Università degli Studi di Cagliari

DOTTORATO DI RICERCA

ECONOMIA

Ciclo XXVII

TITOLO TESI

OPTIMAL LABOR INCOME TAXATION WITH PERKS

Settore/i scientifico disciplinari di afferenza

SECS-P/01, SECS-P/06

Presentata da: UNALI VALERIA

Coordinatore Dottorato PROF. PIRAS ROMANO

Tutor PROF. DEIDDA LUCA GABRIELE

Esame finale anno accademico 2014 – 2015

Abstract

Incentive problems arise in many economic relationships, between workers and firms as well as between those agents and the fiscal authority. If it is well known, in fact, that labor contracts tying wages to performance may mitigate the efficiency costs from unobservable effort, it is also an empirical fact that real word contracts, and the incentives provided wherein, are more frequently based on some sort of non-monetary compensation - fringe-benefits, or perks, as they are called. As a result, the fiscal authority is more often called to decide on the eligibility of those perquisite goods and, therefore, on their taxability.

In spite of their diffusion, however, there is still little consensus among researchers on the reasons why perks should be provided, and what their effect on welfare might me. The standard explanation for perks usually relies on some form of agency problem. However, monitoring concerns are certainly less relevant when dealing with a fiscal authority, that, for the most part, may monitor the provision of perks better than the payment of cash.

Motivated by the concerns about how fringe benefits may restrict the revenue collection problem and affect the redistribution of resources and incentives, this research aims at investigating the relationship between the provision of perks and i) the progressivity/regressivity of the optimal transfer-tax system, and ii) the top-income marginal tax rate.

To the best of our knowledge, this thesis constitutes the first theoretical attempt to deal with the issue at hand. In fact, all existing papers that analytically study the moral hazard problem with perks, either focus on their effect on the effort responsiveness to monetary incentives, or they study the cash-perks substitutability in the optimal labor contract, absent any consideration for the insurance/incentive purposes of income taxation. To some extent, we also contribute to the literature on optimal taxation, by emphasizing the connection between the agents' responses to the fiscal system and the structure of the labor market.

Methodologically, we investigate our research agenda by following two different approaches. In Chapter 2, we develop a static version of a standard stochastic Ramsey's problem with a representative agent and an utilitarian, resource-seeking government who takes the post of the principal and owner of the only firm in the economy. When designing the optimal rules for cash payment and perk provision, the government, who is constrained by his budget balance, takes into account the agent's unobservable reactions to the tax system, and set op-

timal taxes so as to provide either insurance and right incentives.

In Chapter 3, however, we take an alternative perspective, and study a decentralized economy wherein an independent, self-interest firm retains full control of the provision of perks and wage payments, given the fiscal policy announced by the government. Our analysis aims at highlighting the existence of a trade-off between the opposite interests of the government and the employer, that will serve as a rationale for the characterization of the equilibrium marginal tax rate that we derive from a Nash non-cooperative interaction game between the fiscal authority and the agents in the labor market.

Our analysis suggests that, whenever perks are efficiently provided, the government trades off progressivity for perk provision. Our main conclusion, in fact, is that the (centralized) second-best efficient taxation scheme with perks is more regressive compared to its perk-less equivalent. Notably, we also find that, whenever society's preferences for public expenditure and agents' risk aversion are sufficiently low, a regressive top-income marginal tax is also consistent with a positive provision of perks. In a numerical exercise, for example, we show that a marginal tax rate of 30% allows for either i) a positive provision of perks that accounts for 1.6% of the gross income (3.2% of the agent's wage), and ii) a per-capita public expenditure which is 7% of the top-brackets taxpayers' gross income (15% of their earnings) . An equilibrium allocation without perks results, instead, for a 45% marginal tax rate.

However, in spite of the ability of our theoretical framework to capture the relationship between risk aversion and efficient provision of perks, through their effect on the optimal trade-off between the insurance and incentive motives for taxation, our model has its weakness, as to regard its sensitiveness to the stochastic dominance properties of the probability distribution and the level of income in the economy. If, in the former case, deviations from the benchmark distribution are qualitatively important but quantitatively small, in the latter case, the quantitative implications of different income levels are quite significative. A more precise calibration of the models and a better characterization of the results are left to future works.

Contents

Chapter 1

On the Field of Optimal Income Taxation

Incentive problems arise in many economic relationships, between workers and firms as well as between those agents and the fiscal authority. If it is well known, in fact, that labor contracts tying wages to performance may mitigate the efficiency costs from unobservable effort, it is also an empirical fact that real word contracts, and the incentives provided wherein, are more frequently based on some sort of non-monetary compensation - fringe-benefits, or perks, as they are referred to.

As a result, the fiscal authority is more often called to decide on the eligibility of those perquisite goods and, therefore, on their taxability. According to the Internal Revenue Service (US), for instance, transportation benefits, working space amenities, free cafeterias, along with employee discounts, educational assistance and scholarship for family members, are just some of those "qualified" perks that are defined non-taxable.

Besides, the issue is not far from having a large empirical relevance. As perks may include personal use of corporate jets, membership in selected clubs, financial consulting, estate planning, kindergarten services, airplane tickets and so on, they can reach significant amounts, still representing a small fraction of an individual salary. Furthermore, as some surveys suggest[1], fringe-benefits, though common among chief executives, are certainly not exclusive to them,

---

[1] The Associated Press, for example, use data provided by Equilar for its annual survey of CEOs' pay, consisting of base salary, short- and long-term incentives, employee benefits, perquisites, and compensation protection.

nor they are a phenomenon specific to large and private corporations.

In spite of their diffusion, however, there is still little consensus among current researchers on the reasons why perks should be provided, and what their effect on welfare might me. In the financial literature, for example, the standard explanation for perks relies on some form of agency problems. However, it can be argued that monitoring concerns, which are the key factor of those models, if even less reasonable nowadays (since the disclosure rules issued by the US Securities and Exchange Commission, in 2006), are certainly less relevant when dealing with a fiscal authority, that, for the most part, may monitor the provision of perks better than the payment of cash.

Motivated by the concerns about how fringe benefits may restrict the revenue collection problem and affect the redistribution of resources and incentives, this research aims at merging the macro approach to the optimal income taxation with the micro approach to an efficient labor contract, and studies how perks do affect optimal labor income taxes. More precisely, the purpose of this study is to investigate the relationship between the provision of perks i) and the progressivity/regressivity of the optimal tax-transfer system, ii) and the top income marginal tax rate.

The main contribution of our work is the development of a theoretical framework for the optimal labor income tax, wherein: i) agents' work decisions are unobservable, ii) production technology depends (in a non-deterministic way) on the agent's effort, and iii) labor is optimally repaid by paid-in-advance perks and state-contingent wages.

Technically, this thesis relies on the first-order approach (FOA) to replace the effort incentive compatibility constraint with the first-order necessary condition derived from the agent's maximization problem. Though relying on the FOA means more complex computations, its application is important for at least two reasons. Firstly, it is useful when explaining our dynamics in terms of the responsiveness of the agents's effort to the instruments designed by the principal. Secondly, it allows for a direct comparison with the more recent works on optimal taxation and agency problems.

To the existing literature, we also contribute by emphasizing the connection between the agents' responses to the fiscal system and the structure of the labor market, which was previously absent from the theory of optimal income taxes. In particular, we focus on formally modeling and rigorously deriving the equilibrium allocation of either labor, consumption and tax revenues, when perquisite goods are available to (independent) firms, so as to deal with the

incentive problems that naturally arise in the labor market when the agents'
effort is neither contractable nor observable.

In doing so, however, we neglect some of the points that still concern current
researchers, namely: i) the design of transfers to subside work (the extensive
margin of the labor supply); ii) the differentiated taxation of commodity goods,
and therefore, of perks; iii) the revenue-maximization problem (the so-called
Laffer curve). We also abstract from any issue on collusion between the worker
and the principal, and any misbehavior of the former, and mainly focus on
a class of pure work-related goods, with no productive attributes, which are
uniquely defined according to their complementarity with respect to the agent's
effort[2].

Methodologically, we investigate our research question by following two different
approaches.

In Chapter 2, we develop a static version of a standard stochastic Ramsey's prob-
lem with a representative agent (the worker), where the utilitarian, resource-
seeking government takes the post of the principal and owner of the only firm
in the economy. When designing the optimal rules for cash payment and perk
provision, the government, who is constrained by his budget balance, takes into
account the agent's unobservable reactions to the tax system.

Here, optimal taxes are designed so as to accomplish to either the insurance
and the incentive motive. The key aim of Chapter 2, in fact, is to show if there
exists a room for a "socially" optimal provision of perks, and which restrictions
perks do impose on the optimal taxation, as compared to the the optimal tax
system that would arise in the perk-less economy. To the best of our knowledge,
this chapter constitutes the first attempt to deal with the issue at hand. In
fact, virtually all the existing papers that analytically study the moral hazard
problem with perks, either focus on their effect on the effort responsiveness to
monetary incentives, or they study the cash-perks substitutability in the optimal
labor contract, absent any consideration for the insurance/incentive purposes of
income taxation.

In Chapter 3, however, we take an alternative perspective, and study a decen-
tralized economy wherein an independent, self-interest firm retains full control
of the provision of perks and wage payments, given the fiscal policy announced

---

[2]Though somehow limitative, this restrictive view of the perquisite good is useful for at
least two purposes. First, it allows us to avoid the unnecessary entanglement of mechanisms
that would naturally occur if perks were also used as a consumption good. Second, it makes
easier to apply our analysis to any hierarchical level and any production sector.

by the government. The analysis therein aims at highlighting the existence of a trade-off between the opposite interests of the government and the employer. It serves as a rationale for the characterization of the equilibrium top income marginal tax rate that results from the Nash non-cooperative interaction game between the fiscal authority and the agents in the labor market. It is for seek of tractability that, in defining a social welfare function for that problem, we assume that public revenues, whenever taken as given from the agent, do not directly affect labor supply decisions. What actual revenues do, however, is to affect the equilibrium optimal tax rate.

Though divergent from a standard utilitarian approach, being built on the concept of an "ideal citizen", the analysis in Chapter 3 results a simple formula for the top income marginal tax rate that serves either as a policy recommendation and a renewed testable hypothesis, based on variables whose micro-foundation now relies on more sophisticate, but also more realistic, labor contracts.

In doing so, we stress the need for a deeper connection between real economy and theoretical predictions, and emphasize the loss in economic content when setting the debate on optimal income taxation apart from the richness of the dynamics in the labor-market.

The rest of this chapter is organized as follows. The next section mainly reviews the theory of optimal labor income taxation, by highlighting those theoretical underpinnings needed to place this work in its historical background. In Section 1.2 we propose and discuss a few arguments in favor of the non-taxability of the perquisite good. Section 1.3 discusses our methodology and the technical challenges imposed by the provision of perks. Section 1.4 concludes with an overall view of our results as to regard the effect of perks on the optimal taxation scheme and the top-income marginal tax rate.

1.1 The Theory of Optimal Income Taxation

Though characterized and renewed by the provision of perks, this thesis mainly focuses on the optimal labor income taxation, i.e. on the fair and efficient distribution of the tax burden across individuals with different earnings.

A large academic literature has developed models of tax theory to cast light on the optimal balance between equity and efficiency. Their models typically posit that the government maximizes a social welfare function, subject to his budget constraint, when taking into account the agents' responses to the tax-transfer

system. Since, for sufficiently well-behaved preferences, the social welfare is larger when resources are more equally distributed, but redistributive taxes negatively affect incentives to work and earn income, levying taxes creates a non-degenerated trade-off between equity and efficiency, which is at the core of the optimal labor income taxation problem.

As argued by Diamond and Saez (2011), however, theoretical results in optimal taxation are actually useful for policy recommendations whenever three conditions are met:

- results are based on economic mechanisms that are empirically relevant;

- they are reasonably robust to modeling assumptions and heterogeneity in individual preferences;

- tax policy prescriptions are implementable.

Historically, these conditions have led to two different methodological choices:

1. the "mechanism-design" approach derives the optimal transfer-tax system which is compatible with the information structure. This method, though extremely effective in casting light on the optimal equity-efficiency trade-off, generates tax functions that are complex, and results which tend to depend heavily on the choice of the primitives of the model.

2. the "sufficient-statistics" approach aims at proposing tax formulas that are derived and expressed in terms of estimable statistics. Those include: i) social marginal welfare weights (which capture society's value for redistribution, or society's value for the size of the social state); ii) labor supply elasticities (which capture the efficiency costs of taxation, and, as it is in our setting, its efficiency benefits, for reducing the distortions that arise in the labor market due to the information asymmetry); iii) the endogenous distribution of earnings (or its simplification through the application of Pareto weights).

Since the former approach has recently received new interest from the dynamic public finance literature, because of the major tractability guaranteed by the applicability of the first-order-approach (FOA), which we will discuss below, the model proposed in Chapter 2 constitutes a first attempt to discuss the issue at hand by appealing to the mechanism-design approach. A sufficient-statistics approach will be follows, instead, in Chapter 3, the analysis of the

decentralized economy being necessary to understand the driving forces and underlying mechanisms behind the tax formula therein derived.

To place our thesis in its academic context, the rest of this section briefly reviews the history of the field (distinguished by focus and approach), and discusses some of the theoretical underpinnings of the standard models of optimal income taxation. It also proposes a critical overview of our main contributions to the literature, by highlighting, whenever possible, the advantages/disadvantages of our approach.

