\theta'))) \right.
$$
$$
\left. + \sum_{j=n}^{n+1} \Delta T_{j-1}(\theta') p(\beta_j(\theta')|\theta') \frac{\partial \beta_j(\theta')}{\partial V_n(\theta')} \right] u'(c_n(\theta)) f(\theta') d\theta' \quad \forall n \in \mathcal{N}
$$

where $\varepsilon_n(\theta)$ is the compensated elasticity of income of n-child families with respect to marginal tax rates.[38] □

---

[38] Let $\Delta T_{-1}(\theta) = 0$ and $\Delta T_N(\theta) = 0$.

A ONLINE APPENDICES

A.1 US CHILD TAX PROGRAMS

This subsection introduces how children in families affect the family income taxes. I state 4 main tax programs here.

First, the Child Tax Credit (CTC) program, which is a mean tested program, was enacted as a temporary provision in the Taxpayer Relief Act of 1997. The tax credit per child has gradually increased from $400 to $1,000 from 2001 to 2010 and become refundable for all families by the Economic Growth and Tax Relief Reconciliation Act of 2001.[39] The phase-out income level for married tax payers filing jointly is $110,000. For each $1,000 of income above this threshold, the child tax credits are reduced by $50. Furthermore, the CTC program has become permanent provision by the American Taxpayer Relief Act of 2012.

Second, the Child and Dependent Care Tax Credit (CDCTC) program decreases the tax liability of families by 20% to 35% of childcare expenditures for a qualifying child up to $3,000 for up to two children. Also, $5,000 of a family salary can be excluded from adjusted gross income for childcare if certain regulations are satisfied. The credit is non-refundable, and hence many low-income families do not participate in this program (Refer to Blau (2003) for more details).

Third, families' AGI is reduced by dependent exemptions for each child in families. The exemption amount gradually increased from $2,800 to $3,700 from 2000 to 2011. This program is also a mean-tested transfer and the exemption decreases beginning with phase-out income.

Fourth, the Earned Income Tax Credit (EITC) is another mean-tested transfer program for working families and mostly attained by parents. According to Falk and Crandall-Hollick (2016), %97 of EITC budget is spent for families with children. The maximum credit, the phase in and out rates drastically change with the number of children in families. Table 5 shows the EITC rates for 2011. Families with more children can claim more tax credits (Refer to Hotz and Scholz (2003) for more details).

---

[39]If a family has less tax liability than their child tax credit, they may get the minimum of unclaimed credits and 15% of their income above $3,000.

# of children	earnings ≤	credit rate	max credit	phase-out begins	phase-out rate
0	$6,070	0.08	$464	$12,670	0.08
1	$9,100	0.34	$3,094	$21,770	0.16
2	$12,780	0.40	$5,112	$21,770	0.21
3	$12,780	0.45	$5,751	$21,770	0.21

Table 5: EITC for married tax payers filling jointly for 2011
Source: https://www.irs.gov/irb/2011-02_IRB/ar16.html

A.2 MECHANISM DESIGN: TWO-DIMENSIONAL PRIVATE INFORMATION

In this section, I show the implementability conditions for a two-dimensional private information problem. I approach it similarly to Jacquet, Lehmann, and der Linden (2013) and Kleven, Kreiner, and Saez (2009). I differ from these works in two ways. First, both of these papers consider two groups of households. Yet, families can have an arbitrary number of children in my paper. Second, these works do not consider the time effect of secondary characteristics. However, in this work, any existing child requires parental time, which is perfectly substitutable with market labor.

