Optimal Capital Income Taxation with Heterogeneous Firms

Rodrigo Cerda
Diego Saravia
OPTIMAL CAPITAL INCOME TAXATION WITH HETEROGENEOUS FIRMS

Rodrigo Cerda*
Diego Saravia**

Documento de Trabajo Nº 316

Santiago, Marzo 2007
Versión revisada a enero de 2008

*rcerda@faceapuc.cl
**dsaravia@faceapuc.cl
INDEX

ABSTRACT

1. INTRODUCTION 1

2. THE MODEL 4
   2.1 Firms 4
   2.2 The Household 7
   2.3 The Government 8
   2.4 Equilibrium 9

3. THE RAMSEY PROBLEM AND THE OPTIMAL TAXES 10
   3.1 Optimal Capital Income Taxation 11
   3.2 Optimal Labor Tax and Optimal Profits’ Tax 14
   3.3 Allowing Taxation on the Yields of the Alternative Option 17

4. NUMERICAL EXAMPLES 19

5. CONCLUSIONS 25

APPENDIX 27

REFERENCES 33
Abstract

We study steady state optimal taxation in a context where firms differ in productivity and they decide whether to produce or not after comparing after-tax profits vis-à-vis an outside alternative option. The government taxes capital income, firms' profits and labor income but does not tax the alternative outside option. In this context, taxation might distort the firms' decisions to participate in production (extensive margin) as well as their factor allocations once they decide to produce (intensive margin). We find that the government has incentives to subsidize costs to induce firms into production and tax them using the corporate tax to collect revenues. The optimal capital income tax is negative while the corporate tax rate is positive and the labor income tax is ambiguous.

JEL Classification: H21
1 Introduction

In this paper we study steady state optimal (Ramsey) taxation in a context where firms are heterogeneous in the sense that they differ in their productivity and decide whether to enter into production or not. This decision is taken after comparing the after-tax profits obtained from production vis-à-vis an outside alternative option. The government finances an exogenous expenditure path using three tax instruments: capital income tax, labor income tax and a tax on firms’ profits. Also, the government can issue debt to finance its expenditure but, crucially, we assume that the yields of the alternative option cannot be taxed. So, our context is one of incomplete taxation, where it is not possible to tax all sources of income.

In this context, taxation potentially distorts the firms’ decision to participate in production, i.e. the extensive margin distortion, as well as the factor allocations of the firms already involved in production, i.e. the intensive margin distortion. The optimal combination of tax instruments depends on these distortions. We show that when all firms are involved in production it is optimal to set capital income as well as labor income taxes equal to zero and raise taxes only through the tax on profits. In this case, the tax on profits does not affect either the extensive or the intensive margin, while the other taxes would distort the intensive margin. Consequently, the solution requires to set capital income tax equal to zero, as in Chamley (1986) and Judd (1985) celebrated result, and the tax on labor also equal to zero.

However, when there are firms that decide not to produce, the tax on profits distorts the extensive margin. In this case it is optimal to set a negative capital income tax. By subsidizing capital the social planner induces firms that are in the margin to enter into production and then taxes their profits. A similar intuition is obtained for labor income taxes, but the sign of this tax depends also on labor supply considerations.

In other words, the planner subsidizes costs in an effort to complete the tax system. Since the government cannot tax the outside option, subsidizing costs induce some firms into production and make them taxable. If all firms were already involved in production, it would be pointless to subsidize firms.

To make our point, we also consider the case where the outside option is taxable at the same rate as profits. In this case, the optimal capital and labor income would be zero because it is not necessary to create incentives for firms to participate in production to tax them since the outside
option is already taxable. In this case, taxing profits would not distort any margin and it is optimal to set the other taxes equal to zero.

The paper is related to several other studies in the literature. Chamley (1986) and Judd (1985) find that the optimal capital income taxation is equal to zero in steady state in a competitive environment. After these studies, several papers showed that optimal capital income taxation is different from zero if the context is modified in some ways.

Correia (1996) extends Chamley (1986)’s result to the case where not all factors can be taxed, that is to an environment of incomplete taxation. She finds that taxing capital would be an indirect way of taxing the untaxed factor and the sign of this tax would depend on the complementarity between capital and the untaxed factor. As Correia (1996)’s, our context is one of incomplete taxation since the government does not tax the alternative option but differs in that we analyze an environment where firms are heterogeneous. Firm heterogeneity and the distortions created in the extensive margin are key to our results. In our case, the planner has incentives to subsidize costs making possible the taxation of firms that would not otherwise be taxed. Thus, the optimal capital income taxation is negative.

In the same vein as Correia (1996), other studies find that capital taxes should not be zero as a consequence of incomplete taxation. Jones et al (1993) studied the issue in endogenous growth models showing that including government expenditures as a productive input leads to an optimal tax rate different from zero. The reason is similar to Correia’s explanation since government expenditure is not taxed. Jones et al (1997) also show that the zero income capital tax does no longer hold when there are pure profits generated. Their interpretation of this result is that taxing capital is a way of taxing pure profits in a setting where they cannot be taxed directly.

A second line of research related to our study are Judd (1997), Judd (2002) and Coto-Martinez et al (2007). Judd (1997) and Judd (2002) use a context of monopolistic competition and argue that the optimal capital income tax rate is negative and the tax on profits is positive. Coto-Martinez et al (2007) add entry and exit of firms to the context of Judd’s works where the entrance of new firms augment the general productivity of the economy but imply a waste of resources in the form of a fixed cost. Judd finds that the optimal capital income taxation is negative and the tax on profits is positive. In Coto-Martinez et al (2007) optimal taxes depend on the tax code available.

---

1Correia (1996) suggests that similar results would be obtained if firms present decreasing returns to scale and profits cannot be taxed.
When the available taxes are such that the government can control the number of firms through a tax on profits, it is optimal to subsidize capital to correct the markup distortion as in Judd and set a subsidy or a tax on profits depending on the aggregate returns to specialization. When the tax system does not allow to control the number of firms through profits taxation, they find that the optimal capital income taxation is zero if the returns to specialization are zero. The reason is that, in this case, it is not desirable to subsidize the entrance of new firms since there are only losses (fixed cost) associated with them.

We also find that it is optimal to subsidize capital but in a context of perfect competition (no markup distortions) and without aggregate returns to specialization. Our results also depend on the availability of taxes. In the general case, where the yields of the alternative option cannot be taxed, it is optimal to subsidize capital, and possibly labor, to induce firms that are in the margin into production making possible their taxation. As mentioned above if the alternative option could be taxed at the same rate as profits the labor and capital taxes are zero. Thus, in our case, the subsidy to capital is an effort of the planner to complete the tax system and not the result of distortions created by imperfect competition or the presence of returns to specialization. In this sense it is related to the work of Correia and others mentioned above. In the general case, our context is one of incomplete taxation and the planner has incentives to partially complete them via subsidies.

To the best of our knowledge, the papers most related to ours are the mentioned above. However, there are other papers that find an optimal capital income tax different from zero. Aiyagari (1995) made the point in an economy with borrowing constraints. The reason is that, in these economies, precautionary savings leads to too much capital in steady state. The optimality of taxing capital income was also obtained in OLG models. Recent works using this approach includes Abel (2005) and Erosa and Gervais (2002). Abel (2005) focuses on a context with consumption externalities between generations and shows that taxing capital is a way to correct the no “internalization” of cohorts’ consumption. Erosa and Gervais (2002) derives the optimality of capital taxation as a way of making taxes age-dependant. In a context of private information about agents skills, Golosov et al (2003) show that it is optimal to have a wedge between the marginal benefit and marginal cost of investing, which is consistent with a positive tax in capital income.

The rest of the paper is organized as follows. Section 2 describes the model. Section 3 sets up the
Ramsey problem; subsection 3.1 analyzes capital income taxation, subsection 3.2 sets the optimal labor income and profits tax and subsection 3.3 study the case when the yields of the alternative option is taxable. Section 4 presents numerical examples that confirms our theoretical findings and provides quantitatively impacts in the case of standard utility and production functions. Section 5 concludes the paper.

2 The Model

2.1 Firms

There is a set I of mass one of heterogeneous firms indexed by i that could operate and produce a single good. Firm heterogeneity comes from a productivity parameter, $A_{it}$, which is iid across time $t$ and is distributed across firms with cumulative distribution $G(A_{it})$ with support $[A_l, A_u]$, where $0 < A_l < A_u < \infty$. Let $k_{it}$ and $l_{it}$ be capital and labor used by firm $i$ in the production process in period $t$.

Capital is rented from the representative household each period at the rental rate $r_t$, which, in equilibrium, is the same for all firms. Firms pay an amount $w_t$ as compensation for the use of labor, which is also common to all firms in equilibrium. Both, $r_t$ and $w_t$ are expressed in terms of the consumption good. Each firm production function presents decreasing returns to scale and is given by $A_{it} f(k_{it}, l_{it})$, where $f(k_{it}, l_{it})$ is strictly increasing, strictly concave and satisfy Inada conditions on $k_{it}$ and $l_{it}$. We further assume that $f(k_{it}, l_{it})$ is homogeneous of degree $\theta < 1$ in $(k_{it}, l_{it})$. This last assumption is not important to the main results of the paper but simplifies the exposition.

Let $\phi > 0$ be an outside option, common to all firms, expressed in units of the single good concerning what each firm would get in an alternative project not considered explicitly in the paper.\footnote{This outside option could be interpreted in the same spirit as in Jovanovic (1982) who considered it as a “managerial ability” or “advantageous location” which is also common to all firms in that work.}

In each period of time, firm $i$ must decide between entering into the market to produce or not entering. To take this decision, firms compare the after-tax profits derived from production and the yields of the alternative option. We make the following assumption about the taxes that a firm faces:
Assumption 1. The government taxes firms’ profits at a rate \( \tau_t \) but cannot tax the parameter \( \phi \).

This assumption implies that the tax system is incomplete and it is very important to our results as it will become clear below; we analyze the consequences of dropping it in section 3.3. An implication of this assumption is that we could also interpret \( \phi \) as the return obtained by the firms in an informal sector.

Firms’ profits derived from production are given by the product obtained minus payments to capital and labor. The rental rate of capital and the wage rate are determined in competitive markets and are the same for all firms. Thus, firms deciding to produce must obtain after-tax profits that are at least equal to \( \phi \). Hence, a firm solves the following static problem in period \( t \):

\[
\max \{ \phi \ ; \ V_{it} \}
\]

where \[
V_{it} = \max_{k_{it},l_{it}} (1 - \tau_t)[A_{it}f(k_{it},l_{it}) - r_t k_{it} - w_t l_{it}] = \max_{k_{it},l_{it}} (1 - \tau_t)(1 - \theta)A_{it}f(k_{it},l_{it}),
\]

the last term holds because of the homogeneity of the production function.