1.1.1 Optimal income taxation: Mirrlees (1971)

The modern analysis of optimal income taxation certainly started with Mirrlees (1971). He considered the problem of maximizing a social welfare function based on the agents' utilities, subject to the government budget constraint and the incentive compatibility constraint arising from the individuals' labor supply responses to the tax system.

More precisely, in Mirrlees's model, the agents, by differing in their ability, earn different wage-rates. On account of the social preferences, the government seeks to redistribute from high-income to low-income agents, but he can only do so by setting taxes and transfers based on earnings, rather than on productivity.

The existence of an incentive-compatibility constraint, thus, leads to a non-degenerate equity-efficiency trade-off. The theoretical underpinning of this kind of models is that, under utilitarian social preferences, behavioral responses are the sole factor preventing the complete redistribution of income, that still would result if earnings were fixed.

To state the problem more formally, for a general social welfare function of the form

$$\text{SWF} = \int \omega_i W(u_i(c(y_i))) dF(i),$$

the combination of arbitrary Pareto weights ω_i (which are independent of individual choices), and of the increasing and concave social welfare function $W(\cdot)$ (of the agents' consumption utility u_i at earnings level y_i), pins down the social marginal welfare weight on individual i,

$$\mu_i = \frac{\omega_i W'(\cdot) u_i'(\cdot)}{\lambda}$$

where λ is the social marginal value of government's revenues. Intuitively, μ_i measures the dollar value of increasing individual i's consumption by one unit.

Social preferences for redistribution, then, enter optimal tax formulas through the weights μ_i .

In the class of non-linear tax systems, the optimal marginal tax is set according to

$$\frac{T'(y)}{1 - T'(y)} = \frac{1}{\epsilon_{y\tau}} \left(\frac{1 - F(y)}{yf(y)} \right) \left[1 - \mu(y) \right]$$

where $\mu(y)$ is the average social marginal value of consumption for taxpayers with earnings y_i above y, $\epsilon_y \equiv \dfrac{\partial y}{\partial (1 - T'(y))} \dfrac{1 - T'(y)}{y}$ is the elasticity of earning to the marginal tax rate, and $F(y)$ is the CDF of y, with density $f(y)$ - both endogenous.

Though Mirrlees (1971) had an enormous influence in the development of social contracts based on information asymmetry, its application in practice was very limited. The general solution to this problem is indeed very complex, due to the endogeneity of $\mu(y)$, and provides little intuition. Mirrlees himself carried out several simulations, using a Cobb-Douglas utility function and a log-normal wage distribution under a classical utilitarian SWF, and found that the optimal tax schedule is approximately linear above a certain threshold. Moreover, it is always non-negative.

Appealing to the concepts of aggregate elasticity $\epsilon \equiv \dfrac{\partial y}{\partial (1 - \tau)} \dfrac{1 - \tau}{y}$ and average "normalized" social marginal welfare weight μ, weighted by the pre-tax incomes y_i, the optimal tax rate in the class of linear tax system is given by

$$\tau = \frac{1 - \mu}{1 - \mu + \epsilon}$$

where y is the (endogenous) average income and $\mu = \dfrac{1}{y} \displaystyle\int \mu_i y_i dF(i)$ is also the ratio of the income mean, weighted by the individual social welfare weights μ_i, to the actual average income y.

As a remark, which applies to our setting, we note that, whenever earnings are generated by a partly random process involving luck in addition to effort (as it is in our model), the tax formulas above still hold, as long as the social welfare function $W(\cdot)$ is defined over the individuals' expected utility, $E\big[u_i(\cdot)\big]$.

Stiglitz (1982) discrete version of Mirrlees (1971) model was useful to understand the problem of optimal taxation as a modern information problem. The two-type (high-skilled and low-skilled agent) case highlighted the requirement that the optimal tax system must be set up so as to discourage the high-skilled agent from mimicking the low-skilled by working less. Nonetheless, either the general

model of Mirrlees and Stiglitz's discrete version, have had limited use for actual policy recommendations in practice, since it is hard to obtain tax formulas in terms of "sufficient statistics".

1.1.2 Optimal top income marginal tax

The taxation of high-income earners is a very important aspect of any tax system. Initially, progressive income tax systems were limited to the top of the income distribution. Today, because of the large increase in income concentration, the level of taxation of top incomes (i.e., the top 1%) matters not only for equity reasons but also quantitatively, for revenue-raising needs. Since the top of the income distribution is the one that more is concerned about the provision of the perquisite good, we focus on the optimal top income tax in the last chapter of our thesis, and propose here a brief review of its theoretical derivation.

Standard models of optimal top tax rate assume that the tax rate above a fixed income level $y^\star$ is constant and equal to τ_b. By a perturbation argument, for any small variation $d\tau$, by earning b_i above $y^\star$, individual i mechanically pays $(b_i - y^\star)d\tau$ extra units in taxes. This extra payment creates a social welfare loss equal to $-\mu_i \cdot (b_i - y^\star)d\tau$, where μ_i is the social marginal welfare weight on individual i. The tax change, however, triggers a behavioral response, db_i, so leading to an additional change in taxes equal to $db_i \cdot \tau_b$.

By combining the marginal effect of this small reform on individual i , and aggregating the welfare effects across all the top-bracket taxpayers, so that y_b is the average income of the individuals in the top-bracket, the optimal top tax rate solves

$$\frac{\tau_b}{1 - \tau_b} = \frac{(1 - g_b)(y_b/y^\star - 1)}{\epsilon_b \cdot \left(\frac{y_b}{y^\star}\right)}$$

where g_b is the average social marginal welfare weight (weighted by individuals' income b_i) for the top-bracket individuals, and ϵ_b is the top-bracket individuals' average elasticity of income y_b with respect to the net-of-tax rate $1 - \tau_b$, multiplied by earnings y_b.

Moreover, since the top tail of the income distribution is closely approximated by a Pareto distribution with tail parameter $a = \dfrac{y_b/y^\star}{\left(\dfrac{y_b}{y^\star} - 1\right)} = y_b/(y_b - y^\star) > 1$,

the above formula is generally proposed in its simplified version, that is

$$\tau_b = \frac{1 - g_b}{1 - g_b + a \cdot \epsilon_b}$$

Accordingly, the optimal top tax rate decreases in:

i) g_b. In the limit, when the social marginal weight on top-bracket earners is zero, it simplifies to $\tau_b = \dfrac{1}{1 + a \cdot \epsilon_b}$, which is also the revenue maximizing top tax rate;

ii) ϵ_b, as a higher elasticity leads to larger efficiency losses;

iii) a. Moreover, if $y^\star$ reaches the level of income of the highest income earner, for $a \to 0$, then $\tau_b \to 0$ (unless $\epsilon_b = 0$), which is the famous zero-top-rate result first demonstrated by Sadka(1976) and Seade (1977).

The importance of the top tax rate is in its being the asymptotic limit of the optimal marginal tax rate of the fully non-linear tax problem of Mirrlees (1971). What the literature on either optimal income and optimal top income taxation emphasizes, therefore, is that any policy recommendation on how the optimal tax system should be designed depends on the extent of behavioral responses to taxes, and, hence, on the size of those empirical parameters, such as the elasticity of labor supply and the density of earnings. Different concepts of elasticity have been involved, to take into account either the participation labor supply response (the so-called extensive margin), and the effort's responses (the intensive margin).

Since we focus on the intensive labor supply margin, i.e. on the choice of how much to work, conditional on the compensation scheme, and since wage-rates are set by independent firms, the elasticity concepts that are convenient in our framework are those of an uncompensated (or Marshellian) elasticity of earnings, and that of an uncompensated elasticity of labor supply, both with respect to the tax rate. As we shall see in Chapter 3, in fact, the uncompensated elasticity of labor supply combines with the uncompensated elasticity of earnings (the two been the solutions of two different maximization problems) in order to determine the optimal top tax rate.

Generally speaking, because there are no income effects in our models, in the sense that increasing the (certain amount of) non-labor income of either the agents or the firm does not affect labor supply decisions, uncompensated elasticities are also compensated ones. However, the provision of perks, which is made to the agent by a self-interested firm, is going to slightly change this equivalence. In the spirit of a decentralized economy, the provision of perks that follows to an increase in the tax rate can be looked at as a compensated labor-income increase, since the incentive compatibility constraint still binds. If those "income effects" were large, and perks had a large impact on labor sup-

ply and earnings decisions, distortionary taxes could play even a higher role in balancing insurance/redistribution against incentives.

Finally, our contribution to this literature could be extended to the suggestion of a measure for the density of earnings that is derived from the pricing of wages in the labor market, as a response to the asymmetric information problem. In that sense, we propose a micro-foundation of the income-distribution, that also could serve to rule out the zero top-rate result.

1.1.3   Discrete models of labor income tax

Piketty (1997) and Saez (2002) develop an alternative form of the discrete Mirrlees model with a finite number of possible earnings levels, and therefore, of possible jobs $n = 0, ..., N$, but a continuum of individual types so that the fraction of individuals at each earnings level is a smooth function of the tax rate.

More formally, individual i chooses n so as to maximize a function $u_i(c_n, n)$ where $c_n = y_n - T(y_n)$ is the after-tax consumption level. For a given tax and transfer system which implements $\{c_n\}$, a fraction $p_n(\{c_n\})$ of individuals chooses occupation n, p_n being differentiable.

Assuming that an individual can only work in two adjacent occupations, $n-1$ and n, and that there are no income effects, if we denote by n the occupational choice of individual i, and by $\tau_n = \dfrac{T_n - T_{n-1}}{y_n - y_{n-1}}$ the marginal tax rate between earnings level y_{n-1} and y_n, being $\epsilon_n = \dfrac{1 - \tau_n}{p_n} \dfrac{\partial p_n}{\partial 1 - \tau_n}$ the elasticity with respect to $1 - \tau_n$ of the fraction p_n of individuals in job n, the government chooses $\{T_n\}$, so as to maximize the social welfare function, according to

$$\frac{\tau_n}{1 - \tau_n} = \frac{1}{\epsilon_n} \left[\frac{\sum_{m \geq n}(1 - \mu_m)p_m}{p_n} \right]$$

In spite of the dissimilarities between this model and our model of continuous effort and non-restricted incentive compatibility constraint-set, there exists an important similarity between the above tax formula and the one we derive in Chapter 3.

To make the point clearer, it suffices to note that referring to the probability of realizing different states of income is the same as considering a population of measure one of individuals, fractions of which, at an equilibrium, realize different income levels. The main point is that those fractions, as the discrete probabilities in our model, are endogenous.

1.2   Should fringe-benefits be taxed?

Since it may be argued that, in addition to a non-linear income tax, the government could potentially implement a differentiated taxation on commodity goods, we appeal here to a well established result in optimal taxation theory, to rule out the taxability of the perquisite good in our model, though we do not explicitly derive it.

To clarify the point, consider an economy consisting of N differentiated consumption goods $q = (c, q_1 ..., q_{N-1})$ and pre-tax prices $k = (1, k_1, ..., k_{N-1})$. Individual i derives utility from the consumption bundle q and earnings y, according to an utility function of the form $u_i(y, q)$.

Atkinson and Stiglitz (1976) model of commodity taxation shows that, under the assumptions of weak separability between q and y in $u_i(\cdot, \cdot)$, and homogeneity in the agents' preferences for q, such that $u_i(y, q) = U_i(y, g(q))$, a differentiated commodity taxation is useless as long as earnings are taxed according to a non-linear tax function. As we shall see, conditional on the level of effort, these conditions are equally satisfied in our setting.

The reason behind the result is that, under those assumptions, consumption choices for bundle q provide no information on individuals' ability. If the tax system were more progressive (relative to the standard case), a high ability individual would choose to work less and consume the same bundle q to mimic the income level and preferences of a low-ability agent. Both individuals would pay the same income tax and have the same after-tax consumption, but the high ability agent would end up enjoying more leisure. Differentiated commodity taxes, therefore, create further distortions with no benefit. At the limit, hence, it is optimal to have a uniform zero taxation of all consumption goods, and to attempt redistribution solely through individual, non-linear income taxes alone. However, as pointed out by Diamon and Mirrlees (1971) and by Diamond (1975), with linear income taxes alone, differential commodity taxation can still be useful to non-linearize the tax system. Given this remark, we design our framework so as to implicitly account for a positive transfer to the agents at the bottom of the income distribution, though we do not explicitly solve for it.

A further issue is imposed, however, by the homogeneity of preferences. Since the theorem states that goods for which high-income people have a relatively stronger taste, i.e. goods which are more complementary to leisure, should be taxed more heavily, we model heterogeneity in our models in the form of differences in the production technology, rather than of individuals' utility function.

Moreover, we must remind that we focus on a class of perks that, by construction, are thought of as pure work-related goods, in that being more close substitutes to leisure than wages.

If Atkinson and Stiglitz (1976)'s argument easily applies to our model in Chapter 2, to justify the non-taxability of perks in the decentralized economy studied in Chapter 3, we appeal to Diamond and Mirrlees (1971) model with endogenous prices, and to their result of an optimal fiscal policy maintaining production efficiency. As the latter authors show, despite the need to distort consumption choices to raise revenues (or to redistribute), it is always optimum not to distort the production side, since taxes on those intermediate goods that are not settled for final consumption (i.e. goods the individuals cannot buy on the market), by affecting firms' transactions, would distort production, and therefore, efficiency. Since we assume that the perk good, which is acquired by the firm in the external market, serves as an intermediate good, by ruling out those scenario in which it may serve as a consumption good under perfect information, Diamond and Mirrlees (1971)'s result provides a rationale for imposing no tariff on perks.