Let $\gamma = (\beta, \theta) \in \mathcal{B} \times \Theta = \Gamma$ be the private information of a family. If the family report γ as their characteristics, the government chooses optimal allocation $(c(\gamma), z(\gamma), n(\gamma))$ and associated utility is:

$$\mathcal{U}((c(\gamma), z(\gamma), n(\gamma))) = \Psi\left(u(c(\gamma)) - \theta h\left(\frac{z(\gamma)}{\theta} + b_n\right) + m(n(\gamma), \beta)\right).$$

This mechanism should satisfy the revelation principle, by which any government mechanism can be decentralized by a truthful mechanism $(c(\gamma), z(\gamma), n(\gamma))_{\gamma \in \Gamma}$ such that

$$\mathcal{U}(c(\gamma), z(\gamma), n(\gamma)) \geq \Psi\left(u(c(\gamma')) - \theta h\left(\frac{z(\gamma')}{\theta} + b_n\right) + m(n(\gamma'), \beta)\right) \quad \forall (\gamma \times \gamma') \in \Gamma^2.$$

Let $\mathcal{U}(((c(\gamma'), z(\gamma'), n(\gamma')), \gamma) := \Psi\left(u(c(\gamma')) - \theta h\left(\frac{z(\gamma')}{\theta} + b_n\right) + m(n(\gamma'), \beta)\right)$ be the utility of the family with γ characteristics who reports γ' for its characteristics and gets $(c(\gamma'), z(\gamma'), n(\gamma'))$.

Family characteristics are twofold, and hence a possible mimicking strategy has two dimensions. The possibility of double deviation in the mimicking strategy is handled by the indirect utility of n child families (4) and threshold tastes for children (5) for each n. From the classical mechanism design problem to a pseudo-mechanism design problem, I

first show that the solution to the classical problem can be replaced by a pseudo-problem solution in the next Lemma.

Lemma 3. *Any truthful mechanism $(c(\gamma), z(\gamma), n(\gamma))_{\gamma \in \Gamma}$ can be replaced by a new truthful mechanism $\{c_n(\theta), z_n(\theta)_{\theta \in \Theta}\}_{n \in \mathcal{N}}$, such that $\forall \theta \in \Theta$ and $\forall n \in \mathcal{N}$, there is a $\beta_n(\theta)$ such that if $\beta \in (\beta_n(\theta), \beta_{n+1}(\theta))$,[40] then $\mathcal{U}(c_n(\theta), z_n(\theta), n, \gamma) \geq \max\limits_{\gamma'} \mathcal{U}(((c(\gamma'), z(\gamma'), n(\gamma')), \gamma)$. The new mechanism generates same utility as the original mechanism and the government collects as much taxes as the original mechanism.*

Proof. For each θ, partition the set $\mathcal{B}$ into $N + 1$ sets such that if $\beta \in \mathcal{B}_j$ then $n(\beta, \theta) = j$ for all $j \in \mathcal{N}$. If the family is indifferent between having k children and $k + 1$ children I assume that $n(\beta, \theta) = k + 1$.

For a given θ and $\beta, \beta' \in \mathcal{B}_j$, the truthfulness of the original mechanism implies:

$$u(c(\beta, \theta)) - h\left(\frac{z(\beta, \theta)}{\theta} + b_j\right)\theta + m(j, \beta) \geq u(c(\beta', \theta)) - h\left(\frac{z(\beta', \theta)}{\theta} + b_j\right)\theta + m(j, \beta)$$

$$u(c(\beta', \theta)) - h\left(\frac{z(\beta', \theta)}{\theta} + b_j\right)\theta + m(j, \beta') \geq u(c(\beta, \theta)) - h\left(\frac{z(\beta, \theta)}{\theta} + b_j\right)\theta + m(j, \beta').$$

The first inequality is $\mathcal{U}((\beta, \theta), (\beta, \theta)) \geq \mathcal{U}((\beta, \theta), (\beta', \theta))$ and the second inequality is $\mathcal{U}((\beta', \theta), (\beta', \theta)) \geq \mathcal{U}((\beta', \theta), (\beta, \theta))$. It is easy to see $\mathcal{U}((\beta, \theta), (\beta, \theta)) = \mathcal{U}((\beta', \theta), (\beta', \theta))$, which implies $u(c(\beta, \theta)) - h\left(\frac{z(\beta, \theta)}{\theta} + b_j\right)\theta$ is constant for all $\beta \in \mathcal{B}_j$, and let $V_j(\theta)$ be its value.