Let \( V_{lt} \) and \( V_{ut} \) be the function \( V_{it} \) evaluated at the lowest and highest productivity shocks, \( A_{it} = A_l \) and \( A_{it} = A_u \), respectively.

Assumption 2. \( V_{ut} > \phi \).

This assumption assures entrance of a positive mass of firms into production. The solution to firms problem, equation (1), is stated in the next lemma.

Lemma 1. There exists a threshold technology level \( A^* \) such that firms endowed with technology \( A_{it} \geq A^*_t \) enter into production, while firms endowed with \( A_{it} < A^*_t \) do not enter into production. When \( V_{lt} \leq \phi \), the threshold \( A^*_t \) is interior and uniquely determined by:

\[
\max_{k_{it},l_{it}} (1 - \tau_t)(1 - \theta)A^*_t f(k_{it},l_{it}) = \phi.
\]

When \( V_{lt} > \phi \), the threshold \( A^*_t \) is equal to \( A_t \).

Proof. The function \( V_{it} \) in (1) is an increasing and continuous function of \( A_{it} \). Then, when \( V_{lt} \)

\[\text{By the envelope theorem, } \frac{\partial V_{it}}{\partial A_{it}} = (1 - \tau_t) f(k_{it},l_{it}).\]
is smaller than \( \phi \), there is a unique \( A_{it} \) that makes \( V_{it} \) equal to \( \phi \) given our assumption that \( V_{ut} \) is always higher than \( \phi \). If \( V_{it} \) were larger than \( \phi \), all type of firms would prefer to produce and the threshold \( A^*_t \) will be given by \( A_t \). □

Firms demand capital and labor if they participate in production, however these factors are not needed if the firm is not engaged in production and participate in the alternative option. Thus, factor demands are functions of factor prices and the idiosyncratic shock and are generically given by:

\[
k_{it} = k_{it}(A_{it}, r_t, w_t) \quad if \quad A^*_t \leq A_{it}, \]
\[
k_{it} = 0 \quad if \quad A^*_t > A_{it}
\]

and

\[
l_{it} = l_{it}(A_{it}, r_t, w_t) \quad if \quad A^*_t \leq A_{it} \]
\[
l_{it} = 0 \quad if \quad A^*_t > A_{it}.
\]

Markets are competitive, capital and labor are paid their marginal productivity, and the rental rate and wage rate is the same for all firms. Hence, the rental and wage rates for the economy are given by:

\[
r_t = \frac{\int_{A^*_t}^{A_{it}} A_{it} f_k(k_{it}, l_{it}) dG(A_{it})}{1 - G(A^*_t)}
\]

and

\[
w_t = \frac{\int_{A^*_t}^{A_{it}} A_{it} f_l(k_{it}, l_{it}) dG(A_{it})}{1 - G(A^*_t)}.
\]

where \( 1 - G(A^*_t) \) is the fraction of firms involved in production in period \( t \).

The capital and labor demands for the economy follow from the aggregation of individual factor demands by all the firms that decide to produce; that is, all the firms that get a productivity shock higher than \( A^*_t \). That is:

\[4\text{Each firm equates its marginal productivity of capital and labor to the interest rate and the wage rate respectively, that is, } r_t = A_{it} f_k(k_{it}, l_{it}) \text{ and } w_t = A_{it} f_l(k_{it}, l_{it}). \text{ Equations (5) and (6) follow by aggregation of these expressions among all firms that decide to produce.} \]
\[
K_t^D = \int_{A_t^a} A^u k_{it} dG(A_{it}) \tag{7}
\]
and
\[
L_t^D = \int_{A_t^a} A_{it} dG(A_{it}) \tag{8}
\]

2.2 The household

There is an infinitely lived representative household choosing a consumption path \(\{c_t\}_{t=0}^{\infty}\) and a leisure path \(\{h_t\}_{t=0}^{\infty}\) that maximizes:

\[
\sum_{t=0}^{\infty} \beta^t u(c_t, h_t) \tag{9}
\]

where \(u(.)\) is strictly increasing, strictly concave and three times continuously differentiable in both arguments. We assume also that \(u_{ch} \geq 0\). The household is endowed at time zero with an initial amount of capital \(K_0\) and holds the initial stock of government bonds \(b_0\). Each period, she decides how much to consume, how much to work, how much to invest in capital and government bonds to be held into next period. It is assumed that capital depreciates at rate \(\delta\) while total time available to work and to rest is \(H\). Capital, \(K_t\), is rented to firms in order to be used in the production process at the rental rate \(r_t\). Labor, \(L_t\), is also rented to firms at the rate \(w_t\).

The representative individual receives the after-tax profits that firms obtain in the production process, \(\int_{A_t^a} V_{it} dG(A_{it})\), or alternatively, the returns of the outside option if firms do not engage in production; which are, \(\phi G(A_t^a)\). The government taxes the rental rate at rate \(\tau_{kt}\), the wage rate at the rate \(\tau_{lt}\), firms’ profits at rate \(\tau_{ut}\) and issue one period bonds, which pays a gross interest rate of \(R_t\). Let \(b_t^d\) be the stock of bonds in hands of the representative household. Hence, each period the household faces the following budget constraint:

\[
c_t + i_t + \frac{b_{t+1}^d}{R_t} = \bar{r}_t K_t + \bar{w}_t L_t + \int_{A_t^a} V_{it} dG(A_{it}) + \phi G(A_t^a) + b_t^d \tag{10}
\]

\[
K_{t+1} = (1 - \delta) K_t + i_t \tag{11}
\]

\[
h_t + L_t = H \tag{12}
\]
where following Chamley (1986) we define the variables \( \tilde{r}_t \) and \( \tilde{w}_t \) as \( \tilde{r}_t \equiv r_t(1 - \tau^k_t) \) and \( \tilde{w}_t \equiv w_t(1 - \tau^l_t) \).

Note that the household’s problem does not include explicit expressions concerning uncertainty. In fact, uncertainty in our model arises in the firm sector. While each firm faces an idiosyncratic shock, there is no aggregate uncertainty in this economy as the productivity parameter has the same distribution each period.

The solution to the consumer’s problem yields the standard optimality conditions which include the marginal rate of substitution between present and future consumption (13), the intratemporal rate of substitution between leisure and consumption (14), the non-arbitrage condition (15) and the transversality conditions for capital and government bonds.

\[
\begin{align*}
U_c(t) &= \beta U_c(t+1)(1 + \tilde{r}_{t+1} - \delta) \quad (13) \\
U_h(t) &= U_c(t)\tilde{w}_t \quad (14) \\
R_t &= 1 + \tilde{r}_{t+1} - \delta \quad (15)
\end{align*}
\]

2.3 The government

As is usual in the optimal taxation literature the government collects taxes to finance an exogenous expenditure path \( \{g_t\}_{t=0}^\infty \). We assume that the government expenditure is wasteful; that is it does not provide any utility to the consumer. As noted above, the government finances its expenditure by issuing bonds and levying flat-rate, time-varying taxes on capital income, on labor income and on firms’ profits. To avoid the possibility that the government raises all revenues by taxing initial capital heavily not distorting the economy allocations, we make the standard assumption that the government takes the tax rate on capital income in the first period, \( \tau^k_0 \), as given. We also assume that the government can commit itself to a given policy so we do not analyze commitment issues. Further,

**Assumption 3.** We assume that \( \tau^u < 1 \).

We consider the case where \( \tau^u < 1 \) because when \( \tau^u = 1 \) there would be no firms producing,
making impossible to collect revenues to finance fiscal expenditure.\textsuperscript{5}

Hence the period-\(t\) government’s budget constraint is:

\[
\tau_l^t w_t \int_{A_t^u} A_t f(k_{it}, l_{it}) dG(A_{it}) + \tau_r^t r_t \int_{A_t^u} k_{it} dG(A_{it}) + \tau_u^t \int_{A_t^u} [A_{it} f(k_{it}, l_{it}) - r_{it} k_{it} - w_{it} l_{it}] dG(A_{it}) + \frac{b_{t+1}^u}{R_t} - b_t^u \geq g_t \quad \forall t
\]

The first term on the left hand side is the amount of taxes on labor income, the second is the amount raised from capital income while the third term corresponds to the taxes raised from firms’ profits. Since the firms’ production functions are homogeneous of degree \(\theta\) and using the definitions of \((\tilde{w}_t, \tilde{r}_t)\), the government’s budget constraint can be written as (see appendix):

\[
\tau_l^u (1 - \theta) \int_{A_t^u} A_{it} f(k_{it}, l_{it}) dG(A_{it}) + \theta \int_{A_t^u} A_{it} f(k_{it}, l_{it}) dG(A_{it}) + \frac{b_{t+1}^u}{R_t} - b_t^u \geq g_t \quad \forall t
\]

2.4 Equilibrium

Given the description of our economy, we may state the following definition of equilibrium.

**Definition 1.** A competitive equilibrium is a sequence of allocations \(\{c_t, k_{it}, l_{it}, b_t\}\)\(\infty_{i=0}^{\infty}\), a sequence of prices \(\{r_t, w_t, R_t\}\)\(\infty_{t=0}^{\infty}\), a government policy \(\{\tau^k_t, \tau^l_t, \tau^u_t, g_t\}\)\(\infty_{t=0}^{\infty}\) and a sequence of threshold technology levels \(\{A^*_t\}\)\(\infty_{t=0}^{\infty}\) such that:

1. the household maximizes (9) subject to (10), (11) and (12) taking as given \(K_0\) and \(b_0\),

2. each firm solves (1) conditional on \(A_{it}\),

3. the sequence of threshold technology levels is determined by:

\[
(1 - \tau^u_t) (1 - \theta) A_t^* f(k_t^*, l_t^*) \geq \phi \quad \forall t,
\]

where \(k_t^*, l_t^*\) are the optimal capital stock and labor demanded by the firm endowed with the threshold technology level,

\(\text{In fact, when } \tau^u = 1 \text{ we have } V_{it} = 0 < \phi, \quad \forall i\)
4. the government satisfies (16),

5. the capital market clears, i.e.

\[ K_t = \int_{A_t^u}^{A_t^d} k_{it} dG(A_{it}) \quad \forall t, \quad (17) \]

6. the labor market clears, i.e.

\[ \bar{H} - h_t = L_t = \int_{A_t^u}^{A_t^d} l_{it} dG(A_{it}) \quad \forall t, \quad (18) \]

7. the bonds market clears

\[ b_t^s = b_t^d \quad \forall t, \quad (19) \]