1.3 The first-order approach (FOA)

To some extent, our work is in close relationship with the work by Abraham, and Pavani (2008) on optimal income taxation and hidden savings. Their paper studies the two period version of a dynamic moral hazard model where agents can borrow and save on a risk-free bond market and their asset decisions are not observable to the planner. In spite of the differences between the agent's hidden accumulation of savings, and the principal's provision of perks as substitutes to wages, both settings share the same difficulties of a possible deviation from the efficient consumption plan.

The main contribution of Abraham, and Pavoni (2008) lies in providing sufficient conditions under which the first-order approach (FOA) is valid. The FOA constitutes in replacing the planner's incentive compatibility constraint-set by the corresponding first-order necessary conditions derived from the agents' maximization problem. Rogerson (1985) and Jewitt (1988) provide conditions for the validity of the FOA in the class of static principal-agent models. Certainly, the simplification of the constraint-set becomes even more important in a dynamic environment wherein the principal faces additional information problems, as it happens when the agent has access to an hidden savings technology. Abraham,

and Pavoni (2008)'s strategy is to show that if the agent's problem is globally concave when facing the optimal contract, the first-order conditions are either necessary and sufficient for the optimality of the agent's decisions, so that none potential benefit is derived from jointly decreasing effort and increasing savings. In addition to the conditions that are required in the static and observable savings case, (i.e. the monotonic likelihood ratio property and the concavity of the distribution function), they show that, within the family of non-increasing absolute risk aversion (NIARA) utility functions in consumption, either a strong concavity condition on the distribution function or a strict requirement on the convexity of the disutility of effort guarantee that the FOA is applicable.

As they due, we rely on the first-order approach to characterize the optimal insurance contract with a provision of perks in Chapter 2, and compute how much dispersion the optimal consumption exhibits at the varying of income, when perks are optimally provided. As we shall see, a likewise restriction on the relationship between the Frisch labor elasticity and the concavity of the technology for perks is actually required for the problem being jointly concave in effort and perks.

Furthermore, Abraham and Pavoni (2008)'s main conclusion is that, whenever the agent's utility function exhibits decreasing and convex absolute risk aversion (HARA), the possibility of the hidden asset accumulation makes the optimal consumption a more convex function of income. Hence, under hidden asset accumulation the optimal tax system becomes more regressive (or less progressive) compared to the case where asset accumulation is observable.

In that regard, though we follow a standard utilitarian approach rather than to solve a revenue-maximizing problem, we find that, whenever perks are provided in positive amount, the more regressive tax-transfer system, which is set to support them, could result in welfare losses, if it were implemented at those income levels for which the agents value the provision of insurance more than they value production efficiency. In that, in our model, the coefficient of absolute risk-aversion plays even a greater role than it does in Abraham, and Pavoni (2008).

A final remark is owed, which concerns the relevance and robustness of our predictions. Though our analytical results for the two-income-level model can be easily extended to a framework with more than two states and continuous-income (as long as a monotonic likelihood ratio property is satisfied), closed form solutions for the model in Chapter 3 have been derived only for the special case

where the substitutability/complementarity between perks and wages is ruled out by a restriction on the probability distribution function. Whenever wages are settled by the firm independently from the agents' taste for the perquisite good and its monetary cost, perks do affect welfare only through their direct effect on effort. Absent any substitution between cash payments and fringe-benefits, therefore, our results can be interpreted as a lower-bound of the effect of perks on the economy.

1.4   A sketch of the overall results

To the best of our knowledge, this thesis constitutes the first step toward the theoretical study of how the provision of perks may affect the progressivity of the optimal tax-transfer system and the optimal top income tax rate, in the class of standard moral-hazard models of social insurance (Chapter 2), and strategic interaction games between the fiscal authority and the labor-market agents (Chapter 3).

The two chapters, though clearly related, are independent. If the aim of the first model is to investigate the centralized (socially optimal) allocation and the effect of perks on the designing of the optimal taxation scheme, the analysis proposed in the last chapter focuses on the decentralized equilibrium outcomes that arise in the labor market and on the interests that may lead the optimal tax to discourage or support the provision of perks.

For its static character, the model developed in Chapter 2 may be thought of as the asymptotic steady state of a real economy, indexed by its technology. Our results firstly show that there may be a scope for providing perks under moral hazard that exceeds the one with perfect information. Secondly, we prove that, from an ex-post perspective, the equilibrium marginal-rate of substitution between cash and perks is always, at all levels of consumption and for all states of the world, greater than its first-best value. In spite of this ex-post inefficiency, however, we find that the provision of perks may be socially efficient if their complementarities with the agent's effort allow to offset the negative effects on the labor supply that naturally arise at high income levels.

Moreover, our analysis suggests that, whenever perks are efficiently provided, the government actually trades off progressivity for perk provision. Our main conclusion, in fact, is that the second-best efficient taxation scheme with perks is (more) regressive compared to its perk-less equivalent.

As the agent's risk aversion increases, however, the provision of perks that would have occurred at low income levels according to that logic becomes inefficient. As we show, in fact, at lower states, when the revenues collection problem is more severe, the optimal tax system supporting a positive provision of perks is such that transfers are more progressive, and taxes are less regressive than in the perk-less scenario. Since these changes occur at those states at which the agent values insurance the most, providing perks is actually inefficient.

Notably, in the richness of the results derived in Chapter 3 as an equilibrium solution to the interaction game between the fiscal authority and the agents, a similar prediction can be found. In fact, the main contribution of that chapter consists in showing that, whenever society's preferences for the public expenditure and agents' risk aversion are sufficiently low, a regressive marginal tax rate is consistent with a positive provision of perks if, in relaxing the incentive compatibility constraint of the agent, perks do not make her worse off. On the other hand, however, if agents are sufficiently low risk averse, the equilibrium top-income marginal tax rate that is consistent with a positive provision of perks (and it is so even when preferences for the consumption of the public good are larger) is found to be progressive.

To some extent, we also contribute to the literature based on the sufficient-statistics approach, by proposing a simple tax-formula, which builds on a new theoretical explanation that explicitly accounts for the provision of perks. Its novelty should be found on its replacing the elasticity of the reported income with the combination of the elasticity to the tax rate of effort (which account for the extensive margin of the top-earners' labor supply) and wages (which captures the intensive margin of the labor supply, net, however, the portion of it that is repaid by perks).

Finally, we must mention that, in spite of the ability of our theoretical framework to capture the relationship between risk aversion and efficient provision of perks through their effect on the optimal trade-off between the insurance and incentive motives of taxation, our model has its weakness. In fact, either analytically and by means of numerical simulations, we show how our results are sensitive to the stochastic dominance properties of the probability distribution and the value of the state in the economy. If, in the former case, deviations from the benchmark distribution are qualitatively important but quantitatively small, in the latter case, the quantitative implications of different income levels are quite significative.

A lot of work must still be done, therefore, to assess the ability of our model to

jointly account for the the facts on wages and perk provision, the dynamics of tax revenues and the level of government's expenditure. A precise calibration of our models is at a need, along with a better characterization of the results in terms of the degree of stochastic dominance implied by the income distribution.

Chapter 2

Optimal labor income taxes with perks: a mechanism design approach

That the compensation of employees may contain non-monetary ingredients is a well established fact. These fringe-benefits (or perks, as they are called) though common among chief executives, are certainly not exclusive to them; nor they are a phenomenon specific to large and private corporations.

There is a long-lasting debate on whether agents are over- or under- provided with perks, and whether perks and cash are complements or substitutes in providing incentives. Nonetheless, the literature on non-monetary compensation does not deal, for the most part, with the specific issue considered here: namely, how perks may affect the progressivity of the optimal tax/transfer system. If it is well known, in fact, that top managers receive a large part of their remuneration in terms of perks, it is a far less clear point how the expenditure in perks may affect the design of the optimal labor income tax and, therefore, the distribution of the agents' income and their well-being.

Though the largest part of the literature has addressed the issue on the provision of perks through static versions of a moral hazard problem between a risk-neutral principal (the firm) and a representative agent (the worker), we take here an alternative route, and instead of assuming discrete effort choice and profit-maximizing behavior, we approach the matter by developing a static version of a stochastic Ramsey's problem, where an utilitarian, resource-seeking government

takes the post of the principal and owner of the only firm in the economy. When designing the optimal rules for cash (consumption, so far) and perk provision, the government, who is also constrained by his own budget balance, takes into account the agent's unobservable reaction to the tax system.

It is the different technology for perks and money what gives a novel advantage to the government, while dealing with a risk-averse, utility-maximizer agent, whose effort choices affect - in a non-deterministic way - the realizations of that income on which the government leaves state-contingent distorting taxes.

To investigate our point, we focus on a class of pure work-related goods, with no productive attributes, which we define as any non-taxable, non-monetary compensation that the agent is not allowed to sell. Examples include (i) personal business machines, such as cars, computers, laptops and mobile phones; (ii) workplace amenities and personal services, such as pleasant working environment, coffee machines, meal tickets and gyms; (iii) transportation services, such as automobiles, airplanes, and limousines.

The key assumption in modeling this kind of goods is that there are utility complementarities between perks and effort, in the usual sense that consuming more of the former good increases the utility (here, decreases the disutility) of the latter. No farther distortion, such as the misusing of perks, is taken into account.

To deal with the static nature of our problem, we assume that the government runs a balanced budget in expected terms, as opposite to a state-by-state budget constraint[1]. We justify our choice by assuming that the government - not the agent - has access to an abroad insurance/financial market, that allows him to smooth taxes/consumption across income realizations, when it is optimal to do so. For its static character, this model should be thought of as the asymptotic steady state of a real economy, indexed by its technological parameter, wherein the benefits from perks, if any, shall be found in their decreasing the agent's disutility of effort.

Methodologically, this paper relies on the first-order approach (FOA) to replace the effort incentive compatibility constraint with the first order necessary condition derived from the agent's problem. This allows us to characterize our results in terms of the associated best response functions. To keep things as simple as possible, we assume that preferences are additive separable in consumption and effort, and let the agent's labor disutility be strictly convex. We also assume a

---

[1]Though we limit ourself to the case of just two states of the world, our analysis extends to any multi-state problem and, in the limit, to the continuum case.

linear stochastic process for income realizations, either in effort and in productivity, and see how the (weak) income and substitution effects propagate through the agent's optimal effort responses to either the government's instrument. with respect to the government's instruments. Though sufficient conditions for the validity of the FOA in this class of models have been recently identified by Abraham and Pavoni (2008), and those are also met in our setup, we argue that the existence of perks, and the weak separability between perks and effort, imposes an even stronger restriction on the convexity of effort which, though related to the concavity of the perk technology, is not affected by the cost of perks.

Although relying on the FOA means more complex calculations than the discrete effort-choice option does, it allows us to analyze the issue at hand by addressing a few questions that naturally arise from observing the existence of perk. Namely, if firms are willing to pay in perks, workers are willing to accept them and the financial authority is willing to regulate them, how strongly is their provision tied to social welfare? How does the optimal provision of perks depend upon the cost of perks and the government resource-requirement? What is the relationship between the optimal amount of perk and the the agent's labor supply sensitiveness to taxes? Do (not cost-minimizing) perks help to overcome the inefficiency of distorting taxes?

Before discussing the optimal mechanism in presence of moral hazard, we provide a benchmark model with perfect information, where perks may be used, along with cash, to reward the agent's observable effort. Our analysis suggests that, when effort is observable, the government reaches his objective by letting the agent be the full-residual claimant of all her expected income, above the public expenditure requirement. The system of optimal transfers and income taxes fully insures the agent against the systemic risk. As to regard to perks, we argue that, even in a world with perfect information, perks might in principle be provided in positive amounts because of their "being a different good". Though our results replicate the well-established equality, at an optimum, between relative price and marginal rate of substitution between cash and perks, we discuss the conditions under which there may be a room for the provision of perks with moral hazard, that there is not with perfect information, and see that, whenever there is a scope for providing perks in the first best, there must be then also a scope for providing perks under asymmetric information.

When applied to the moral hazard setting, our analysis firstly shows that, from an ex-post perspective, the equilibrium marginal rate of substitution between money and perks is always, at all levels of consumption and for all states of

the world, greater than its value of first best. From a normative prospective, however, we argue that, because of the complementarities that arise between cash and perks when taking into account the agent's optimal responses, there exists none income tax (even non-proportional) nor excise on perks such that we can replicate the result of first-best.

Secondly, we notice that the agency problem causes an additional distortion, which makes it less costly for the government to implement high effort by perks. Indeed, the substitutability between perks and leisure, and the complementarity between perks and cash that arise from the agent's optimal response functions are the driving forces of our model. However, this latter force is counterbalanced by an opposite force, which requires the government's budget constraint to bind. Though giving more perks increases the responsiveness of effort to cash payments and, therefore, the efficiency loss from distorting taxes, it also reduces the ex-post disposable resources, so enhancing a negative income effect that makes the agent more willing to work on her-own. If such a negative income effect is large enough, an increase in the cost of perks, for example, by increasing the expected level of income, could also allow for a more progressive tax/transfer system.