Note that at least as much taxes should be collected with the new mechanism. Let $Z_j(\theta) = \{z(\beta, \theta) | \beta \in \mathcal{B}_j(\theta)\}$. Define $t_j = \sup_{z \in \overline{Z}_j(\theta)} z - u^{-1}\left(V_j(\theta) + h\left(\frac{z}{\theta} + b_j\right)\theta\right)$. Note that $z - u^{-1}\left(V_j(\theta) + h\left(\frac{z}{\theta} + b_j\right)\theta\right)$ is a weakly concave function in z and reaches maximum for a $\tilde{z}$ value and goes to $-\infty$ when $z \to \infty$. So there is a $z_j(\theta) \in \overline{Z}_j(\theta)$ such that $t_j = z_j(\theta) - u^{-1}\left(V_j(\theta) + h\left(\frac{z_j(\theta)}{\theta} + b_j\right)\theta\right)$.[41] Define $c_j(\theta) := u^{-1}\left(V_j(\theta) - h\left(\frac{z_j(\theta)}{\theta} + b_j\right)\theta\right)$. Note that $(c_j(\theta), z_j(\theta))$ maximizes the taxes over the closure of the set $\{c(\beta, \theta), z(\beta, \theta)\}_{\beta \in \mathcal{B}_j(\theta)}$. These procedures can be followed for all $j \in \mathcal{N}$.

Finally, I define $\beta_n(\theta) := M_n^{-1}(V_n(\theta) - V_{n+1}(\theta))$ where $M_n(\beta) := m(n + 1, \beta) - m(n, \beta)$ for all $n \in \mathcal{N}$.[42] $\beta_n(\theta)$ are the threshold tastes for children for each θ and for each n. Note that truthfulness of original mechanism implies: for all $\beta \in \mathcal{B}_j(\theta)$ the family chooses $n = j$ and $(z_j(\theta), c_j(\theta))$, i.e. $V_j(\theta) + m(j, \beta) \geq V_{j'}(\theta) - m(j', \beta)$ for all $j' =$

---

[40]I let $\beta_0 = \underline{\beta}$ and $\beta_{N+1} = \overline{\beta}$.
[41]$\overline{Z}_j(\theta)$ is the closure of the $Z_j(\theta)$
[42]Let $\beta_0 = \underline{\beta}$ and $\beta_{N+1} = \overline{\beta}$.

$0, 1, \ldots, N$. Pick $j' = j-1$ and $j' = j+1$. Then it is easy to see that $\underbrace{M_n^{-1}(V_j(\theta) - V_{j+1}(\theta))}_{\beta_j(\theta)} \geq$

$\beta \geq \underbrace{M_n^{-1}(V_{j-1}(\theta) - V_j(\theta))}_{\beta_{j-1}(\theta)}.$[43] Therefore $\mathcal{B}_j(\theta) = (\beta_{j-1}(\theta), \beta_j(\theta))$.

All is left to show the new mechanism $\{(c_n(\theta), z_n(\theta))_{\theta \in \Theta}\}_{n \in \mathcal{N}}$ is truthful. First I show it is truthful within families with the same number of children: For all $\theta, \theta', \beta \in \mathcal{B}_j(\theta), \beta' \in \mathcal{B}_j(\theta')$:

$$\mathcal{U}(c_j(\theta), z_j(\theta), (\beta, \theta)) = \Psi(V_j(\theta) + m(j, \beta)) \geq \Psi\left(u(c(\beta', \theta')) - h\left(\frac{z(\beta', \theta')}{\theta}\right) - m(j, \beta)\right)$$

where the inequality is from the truthfulness of the initial mechanism.[44] As a result,

$$\mathcal{U}(c_j(\theta), z_j(\theta), (\beta, \theta)) \geq \mathcal{U}(c_j(\theta'), z_j(\theta'), (\beta, \theta)).$$