8. the good market clears

\[ c_t + g_t + K_{t+1} = \int_{A_t^u}^{A_t^d} A_{it} f(k_{it}^d) dG(A_{it}) + \phi G(A_t^*) + (1 - \delta)K_t \quad \forall t \quad (20) \]

3 The Ramsey Problem and the Optimal Taxes

Our goal is to characterize the tax rates that are consistent with the allocations in a second best steady state, assuming that the economy converges to this steady state in the long run. As is standard in the literature, the social planner will choose among the set of competitive equilibria available the one that maximizes the representative individual utility. The planner chooses the allocations, tax rates and threshold technologies subject to good market clearing, consumer budget constraints, government budget constraints and individual’s and firms’ optimality conditions. Therefore, the planner solves the following problem:
\[ L = \max_{\{c_t, \tau^u_t, k_{it}, l_{it}, b_t, \tilde{w}_t, \tilde{r}_t, \tilde{A}_i^*, \tilde{A}_i^+\}} \sum_{t=0}^{\infty} \beta^t \left\{ u \left( c_t, H - \int_{A_i^*}^{A_u} l_{it} dG(A_{it}) \right) \right. \]

\[ + \lambda_1 \left[ \int_{A_i^*}^{A_u} A_{it} f(k_{it}, l_{it}) dG(A_{it}) + \phi G(A_{it}^*) + (1 - \delta) \int_{A_i^*}^{A_u} k_{it} dG(A_{it}) - c_t - g_t - \int_{A_i^+}^{A_{it+1}} k_{it+1} dG(A_{it+1}) \right] + \]

\[ + \lambda_2 \left[ \tau^u_t \int_{A_i^*}^{A_u} (1 - \theta) A_{it} f(k_{it}, l_{it}) dG(A_{it}) + \int_{A_i^*}^{A_u} \theta A_{it} f(k_{it}, l_{it}) dG(A_{it}) \right] - \tilde{r}_t \int_{A_i^*}^{A_u} k_{it} dG(A_{it}) - \tilde{w}_t \int_{A_i^*}^{A_u} l_{it} dG(A_{it}) + \frac{b_{t+1}}{1 + \tilde{r}_{t+1} - \delta} - b_t - g_t \left. \right] + \]

\[ + \lambda_3 \left[ u_c(t) - \beta u_c(t + 1)(1 + \tilde{r}_{t+1} - \delta) \right] + \lambda_4 \left[ u_h(t) - u_c(t) \tilde{w}_t \right] + \lambda_5 \left[ (1 - \tau^u_t)(1 - \theta) A_{it}^* f(k_{it}^*, l_{it}^*) - \phi \right] \}

(21)

Note that the above Ramsey problem is written as in the “dual approach”, similar to many papers in the literature. In the problem, we followed Chamley (1986) by including \( \tilde{r}_t \) and \( \tilde{w}_t \). Note that these expressions do not represent prices but replace the capital income tax and the labor income tax, respectively.

The first constraint in this problem is the good market clearing, (20). The second one is the government budget constraint (16) taking into account the non-arbitrage condition (15), while the third one is the intertemporal consumption Euler equation (13). The fourth restriction is the intratemporal marginal rate of substitution between consumption and leisure (14). The last restriction indicates that the marginal firm (i.e. the less productive one that decides to operate) must earn after-tax profits at least as large as the outside option in the alternative activity.\(^6\) For expositional simplicity we post in the appendix the optimal conditions of the Ramsey problem.

### 3.1 Optimal capital income taxation

We will next state the planner’s optimal condition concerning capital stock in the \( i^{th} \) firm, \( k_{it} \). The optimality condition evaluated in steady state is:\(^7\)

\(^6\)We do not post the consumer budget constraint because it is redundant by Walras law.

\(^7\)Follows from equation (53) in the appendix, dividing both sides by \( \beta^t \) and taking out time indexes (since we analyze steady state).
\[
\lambda_1 \left[ A_i f_k(k_i, l_i) + (1 - \delta) - \frac{1}{\beta} \right] g(A_i) + \lambda_2 \left[ \tau^u (1 - \theta) A_i f_k(k_i, l_i) + \theta A_i f_k(k_i, l_i) - \tilde{r} \right] g(A_i)
\]
\[
+ \lambda_5 (1 - \tau^u) (1 - \theta) A^* f_k(k^*, l^*) \mathbf{1} [A_i = A^*] = 0,
\] (22)

where \( \mathbf{1} [A_i = A^*] \) equals to one if the firm \( i \) is the marginal one and is zero otherwise.

The first term of this optimality condition indicates the marginal social value of the increase in output derived from the marginal increase in capital by firm \( i \) net of investment cost. The second term is the social valuation of the increase in tax revenues derived from the increase in capital; while the last term indicates that if the firm is the marginal one, there is an additional consideration that measures its incentive to enter into production.

Similarly, the first order condition with respect to \( \tau^u \), evaluated in steady state, yields the following expression:

\[
\lambda_2 \int_{A^*}^{A_u} A_i (1 - \theta) f(k_i, l_i) dG(A^i) = \lambda_5 A^* (1 - \theta) f(k^*, l^*). \quad (23)
\]

This expression highlights the extensive margin distortion produced by a change in \( \tau^u \). It balances the marginal social cost of raising \( \tau^u \), which is given by the social value of displacing the marginal firm out of production –the term on the right hand side containing \( \lambda_5 \)–, with the social value of raising government revenues through the increase in this tax rate.

Integrating (22) with respect to all the firms involved in production and using (23), we obtain the following expression for \( \tau^k \) (see appendix):

\[
\tau^k = \left[ (1 - \theta)(1 - \tau^u) \right] \left( \frac{\lambda_5}{\frac{\lambda_1}{\lambda_1 + \lambda_2}} \right) [S_Y - 1], \quad (24)
\]

where \( S_Y \equiv \frac{(1 - \theta) A^* f(k^*, l^*)}{(1 - \theta) \int_{A^*}^{A_u} A_i f(k_i, l_i) \frac{dG(A^i)}{dG(A^*)}} \) is the share of production (profits) of the marginal firm in total production (profits). Note that \( S_Y \) is positive but less than one since the marginal firm has a lower production (profits) than the rest of the firms involved in production.

Equation (24) is not a reduced form expression for \( \tau^k \). In fact, it depends on other endogenous
variables such as $A^*, \frac{\lambda_5}{\lambda_1 + \lambda_2}$ and $S_Y$. However, it allows us to obtain some intuition about the sign of $\tau^k$. Firstly, note that $\tau^k \leq 0$. This result holds because (1) $\tau^u < 1$ -if no firm would be involved in production-, (2) $\lambda_1, \lambda_2$ and $\lambda_5$ are non negative and (3) $S_Y < 1$.

The direct implication of setting $\tau^k$ less than zero is that the steady state rental rate faced by firms, $r$, is depressed.\(^8\) This is in fact a capital income subsidy that provides incentives to firms that are not producing, but in the neighborhood of doing so, to enter into production and allows the government to obtain revenue from them using the tax on profits. This would be the case if $\tau^u > 0$, a result that will be shown below.

Secondly, note that $\tau^k$ is zero when all firms are involved in production.\(^9\) In this case $\lambda_5 = 0$ by complementary slackness. We can relate these results to the incompleteness of the tax system. If some firms are not involved in production, it is impossible to obtain fiscal revenues from their profits and thus the tax system is incomplete. The planner reacts by setting a subsidy to production costs, throughout a lower interest rate faced by firms, inducing some of them into production and completing the tax system, at least partially. In the case that all firms are producing, the tax system is complete and there is no further need to subsidize capital to induce firms into production.

Note that when there is no heterogeneity between firms, i.e. $S_Y = 1$, the optimal capital tax rate is zero. However, this situation is considered in the above discussion since in this case all firms would be producing; if all firms were in the alternative option there would not be taxation and the analysis loses relevance.

We can summarize the findings of this section in the following proposition:

**Proposition 1.** It is optimal to set $\tau^k$ less than zero if and only if not all firms are producing. In the case that all firms are involved in production, the optimal $\tau^k$ is zero.

**Proof.** See the discussion above.

\(^8\)By the consumer’s Euler condition, equation (13), in a steady state, we have:

$$1 = \beta(1 + r(1 - \tau^k) - \delta).$$

\(^9\)An example where we would have this situation is when $\phi = 0$ and $A_l > 0$. 

13
3.2 Optimal labor tax and optimal profits’ tax

In this section, we focus on the labor and profits’ tax chosen by the planner. As in the case of the expression concerning $\tau^k$, the expressions we obtain next are not reduced form solutions for those taxes as they will depend on other endogenous variables. However, as above, we will be able to obtain the signs and intuition about the economic determinants involved.

The optimal labor tax is obtained as follows. Similarly to the optimality condition of capital stock, equation (22), we may obtain the optimality condition with respect to the allocation of labor in the $i^{th}$ firm; which evaluated in steady state yields:

$$
\left[ -u_h + \lambda_3 u_{ch}(\bar{r} - \delta) - \lambda_4 u_{hh} + \lambda_4 u_{ch} \bar{w} \right] g(A_i) \\
+ \lambda_1 A_i f_l(k_i, l_i) g(A_i) + \lambda_2 \left[ \tau^u (1 - \theta) A_i f_l(k_i, l_i) + \theta A_i f_l(k_i, l_i) - \bar{w} \right] g(A_i) \\
+ \lambda_5 (1 - \tau^u)(1 - \theta) A_i f_l(k_i^*, l_i^*) 1[A_i = A^*] = 0,
$$

(25)

where, as before, $1[A_i = A^*]$ takes the value of one if the firm $i$ is the marginal one and zero otherwise.

The social planner balances the social benefit of increasing output through a marginal increase in labor (the term involving $\lambda_1$), the social value of increasing tax revenues (the term involving $\lambda_2$) and the social value of the change in marginal firms’ profits (the term involving $\lambda_5$) with the marginal social costs of increasing labor that is given by the direct effect in the utility of the consumer and the effects in the marginal rates of substitution (first term of (25)). Similar effects were present in the derivation of (22), with the exception of the last one.