In a quantitative exercise we test the ability of our model to keep track of all these dynamics. Imposing the restrictions on parameter values we derived from our analytical analysis, we compute the optimal allocation with perks and compare it to the solution for the perk-less economy. Our simulations suggest that the provision of perks reduces the progressivity of the tax system supporting the equilibrium consumption levels. The main economic intuition behind our result is as follows. The utility complementarities between perks and effort have an incentive effect, which also causes an increase in the sensitivity of labor supply to monetary payments. The higher that elasticity is, the greater is the distorting effect of any progressive income tax, which in turn explains why, in an economy with perks, the supporting tax system is more regressive. Moreover, as long as perks are optimally provided the optimum amount of perks depends on either i) the differential of equilibrium consumption levels (inversely), and ii) the the consumption dispersion (directly). Put it differently, from a social point of view, it is optimum for the government to trade off progressivity for perk provision.

The rest of the chapter is organized as follows. Section 2.1 briefly reviews some of the most recent works on perk provision and optimal compensation schemes. Section 2.2 describes the setup of the model and introduce the concept of inter-

and intra- state progressivity of the tax system. Section 2.3 characterizes the optimal allocation (with and without perks) of first-best and provides some useful comparative statics results. Section 2.4 discusses the optimal consumption plan and provision of perks under asymmetric information. It presents the main analytical results of the paper as to regard the channels through which perks may affect the progressivity of the optimal income tax. Section 2.5 puts the model at work in a quantitative exercise. Section 2.6 concludes.

2.1 Perks in the literature

For the most part, the literature on perks does not deal with the specific issue considered here, namely, how perks do affect optimal income taxation. Neither does the literature on optimal taxation and multiple goods, because of the technology differences that there exist between any commodity good, either used as a consumption good or a production factor, and the perquisite good.

Moreover, the question at hand is far from having no empirical relevance, as recent surveys suggest. A research conducted by Equilar, for the Mercury News, reports on the compensations received by several Silicon Valley's chief executives, in 2006. It indicates that Larry Ellison, Oracle's CEO, received almost two millions in 'perks and other compensations', and that Meg Whitman (eBay) landed more than one million, while several other CEOs received $500,000 or more. In Grinstein et al. (2009), of a random sample of 361 firms belonging to the S&P 1500 Index for the years 2006–2007, 90% has been found to provide perks to their top five executives, with a mean-annual value of $296,300.

If empirical work have focused on the compensation of high profile CEOs and its effects on market capitalization and market values (see Yermack (2006), Jensen and Murphy (1990), Rajan and Wulf (2006)), others have approached the problem theoretically, by developing different versions of a standard moral hazard problem.

In the financial literature, the standard explanation for perks relies on some form of agency problem (see Jensen and Meckling (1976) and Bebchuk and Fried (2003)). Perks are used by managers to appropriate some of the surplus generated by the firm, in a way that is neither approved nor acknowledged by shareholders. This is beneficial for them, but detrimental for welfare. We think, however, that monitoring concerns, which are the key factor of those models, are less reasonable when dealing with a fiscal authority, that, for the most part,

may monitor the provision of perks better than the provision of cash.

A second group of researchers consider perks as productive goods. This view suggests that perks are useful instruments to align the objectives of principal and agents, so that the investment in them is efficient. Other authors, however, while thinking of perks as goods that, in Rosen's (2000) terminology, have no "productive" attributes but are uniquely defined by their complementarity with an agent's effort, do focus on the issue of perk optimal provision and ex-post efficiency.

To this group belongs the work by Bennardo, Chiappori and Song (2010). They show that whenever perks are substitute to leisure (or even less complementary to leisure than money), the optimal incentive scheme involves over-provision of such commodities, in the sense that the agent consumes more of them than what she would decide to if given a choice between money and perks at the current market price. Such perks can profitably be used for pure incentive purposes, even when they generate no productivity gains. Our paper, though building on the same assumption, goes a little farther and tries to relate that ex-post inefficiency in providing perks to the behavior of the agent's response functions. On the same ground stands the paper by Weinschenk (2013), according to which the principal may either underinvest or over-invest in perks, depending on whether the work-relatedness of perks does exceed or not a measure of the performance precision. Most significantly, the author finds that if perks harm the agent, the principal never underinvests and may over-invest in perks. The opposite happens if perks increase the expected utility of the agent.

Despite the differences between the last two models, they share the same assumption that principal and agents are opposite one the others, and that all the bargaining power is held by the former. In the present paper, we show that excessive perks can appear, from an ex-post perspective, even when the provision of perks is decided by an utilitarian principal. We are more concerned, however, about understanding how that phenomenon is linked to the agent's optimal behavior as to regard taxation. We argue that the agent may like perks less than money and yet she may found them of some advantage if they glean the efficiency loss that arises in presence of an agency problem.

Among those that approach perks as productive goods is the work developed by Marino and Zabojnik (2008). They study a static moral hazard model with technological perks and an liability-unrestricted, risk-averse agent. Differently from our model, where perks play no role in the production function, in their model perks are either a consumption good and a productivity enhancement

tool. Moreover, due to their specification, though perks do not affect (the monetary equivalent of) the agent's effort disutility (as we assume), effort does increase (the monetary equivalent of) the utility from consuming the perk good. The complementarities between perks and effort, either in the production function and in the consumption utility, capture the controversial idea that an employee is likely to derive either a larger output and a greater utility the more intensively the perks are used in the production process. Those complementarities have an incentive effect that allows the principal to decrease the pay-performance sensitivity and the uncertainty in the agent's income, thus leading to a lower expected pay. However, the resulting substitutability between perks and wages that they derive is quite mechanical, being driven by the way volatility comes along in the model. Moreover, due to the agent's unlimited liability, the principal's problem falls into one of maximizing the (certain equivalent of the) total expected surplus, so that the provision of perks is always socially optimal and neutral for an agent's well being. As a corollary, they also find that the agent shall never be charged with the cost paid for perks by the principal. However, we argue that, if it is the case, it is so because no resource constraint is assumed to hold on the principal's side, the principal being the residual claimant of the agent's income. It is, therefore, on the account of these dissimilarities from our model, that we discussed the effect of changes in the cost of perks for the agent's well-being.

For its understanding of perks as a disutility-reducing device, Kvaloy, and Schottner (2012) is the work most related to ours. They characterize a static moral hazard model, with a risk-neutral agent and continuous effort, in which the principal can take a costly motivational action in order to reduce the worker's effort cost. They distinguish two cases: in the first case, the principal itself chooses the amount of motivation and bears its cost; in the second one, the principal delegates the provision of motivational effort to another agent (a senior or motivator), charges the motivator by the motivational cost and pays him back a bonus, depending on the performance of the subordinate, low-ranking worker. Their main contribution is a complete discussion of the conditions (on the concavity of the agent's disutility of effort) under which monetary payments and perks are substitutes in reducing the agent's cost of effort, and the characterization of a solution where banding the objective of two different subjects helps to by-pass the course of asymmetric information.

Aiming at understanding the role of perks when such an asymmetric information there does not exist, Carrothers, Han and Qiu (2012) develop an equilibrium

matching model for a competitive CEO market, where CEOs' wages and perks are both determined by bargaining between firms and CEOs. In their model, (i) the production technology is firm-agent specific: it depends upon the firm's size, the CEO's talent, while it may account or not for perks as a production input; (ii) CEOs have Cobb-Douglas preferences over cash and perks, and they are liability constrained; (iii) there is neither moral hazard, nor asymmetric information in the model. As it would be expected, their results show that firm size, wage, perks and talent are all positively related. Interestingly, they also found that perks are more sensitive than wage to changes in firm size when there are economies of scale in the cost of providing them. This suggests that the designing of perk cost plays a big role in characterizing the result even under perfect information, regardless of the interaction between perks and effort.

2.2 The stochastic economy

Our model builds on a typical static moral hazard problem, with a risk-neutral principal (the government) and a risk-averse agent (the worker), and assumes that consumption of perks occurs at the beginning of the period (simultaneously with the effort decision), while the consumption of the numeraire good (cash) only occurs at the end of it, when the state of the economy is realized and uncertainty is resolved. Moreover, we assume that either type of consumption cannot be negative.

We can interpret this extended setup with perks in several different ways[2]. Our most preferred interpretation is that of an optimal tax/transfer provision problem (social insurance), where the principal is a benevolent social planner whose objective is to maximize the welfare of the representative citizen. The planner offers a tax/transfer system supplemented by the provision of perks to insure her against the systemic risk and, at the same time, provide her appropriate incentives for working hard.

---

[2]First, our framework may describe a private franchising agreement, where the agent, as a franchisee, can affect future outcomes by exerting unobservable effort. The franchisor is the planner, who offers a contract which implies for the franchisee an initial service displacement from the franchisor (the provision of perks) against the payment of a fee, which depends on the realized state. Another interpretation is as a two-stage compensation contract, where the difference in the income states plays the role of the surplus the principal and the worker may be able to share. In this case, $c_s(y_s)$ is the wage-rate for state s. Wages have to provide the right incentives for the agent to exert the right effort. Moreover, since punishments (lower wages) for state-0-income realization cannot be implemented, the principal can incentivize effort through the payment of perks, if he finds optimal to do so.

2.2.1 Technology, information set and markets. Consider a small open economy, consisting of a one-period-living risk-averse representative agent (She) and an infinitely living risk-neutral government (He). The productivity of this economy is denoted by $\theta \in \Theta \equiv [\underline{\theta}, \overline{\theta}] \subset \mathcal{R}$, which we assume common knowledge of either the agent and the government. The agent likes consumption $c \in [0, \bar{y}]$ and perks $q \in [0, \bar{y}]$, and dislikes effort $n \in [0, \bar{n}]$, with both $\bar{y}$ and $\bar{n}$ being non-negative and finite. From her effort, the agent earns a gross income $\mathbf{y} = \{y_0, y_1\}$, with $0 \leq y_0 < y_1 \leq \bar{y}$, that can be either consumed by the agent or used by the government to finance his expenditures. We call y_s the state of the economy. We maintain that income realizations are publicly observable, verifiable and, therefore, contractable.

The technology of this economy consists in a collection of kernels $\langle p_0, p_1; \forall \theta \in \Theta \rangle$ that maps the agent's effort n into a probability measure over the set of income realizations. We denote by $p_s(n, \theta)$ the conditional probability that the agent in the economy θ realizes income y_s, i.e. $p_s(n, \theta) \equiv \text{Prob}\{y = y_s | n, \theta\}$.

Throughout all this chapter we make the following assumption about the income generating process.

Assumption(PD, Probability Distribution): The discrete density functions are such that

- $p_1(n, \theta) = \alpha(n + \theta)$, with $\alpha \in (0, 1]$

- $0 < p_1(n, \theta) < 1$, for all θ and $n \in [0, \bar{n})$, where $\bar{n} := \{\tilde{n} | \ p_1(\tilde{n}, \theta) = 1\}$, and $p_0(n, \theta) = 1 - p_1(n, \theta)$

It follows that $p'_1 \equiv \dfrac{\partial p_1(n, \theta)}{\partial n} = \alpha > 0$, while $p'_0 \equiv \dfrac{\partial p_0(n, \theta)}{\partial n} = -\alpha < 0$.

That is, the probability of realizing any y_s is either linearly increasing or linearly decreasing in n. Moreover, for $\theta > 0$ and all $s = \{0, 1\}$, p_s is linearly increasing (or decreasing) in θ as it is in effort. The random factor α, which - of the two likelihood ratios (LR) $\dfrac{p'_s}{p_s}$ - only affects $\dfrac{p'_0}{p_0}$, also measures the effect of changes on n and θ on the probability function p_s.

Sufficient restrictions will be imposed on the parameter values, such that, given θ, $p_1 \in (0, 1)$ is satisfied at an optimum. To simplify notation, we drop productivity as an argument and denote by $p_s(n)$ what shall be understood as $p_s(n, \theta)$.

We also assume that the agent has not access to financial markets: she cannot insure herself against the idiosyncratic risk (living in an economy whose productivity is θ), nor against the systemic risk (α), unless she participates to the government's system.

2.2.2 The government. The government finances his constant purchases G, with $y_0 < G < y_1$, by levying type-and-state contingent taxes on labor income alone, $T_s(y_s) \equiv y_s - c_s$. Perks are not taxable.

The government is forced to run a balanced budget that only holds in expected terms:

$$\sum_s p_s(n)\ (y_s - c_s) - G - kq = 0 \qquad \text{(GBC)}$$

Here, c and q should be understood as the equilibrium quantities of the numeraire good (money or consumption) and perks, respectively. We assume that one unit of real value paid by perks costs $k > 0$ to the government, with k being the relative price of q in terms of money. k can be thought of as the marginal cost of the perk good, added any efficiency loss that may be due to either i) transformation costs, and ii) bargaining with the agent about their provision.

It is wort noting that a consumption plan that makes expression (GBC) negative would not be possible, since the government would not find anybody willing to be on the lending side of the financial transactions that serve him as an insurance device. To keep things as simple as possible, we neglect the costs associated with the government financial activity or pretend they are already included in G.

In designing the optimal fiscal policy, the government - who maximizes a social welfare function defined over the agent's preferences for cash, perks and effort - is constrained from a non-negative consumption requirement:

$$\forall s, \qquad c_s \geq 0 \qquad \text{(LL)}$$

We also notice that, whenever (GBC) binds in equilibrium, the expected income tax is always equal to the government's expenditure, $G + kq^\star$.

2.2.3 The tax system. In our setting, income taxation obeys two purposes. Firstly, since the agent is risk averse, we expect the benevolent social planner to smooth some of the income risk the agent faces. This is the insurance motive of taxation. Though from Abraham and Pavoni (2008)'s moral-hazard model

we already know that the constraint efficient allocation varies from that in a "pure bond economy", i.e. from the allocation the agent can obtain by insuring herself through borrowing and lending, the question still remain, in our optimal taxation setting, on whether the planner can provide additional insurance by means of the perk good. This result would imply that perks have a role in enhancing welfare compared to a pure perk-less economy.