I also show the mechanism is truthful cross-sectionally: for all $\theta, \theta', \beta \in \mathcal{B}_j(\theta), \beta' \in \mathcal{B}_{j'}(\theta')$:

$$\mathcal{U}(c_j(\theta), z_j(\theta), (\beta, \theta)) = \Psi(V_j(\theta) + m(j, \beta)) \geq \Psi(V_{j'}(\theta) + m(j', \beta))$$

$$\geq \Psi\left(u(c(\beta', \theta')) - h\left(\frac{z(\beta', \theta')}{\theta}\right) + m(j', \beta)\right)$$

where the first inequality comes from the definition of β_n and the second inequality is satisfied by the truthfulness of the original truthful mechanism. Hence:

$$U(c_j(\theta), z_j(\theta), (\beta, \theta)) \geq U(c_{j'}(\theta'), z_{j'}(\theta'), (\beta, \theta)).$$

This procedure can be followed for any $j \in \mathcal{N}$. $\qquad\square$

This lemma allows me to move from $\{c(\beta, \theta), z(\beta, \theta), n(\beta, \theta)\}_{(\beta,\theta) \in \mathcal{B} \times \Theta}$ schedule to $\{(c_n(\theta), z_n(\theta))_{\theta \in \Theta}\}_{n \in \mathcal{N}}$ schedule. I directly use the one-dimensional implementation requirement as long as the single-crossing condition is satisfied.

Definition 2. $\{z_n(\theta)_{\theta \in \Theta}\}_{n \in \mathcal{N}}$ is implementable if and only if there exist transfer functions $\{c_n(\theta)_{\theta \in \Theta}\}_{n \in \mathcal{N}}$ such that $\{(c_n(\theta), z_n(\theta))_{\theta \in \Theta}\}_{n \in \mathcal{N}}$ is a truthful mechanism.

In the following lemma, I prove that a one dimensional requirement is sufficient for the two-dimensional problem in this framework:

---

[43]Note that I let m to be concave in its first dimension and therefore $\beta_{j-1}(\theta) < \beta_j(\theta)$.

[44] Note that $(c_j(\theta'), z_j(\theta'))$ is in the closure of the set $(c(\beta, \theta), z(\beta, \theta))_{\beta \in \mathcal{B}_j(\theta)}$.

Lemma 4. *The income profile $\{z_n(\theta)_{\theta \in \Theta}\}_{n \in \mathcal{N}}$ is implementable if and only if $\dot{z}_n(\theta) := \frac{\partial z_n(\theta)}{\partial \theta} \geq 0$.*

Proof. Note that $u(c) - h\left(\frac{z}{\theta} + b_n\right)\theta$ satisfies the classic single crossing condition. The one-dimensional implementability condition is that: $\dot{z} \geq 0$ if and only if there is $c(\theta)$ such that $u(c(\theta)) - h\left(\frac{z(\theta)}{\theta} + b_n\right)\theta \geq u(c(\theta')) - h\left(\frac{z(\theta')}{\theta} + b_n\right)\theta$ for all θ, θ'.

For the "if" side of the lemma, I directly apply the one-dimensional implementability condition: for all $n \in \mathcal{N}$, let $\{z_n(\theta)_{\theta \in \Theta}\}_{n \in \mathcal{N}}$ is implementable. Then for a particular n, truthfulness implies $u(c_n(\theta)) - h\left(\frac{z_n(\theta)}{\theta} + b_n\right)\theta \geq u(c_n(\theta')) - h\left(\frac{z_n(\theta')}{\theta} + b_n\right)\theta$ for all θ, θ'. As a result, the one-dimensional result suggests that for each $n \in \mathcal{N}$ income is non-decreasing: $\dot{z}_n \geq 0$.