Integrating this expression for all the firms involved in production and using the first order conditions with respect to $c_t$, $\bar{w}_t$ and $\bar{r}_t$ in equations (55) to (57) in the appendix, we obtain (see the appendix):

$$
\tau^I \left[ \lambda_2 (1 + \sigma_{hh} + \sigma_{ch}) + u_c \right] = \lambda_2 \left[ \frac{(1 - \theta)(1 - \tau^u)}{S_Y} (S_Y - 1) + (\sigma_{hh} + \sigma_{cc} + \sigma_{ch}) \right]
$$

(26)

\footnote{Follows from dividing both sides of equation (54) in the appendix by $\beta^t$ and dropping the time indexes since we are in steady state.}
where $\sigma_{cc} = -\frac{u_{cc}}{u_c}$, $\sigma_{ch} = \frac{u_{ch}}{u_h}$, $\sigma_{hh} = -\frac{u_{hh}(1-h)}{u_h}$ and $\tilde{c} = \left[ \tilde{r} \left( \int_{A_i}^{A_u} k_i dG(A_i) + b_{ss}/\beta \right) + \tilde{w} \int_{A}^{A_u} l_i dG(A_i) \right]$

is total individual’s income, excluding firm’s transfers, in steady state.\(^{11}\) Note that $\sigma_{cc}, \sigma_{ch}, \sigma_{hh} > 0$.

The term in parenthesis that multiplies $\tau^l$ is positive, so the sign of this tax depends on the sign of the right hand side of (26). The first term on the right hand side has similar components to the expression obtained for $\tau^k$ in (24) and it is negative. The second term is the sum of $\sigma_{cc}, \sigma_{ch}$ and $\sigma_{hh}$, which is related to the concavity of the utility function and is positive.

Equation (26) shows that there are two forces involved in the determination of the sign of the optimal labor tax rate. On the one hand, and similar to the case of $\tau^k$, there is an incentive to subsidize firms’ production costs (the first part of the right hand side on 26) to induce firms into production, which allows to collect fiscal revenue from these additional firms by using the corporate tax. On the other hand, there is a second term that considers the impact on the individual’s utility of distorting leisure that shows that the more concave is the utility function, the more likely the optimal labor tax to be positive.

Further, note that if the whole set of firms is involved in production $\lambda_5 = 0$ and, from equation (23), $\lambda_2 = 0$. It follows from equation (26) that the optimal $\tau^l$ is zero, as in the case of capital taxation. The intuition is that as the whole set of firms is already involved in production, the planner’s incentives to subsidize firms’ costs disappear. Fiscal revenue will be obtained from firms’ profits, as it will be shown below. As a result, the optimal policy is to set $\tau^l = 0$ to avoid distortions in the marginal rates of substitution.

We next focus in obtaining the optimal profits’ tax $\tau^u$. We will initially analyze the case in which there are firms involved in production while others obtain $\phi$ in the alternative outside option. In this case $A^*$ is interior, i.e. $A_l < A^* < A_u$ and $\lambda_5 > 0$.

To obtain $\tau^u$ note that using equations (25) and the first order condition with respect to $A^*$ (equation (58) in the appendix) yields (see the appendix):

\(^{11}\) Also, note that

$$\beta b_{ss} = \left( 1 + \tilde{r} - \delta \right) b_{ss}$$

is the gross return on bonds in steady state and

$$\left( 1 - \beta \right) b_{ss} = \frac{\tilde{r} - \delta}{1 + \tilde{r} - \delta} b_{ss} = \frac{R - 1}{R} b_{ss}$$

is the bonds’ interest payment expressed in units of this period.
\[
\frac{\tau^u}{1 - \tau^u} = \frac{\lambda_5}{\lambda_1 + \lambda_2} \left[ \frac{\theta(1 - S_Y)}{1 - G(A^*)} + \frac{1}{A^*g(A^*)} \right]
\]  

(27)

Again, this expression is not a closed form solution since it depends on other endogenous variables. However, it is enough to determine the sign of \(\tau^u\) which is positive since the right hand side of (27) is positive for the reasons mentioned above.

Let’s analyze next the case in which \(A^*\) is not interior, i.e. \(A_l = A^*\). In this case equation (27) cannot be applied in the analysis since it was obtained using the first order condition with respect to \(A^*\), equation (58) in the appendix, that is no longer valid in the case that \(A^* = A_l\). To obtain \(\tau^u\) in this corner case note that we must satisfy the government budget constraint which in steady state is:

\[
\tau^u \int_{A_l}^{A_u} (1 - \theta)A_i f(k_i, l_i) dG(A_i) + \frac{b_{ss}}{1 + \bar{r} - \delta} - b_{ss} = g
\]

where \(b_{ss}\) are the level of government bonds in steady state. It follows:

\[
\tau^u = \frac{g + (1 - \beta)b_{ss}}{(1 - \theta) \int_{A_l}^{A_u} A_i f(k_i, l_i) dG(A_i)}
\]  

(28)

Note that in the case where all firms are involved in production, the sign of \(\tau^u\) depends on fiscal expenditure, \(g\), plus bond interest payments in steady state, \((1 - \beta)b_{ss}\). We can summarize the preceding discussion in the following proposition.

**Proposition 2.** In the case that some firms are not involved in production, the optimal tax on profits is positive while the sign of the optimal labor tax remains ambiguous. However, in the case that all firms are involved in production, the optimal tax on profits is different from zero while the labor tax is zero.

\(^{12}\)When \(\tau^k = \tau^l = 0\).
3.3 Allowing Taxation on the Yields of the Alternative Option.

We have analyzed a set up where the planner faces a problem in which (1) there are heterogeneous firms and (2) there is an alternative outside option to firms which return is not taxable. In this setup, we have shown that \( \tau^k \leq 0 \) and \( \tau^u > 0 \) while the sign of \( \tau^l \) is ambiguous. We have also shown that in the case that all firms choose to be involved in production, the optimal taxes are \( \tau^k = \tau^l = 0 \) and \( \tau^u \neq 0 \).

We will argue next that these results depend crucially in the absence of taxation of the outside option at the same rate as profits derived from production. To understand the importance of this assumption, we will allow next for taxation of \( \phi \) at the same rate that is applied to profits obtained in production, \( \tau^u \). In that case, our Ramsey problem would be modified as follows:

\[
L = \max_{\{c_t, r_t, k_t, l_t, b_t, \tilde{w}_t, \tilde{r}_t, A_{it}^*\}} \sum_{t=0}^{\infty} \beta^t \left\{ u\left( c_t, H - \int A_{it} f(k_{it}, l_{it})dG(A_{it}) \right) \right. \\
\quad + \mu_1^1 \left[ \int_{A_{it}^*} A_{it} f(k_{it}, l_{it})dG(A_{it}) + \phi G(A_{it}^*) + (1 - \delta) \int_{A_{it}^*} k_{it} dG(A_{it}) - c_t - g_t - \int_{A_{it}^*} k_{it+1} dG(A_{it+1}) \right] + \\
\quad + \mu_1^2 \left[ \tau_t^u \phi G(A_{it}^*) + \tau_t^u \int_{A_{it}^*} (1 - \theta) A_{it} f(k_{it}, l_{it})dG(A_{it}) + \int_{A_{it}^*} \theta A_{it} f(k_{it}, l_{it})dG(A_{it}) \right] \\
\quad - \tilde{r}_t \int A_{it} k_{it} dG(A_{it}) - \tilde{w}_t \int A_{it} l_{it} dG(A_{it}) + \frac{b_{t+1}}{1 + \tilde{r}_{t+1} - \delta} b_t - g_t \right\} \\
\left. \quad + \mu_3^3 \left[ u_c(t) - \beta u_c(t+1) (1 + \tilde{r}_{t+1} - \delta) \right] + \mu_4^4 [u_h(t) - u_c(t) \tilde{w}_t] + \mu_5^5 (1 - \tau_t^u) [(1 - \theta) A_{it}^* f(k_{it}^*, l_{it}^*) - \phi] \right\} \\
\quad (29)
\]

Problem (29) differs from problem (21) in two ways. First, the government budget constraint includes the taxation of the yields from the outside option, \( \tau_t^u \phi G(A_{it}^*) \), and second, the marginal firm’s entry decision differs as a firm obtains a return \( (1 - \tau_t^u) \phi \) if it chooses not to participate in production. We will next obtain the optimal taxes in this case. The next proposition states the results.

**Proposition 3.** If we allow taxation of the outside option at the rate \( \tau^u \), the optimal tax rates are \( \tau^k = \tau^l = 0 \) while \( \tau^u \neq 0 \).
Proof. We focus initially in the case of interior solution in $A_t^*$. The first order condition with respect to $\tau^u$, evaluated in steady state is:\(^{13}\)

$$\mu^2 \left[ \phi G(A_t^*) + \int_{A_t^*}^1 (1 - \theta) A_{it} f(k_{it}, l_{it}) dG(A_{it}) \right] = \mu^5 \left[(1 - \theta) A_t^* f(k_t^*, l_t^*) - \phi \right]$$  (30)

Since in an interior solution $(1 - \theta) A_t^* f(k_t^*, l_t^*) = \phi$ it follows that $\mu_2 = 0$. Using this condition, and the optimality condition with respect to capital stock, we obtain:

$$\tau^k = - \frac{(1 - \theta)(1 - \tau^u) \mu_5}{1 - G(A_t^*)} \mu_1 \leq 0$$  (31)

On the other hand, using the optimality condition with respect to $A_t^*$, the optimality condition on labor and the result concerning $\mu_2$, we get:

$$\tau^k = \frac{(1 - \theta)(1 - \tau^u) \mu_5}{r k^*} \mu_1 \left[ \frac{w l^*}{1 - G(A_t^*)} + \frac{A_t^* f(k^*, l^*)}{A_t^* g(A_t^*)} \right] \geq 0$$  (32)

Note that (31) implies $\tau^k \leq 0$ while (32) implies $\tau^k \geq 0$. It follows that $\tau^k = 0$. Further, the optimality condition with respect to to labor implies:

$$\tau^l (u_c + \mu_2(1 + \sigma_{hh} + \sigma_{ch})) = \frac{(1 - \theta)(1 - \tau_u)}{S_Y} [S_Y - 1] + \mu_2 [\sigma_{hh} + \sigma_{cc} + \sigma_{ch}]$$  (33)

But since $\mu_2 = 0$, it follows that (33) implies $\tau^l = 0$. Finally, to satisfy the government budget constraint we require:

$$\tau^u = \frac{g + (1 - \beta) b_{ss}}{(1 - \theta) \int_{A_t^*}^1 A_{it} f(k_{it}, l_{it}) dG(A_{it})}$$  (34)

where $b_{ss}$ are the level of government bonds in steady state.