The second motive for taxation builds on incentives. In offering the tax/transfer system, the social planner also tries to provide the agent with appropriate incentives for exerting high effort. Though working hard may be beneficial for the agent as well, it is certainly necessary for the planner himself, since he is constrained by his budget balance. When high punishments (low consumption levels) for low income realizations cannot be imposed, providing perks may be a valuable alternative in order to relax the agent's incentive compatibility constraint.

Intuitively, the full support assumption we made plays a major role in balancing the two motives for taxation. Under this condition, in fact, the agent has incomplete control over income realizations. Hence, the government can implement schemes that impose a tax payment in one state and a transfer in the other in such a way that the agent is not able to avoid paying taxes with certainty. Still, however, the agent may be allowed to enjoy - with certainty - the provision of perks.

No restrictions are imposed on the functional form of $T_s(\mathbf{y})$, that will be derived from the optimal plan for consumption. Clearly, the curvature of the c_s functions is closely related to the progressivity of the tax system. Thus, in order to study the pattern of the optimal tax system, we allow y_1 to increase - from $y_{1,i}$ to $y_{1,i+1}$- and look at the curvature of the optimal consumption allocation.

Following Abraham, and Pavoni (2008), we define the following,

Definition(TS, Tax System): We say that the tax scheme is:

- "inter-state" progressive (regressive) if $\dfrac{c_{1,i} - c_{0,i}}{y_{1,i} - y_0}$ is decreasing (increasing) in $y_{1,i}$;

- "intra-state" progressive (regressive) if $\dfrac{c_{1,i+1} - c_{1,i}}{y_{1,i+1} - y_{1,i}}$ is decreasing (increasing) in $y_{1,i+1}$.

This definition implies that whenever consumption is a convex (concave) function of income we have a regressive (progressive) tax system supporting it. In

a progressive tax system taxes are increasing faster than income does. At the same time, for the state in which the agent is receiving a transfer, transfers are increasing slower than income decreases. The opposite happens with a regressive tax scheme.

If the scheme is progressive, incentives are provided by imposing "large penalties" at low income levels. On the other hand, if the scheme is regressive, incentives are provided by larger rewards for high output realizations. If the scheme is proportional these rewards and punishments are balanced.

The next proposition reports on a well-establish result on the literature on optimal taxation and first-order approach

Proposition: (AP, Proposition6) Assume the FOA is justified, and that the optimal $\langle c_0, c_1 \rangle$ is interior at $q = 0$. Then, if it holds that i) the likelihood ratio (LR) is monotone and convex (concave), and ii) $\dfrac{1}{u_c(\cdot)}$ is concave (convex) in c, and that iii) the absolute risk aversion coefficient (ARA) is decreasing and convex (constant), then the optimal tax scheme is regressive (progressive)[3].

2.2.4 The agent's preferences and strategy. The representative agent chooses $n \in [0, \bar{n}]$ in order to maximize her expected utility

$$U^{\langle p \rangle}(\mathbf{c}, q; n) = \sum_s p_s(n) \, u(c_s) - g(n, q) \qquad \text{(AEU)}$$

subject to:

$$\forall s, \qquad c_s = y_s - T_s(y_s)$$

where $u(c) : [0, \bar{y}] \to \mathcal{R}$ is any non-decreasing, strictly concave, twice-continuously differentiable function of c, that is $u_c(\cdot) > 0 > u_{cc}(\cdot)$; $g : [0, \bar{n}] \times [0, \bar{y}] \to \mathcal{R}$ is strictly increasing and strictly convex in its first argument, n, and strictly decreasing and convex in its second arguments, q. That is, $g_n > 0 < g_{nn}$ and $g_q < 0 < g_{qq}$, for all $n > 0$ and $q \geq 0$. For $n = 0$, $g(0, q) = g_n(0, q) = 0$.

The complementarity between q and n, which is the key feature of our model, is what makes of perks a purely work-related good.

---

[3]When the LR is convex, CRRA with $\eta \leq 1$ induces regressive schemes since ARA is decreasing and convex and $1/u_c(\cdot)$ is concave. This case includes the logarithmic utility function, for which linear LR, rather than convex LR, leads to proportional schemes.

2.2.5 Allocation and Ramsey's problem

Definition(Feas, Feasibility): A feasible allocation is a collection $\langle c_0, c_1, n, q; \forall\theta \rangle$ such that expressions (GBC), (LL) and $q \geq 0$ are satisfied.

Definition(FP, FiscalPolicy): A government's policy is a collection of functions $\langle T_0(y_0), T_1(y_1); \forall\theta \rangle$ such that $c_s \equiv y_s - T_s(y_s)$, for all $s \in \{0,1\}$.

Definition(FB, FirstBest): A FB optimal equilibrium allocation is a feasible allocation, a government's policy, a price and technology system $\{\langle k, G\rangle, \langle p_s\rangle\}$ such that: i) given the government's policy, the price and technology system, the allocation maximizes expression (AEU); ii) given the allocation, the price and technology system, the government policy satisfies expressions (GBC) , (LL) and $q \geq 0$.

There are many equilibrium allocations, indexed by different government expenditures, technologies and productivities. Following Lucas and Stokey's (1983) argument, we let the government choose directly the optimal feasible allocation, and discuss the progressivity of the optimal tax system by looking at the equilibrium consumption levels.

We call the optimal taxation problem a Ramsey problem. The next section states the problem and characterizes the solution for the first best. Before doing so, however, we assume the following functional form for the agent's disutility of effort,

Assumption(N, Disutility of effort): $g(n, q) = \dfrac{\delta n^{1+\sigma}}{1+\sigma} \dfrac{1}{(1+aq)^{\gamma}}$.

- $\{\gamma, \delta, a\} \gg 0$ and $\sigma \geq 1$.

Accordingly, $g(n, 0) = \dfrac{\delta n^{1+\sigma}}{1+\sigma}$, which is the standard isoelastic disutility of effort. Because of this, any comparison with the optimal allocation in the perk-less economy becomes straightforward. Furthermore, by this specification form, we aim at shaping the curvature of the perk-technology by means of two parameters, a and γ, that are directly comparable with δ and σ. Here, in fact, a accounts for the work-relatedness of perks and, along with γ, it controls for their degree of diminishing returns.

2.3 The Ramsey's problem with PI

When effort is observable, the government's problem can be stated as follows:

$$(\text{P1}) \qquad \max_{q,n,\{c\}} \sum_s p_s(n) \, u(c_s) - g(n,q)$$

subject to:

$$\sum_s p_s(n) \, (y_s - c_s) - G - kq = 0 \qquad\qquad (\text{GBC})$$

$$\forall s, \quad c_s \geq 0 \qquad\qquad (\text{LL})$$

$$q \geq 0$$

The Lagrangian to (P1) is therefore as follows

$$\mathcal{L} = \max_{n,q,\{c\}} \left\{ -\frac{\delta n^{1+\sigma}}{1+\sigma} \frac{1}{(1+aq)^\gamma} + \sum_{s=0}^1 p_s(n) \left[u(c_s) + \lambda \, (y_s - c_s) \right] \right.$$

$$\left. -\lambda(G + kq) + \sum_{s=0}^1 \chi_s c_s + \xi q \right\}$$

where λ is the non-negative Lagrange multiplier on (GBC), and $\{\chi_s\}_{s=0}^1$ and ξ are the non-negative Lagrange multipliers associated to (LL) and $q > 0$, respectively. For fixed G and $\langle p_s \rangle$, we maximize $\mathcal{L}$ with respect to $\langle n, , q, c_0, c_1 \rangle$.

2.3.1 The analysis

The first-order conditions to problem (P1) state

$c_s:$ $\qquad p_s(n) \left[u_c(c_s) - \lambda \right] \leq 0, \ \forall s \geq 0 \qquad$ if $<$, then $c_s^\star = 0$

$q:$ $\qquad \dfrac{\delta n^{1+\sigma}}{1+\sigma} \dfrac{\gamma a}{(1+aq)^{\gamma+1}} - \lambda k \leq 0, \qquad$ if $<$, then $q^\star = 0$

$n:$ $\qquad \sum_s p_s' \left[u(c_s) + \lambda(y_s - c_s) \right] - \dfrac{\delta n^\sigma}{(1+aq)^\gamma} \geq 0, \qquad$ if $>$, then $n^\star = \bar{n}$

$\lambda:$ $\qquad \sum_s p_s(n) \, (y_s - c_s) = G + kq$

To interpret these conditions, we first note that the Lagrange multiplier λ equals the rate of change of the maximum value of the agent's expected utility with respect to changes in the value of the public expenditures. In other words, λ is the marginal social cost of resources. The first order condition with respect to c_s states that, for any $s: \ c_s > 0$, this marginal cost shall equate the marginal gain

from consuming c_s. With perfect information, the features of the distribution function do not affect this mechanism. Actually, neither the limited liability does. To see why, let denote by $\hat{s}$ the state $s: c_s > 0$, and let be $\bar{c} \equiv c_{\hat{s}}$. The focs with respect to $\bar{c}$ and c_s jointly state:

$$\frac{1}{u_c(c_s)} \geq \frac{1}{u_c(\bar{c})}, \qquad \text{if } >, \text{ then } c_s^\star = 0$$

This condition, that holds whatever $p_s(n)$ is, implies that, as long as there is at least one state for which the agent is worthy to be paid a positive amount of cash, it is optimal to supply her with a constant consumption stream. When it is the case, the optimal $\bar{c}$ is uniquely defined from expression (GBC), i.e.:

$$\bar{c}^\star = \sum_s p_s(n^\star) y_s - G - kq^\star$$

Therefore, whether perks are provided or not does no affect the necessary condition for optimality with respect to c_s, but just the level of the optimal constant consumption level, throughout the government's budget constraint.

Since the optimal consumption is bounded, either from below and from above, so it is the optimal tax/transfer. That is, the agent is full-insured against the systemic risk, she being the full-residual claimant of all her expected income above the public expenditure, in either G and q.

From the foc with respect to effort, we see that a necessary condition for a strictly positive solution to exist is that there is at least one state $y_s > 0$, whose probability increases with effort. The optimal effort is then the one that maximizes the agent's expected utility at the constant consumption level $\bar{c}$. The following proposition states it more formally.

Proposition1: Fix problem (P1), and let Assumption(PD) holds. Then, the optimal first-best allocation has effort $n^\star > 0$ and consumption $c_s = \bar{c} \geq 0$, $\forall \ s = \{0,1\}$, where $\bar{c}$ is the expected income, net of the government's equilibrium expenditure in either G and perks.

The system of optimal transfer and income taxes, which is bounded from above, with bound $\bar{T}_1 = y_1 - \bar{c}$, either i) fully insures the agent against the systemic risk, ii) it makes the agent be the full-residual claimant of all her expected income above $G + kq^\star$.

With $c_0 = c_1 = \bar{c}$, the (LL) constraint reduces to $\bar{c} \geq 0$, with multiplier $\chi \geq 0$. We can immediately rule out $\bar{c} = q = 0$, since this implies government's finances

are in surplus. Since the utility function is strictly increasing, all available income will be spent at the optimal point (no interior solution). Therefore, the budget constraint must hold with equality. For any utility function that satisfies Inada's conditions, moreover, if it were $\bar{c} = 0$, then $\lambda = +\infty$ and $n = \bar{n}$ would still imply $q = 0$, which is a violation of optimality. At an optimum, therefore, either the set of effective constraint is $E_1 = \{(GBC)\}$, or it is $E_2 = \{(GBC), q = 0\}$.

2.3.1.1 The optimality condition for perks

With $\lambda > 0$, the foc with respect to q says that, if the maximum occurs at any positive amount of q, the marginal return of perks, i.e. the marginal reduction in the agent's disutility of effort, must be equal to the marginal cost of perks, as expressed in terms of consumption utility, i.e λk. By combining the two first-order conditions, the one with respect to $\bar{c}$ and the other with respect to q, we obtain the well known equality result between the marginal rate of substitution of $\bar{c}$ in terms of q and their relative price.

$$\frac{u_c(\bar{c})}{U_q} = \frac{1}{k}$$

To look at it differently, from the first-order conditions (taken with equality) and the government's budget constraint, we can write

$$\lambda = \frac{u_c(\bar{c}) \cdot \bar{c} + |g_q| \cdot q}{\sum_s p_s(n) y_s - G}$$

The first relation means that at any optimal constrained allocation the marginal utility *per dollar* must be the same at both margins, i.e. for $\bar{c}$ and q. The same marginal utility per dollar must occur when the expected *extra revenue*, i.e. $\sum_s p_s(n) y_s - G$, is spread out over both consumption and perks, as when it is spent in either good. That is, the rate of change of utility with respect to any additional revenue must be the same at each margin and at either margin simultaneously.

This non-surprising result is due to centralization and perfect information. Under these assumptions, perks are certainly not a "prior good" with respect to cash. Only their relative price is what matters (to either the worker and the government), and it is their relative price that determines their marginal value and, therefore, the marginal rate of substitution.