Now let $\dot{z}_n \geq 0$. Similarly, using the one-dimensional result, there is $c_n(\theta)$ such that $u(c_n(\theta)) - h\left(\frac{z_n(\theta)}{\theta} + b_n\right)\theta \geq u(c_n(\theta')) - h\left(\frac{z_n(\theta')}{\theta} + b_n\right)\theta$ for all θ, θ'.

Within sections, the one-dimensional condition is directly applicable, as shown above. All that is need to be shown is that cross-sectional truth-telling is satisfied. Note that the steps are same in the proof of previous lemma where I show that cross-sectional deviation is not profitable. $\square$

A.3 FAMILY INCOME ELASTICITY

Let $\varepsilon_m := \frac{\partial \log z_m}{\partial \log(1-\tau)}$ be the elasticity of male income with respect to net marginal tax rates. Similarly, let ε_f represents the female income elasticity. In this work, I focus on married households who file tax returns jointly. According to the US tax code, the last dollar earned by a family member is marginally taxed unconditionally. Therefore, the family income is the sum of earnings of partners, i.e. $z = z_m + z_f$, and the family income elasticity is:

$$\varepsilon := \frac{\partial \log z}{\partial \log(1-\tau)} = \frac{(1-\tau)}{z}\frac{\partial z}{\partial(1-\tau)} = \frac{(1-\tau)}{z_f + z_m}\frac{\partial(z_f + z_m)}{\partial(1-\tau)} = \frac{z_f}{z_f + z_m}\varepsilon_f + \frac{z_m}{z_f + z_m}\varepsilon_m.$$

This implies that the family income elasticity is a convex combination of individual income elasticities. To figure out family income elasticity, I need ε_f, ε_m, and the share of female earnings of family income. Note that the utility function is quasi-linear in consumption and hence elasticity of income with respect to net marginal tax rates is equal to the Frisch elasticity of labor supply. Therefore I look at the literature on Frisch elasticity.

There is a voluminous literature on elasticity of labor supply of males. Pencavel (1986) and Keane (2011) are excellent surveys of labor responses and taxes. According to the former work, the elasticity range is from zero to 0.5 and according to the latter, the range is

from zero to 0.7. Yet, both works state that the median value is 0.2 for the Frisch elasticity of males. The elasticity ranges are calculated by the estimates of many studies including the ones which are not estimated using the US data. Hence, I look particularly at French (2005) and Ziliak and Kniesner (2005) who use Panel Study of Income Dynamics (PSID) data. The former estimates the Frisch elasticity of men at 0.3 and the latter estimates around 0.5. I take the average value $\varepsilon_m = 0.4$ in my setup. This value is also in line with Blundell, Pistaferri, and Saporta-Eksten (2012) who finds an estimate of 0.4 for elasticity of male labor supply.

The literature on the Frisch elasticity of females is not as large as on male elasticities. Blundell, Pistaferri, and Saporta-Eksten (2012) in an exception. They find that the elasticity of married women lies between 0.8 to 1.1. When the household's utility is additive separable, the estimate is 0.8, and I pick $\varepsilon_f = 0.8$.

Next, I calculate $\frac{z_f}{z_f + z_m}$ by using the sample of the CPS data with restrictions stated in Section 4.2. Females earn around 40% of the family income (see Figure 11). Note that this ratio suggests that the gender gap for this sample is 0.67, which is quite close to the actual gender gap in the US (0.7).
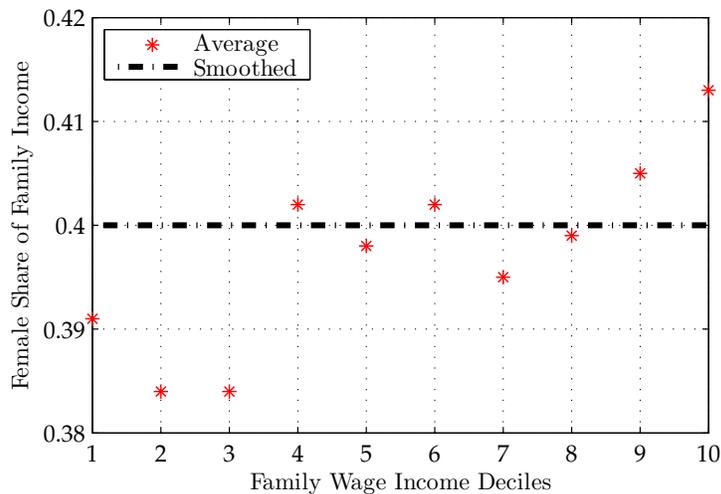