---

\(^{13}\)We do not post in the appendix the derivations since they are similar to the ones obtained in the case where $\phi$ is not taxable.
We will now establish the results in the case of no interior solution, i.e. $A^* = A_1$. In this case $\mu_5 = 0$ and by the optimality condition on $\tau^u$, we have $\mu_2 = 0$. Trivially by (32) and (33), we obtain $\tau^k = \tau^l = 0$ and by (34), we get $\tau^u \neq 0$. □

Proposition 3 illustrates the importance of the impossibility of taxing firms’ outside option at the same rate as profits in our results: if we allow taxation of the proceeds of the alternative activity at the same rate as profits, capital income and labor taxes would be zero. In this case the tax on profits, $\tau^u$, is non distortive since it does not affect either the intensive or the extensive margins because in both sectors there is a common tax rate.\(^{14}\)

In the case that it is not possible to tax $\phi$ and not all firms are involved in production, the corporate tax rate distorts the extensive margin and we obtain capital income and possibly labor taxes less than zero. However, if all firms are involved in production, the tax on profits is non distortive and capital and labor taxes are zero.

These results indicate that the planner would use profit taxation as the only source of tax income when it is non distortive. However, when the tax on profits is distortive because there are firms in a non-taxable outside option, the planner has incentives to subsidize costs inducing firms that are in the margin into production. In this way, the planner completes, at least partially, the tax system.

4 Numerical Examples

We now use numerical methods to simulate calibrated versions of the model we have analyzed. We will use these results to confirm the validity of our analytical expressions and obtain additional light on our results. We use the following standard functional form for the utility function:

$$U = \frac{(c_t^{\Phi^2}(H - l_t)^{1-\Phi^2})^{\Phi^3}}{\Phi^3}$$  \hspace{1cm} (35)

In another hand, the production function of a firm operating:

\(^{14}\)Note that if the tax on the alternative option, call it $\tau^\phi$, were different from the tax on profits a change in this tax would affect the extensive margin. However, our analysis of non taxation of the alternative option holds if we redefine the tax on profits as $1 - \tau^\phi = \frac{1 - \tau^v}{1 - \tau^\phi}$.\)
General Method

In our simulations, we consider the long-run steady state of the Ramsey equilibrium. To obtain our results, we follow the procedure developed in Schmidt-Gohré and Uribe (2006). In the notation that follows, we eliminate the time index since we analyze the steady state. Let $\mathcal{F}(x, \Gamma)$ be the first order condition of the Ramsey problem defined in (21), where $x$ are the variables and $\Gamma$ are the parameters of the problem. Optimality requires $\mathcal{F}(x, \Gamma) = 0$. Our goal is to obtain $x_{ss}$ -where ss indicates steady state- such that $\mathcal{F}(x_{ss}, \Gamma) = 0$. To obtain the solution, we use the symbolic Matlab toolbox and we follow the next algorithm:

1. Guess an initial candidate vector $x_{ss}^j$ and choose criteria $\delta > 0, \nu > 0$, where $j$ indicates iteration.

2. Compute the direction $s^j$ to modify the initial candidate vector $x_{ss}^j$. The direction is chosen as in the steepest decent method,\textsuperscript{15}

$$s^j = -\nabla \mathcal{F}(x_{ss}^j, \Gamma)'$$

where $\nabla \mathcal{F}(x_{ss}^j, \Gamma)$ is the Jacobian of $\mathcal{F}(x, \Gamma)$ evaluated at $x_{ss}^j$ and $'$ indicates transpose.

3. Solve for the line-step criterion, $\lambda_j$, as in $\lambda_j = \arg \min_{\lambda} \mathcal{F}(x_{ss}^j + \lambda s^j, \Gamma)$.

4. Compute the update $x_{ss}^{j+1}$ as in $x_{ss}^{j+1} = x_{ss}^j + \lambda_j s^j$.

5. If $\|x_{ss}^j - x_{ss}^{j+1}\| < \nu (1 + \|x_{ss}^j\|)$ continue to next step, otherwise go back to step 2.

6. If $\|\nabla \mathcal{F}(x_{ss}^{j+1}, \Gamma)\| < \delta (1 + \|\mathcal{F}(x_{ss}^{j+1}, \Gamma)\|)$ stop and report success, otherwise report failure.

\textsuperscript{15}We alternatively used the Broyden-Fletcher-Goldfarb-Shannon method but we found no significative differences in our results.
**Calculation of the guess \( x_{ss}^j \)**

To implement the numerical procedure, we require a candidate steady state vector in the \( j^{th} \) iteration, \( x_{ss}^j \), which is calculated as follows. We discretize the number of firms and we include as initial guesses tax rates plus labor supply, i.e. \((\tau_{ss}^u, \tau_{ss}^l, \tau_{ss}^k, l_{ss}^j)\). The rest of the variables evaluated in the steady state are obtained by using the following algorithm in iteration \( j \). Using the consumer’s intertemporal optimality condition evaluated in steady state we get:

\[
\bar{r}_{ss}^j = \frac{1}{\beta} - (1 - \delta) \tag{37}
\]

We next obtain labor demand by each firm. Labor market clearing condition requires:

\[
l_{ss}^j = \sum_{i=1}^{N} l_i^j \tag{38}
\]

where \( N \) is the discrete number of total firms in the economy. Further, note that each firm’s labor demand -if it decides to operate- is:\(^{16}\):

\[
l_i^j = \left[ \frac{A_i \alpha \theta}{r} \right]^{\frac{1}{1-\theta}} \left( \frac{k_i}{l_i} \right)^{\frac{\alpha - 1}{1 - \theta}} = \left[ \frac{A_i \alpha \theta}{r} \right]^{\frac{1}{1-\theta}} \left( \frac{\alpha}{1 - \alpha \cdot \frac{w}{r}} \right)^{\frac{\alpha - 1}{1 - \theta}} \tag{42}
\]

Note that equations (38) and (42) provide labor demand as function of the labor supply guess and productivity parameters:

\(^{16}\)In fact, each firm’s optimality conditions are

\[
A_i \alpha \theta k_i^{\alpha - 1} l_i^{(1 - \alpha) \theta} = r \tag{39}
\]

\[
A_i (1 - \alpha) \theta k_i^{\alpha - 1} l_i^{(1 - \alpha) \theta - 1} = w \tag{40}
\]

which imply:

\[
\frac{k_i}{l_i} = \frac{\alpha \cdot w}{1 - \alpha \cdot \frac{w}{r}} \tag{41}
\]

Replacing equation (41) in (39) is (42).
\[
l_{i,ss}^j = \left[ \frac{A_i^{1-\theta}}{\sum_{z=1}^{N} A_z^{1-\theta}} \right] l_{ss}^j \quad i = 1, \ldots, N \tag{43}
\]

Similarly capital demand, if the firm operates, is:

\[
k_{i,ss}^j = \frac{\tilde{r}_{i,ss}^{j}}{1-r_{i,ss}^{j}} \frac{1}{(A_i \theta \alpha(l_{i,ss}^j)^{\theta-1})^{1-\alpha}} \tag{44}
\]

We define next an indicator function equal to one when the firm decides to operate by using:

\[
i_i(operation)^j = 1 - \frac{\left( \arctan \left( \frac{-(1-\tau^u) A_i (k_{i,ss}^j l_{i,ss}^j)^{\theta-1} + \phi}{\epsilon} \right) + \pi/2 \right)}{\pi} \tag{45}
\]

where \(i_i(operation)^j\) is the indicator function which is equal to one if the \(i^{th}\) firm operates and zero otherwise, while \(\epsilon > 0\) is a parameter. Note that in the above indicator function \(k_{i,ss}^j\) and \(l_{i,ss}^j\) represent capital and labor demand if the \(i^{th}\) firm operates. These demands are defined in (43) and (44). The parameter \(\epsilon\) determines shape of the function, the smaller is \(\epsilon\) the less smooth is the function. An example of this function is shown in figure (1). It shows the case of a firm that faces \((\tau^u = 0.1, \phi = 1)\). In the figure we treat as exogenous capital and labor demand. Obviously in our model, these two last variables are endogenous. However, in the figure we treat them as exogenous to describe the way the function works out. In the figure, \(A_i \approx 1.1\) is a threshold level: if the firm draws a productivity parameter larger than the threshold level, the firm’s after-tax profits are larger than \(\phi\) and the firm operates. Clearly, the smaller is \(\epsilon\) the more the function resembles an indicator function.

[Insert figure 1 about here]

Note that if a firm does not operate, it does not demand either labor or capital. Therefore, we update labor and capital demand by firms as in:

\[
k_{i,ss}^j = \begin{cases} k_{i,ss}^j & i_i(operation)^j = 1 \\ 0 & i_i(operation)^j = 0 \end{cases} \tag{46}
\]
\[ l_{i,ss}^j = \begin{cases} l_{i,ss}^j & \textbf{if } (operation)_i^j = 1 \\ 0 & \textbf{if } (operation)_i^j = 0 \end{cases} \]  \hspace{1cm} (47)

Using the clearing of the goods market, we obtain consumption in steady state:

\[ c_{ss}^j = \sum_i 1_i(\text{operation})A_i(k_i^\alpha l_i^{1-\alpha})^\theta + \sum_i (1 - 1_i(\text{operation}))\phi - \delta \sum_i 1_i(\text{operation})k_i - G \]  \hspace{1cm} (48)

where \( G \) is fiscal expenditure. Next note that marginal utilities of consumption and leisure in steady state are:

\[ U_c^j = \Phi \left( (c_{ss}^j)^{\Phi 2} (\bar{H} - l_{ss}^j)^{1-\Phi 2} \right)^{\Phi 3-1} \left( (c_{ss}^j)^{\Phi 2} (\bar{H} - l_{ss}^j)^{1-\Phi 2} \right) \]  \hspace{1cm} (49)

\[ U_h^j = (1 - \Phi 2) \left( (c_{ss}^j)^{\Phi 2} (\bar{H} - l_{ss}^j)^{1-\Phi 2} \right)^{\Phi 3-1} \left( (c_{ss}^j)^{\Phi 2} (\bar{H} - l_{ss}^j)^{1-\Phi 2} \right) \]  \hspace{1cm} (50)

It follows that to satisfy the intratemporal marginal rate of substitution between consumption and leisure, total labor supply is\(^{17}\):

\[ l_{ss}^j = \bar{H} - \frac{c_{ss}^j}{\bar{w}^j} \left( \frac{1 - \Phi 2}{\Phi 2} \right) \]  \hspace{1cm} (51)

Equation (51) allows us to update our guess on labor supply. Let \( l_{ss}^{j,\text{updated}} \) be the update obtained from (51). If \( \| l_{ss}^j - l_{ss}^{j,\text{updated}} \| < \vartheta \), where \( \vartheta > 0 \) is a convergence criterion, we have computed the candidate vector \( x_{ss}^j \). Otherwise, we go back to (37) and we recompute the steady state variables in the \( j^{th} \) iteration, using the updated labor supply.