Substituting out the optimal consumption plan and the government's budget constraint, problem (P1) reduced to

$$\text{(P1-R)} \qquad \mathcal{L} = \max_{n,q} \ u \left[y_0 + \alpha(n + \theta)(y_1 - y_0) - G - kq \right]$$

$$- \frac{\delta n^{1+\sigma}}{1 + \sigma} \frac{1}{(1 + aq)^\gamma} + \xi q$$

The focs to this problem are, therefore,

$$\mathcal{L}_n : \qquad u_c(\bar{c})\alpha(y_1 - y_0) - \frac{\delta n^\sigma}{(1 + aq)^\gamma} \geq 0 \qquad \text{if } >, \text{ then } n^\star = \bar{n}$$

$$\mathcal{L}_q : \qquad -u_c(\bar{c})k + \frac{\gamma a \delta n^{1+\sigma}}{(1 + \sigma)(1 + aq)^{\gamma+1}} \leq 0 \qquad \text{if } <, \text{ then } q^\star = 0$$

From the government's budget constraint, and Assumption(PD), we already know that

$$\bar{c} = y_0 + \alpha(n + \theta)(y_1 - y_0) - G - kq$$

Together, $\mathcal{L}_n$ and $\mathcal{L}_q$ imply that the optimal amount of perks is adjusted depending on the cut-off level

$$k_1(n_0) \equiv \frac{\gamma a \alpha(y_1 - y_0) n_0}{1 + \sigma}$$

Therefore, if $k \geq k_1(n_0^\star)$, where $n_0^\star = n^\star|_{q=0}$, then, $q^\star = 0$ and perks are never part of a first-best solution. Unfortunately, this condition involves all the parameters of the model, so that it can be argued that the result for perks it is a matter of parametrization. Indeed, it is. Besides, the scope from providing perks with perfect information, which is strongly related to the decreasing marginal returns of consumption, may fail to catch the full content of their provision. We defer, however, to next section the discussion of the rationale behind the provision of perks, while dealing with the asymmetric information problem. Nonetheless, it is worth noting that, if $n_0 > \dfrac{1 + \sigma}{\gamma a}$, then $k_1 > \alpha(y_1 - y_0)$ and for $k \in (\alpha(y_1 - y_0), k_1(n_0))$ perks may be provided even if they are technologically inefficient. Moreover, by the dependence of k_1 from n_0, it is always possible to find a value of δ such that either case holds[4].

Proposition2: Fix problem (P1-R), and let Assumptions (PD) and (N) hold.

---

[4] At $q^\star = 0$, the equilibrium values of $\bar{c}$ and n solve the following system of equations

$$\begin{cases} \bar{c} = \alpha\theta(y_1 - y_0) - (G - y_0) \quad + \alpha(y_1 - y_0)n \\ \bar{c} = u_c^{-1} \left[\dfrac{\delta n^\sigma}{\alpha(y_1 - y_0)} \right] \end{cases}$$

where $u_c^{-1}[\cdot]$ denotes the inverse of the marginal utility of consumption.

1. If $q^\star_{FB} = 0$, then it holds that $k \geq k_1 \equiv \dfrac{\alpha(y_1 - y_0)a\gamma}{1 + \sigma} n^\star_0$;

2. If $q^\star_{FB} > 0$, with $q^\star = \max\left\{ \left(\dfrac{\delta n^{\star 1 + \sigma}}{a(1 + \sigma)u_c(\bar{c}^\star)k} \right)^{\frac{1}{1+\gamma}} - \dfrac{1}{a}, \ 1 \right\}$, then $k < k_1$.
   Moreover, at any interior solution, perks are adjusted so that the marginal rate of substitution between cash and perks is equal to their relative prices, $1/k$.

3. If $\gamma < \sigma$, the above necessary conditions are also sufficient.

4. If the disutility of effort and the perk technology are such that the optimal effort in the perk-less economy satisfies $n_0 > \dfrac{1 + \sigma}{\gamma a}$, then, for all $k \in (\alpha(y_1 - y_0), k_1(n_0))$ perks may be provided in the first best even if they are technologically inefficient.

The proof of the statement in point 3 of Proposition2 is discussed in Appendix 2A. It builds on the negativeness of the Hessian matrix to problem (P1-R) and it ensures that the government's problem is jointly concave in either effort and perks, at any interior solution for perks. The economic content of point 4, instead, is the following. As long as our welfare criterium is utility, the decreasing marginal returns of consumption leave rooms for a positive provision of perks, even if they are not a cost-minimizing tool. Nonetheless, for this being the case, the optimal effort implementable without perks must be already sufficiently high. In that, we say that perks may be more effectively used as a consumption good.

2.3.2 Comparative statics results

The complete discussion of the comparative statics analysis may be found in Appendix 2A. Here, we limit ourselves to a brief statement of the main results, that will serve as a reference point when discussing the dynamics that lead to the second-best optimal allocation.

Corollary1: Fix problem (P1-R) and let Assumptions (PD) and (N) hold. The optimal first-best allocation is such that:

1. if perks are optimally not provided, then the optimal effort decreases in either y_1, α and θ;

2. if $q^\star > 0$, then the equilibrium effort level, $n^\star$ may either decrease or increase in both y_1 and α, depending on i) the agent's absolute risk aversion, ii) the agent's preferences for perks. $q^\star$ always increases in either y_1 and α, while both effort and perks decrease in θ and k.

3. If preferences and technology are such that both effort and perks show high decreasing marginal returns, i.e. if $|\mathcal{L}_{nn}|, |\mathcal{L}_{qq}| > \mathcal{L}_{nq}$ for all n, q, then at any interior solution, $\dfrac{\partial n}{\partial y_1} - \dfrac{\partial q}{\partial y_1} < 0$, $\dfrac{\partial n}{\partial \theta} - \dfrac{\partial q}{\partial \theta} < 0$ and $\dfrac{\partial n}{\partial k} - \dfrac{\partial q}{\partial k} > 0$.

To understand the strength of Corollary1, we must keep in mind that in this framework, an increase in either y_1, α and θ does not affect the relative price of consumption with respect to perks. This is so because, with perfect information, the agent receives a constant payment that is the same in either state. Variations in the optimal allocation, therefore, results from a pure income effect, which is negative for effort and positive for perks. As the cost of perks increases, the substitution effects that push for providing less perks and more consumption, in reason of the complementarity of perks and effort, are the leading forces in explaining the drop in effort. In fact, either the income effect and the substitution effect that follow an increase in k would have act so as to reduce leisure.

Clearly, when effort is observable, the planner, by choosing either the optimal effort and the optimal amount of perks simultaneously, takes advantage of the complementarities between effort and perks in order to offset the income effect. However, when the convexity of effort is sufficiently high and the technology for perks show very high decreasing marginal returns those income effects seem to dominate.

2.4   The Ramsey's problem with MH

This section extends our analysis of the optimal taxation problem to the case where the agent's effort is neither observable, nor contractable. The governments wants to maximize the agent's expected utility, subject to his budget balance and the agent's incentive compatibility constraint. This constraint requires that the effort the planner wants to induce results from the agent's utility-maximizing behavior, given the terms of the social contract. That is, since the government cannot observe the agent's effort, he must design a mechanism such

that it is in the agent's self-interest to apply the effort that she is meant to implement.

As before, we assume that the agent, who is cut off from all insurance/asset markets, may be provided with either cash and any non-negative amount of the single, non-taxable, work-related good that is called perk.

We analyze the problem with moral hazard in three steps. In step 1 (Section 2.4.1), we assume that the contract $\langle c_0, c_1, q \rangle$ is fixed, and study the utility-maximization behavior of the agent given the contract. In step 2, we derive the optimal fiscal policies as to regard consumption (Section 2.4.2) and perks (Section 2.4.3). Precisely, we will highlight the mechanisms that may affect the curvature of the optimal consumption fact, by deriving a simple expression for our index of progressivity. Then we will discuss the conditions under which perks may be provided with moral hazard when they are not under perfect information, and characterize the shape of the agent's indifference curves at any equilibrium allocation. Unfortunately, because of the non-linearity of our problem, we cannot characterize the optimal contract much further analytically. Numerical simulations are the scope of Section 2.5. There, we combine the payment rules with the agent's incentive compatibility constraint and the government's budget balancing order to derive the optimal incentive-compatible allocation $\{\langle c_0, c_1, q \rangle, n\}$ as an equilibrium outcome.

2.4.1 Step 1: The agent's problem

The agent has preferences over streams of consumption levels, $\mathbf{c} = \{c_0, c_1\}$, with $0 \leq c_0 \leq c_1 \leq \bar{y}$, perks $q \in [0, \bar{y}]$, and effort choices, $n \in [0, \bar{n}]$, that are described by an expected utility function of the form

$$U^{\langle p \rangle}(\mathbf{c}, q; n) = \sum_s p_s(n)\, u(c_s) - \frac{\delta n^{1+\sigma}}{1+\sigma} \frac{1}{(1+aq)^\gamma}$$

where $u(\cdot)$ is a strictly increasing, strictly concave and twice-continuously differentiable in class of CRRA utility function with relative risk aversion coefficient $\eta > 0$, and the parameters are such that $\{\gamma, a, \delta\} \gg 0$ and $\sigma \geq 1$.

The agent's maximization problem

$$\max_n U^{\langle p \rangle}(\mathbf{c}, q; n)$$

implies the following first- and second- order conditions

$\mathcal{L}_n:$ $\qquad \sum_s p'_s u(c_s) - \dfrac{\delta n^\sigma}{(1+aq)^\gamma} \geq 0,$ $\qquad\qquad$ if $>$, then $n = \bar{n}$

$\mathcal{L}_{nn}:$ $\qquad -\dfrac{\sigma \delta n^{\sigma-1}}{(1+aq)^\gamma} < 0$ $\qquad$ for all feasible n, q

In the spirit of the first-order approach (FOA), the first inequality replaces the agent's incentive compatibility constraint (IC). Implicitly, it defines the agent's effort as a function of either monetary and non-monetary incentives, i.e. $n = n^\star(\mathbf{c}, q)$. Standard comparative statics analyses yield to the following functions,

$$\frac{\partial n}{\partial c_s} = \frac{p'_s u_c(s)}{\sigma \delta n^{\sigma-1}}(1+aq)^\gamma \qquad \mathrm{sign}\Big(\frac{\partial n}{\partial c_s}\Big) = \mathrm{sign}(p'_s)$$

$$\frac{\partial n}{\partial q} = \frac{\gamma a}{\sigma}\frac{n}{(1+aq)} \qquad > 0 \qquad \text{for all feasible } n, q$$

The optimal response to consumption, $n_s \equiv \dfrac{\partial n}{\partial c_s}$, therefore, is higher, in magnitude, the lower is the convexity of effort and the higher are q and the marginal utility of consumption in state s, the latter being larger the smaller are either c_s and η. However, for larger q, given n, $n_q \equiv \dfrac{\partial n}{\partial q}$ is smaller. By simple computation, we derive

$$\mathrm{sign}\Big(\Big|\frac{\partial n}{\partial c_s}\Big| - \frac{\partial n}{\partial q}\Big) = \mathrm{sign}\Big((1+aq)^{\gamma+1} - \frac{\gamma a \delta n^\sigma}{\alpha u_c(c_s)}\Big)$$

Because of the dependence of n from either c_s and q, it is in general difficult to say what the sign of the difference $|n_s| - n_q$ may be. However, some sufficient conditions can be found whenever $\bar{n} \leq 1$ holds. We state them in terms of Lemmas.

Lemma1: If $\alpha u_c(\bar{y}) > \gamma a \delta$, then $n_c > n_q$, $\forall q \geq 0$.

For $u(\cdot)$ in the class of CRRA functions, if we also assume $\bar{y} = 1$, the above condition reduces to $\dfrac{\alpha}{\delta} > \gamma a$. Thai is, when a net measure of the profitability of effort, namely α/δ, is larger that the direct marginal effect of the first unit of perks, then the effort responsiveness to cash payments is higher than the one to perks.

The reversed inequality, however, do not imply that $n_q > n_c$ for all or some $q \geq 0$. Nonetheless, a simple case arises when $\gamma = 1$. Indeed,

Lemma2: If $\gamma = 1$, then $\Big|\dfrac{\partial n}{\partial c_s}\Big| > \dfrac{\partial n}{\partial q}$ if and only if $q^2 + \dfrac{2}{a}q + \dfrac{1}{a^2}\Big(1 - \dfrac{\gamma a \delta n^\sigma}{\alpha u_c(c_s)}\Big) >$
$\qquad$ 0. Therefore, if $\alpha u_c(c_{s(0)}) > g_{nq}(n_0, 0) = \gamma a \delta n_0^\sigma$, then $n_c > n_q$ for all

$q \geq 0$. If $\alpha u_c(c_{s(0)}) < \gamma a \delta n_0^\sigma$, then $\exists \bar{q} > 0 \; : \; n_q \geq n_c \; \forall q \in [0, \bar{q}] \; \& \; n_c > n_q \; \forall q > \bar{q}$, with $\bar{q} = \dfrac{1}{a}\left[\sqrt{\dfrac{a\delta n^\sigma(\bar{q})}{\alpha u_c(c_s(\bar{q}))}} - 1\right]$ endogenous.

Though the first-order derivatives suffice to characterize a candidate solution to our problem, we need the second-order partial derivatives of the agent's choice-function to study the global concavity of the government's problem. We do it in Appendix 2.C. Nonetheless, before going farther with our analysis, it is worthy to stress out the behavior of the optimal n as regard to q.

Lemma3: $n_{qn} \equiv \dfrac{\partial^2 n}{\partial q \partial n} > 0$ and $n_{qq} \equiv \dfrac{\partial^2 n}{\partial q^2} < 0$, while sign$\left[n_{sq}\right]$ = sign $\left(p'_s\right)$, for all $\mathbf{c}, q, n$, with $n_{sq} \equiv \dfrac{\partial^2 n}{\partial c_s \partial q}$.