Figure 11: Female Share of Family Income

As a result, I use these values to calculate family income elasticity: $\varepsilon = 0.6 \times 0.4 + 0.4 \times 0.8 = 0.56$.

A.4 ADDITIONAL ROBUSTNESS ANALYSIS

This subsection provides additional robustness analysis for U-shaped tax credits. Figure 12a and 12b show the optimal tax credits for different values of elasticity of having second child with respect to her marginal cost, $\lambda = 0.35$ and $\lambda = 0.5$, respectively. The optimal child tax credits are U-shaped in income.

Figure 12c shows that the optimal tax credits are decreasing in income when the child-rearing requires only goods and the goods amount are set to $e_1 = \$11,000$ and $e_2 = \$19,000$. On the other hand, when child-rearing requires only 5% of parental working time for each child as Knowles (1999) suggests, the credits are increasing in income almost everywhere.

Next, when I relax the age restriction of spouses in families from 35-45 to 25-45, the optimal tax credits are still U-shaped (see Figure 12e).

Finally, I consider a case where families can have up to three children. Figure 12f shows that optimal tax credits are U-shaped in income and decreasing by family size.[45]

A.5 NUMERICAL SOLUTION

Hamiltonian requires calculations for inner integrals. I first theoretically calculate the following integral:

$$
\int \frac{(V - A\beta)^{1-\sigma}}{1 - \sigma} \lambda \beta^{-\lambda-1} d\beta = \frac{\beta^{-\lambda} V^{-\sigma}}{(1-\sigma)(1-\lambda)} \times
$$
$$
\left((\lambda - 1) V {}_2F_1(-\lambda, \sigma; 1 - \lambda; \frac{A\beta}{V}) - A\lambda\beta {}_2F_1(1 - \lambda, \sigma; 2 - \lambda; \frac{A\beta}{V}) \right)
$$

where ${}_2F_1$ is Gauss hypergeometric function. Matlab's command for hypergeometric functions is very slow. Instead, I use ${}_2F_1$'s definition for calculations:

$$
{}_2F_1(a, b; c; y) = 1 + \frac{ab}{1!c} y + \frac{a(a+1)bb(b+1)}{2!c(c+1)} y^2 + \ldots = \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{y^n}{n!}.
$$

I create a loop in which each step I add $\frac{(a)_n (b)_n}{(c)_n} \frac{y^n}{n!}$ to the value of integral. When an additional term is less than 10^{-15}, I stop the loop. This calculation is very accurate and quick that has been used many times to calculate social welfare in the government's problem.
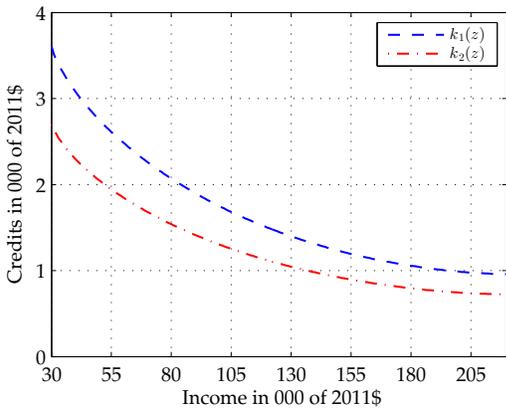
---