Equations (37), (43) to (51) and the tax guesses allow us to obtain our candidate vector \( x_{ss}^j \) in the \( j^{th} \) iteration.

\(^{17}\) In this step, we require \( \bar{w}^j \) which is computed as:

\[ \bar{w}^j = A_N(1 - \alpha)\theta^\alpha l_N^{1-\alpha} l_N^{1-\alpha}(1 - \tau_{ss}^j) \]

where \( N \) indicates the firm with the larger productivity parameter which by assumption (2) is always involved in production.
Parameters and results

We set the following parameters: $\beta = 0.9906$, $\delta = 0.05$, $\alpha = 0.36$, $\Phi_2 = 0.75$, $\Phi_3 = 1$, $\epsilon = 10^{-100}$. These parameters are consistent with the values reported Schmidt-Gohré and Uribe (2006). We discretize productivity in 50 equidistant points in the range $[1, 5]$, each of the points with 2% probability. Hence, our numerical exercises will have 50 different types of firms.

Figures (2) to (4) shows the results concerning after-tax profits, labor demand and capital demand per type of firm for $\theta = \{0.7; 0.85; 0.9\}$ and $\phi = 1.5$. In the figures, when a firm’s after-tax profit equals to $\phi$, the firm is not involved in production. As shown in the figures, optimal tax rates differ. In line with our theoretical discussion, we obtain in the three cases $\tau^u > 0, \tau^k, \tau^l < 0$. In general, the larger is $\theta$ the larger is $\tau^u$ and the larger are the subsidies to the capital and labor income, $\tau^k$ and $\tau^l$. Further, figure (2) shows also that the larger is $\theta$, the larger is the fraction of firms with after-tax profits equal to $\phi$, i.e. the larger is the fraction of firms not involved in production. Figures (3) and (4) show respectively labor and capital demand per firm. The larger is $\theta$, the larger is the increase in labor and capital demand as productivity raises, conditional on the firm being in operation. These result hold because the larger is $\theta$, the more elastic is the marginal cost of the firm and therefore the larger is the output and factor demands responses to the change in productivity.

[Insert figures (2) to (4) about here]

Table (1) shows the results for different values of $\phi$. The first three columns of the table present the firm’s exogenous parameters, the next three columns show the preference parameters and the last four columns show the results, including the fraction of firms involved in production, $1 - G(A^*)$. In the table, fiscal expenditure is set at a level of 500, which corresponds to 7% of output when $(\phi = 0, g = 0)$, i.e. the case in which all firms are involved in production and there are no fiscal distortions.

The table shows that, conditional on $\phi$, a larger $\theta$ in most of the cases is associated with larger corporate tax rate and larger capital and labor subsidy. Similarly, the fraction of firms involved in production decreases. These results are in line with the results obtained in figures (2) to (4) but as shown in the table, they also apply to the cases of $\phi = \{1; 1.5; 2\}$. Intuitively, the larger is $\theta$, holding constant other parameters, the lower are firms’ profits and the larger must be the corporate tax rate.
to raise revenues. In another hand, larger subsidies to production cost are required to complete the tax system, i.e. provide incentive to firms to produce, the larger is $\theta$. Finally, the result concerning a lower fraction of firms involved in production is easily explained because a larger $\theta$ is associated with larger factor payments and lower after-tax pure firms’ profits.

Table (2) provides sensibility analysis. In the table we use as benchmark the cases: $(\phi, \theta) = (1, 0.85)$ and $(\phi, \theta) = (1, 0.9)$. We initially include in the table a larger fiscal expenditure. We set fiscal expenditure at 600 which corresponds approximately to 8.5% of output when all firms are involved in production and there are no fiscal distortions. While the signs of the tax rates continue to be $\tau^u > 0, \tau^k, \tau^l < 0$, there is not a unique response of the taxes rates to the increase in fiscal expenditure. In one hand when $\theta = 0.85$, the magnitudes of both the subsidies and the corporate tax rate decrease and as a result, the fraction of firms become larger. In that case, the base of collection, in terms of the number of firms, increases. In another hand when $\theta = 0.9$, the corporate tax rate marginally increases while the labor tax rate becomes larger and the capital income tax rate approaches zero. In this case, the fraction of firms involved in production remained stable while the increase in labor subsidy is small compared to the decrease in capital income tax rate, i.e. the increase in tax revenues is obtained holding constant the base of the tax collection and decreasing net subsidies.

We next set $\alpha = 0.4$, i.e. we increase the capital share in the production function. In this case, the optimal subsidy in capital income becomes larger, both in the case of $\theta = \{0.85; 0.9\}$ while the labor subsidy and the corporate tax rate approach to zero. As a result, the number of firms involved in production is larger. In this case, the firm’s capital demand becomes more elastic providing more incentives to the planner to depress the rate of return faced by firms to induce more firms into production. Since the number of firms involved in production -which is a component of the tax base- raises, the planner might depress the labor subsidy and the corporate tax rate.

Finally, we provide a sensibility analysis of the response of optimal tax rates vis-à-vis the parameters of the utility function, $(\Phi_2, \Phi_3)$. In one hand, when we increase $\Phi_2$, the triplet of taxes approaches zero while the fraction of firms involved in production increases. In this case, distortions in leisure are more relevant, and the planner reacts by setting a smaller subsidy to labor
income. To satisfy the government budget constraint the planner reacts by providing incentives to new firms to enter into production by setting lower distortions -through lower corporate tax rate- in the extensive margin decision. In another hand, when we set $\Phi_3 = 0.75$, i.e. the utility function becomes more concave, we obtain mix results. In the case of $\theta = 0.85$, the capital income subsidy approaches zero while the magnitude of the labor income subsidy and the magnitude of the corporate tax rate increase and subsequently the fraction of firms involved in production decreases. When $\theta = 0.9$, the contrary holds. The consequence of this last set of results is that the planner uses the optimal tax rates such that fluctuations in the fraction of firms involved in production is diminished, as a way of decreasing fluctuations in consumption and labor supply (leisure).

[Insert table (2) about here]

5 Conclusions

In this paper we have introduced heterogeneous firms to study optimal taxation. Firms differ in their productivity and have to decide if they want to produce in each period after comparing their expected profits vis-à-vis an outside option that is not taxable, i.e. our environment is one of incomplete taxation (Correia (1996)). To finance its expenditure the government relies on capital income, labor and profits taxation. The presence of heterogeneous firms implies two kind of possible distortions from taxes. They may affect the extensive margin decisions (i.e. whether to produce or not) and the intensive margin decisions (i.e. optimal allocation given that they decide to produce).

We have shown that the results depend on whether all firms decide to produce or the less productive firms decide not to produce. In the first case, there is no distortion in the extensive margin and the social planner will not tax capital, which replicates Chamley (1986) and Judd (1985) results of not taxing capital income in the long run. Also, in this case, it is optimal not to tax labor and leaving the tax on profits as the only source of fiscal revenues.

However, when there are firms that are not so productive or when the outside option is high enough such that some firms prefer not to produce, it is optimal to subsidize capital. The sign of the tax on labor income is ambiguous depending on the distortions that creates in the labor supply.

The intuition of the subsidy is related to the government’s impossibility to tax the firms’ alternative option. By subsidizing production costs, the government induce firms to produce making
possible to tax them through their profits. That is, in the second best, the planner is partially completing the tax instruments by taxing firms that they would not be taxed if they remain in the alternative option. In this respect, firm heterogeneity is key to the results.

We have also analyzed the case when the government can tax the yields from the alternative option at the same rate that taxes profits from firms that are involved in production. In this case the tax system is complete and there is no need to induce firms into production by subsidizing. Tax on profits are now non distorting and it is optimal to set tax on capital and labor income equal to zero.
Appendix

• Derivation of equation 16

Since the production functions are homogeneous of degree $\theta$, the government budget constraint is:

$$
\tau_l^t w_t \int_{A_t^*}^A u l_{it} dG(A_{it}) + \tau_k^t r_t \int_{A_t^*}^A k_{it} dG(A_{it}) + \tau_u^t \int_{A_t^*}^A A_{it} (1 - \theta) f(k_{it}, l_{it}) dG(A_{it}) + \frac{b_{t+1}^s}{R_t} - b_t^s \geq g_t \ \forall t
$$

Further note that

$$
\tau_l^t w_t \int_{A_t^*}^A u l_{it} dG(A_{it}) = w_t \int_{A_t^*}^A u l_{it} dG(A_{it}) - \tilde{w}_t \int_{A_t^*}^A u l_{it} dG(A_{it})
$$

$$
\tau_k^t r_t \int_{A_t^*}^A k_{it} dG(A_{it}) = r_t \int_{A_t^*}^A k_{it} dG(A_{it}) - \tilde{r}_t \int_{A_t^*}^A k_{it} dG(A_{it}).
$$

Replacing in the government budget constraint,

$$
\int_{A_t^*}^A (w_t l_{it} + r_t k_{it}) dG(A_{it}) - \tilde{w}_t \int_{A_t^*}^A u l_{it} dG(A_{it}) - \tilde{r}_t \int_{A_t^*}^A u k_{it} dG(A_{it}) + \tau_u^t (1 - \theta) \int_{A_t^*}^A A_{it} f(k_{it}, l_{it}) dG(A_{it}) + \frac{b_{t+1}^s}{R_t} - b_t^s \geq g_t \ \forall t
$$

Finally, since $w_t l_{it} + r_t k_{it} = \theta A_{it} f(k_{it}, l_{it})$, we get equation (16) in the text.
Optimality conditions of Ramsey problem in equation (21).