The economic content behind Lemma 3 is the following. The contribution of perks as an incentive-device is higher the higher is the effort that the agent is asked to implement, though this effect shows diminishing returns.
In spite of the lack of any complementarity between effort and consumption and between consumption and perks in the agent's utility function, because of the stochastic nature of our problem, the agent's responses to monetary incentives also depend upon the amount of non-monetary ones. In the literature on perks, to $n_{1q} > 0$ corresponds a hidden benefit of reward, since monetary incentives complement and enhance the effect of perquisite goods on effort and vice-versa. We argue, however, that when an increase of q makes labor supply more elastic as to regard cash, income taxation (whatever its purpose might be) is more distorting.

Understanding the dynamics implied by Lemma 2 and 3 is quite important, since they are relevant in driving our results. In fact, the following lemma and proposition link the scope for providing incentives (the tightness of the moral hazard problem) to the optimal fiscal policy through the the effort responsiveness to either consumption and perks.

Lemma4: Fix the agent's incentive compatibility constraint of problem (P2). The partial derivatives of $n^\star$ with respect to the government's instruments also imply that the total derivative of n with respect to y_1 is given by

$$\frac{dn}{dy_1} = \Phi(\mathbf{c}, n, q)\left[\frac{dc_1}{dy_1} - \frac{u_c(c_0)}{u_c(c_1)}\frac{dc_0}{dy_1}\right] + {}_{q>0}\Psi(n, q)\frac{dq}{dy_1}$$

where $\Phi(\mathbf{c}, n, q) = \dfrac{(1+aq)^\gamma \alpha u_c(c_1)}{\delta \sigma n^{\sigma-1}}$ and $\Psi(n, q) = \dfrac{n\gamma a}{\sigma(1+aq)}$.

Here, the term $\dfrac{u_c(c_0)}{u_c(c_1)}$ depends on either i) the curvature of the $u(\cdot)$ function, and ii) the degree of insurance provided by the contract; while the total derivatives $\dfrac{dc_s}{dy_1}$ capture the progressivity/regressivity of the supporting tax system. To see why, let first notice that in our setup, as it is in the literature on optimal taxation with moral hazard, but neither perks nor public expenditure expenditure, any progressive (respectively, regressive) tax system is such that, by definition[5], there exist an income level $\tilde{y} \in [y_0, \bar{y}]$ for which the following hold:

a) for $y > \tilde{y}$, either $\dfrac{dc_1}{dy_1} < 0 \ (> 1)$ and $0 < \dfrac{dc_0}{dy_1} < 1 \ (> 1)$; and

b) for $y < \tilde{y}$, $\dfrac{dc_0}{dy_1} > 1 \ (< 1)$.

Nonetheless, along with this, we must account for the compliances that naturally arise from assuming the existence of a government's budget constraint. In fact, if it is clear that any increase in state-s consumption, by reducing the marginal utility in that state, relaxes the incentive compatibility constraint and, therefore, modifies how the gains of allocating utility vary across states, it is also obvious that any relaxation of the incentive compatibility constraint does change the shadow cost of the government's revenues. At low income levels, therefore, when the marginal social value of wealth is indeed larger, the optimal tax system shows - all other things equal - a higher degree of regressivity. Intuitively, we derive our result from the feature of the utility function which, by satisfying the Inada's conditions, prevents large (potentially infinite) punishments and also makes the provision of incentives at low consumption levels less-costly.

For a progressive tax system, the existence of a wasteful public expenditure implies an even more negative responsiveness of c_1 to y_1, at low income levels. On the other hand, for a regressive tax system, for which that responsiveness can be positive even at low income gap, our theory only implies that $\dfrac{dc_1}{dy_1}\Big|_{y=y_0} < 0$. Nonetheless, by the continuity of the solution, there must be an interval $\mathcal{I}$ such that $\dfrac{dc_1}{dy_1} < 0$ for at least all $y \in [y_0, G] \subseteq \mathcal{I}$.

The following Proposition, by summarizing our result, completes the proposition discussed in Section 2.2.3.

---

[5]See AP (2006) and the discussion in Section 2.2.3 of this paper.

Proposition3: Assume the FOA is justified and that the optimal $\langle c_0, c_1 \rangle$ is interior at $q = 0$. Then,

1. If it holds that i) the likelihood ratio (LR) is monotone and convex (concave), and ii) $\dfrac{1}{u_c(\cdot)}$ is concave (convex) in c, and that iii) the absolute risk aversion coefficient (ARA) is decreasing and convex (constant), then the optimal tax scheme is regressive (progressive)

2. Whenever the government's expenditure is such that $G > y_0$ and it is wasteful, all conditions i), ii) and iii) in point 1 being satisfied, the regressivity of the optimal tax scheme increases - as compared to the case where $G = 0$ -, at those income levels for which the marginal social value of wealth is positive.

At low income levels, i.e. at those states at which the marginal value of resources is higher for the principal, the more progressive the tax system is, the more negative is the effect on implementable effort of increments in the income gap. Put it differently, if the agent were facing a progressive tax system, the government could still be able to offset the negative income effect on effort by providing her with perks, though the effect of increasing perks can be subject to diminishing returns. However, at higher income levels, for which the progressivity of the tax system implies that $\dfrac{dc_s}{dy_1} \to 0$ for all s, since those negative effects are reduced in magnitude, the problem of diminishing returns in perks is less severe.

The same logic applies when the agent faces a regressive tax system. In this case, the negative effects of income are mechanically reduced on account of the higher weight that the increased ratio of marginal utility puts on the reduced sensitivity of c_0 to y_1 at low income levels. To counteract the reduction in effort by providing perks is then less costly, since less perks are required, and those are less subject to diminishing returns. Moreover, at high income levels, the regressive tax system is such that $\dfrac{dc_0}{dy_1} \to 0$, while $\dfrac{dc_1}{dy_1} > 1$. On account of that, the principal can obtain a positive sensitivity of effort to y_1, by adjusting optimally the provision of either state-1 consumption and perks, both having a positive effect on $\dfrac{dn}{dy_1}$.

Proposition4: Fix the agent's incentive compatibility constraint of problem (P2). If the tax system were

1. progressive, then the principal may counteract the negative effect of income on effort by providing the agent with perks. Moreover, the amount of perks needed to offset that effect is decreasing in y_1. However, at low y_1, the positive effect of perks is humbled by their diminishing marginal returns.

2. regressive, then the negative effects of income on effort are reduced at low income levels and they cancel out for y_1 sufficiently high. Therefore, the principal can implement a positive sensitivity of effort to y_1 by optimally adjusting the provision of either perks and state-1 consumption.

To conclude this section, we notice that, in the present set-up, the effect of effort on the marginal utility of perks is always positive, being $U_{qn} > 0$ for all $n > 0$. That is, except from moral hazard, there are no further distortions, such as misusing of perks. Though it can be argued that such misbehavior is part of any work-related good, we abstract from this problem to focus on the intrinsic ability of perks of moving the agents closer to the Pareto-optimum when asymmetric information further constrains the government's problem of collecting resources and providing insurance.

2.4.2 Step 2: The government's problem w.r.t. consumption

In this second step of our analysis, we assume that the government chooses the values of the contract $\langle c_0, c_1, q \rangle$ so as to maximize the agent's expected utility, subject to: i) budget balance, ii) the agent's incentive compatibility constraint, iii) limited liability, and iv) the non-negative constraint on perks. The government's problem can be stated as follows,

$$(\text{P2}) \qquad \max_{q,\mathbf{c}} \sum_s p_s(n^\star)\, u(c_s) - \frac{\delta n^{\star 1+\sigma}}{1+\sigma} \frac{1}{(1+aq)^\gamma}$$

subject to:

$$\sum_s p_s(n^\star)\, (y_s - c_s) - G - kq = 0 \qquad\qquad (\text{GBC})$$

$$n^\star = \left\{ n : \sum_s p_s' u(c_s) - \frac{\delta n^\sigma}{(1+aq)^\gamma} = 0 \right\} \qquad\qquad (\text{IC})$$

$$\forall s, \qquad c_s \geq 0 \qquad\qquad (\text{LL})$$

45

$$q \geq 0$$

After substituting out the agent's choice function (so dropping expression (IC) from the set of constraints), the Lagrangian to this problem is

$$\mathcal{L} = \max_{c_0, c_1, q} \left\{ \sum_s p_s(n^\star) \, u(c_s) - \frac{\delta n^{\star 1+\sigma}}{1+\sigma} \frac{1}{(1+aq)^\gamma} + \lambda \left[\sum_s p_s(n^\star) \, (y_s - c_s) - G - kq \right] \right.$$

$$\left. + \sum_s \chi_s c_s + \xi q \right\}$$

where $\lambda \geq 0$ represents the Lagrange multiplier associated with the government's budget constraint, and χ_s and ξ are the non-negative multipliers associated with (LL) and $q > 0$, respectively. The choice variables of problem (P2) are $\langle c_0, c_1, q \rangle$.

Definition(SB, Second Best): A SB optimum equilibrium is a feasible allocation, a fiscal policy, a price and a technology system such that: i) given $\{T_s(\cdot)\}$ and $\{\langle k, G \rangle, \langle \mathbf{p} \rangle\}$, the allocation maximizes (AEU); given the allocation,$\langle k, G \rangle$ and $\langle \mathbf{p} \rangle$, ii) $\{T_s(\cdot)\}$ satisfies (GBC), (LL) and $q \geq 0$, iii) the compensation scheme $\mathcal{W} \equiv \langle c_0, c_1, q; \forall \theta \rangle$ implements the effort policy $n^\star(\mathbf{c}, q; \theta)$ as implied by (IC).

2.4.2.1 The optimal consumption plan

Using the comparative statics results derived earlier, and by defining $\mathcal{L}_n \equiv \lambda \left[\sum_s p'_s \, (y_s - c_s) \right]$, the first order conditions with respect to c_s and q are,

$$\mathcal{L}_c: \qquad \mathcal{L}_n \frac{\partial n}{\partial c_s} + p_s \left[u_c(s) - \lambda \right] \leq 0, \qquad \text{if } <, \, c_s = 0$$

$$\mathcal{L}_q: \qquad \mathcal{L}_n \frac{\partial n}{\partial q} + \frac{\delta n^{1+\sigma}}{1+\sigma} \frac{a}{(1+aq)^{\gamma+1}} - \lambda k \leq 0, \qquad \text{if } <, \, q = 0$$

$$\mathcal{L}_\lambda: \qquad \sum_s p_s(n) \, (y_s - c_s) = G + kq$$

Though the marginal utility derived from consuming c_s is a function of c alone, because of the absence of any complementarity between cash and perks, the optimality condition with respect to c_s is affected by q throughout the impact of perks on the agent's reaction function with respect to c_s. In order to maximize his objective, therefore, the government takes into account how each of his instrument affects (simultaneously) the agent's behavior.

As we are interested in any equilibrium at which $n^* > 0$, we can conveniently neglect the non-negative constraint on c_1. In the following, we assume that an interior solution for c_0 does exist. [6] Accordingly, there are only two sets of all the possible constraints that are effective at an optimum, either $E_1 = \{(GBC)\}$ or $E_2 = \{(GBC), q = 0\}$. The following analysis applies to both.

Let be $\Delta \equiv \sum_s p'_s(y_s - c_s)$. Then, $\mathcal{L}_n = \lambda \Delta$. Optimality conditions for c_1 and c_0 thus imply,

$$\frac{1}{\lambda} = \frac{1}{u_c(c_1)} - \Delta \frac{(1 + aq)^\gamma}{\sigma \delta n^{\sigma - 1}} \frac{p'_1}{p_1}, \qquad (1)$$

and

$$\frac{1}{u_c(c_1)} - \frac{1}{u_c(c_0)} - \Delta \frac{(1 + aq)^\gamma}{\sigma \delta n^{\sigma - 1}} \left(\frac{p'_1}{p_1} - \frac{p'_0}{p_0} \right) \leq 0, \qquad \text{if } <, c_0 = 0 \qquad (2)$$

The following Lemmas can be derived.

Lemma5: At a SB optimal allocation, the equilibrium level of effort is inefficient. Moreover, perfect risk shearing can not be achieved.

To prove the statement, firstly we notice that the efficient level of effort, the one at which $\mathcal{L}_n = 0$, requires either $\lambda = 0$ or $\Delta = 0$. Since of the two requirements, the former does violate optimality, through its implying $GBC > 0$, it remains to discuss if at an optimal allocation, $\Delta = 0$ may be an equilibrium outcome. However, from the first-order conditions with respect to either c_0 and c_1, if $\Delta = 0$, then $u_c(c_s) = \lambda$ for all s such that $c_s > 0$. Therefore, by (2), for either realized income level, either the corresponding cash payment is constant, or $c_1 = y_1 - y_0$ and $c_0 = 0$. In the former hypothesis, the incentives compatibility constraint would be only binding at $n^* = 0$; in the latter, $GBC = 0$ would be violated.

In other words, when the government needs to provide agents with incentives, to choose a positive effort level, risk sharing cannot be his only consideration. With both λ and Δ strictly positive, the agent's consumption will generally vary with income, trading off some risk-sharing benefits for incentive provision. As a

---

[6]It is precisely the case when $u(\cdot)$ is CRRA with parameter $\eta \geq 1$. In this case, in fact, if it were $c_0 = 0$, (IC) whould imply that, for any feasible c_1 and q, the agent chooses $n = \bar{n}$. Moreover, since, by construction it must hold that $p_1(\bar{n}, \cdot) < 1$, the agent's expected utility is $U(0, c_1, q, \bar{n}) = -\infty$, which is infinitely less than what she can get for any $c_0 > 0$. Therefore, we would expect a benevolent government to set $c_0 > 0$.