[45]I numerically solve the government's problem (PMDP), using the numerical solver DIDO version 7.3.7. For details on the solution algorithm, refer to Ross and Fahroo (2003).
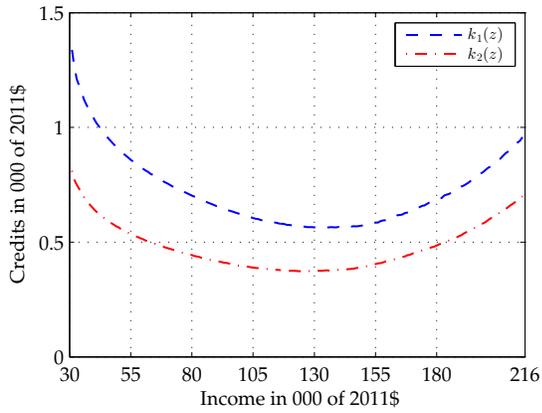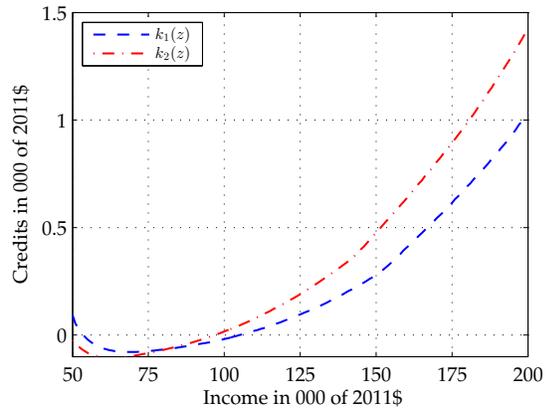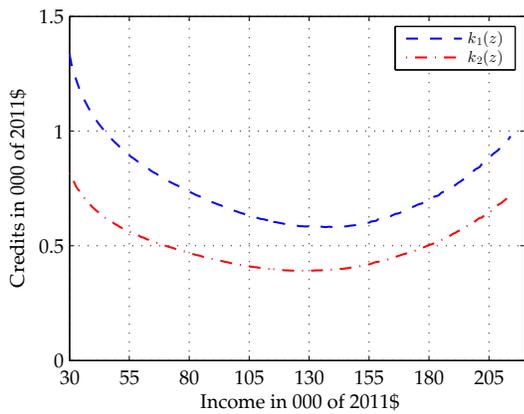
(a) Elasticity of Having Children
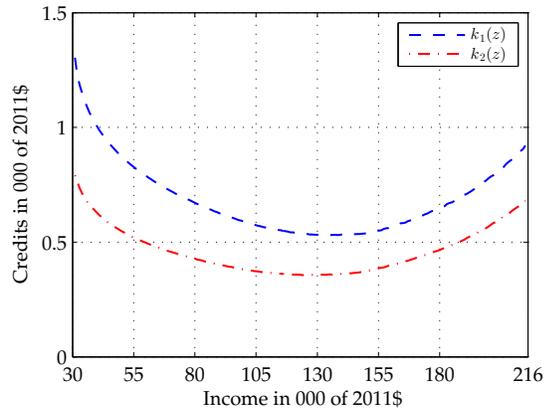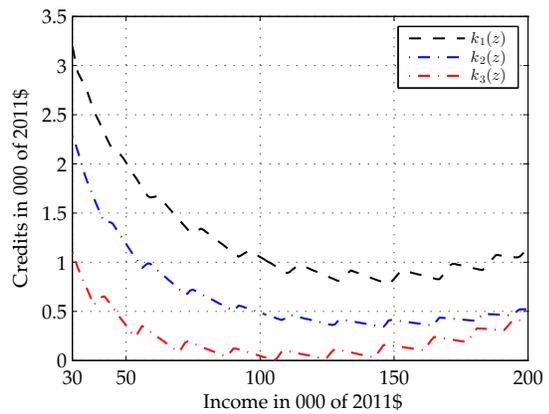
(b) Elasticity of Having Children

(c) Only Goods Cost

(d) Only Time Cost

(e) Age Restriction of Spouses

(f) $N = 3$ Children