The optimality condition with respect to \( k_{it} \) is:

\[
\lambda_{it} \beta^t [A_{it} f_k(k_{it}, l_{it}) + (1 - \delta)] g(A_{it}) + \lambda_{2t} \beta^t [\tau_t^u(1 - \theta) A_{it} f_k(k_{it}, l_{it}) + \theta A_{it} f_k(k_{it}, l_{it}) - \tilde{r}_t] g(A_{it}) \\
+ \lambda_{5t} \beta^t (1 - \tau_t^u)(1 - \theta) A_t^* f_k(k_t^*, l_t^*) \mathbb{1} [A_{it} = A_t^*] = \beta^{t-1}
\] (53)

The optimality condition with respect to \( l_{it} \) is:

\[
\beta^t [-u_h(t) + \lambda_{3t} u_{ch}(t)(\tilde{r}_t - \delta) - \lambda_{4t} u_{hh}(t) + \lambda_{1t} u_{ch}(t)\tilde{w}_t] g(A_{it}) \\
+ \beta^t \lambda_{1t} A_{it} f_l(k_{it}, l_{it}) g(A_{it}) + \beta^t \lambda_{2t} [\tau_t^u(1 - \theta) A_{it} f_l(k_{it}, l_{it}) + \theta A_{it} f_l(k_{it}, l_{it}) - \tilde{w}_t] g(A_{it}) \\
+ \beta^t \lambda_{5t} (1 - \tau_t^u)(1 - \theta) A_t^* f_l(k_t^*, l_t^*) \mathbb{1} [A_{it} = A_t^*] = 0
\] (54)

The optimality condition with respect to \((c_t, A_t^*, \tilde{w}_t, \tilde{r}_t)\) evaluated in steady state are:

\[
t^{c} : \quad u_c(c, h) - \lambda_1 + u_{cc}(c, h) \left[ \lambda_3 (-\tilde{r} + \delta) - \lambda_4 \tilde{w} \right] = 0 \\
t^{\tilde{w}} : \quad -\lambda_2 \int_{A_t^*}^A l_t dG(A_t) - \lambda_4 u_c(c, h) = 0 \\
t^{\tilde{r}} : \quad -\lambda_2 \int_{A_t^*}^A k_t dG(A_t) - \lambda_2 \frac{b}{1 + \tilde{r} - \delta} - \lambda_3 u_c(c, h) = 0 \\
t^{A} : \quad [u_h - \lambda_3 u_{ch}(\tilde{r} - \delta) - \lambda_4 u_{hh}(\tilde{w} + \lambda_4 u_{hh})]^* g(A^*) + \lambda_1 \left[ -1 + (1 - \tau^u)(1 - \theta) \right] A^* f(k^*, l^*) g(A^*) + \\
+ \lambda_1 \tilde{r} k^* g(A^*) + \lambda_2 [-\tau^u(1 - \theta) - \theta] A^* f(k^*, l^*) g(A^*) + \lambda_2 [\tilde{r} k^* + \tilde{w} l^*] g(A^*) \\
+ \lambda_5 (1 - \tau^u)(1 - \theta) f(k^*, l^*) = 0
\] (55, 56, 57, 58)
• Derivation of equation (24):

Integrating (22) with respect to the firms involved in production, we have:

\[
\lambda_1 \left[ \int_{A_i^*}^{A_i} A_i f_k(k_i, l_i) dG(A_i) + \left( (1 - \delta) - \frac{1}{\beta} \right) (1 - G(A_i^*)) \right] \\
+ \lambda_2 \left[ \tau^u (1 - \theta) \int_{A_i^*}^{A_i} A_i f_k(k_i, l_i) dG(A_i) + \theta \int_{A_i^*}^{A_i} A_i f_k(k_i, l_i) dG(A_i) - \tilde{r}(1 - G(A_i^*)) \right] \\
+ \lambda_5 (1 - \tau^u)(1 - \theta) A^* f_k(k^*, l^*) = 0 \quad (59)
\]

Adding and subtracting \( \lambda_2 \int_{A_i^*}^{A_i} A_i f_k(k_i, l_i) dG(A_i) \) and dividing by \( 1 - G(A_i^*) \), we obtain:

\[
\lambda_1 \left[ \frac{\int_{A_i^*}^{A_i} A_i f_k(k_i, l_i) dG(A_i)}{1 - G(A_i^*)} + \left( (1 - \delta) - \frac{1}{\beta} \right) \right] \\
+ \lambda_2 \left[ (\tau^u - 1)(1 - \theta) \frac{\int_{A_i^*}^{A_i} A_i f_k(k_i, l_i) dG(A_i)}{1 - G(A_i^*)} + \frac{\int_{A_i^*}^{A_i} A_i f_k(k_i, l_i) dG(A_i)}{1 - G(A_i^*)} - \tilde{r} \right] \\
+ \frac{\lambda_5}{1 - \tau^u}(1 - \theta) A^* f_k(k^*, l^*) = 0 \quad (60)
\]

Since \( \beta(1 + \tilde{r}_i - \delta) = 1 \) and using (5):

\[
\lambda_1 [r - \tilde{r}] + \lambda_2 [(\tau^u - 1)(1 - \theta)r] + \lambda_2 [r - \tilde{r}] + \frac{\lambda_5}{1 - G(A_i^*)}(1 - \tau^u)(1 - \theta)r = 0 \quad (61)
\]

Dividing by \( r \):

\[
\lambda_1 \tau^k + \lambda_2 [(\tau^u - 1)(1 - \theta)] + \lambda_2 \tau^k + \frac{\lambda_5}{1 - G(A_i^*)}(1 - \tau^u)(1 - \theta) = 0 \quad (62)
\]

Using (23) in (62):

\[
\tau^k = \left[ \frac{\lambda_5}{\lambda_1 + \lambda_2} \right] \left[ \frac{(1 - \tau^u)(1 - \theta)}{1 - G(A_i^*)} \right] [S_Y - 1] \quad (63)
\]
Where (63) is equation (24) in the text.

- Derivation of equation (26):

Integrating (25) with respect to the firms involved in production, we have:

\[
-uh + \lambda_3 u_{ch}(\bar{r} - \delta) - \lambda_4 u_{hh} + \lambda_4 u_{ch}\bar{w} \right] (1 - G(A^*)) + \lambda_1 \int_{A^*}^A A_i f_i(k_i, l_i) dG(A_i)
\]

\[
+ \lambda_2 \left[ \tau^u(1 - \theta) \int_{A^*}^A A_i f_i(k_i, l_i) dG(A_i) + \theta \int_{A^*}^A A_i f_i(k_i, l_i) dG(A_i) - \bar{w}(1 - G(A^*)) \right]
\]

\[
+ \lambda_5 (1 - \tau^u)(1 - \theta) A^* f_l(k^*, l^*) = 0
\]

Adding and substracting \(\lambda_2 \int_{A^*}^A A_i f_i(k_i, l_i) g(A_i)\) and dividing by \(1 - G(A^*)\), we obtain:

\[
-uh + \lambda_3 u_{ch}(\bar{r} - \delta) - \lambda_4 u_{hh} + \lambda_4 u_{ch}\bar{w}
\]

\[
+ \lambda_1 \int_{A^*}^A A_i f_i(k_i, l_i) g(A_i) \left[ (\tau^u - 1)(1 - \theta) \int_{A^*}^A A_i f_i(k_i, l_i) dG(A_i) \right]
\]

\[
+ \lambda_2 \left[ \frac{\int_{A^*}^A A_i f_i(k_i, l_i) g(A_i) \left[ \tau^u - 1 \right](1 - \theta) A^* f_l(k^*, l^*) \right] \right]
\]

\[
+ \lambda_5 (1 - \tau^u)(1 - \theta) A^* f_l(k^*, l^*) = 0
\]

Using (6):

\[
[uh - \lambda_3 u_{ch}(\bar{r} - \delta) + \lambda_4 u_{hh} - \lambda_4 u_{ch}\bar{w}]
\]

\[
= \lambda_1 w + \lambda_2 [(\tau^u - 1)(1 - \theta) w] + \lambda_2 w\tau l + \lambda_5 \frac{1 - G(A^*)}{1 - \tau^u(1 - \theta) w}
\]

(66)
It follows that:

\[
\tau_l = \frac{[u_h - \lambda_3 u_{ch}(\bar{r} - \delta) + \lambda_4 u_{hh} - \lambda_4 u_{ch}\bar{w}]}{\lambda_2 w} - \frac{\lambda_1}{\lambda_2} + \left[1 - \frac{\lambda_5}{\lambda_2(1 - G(A^*))}\right] (1 - \tau^u)(1 - \theta)(67)
\]

Using (23) in (67):

\[
\tau_l = \frac{[u_h - \lambda_3 u_{ch}(\bar{r} - \delta) + \lambda_4 u_{hh} - \lambda_4 u_{ch}\bar{w}]}{\lambda_2 w} - \frac{\lambda_1}{\lambda_2} + \left\{\frac{S_Y - 1}{S_Y}\right\} (1 - \tau^u)(1 - \theta) \tag{68}
\]

Note that using (55) to (58):

\[
\frac{[u_h - \lambda_3 u_{ch}(\bar{r} - \delta) + \lambda_4 u_{hh} - \lambda_4 u_{ch}\bar{w}]}{\lambda_2 w} - \frac{\lambda_1}{\lambda_2} = \frac{u_h}{\lambda_2 w} + \frac{u_{ch}}{u_c w} \left(\int_{A^*}^{A_u} k_i dG(A_i) + \frac{b}{1 + \bar{r} - \delta}\right) (\bar{r} - \delta)
\]

\[ - \frac{u_{hh} \int_{A^*}^{A_u} k_i dG(A_i)}{u_c w} + \frac{u_{ch} \int_{A^*}^{A_u} k_i dG(A_i) \bar{w}}{u_c w}
\]

\[ - \frac{u_c}{\lambda_2} - \frac{u_{cc}}{u_c} \left(\int_{A^*}^{A_u} k_i dG(A_i) + \frac{b}{1 + \bar{r} - \delta}\right) (\bar{r} - \delta)
\]

\[ - \frac{u_{cc}}{u_c} \left(\int_{A^*}^{A_u} l_i dG(A_i)\right) \bar{w} \tag{69}
\]

Since \(u_c \bar{w} = u_h\), we have:

\[
\frac{[u_h - \lambda_3 u_{ch}(\bar{r} - \delta) + \lambda_4 u_{hh} - \lambda_4 u_{ch}\bar{w}]}{\lambda_2 w} - \frac{\lambda_1}{\lambda_2} = -\tau^l \frac{u_c}{\lambda_2}
\]

\[
- \frac{u_{cc}}{u_c} \left[\left(\int_{A^*}^{A_u} k_i dG(A_i) + \frac{b}{1 + \bar{r} - \delta}\right) (\bar{r} - \delta) + \left(\int_{A^*}^{A_u} l_i dG(A_i)\right) \bar{w}\right]
\]

\[ + \frac{u_{ch}}{u_c} \left[\left(\int_{A^*}^{A_u} k_i dG(A_i) + \frac{b}{1 + \bar{r} - \delta}\right) (\bar{r} - \delta) + \left(\int_{A^*}^{A_u} l_i dG(A_i)\right) \bar{w}\right] (1 - \tau^l)
\]

\[ - \frac{u_{hh} \int_{A^*}^{A_u} l_i dG(A_i)}{u_h}(1 - \tau^l)\]

Further \(\beta(1 + \bar{r_l} - \delta) = 1\) implies \(\frac{b}{1 + \bar{r} - \delta} = b/\beta\). Let \(\sigma_{cc} = -\frac{u_{cc}}{u_c}\), \(\sigma_{ch} = \frac{u_{ch}}{u_h}\), \(\sigma(h) = \frac{u_{hh}(1-h)}{u_h}\) where \(\bar{c} = \int_{A^*}^{A_u} k_i dG(A_i) + b_{ss}\beta + \bar{w} \int_{A^*}^{A_u} l_i dG(A_i)\) is total individual’s income - excluding firm’s
transfers in steady state. It follows:

\[
\frac{[u_h - \lambda_3u_{ch}(\tilde{r} - \delta) + \lambda_4u_{hh} - \lambda_4u_{ch}\tilde{w}]}{\lambda_2w} - \frac{\lambda_1}{\lambda_2} = -\tau^l\frac{u_c}{\lambda_2} + \sigma_{cc} + \sigma_{ch}(1 - \tau^l) + \sigma_{hh}(1 - \tau^l) \tag{70}
\]

Replacing (70) in (68):

\[
\tau^l = -\tau^l\frac{u_c}{\lambda_2} + \sigma_{cc} + \sigma_{ch}(1 - \tau^l) - \sigma_{hh}(1 - \tau^l) + \left[\frac{S_Y - 1}{S_Y}\right](1 - \tau^u)(1 - \theta)
\]

\[
\Rightarrow \tau^l \left(1 + \frac{u_c}{\lambda_2} + \sigma_{ch} + \sigma_{hh}\right) = (\sigma_{cc} + \sigma_{ch} + \sigma_{hh}) + \left[\frac{S_Y - 1}{S_Y}\right](1 - \tau^u)(1 - \theta) \tag{71}
\]

This is equation (26) in the text.

**Derivation of equation (27):**

Replacing (66) in (58):

\[
\lambda_1wl^*g(A^*) + \lambda_2[(\tau^u - 1)(1 - \theta)wl^*g(A^*)] + \lambda_2\tau^lwl^*g(A^*) + \frac{\lambda_5}{1 - G(A^*)}(1 - \tau^u)(1 - \theta)wl^*g(A^*)
\]

\[
+ \lambda_1\left[-1 + (1 - \tau^u)(1 - \theta)\right]A^*f(k^*, l^*)g(A^*) + \lambda_1\tilde{r}k^*g(A^*)
\]

\[
+ \lambda_2[-\tau^u(1 - \theta - \theta)]A^*f(k^*, l^*)g(A^*) + \lambda_2[\tilde{r}k^* + \tilde{w}l^*]g(A^*) + \lambda_5(1 - \tau^u)(1 - \theta)f(k^*, l^*) = 0 \tag{72}
\]

Using \(\tilde{r} = r(1 - \tau^k)\) and additional steps of algebra:

\[
(\lambda_1 + \lambda_2)[wl^* + rk^*]g(A^*) + \lambda_2[(\tau^u - 1)(1 - \theta)wl^*g(A^*)] + (\lambda_1 + \lambda_2)[-1 + (1 - \tau^u)(1 - \theta)]A^*f(k^*, l^*)g(A^*) - (\lambda_1 + \lambda_2)r\tau^k k^*g(A^*)
\]

\[
+ \lambda_5(1 - \tau^u)(1 - \theta)f(k^*, l^*) + \frac{\lambda_5}{1 - G(A^*)}(1 - \tau^u)(1 - \theta)wl^*g(A^*) = 0 \tag{73}
\]
It follows:

\[(\lambda_1 + \lambda_2)\theta A^* f(k^*, l^*) g(A^*) + \left[ -\lambda_2 + \frac{\lambda_5}{1 - G(A^*)} \right] [(1 - \tau^u)(1 - \theta)wl^* g(A^*)] \]

\[+ \quad (\lambda_1 + \lambda_2) [-1 + (1 - \tau^u)(1 - \theta)] A^* f(k^*, l^*) g(A^*) - \frac{\lambda_5}{1 - G(A^*)} (1 - \tau^u)(1 - \theta)(S_Y - 1)rk^* g(A^*) \]

\[+ \quad \lambda_5 (1 - \tau^u)(1 - \theta) f(k^*, l^*) = 0 \]

Since (23) implies \(\lambda_2 = \frac{\lambda_5}{1 - G(A^*)} S_Y\), we have:

\[-(\lambda_1 + \lambda_2)\tau^u(1 - \theta) A^* f(k^*, l^*) g(A^*) + \frac{\lambda_5}{1 - G(A^*)} (1 - \tau^u)(1 - \theta) [1 - S_Y] \theta A^* f(k^*, l^*) g(A^*) \]

\[+ \quad \frac{\lambda_5}{A^* g(A^*)} (1 - \tau^u)(1 - \theta) A^* f(k^*, l^*) = 0 \]

Dividing by \((\lambda_1 + \lambda_2)(1 - \tau^u)(1 - \theta)A^* f(k^*, l^*)\), we finally obtain:

\[\frac{\tau^u}{1 - \tau^u} = \frac{\lambda_5}{\lambda_1 + \lambda_2} \left[ \theta(1 - S_Y) + \frac{1}{A^* g(A^*)} \right] \quad (74)\]

which is (27) in the text.

References


Figures and tables

Figure 1: Indicator function, \( 1_i(\text{operation}) = 1 - \frac{(\arctan(-\frac{1-\tau^u}{k_ih_i^{1-\alpha}}\theta + \phi) + \pi/2)}{\pi} \)

\( k_i = 1.5, h_i = 0.7, \phi = 1, \tau^u = 0.1, \alpha = 0.36, \theta = 0.4 \)
Figure 2: Steady state after-tax profits per firm
Figure 3: Steady state labor demand per firm
Figure 4: Steady state capital demand per firm

![Graph showing steady state capital demand per firm with different productivity levels and capital demands.](image-url)
Table 1: Optimal policies in steady state

<table>
<thead>
<tr>
<th>φ</th>
<th>θ</th>
<th>α</th>
<th>Φ₂</th>
<th>Φ₃</th>
<th>g</th>
<th>1 − G(Å*)</th>
<th>τᵏ</th>
<th>τᵢ</th>
<th>τᵤ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.7</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.66</td>
<td>-0.0358</td>
<td>0.0021</td>
<td>0.0166</td>
</tr>
<tr>
<td>1</td>
<td>0.75</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.60</td>
<td>-0.233</td>
<td>-0.1563</td>
<td>0.183</td>
</tr>
<tr>
<td>1</td>
<td>0.8</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.54</td>
<td>-0.3294</td>
<td>-0.0655</td>
<td>0.1561</td>
</tr>
<tr>
<td>1</td>
<td>0.85</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.44</td>
<td>-0.3935</td>
<td>-0.2635</td>
<td>0.2812</td>
</tr>
<tr>
<td>1</td>
<td>0.9</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.30</td>
<td>-0.3816</td>
<td>-1.0553</td>
<td>0.5319</td>
</tr>
<tr>
<td>1.5</td>
<td>0.7</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.60</td>
<td>-0.0731</td>
<td>-0.0041</td>
<td>0.0492</td>
</tr>
<tr>
<td>1.5</td>
<td>0.8</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.48</td>
<td>-0.2945</td>
<td>-0.0616</td>
<td>0.1484</td>
</tr>
<tr>
<td>1.5</td>
<td>0.85</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.40</td>
<td>-0.399</td>
<td>-0.0855</td>
<td>0.2109</td>
</tr>
<tr>
<td>1.5</td>
<td>0.9</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.30</td>
<td>-0.4242</td>
<td>-0.3199</td>
<td>0.2822</td>
</tr>
<tr>
<td>2</td>
<td>0.7</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.54</td>
<td>-0.0878</td>
<td>0.0188</td>
<td>0.0679</td>
</tr>
<tr>
<td>2</td>
<td>0.85</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.36</td>
<td>-0.2483</td>
<td>-0.5221</td>
<td>0.2147</td>
</tr>
<tr>
<td>2</td>
<td>0.9</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.28</td>
<td>-0.3197</td>
<td>-0.0866</td>
<td>0.1728</td>
</tr>
</tbody>
</table>
Table 2: Optimal policies in steady state, sensibility analysis

<table>
<thead>
<tr>
<th>φ</th>
<th>θ</th>
<th>α</th>
<th>(\Phi_2)</th>
<th>(\Phi_3)</th>
<th>(g)</th>
<th>(1 - G(A^*))</th>
<th>(\tau^k)</th>
<th>(\tau^l)</th>
<th>(\tau^u)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.85</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.44</td>
<td>-0.3935</td>
<td>-0.2635</td>
<td>0.2812</td>
</tr>
<tr>
<td>1</td>
<td>0.9</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.30</td>
<td>-0.3816</td>
<td>-1.0553</td>
<td>0.5319</td>
</tr>
<tr>
<td>1</td>
<td>0.85</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>600</td>
<td>0.46</td>
<td>-0.0109</td>
<td>-0.0514</td>
<td>0.0219</td>
</tr>
<tr>
<td>1</td>
<td>0.9</td>
<td>0.36</td>
<td>0.75</td>
<td>1</td>
<td>600</td>
<td>0.30</td>
<td>-0.2141</td>
<td>-1.1232</td>
<td>0.565</td>
</tr>
<tr>
<td>1</td>
<td>0.85</td>
<td>0.4</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.48</td>
<td>-0.4368</td>
<td>-0.1017</td>
<td>0.2317</td>
</tr>
<tr>
<td>1</td>
<td>0.9</td>
<td>0.4</td>
<td>0.75</td>
<td>1</td>
<td>500</td>
<td>0.36</td>
<td>-0.3923</td>
<td>-0.2923</td>
<td>0.3191</td>
</tr>
<tr>
<td>1</td>
<td>0.85</td>
<td>0.36</td>
<td>0.85</td>
<td>1</td>
<td>500</td>
<td>0.46</td>
<td>-0.2905</td>
<td>-0.0546</td>
<td>0.1434</td>
</tr>
<tr>
<td>1</td>
<td>0.9</td>
<td>0.36</td>
<td>0.85</td>
<td>1</td>
<td>500</td>
<td>0.34</td>
<td>-0.3349</td>
<td>-0.2621</td>
<td>0.2708</td>
</tr>
<tr>
<td>1</td>
<td>0.85</td>
<td>0.36</td>
<td>0.75</td>
<td>0.75</td>
<td>500</td>
<td>0.40</td>
<td>-0.1416</td>
<td>-2.5552</td>
<td>0.4611</td>
</tr>
<tr>
<td>1</td>
<td>0.9</td>
<td>0.36</td>
<td>0.75</td>
<td>0.75</td>
<td>500</td>
<td>0.34</td>
<td>-0.4456</td>
<td>-0.2417</td>
<td>0.2637</td>
</tr>
</tbody>
</table>