For all $u(\cdot)$ such that $\eta < 1$, instead, the infinite marginal utility of consumption at zero also implies a violation of optimality.

by-product, the equilibrium effort, though optimum, is below the level at which it maximizes the government's objective, that is $\mathcal{L}_n > 0$.

Lemma6: Whenever $\dfrac{p'_s}{p_s} \neq \dfrac{p'_{\hat{s}}}{p_{\hat{s}}}$, the optimal payment to the agent increases with the value of the likelihood ratio, weighted by a positive factor, which depends on n, q and c. Furthermore, the higher is c_0, the larger is c_1 of equilibrium.

The proof of the statement follows from (2). As we see, the difference between the inverse marginal utilities is negatively related to the difference in the likelihood ratios. Moreover, since $p'_0 < 0$, optimality conditions imply that $u_c(c_0) > u_c(c_1)$, and, therefore, by the concavity of $u(\cdot)$, that $c_0 < c_1$.

With bounded resources, this consumption structure reflects the logic of statistical inference altogether with that of incentive provision. Rising n by rising c_1 contributes to greater efficiency. On the other hand, a contract with a greater monetary incentive component will tend to lower the expected utility because of diminishing returns in consumption, leaving room for the provision of perks.

Finally, we notice that, by a revealed preference argument, using additional signals (so as the productivity) will increase the government's objective as long as these signals have an independent effect on the likelihood ratio (as it is when $p_s(n, \theta) = \alpha_s(n + \theta)$). Moreover, since the likelihood ratio decreases in θ we will aspect the optimal monetary compensation to be affected by two forces: the first one decreases c_1 in θ. It holds anytime the rationale for monetary compensation is to incentivize work, and effort and productivity have no positive cross-effect. The second one pushes for increasing c_1 in θ, as a consequence of a weakly income effect.

The following Proposition summarizes the results as to regard the optimal consumption plan.

Proposition5: Fix Problem (P2) and let Assumption (SD) hold. Then, the stream $\{c_s\}$ will vary with $\{y_s\}$, trading off risk-sharing for incentive provision. The optimal c_s increases with the equilibrium value of the likelihood ratio $\dfrac{p'_s}{p_s}$. Moreover,

- the difference in the inverse marginal utilities (taken in pairs) follows the difference in the likelihood ratios, as weighted by a factor which depends on $q^\star$.

- since the likelihood ratio decreases in θ, any c_s is affected by two forces: statistical inference vs. income effect.

2.4.2.2 About the convexity of the optimal consumption plan

This subsection introduces our index of productivity. We derive it from the first-order condition with respect to c_0, and the government's budget constraint. Altogether, those imply that, at any interior solution for c_0, the planner set the optimal consumptions according to the following system

$$
\begin{cases}
c_0 = y_1 p_1 \nu + y_0 (1 - p_1 \nu) - G - kq \\
c_1 = y_1 p_1 \mu + y_0 (1 - p_1 \mu) - G - kq
\end{cases}
$$

where ν and μ measure the direct equilibrium responsiveness of , respectively, c_0 and c_1 to expected income $p_1 y_1$, and are define as follows,

$$
\nu \equiv \frac{\delta \sigma n^{\sigma-1} p_1 p_0}{\delta \sigma n^{\sigma-1} p_1 p_0 + \alpha^2 (1 + aq)^\gamma} \qquad \mu \equiv \frac{\delta \sigma n^{\sigma-1} p_1 p_0 + \alpha^2 (1 + aq)^\gamma / p_1}{\delta \sigma n^{\sigma-1} p_1 p_0 + \alpha^2 (1 + aq)^\gamma}
$$

With $0 < \nu < 1$ and $\mu > 1$, it holds that the partial derivatives are such that $\dfrac{\partial c_0}{\partial y_1} < \dfrac{\partial c_1}{\partial y_1}$, for any $n^\star$.

Since ν and μ are linked, we define a unique variable $\mathcal{M} \equiv \dfrac{\alpha^2 (1 + aq)^\gamma}{\delta \sigma n^{\sigma-1} p_1 p_0}$, such that $\nu = \dfrac{1}{1 + \mathcal{M}}$ and $\mu = \dfrac{1 + \frac{\mathcal{M}}{p_1}}{1 + \mathcal{M}}$. By simple substitution in the system of equations, we obtain an expression for our index of interstate progressivity[7]. Namely,

$$
\frac{c_1 - c_0}{y_1 - y_0} = \frac{\mathcal{M}}{1 + \mathcal{M}}
$$

and, therefore,

$$
\frac{c_{1,i} - c_{0,i}}{y_{1,i} - y_0} = \frac{\mathcal{M}(n(y_{1,i}), q(y_{1,i}))}{1 + \mathcal{M}(n(y_{1,i}), q(y_{1,i}))}
$$

---

[7] A similar expression for the index of intrastate progressivity can be obtained by computing $c_{1,i+1}$ and subtracting to it the same expression for $c_{1,i}$. With $\tilde{p}_{1,j} = p_{1,j} \mu_j$, for $j = \{i, i+1\}$, we have that

$$
\frac{c_{1.i+1} - c_{1,i}}{y_{1,i+1} - y_{1,i}} = \tilde{p}_{1,i+1} + \frac{y_{1,i} - y_0}{y_{1,i+1} - y_{1,i}} \left(\tilde{p}_{1,i+1} - \tilde{p}_{1,i} \right) - \frac{k}{y_{1,i+1} - y_{1,i}} \left(q_{i+1} - q_i \right)
$$

We notice, first of all, that the effects on our index of changes in y_1 are all of second-order. That is, since y_1 does not enters the expression for $\mathcal{M}$ directly, progressivity is measured only through the effect of income variations on both the provision of perks and the implementable effort. Moreover, other than from the direct effect on effort, that only accounts if $\sigma > 1$, progressivity is also affected by the features of the probability distribution, since the effort responsiveness of $p_1 p_0$, which can be either positive or negative, depends on n, and on α and θ as well. This complicates our analysis. Nonetheless, for $\theta = 0$, $\mathcal{M} = \dfrac{(1+aq)^\gamma}{\delta\sigma n^\sigma\left[\frac{1}{\alpha}-n\right]}$. For $\theta = 0$, therefore, we obtain

$$\frac{\partial}{\partial y_{1,i}}\left[\frac{c_{1,i}-c_{0,i}}{y_{1,i}-y_0}\right] = \frac{1}{(1+\mathcal{M})^2}\frac{d\mathcal{M}}{dy_{1,i}}$$

where $\dfrac{d\mathcal{M}}{dy_{1,i}} = \dfrac{\partial\mathcal{M}}{\partial q}\dfrac{dq}{dy_{1,i}} + \dfrac{\partial\mathcal{M}}{\partial n}\dfrac{dn}{dy_{1,i}} =$

$$= \frac{(1+aq)^{\gamma-1}}{\delta\sigma n^\sigma\left[\frac{1}{\alpha}-n\right]}\left[\frac{\gamma a}{\sigma}\frac{dq}{dy_{1,i}} - \frac{(1+aq)\left[\frac{1}{\alpha}-\frac{1+\sigma}{\sigma}n\right]}{n\left[\frac{1}{\alpha}-n\right]}\frac{dn}{dy_{1,i}}\right]$$

From the above expression, it is clear how either the sign and the magnitude of the interstate progressivity depend on the responsiveness of both perks and effort to income increases. Unfortunately, since the pattern of the latter changes with the magnitude of the income gap, through the relevance of the resources collecting problem and the tightness of the moral hazard, we cannot characterize the optimal social contract much further analytically without having a better understanding of the dynamics that lead to a positive provision of perks, or without imposing heavy restrictions on the parameters.

The exercise we propose in Section 2.5 aims at exploring the quantitative importance of the provision of perks in assessing the progressivity that arises in equilibrium.

2.4.3 Step 2: The government's problem w.r.t. perks

The optimality condition for the perk good, q , states that

$$\mathcal{L}_q : \qquad \frac{\delta n^{1+\sigma}}{1+\sigma}\frac{a\gamma}{(1+aq)^{\gamma+1}} \leq \lambda\left[k-\Delta\frac{\partial n}{\partial q}\right], \qquad \text{if } <, q^\star = 0$$

with $\Delta = \sum_s p_s'(y_s - c_s)$. Substituting out λ, we can see how the provision of perks is linked to the agent's responses. Let be $U_q \equiv \dfrac{\delta n^{1+\sigma}}{1+\sigma}\dfrac{\gamma a}{(1+aq)^{\gamma+1}}$. Then,

$$\mathcal{L}_q: \qquad \frac{U_q}{u_c(c_s)} \leq \frac{k - \Delta \dfrac{\partial n}{\partial q}}{1 - \dfrac{\Delta}{p_s} \dfrac{\partial n}{\partial c_s}}, \qquad \text{if} <, q^\star = 0$$

which holds for all s, such that $c_s > 0$. Because of its relevance, we state the above result in terms of a Proposition and its Corollary.

Proposition6: Let define $\hat{k} \equiv \left[k - \Delta \dfrac{\partial n}{\partial q} \right] \Big/ \left[1 - \dfrac{\Delta}{p_s} \dfrac{\partial n}{\partial c_s} \right]$. Then, at an optimum perks are adjusted so that their marginal social value in terms of forgone consumption c_s, i.e the marginal rate of substitution MRS_{qc_s}, is equal to (or less than, if $q^\star = 0$) the relative price $\hat{k}$ which is *relevant* when taking into account the efficiency costs from the moral hazard.

By rearranging $\mathcal{L}_q$, we can see how deeply the provision of perks is linked to the agent's responses. Namely,

$$\mathcal{L}_q: \qquad MRS_{sq} \geq \frac{1}{k} - \frac{\Delta}{k} \left[\frac{1}{p_s} \frac{\partial n}{\partial c_s} - \frac{u_c(c_s)}{U_q} \frac{\partial n}{\partial q} \right], \qquad \text{if} >, q^\star = 0$$

That is, whether the MRS_{sq} (between c_s and q) of second-best is larger or smaller than its value of first-best (namely, $\dfrac{1}{k}$), it depends on the difference, at $q^\star$, $n^\star$ and $c^\star$, of the weighted derivatives of n with respect to c_s and q. In principal, therefore, both scenarios may apply. Because of our specification form, however, we can show that

$$MRS_{sq} \geq \frac{1}{k} + \frac{\displaystyle\sum_s \alpha(y_s - c_s)}{k\delta\sigma n^\sigma} \frac{(1 + aq)^\gamma}{u_c(c_s)^{-1}} \left[(1 + \sigma) - \frac{np_s'}{p_s} \right], \qquad \text{if} >, q^\star = 0$$

That is, given Assumption(PD), provided that $\theta > -\dfrac{\sigma n^\star}{1 + \sigma}$, at any optimal solution (even at $q^\star = 0$) the marginal rate of substitution between c_s and q is larger than the MRS_{sq} of first-best. More generally, the smaller (or negative) is the likelihood ratio evaluated at $n^\star$, the higher is the MRS_{sq} of equilibrium. The optimal second best contract, therefore, is such that at any interior solution for perks, the perk good is over-consumed, in the sense that should the agent have received a cash amount equal to the value of the optimal provision of perks, she would have purchased less perks. In fact, the equilibrium marginal utility of consumption, at any state, is strictly larger than the marginal utility of one dollar of expenditure in perks. That is, from an ex-post perspective, after the realization of either state, the agent regrets her consumption decisions: she

would have liked to receive less perks in the first stage and more of the numeraire good afterwards. Nonetheless, from an ex-ante perspective, efficiency requires perks to be provided in kind and fully consumed under uncertainty, though such an amount results in more perks than what the agent would have freely purchased on the market if that uncertainty were resolved.

Up to this point, however, we cannot infer whether the provision of perks is (even ex-post) harmful to the agent, but just that it is socially optimum (ex-ante) to provide them in larger amount than it is in the first-best. We summarize the previous result in form of a corollary to Proposition6.

Corollary2: Because of the agency problem (i.e. whenever $\mathcal{L}_n > 0$), for any value of the effort functions (such that n_q and n_c are either different from zero) and for all $\theta \geq 0$ or θ negative but sufficiently high, it holds that $MRS_{qs} > k$ and the provision of perks in the second-best is ex-post inefficient. Moreover, if $n_{qc} \neq 0$, then there is no income tax, even non-proportional, nor excise on q, such that the optimality condition of FB, namely $MRS_{qs} \leq k$, is achieved.

The effect of increasing k. Suppose an interior solution for perks do exist. What is the effect on the $MRS_{q,1}$ of increasing k?

We write the adjusted relative price as $\hat{k} = \dfrac{k/\Delta - \dfrac{\partial n}{\partial q}}{p_1(n)/\Delta - \dfrac{\partial n}{\partial c_1}}$. Then, an increase in k raises the opportunity cost of perks with respect to the opportunity cost of paying c_1 (namely, 1) with probability $p_1(n)$. On that account alone, we should expect the government to substitute away from perks, toward more monetary incentives. However, as k increases, disposable resources also decreases. Since leisure is a normal good, we might expect the agent to be willing to work more on herself[8]. But, fixed q, the higher is n, the lower is the true opportunity cost of perks, because either i) $p_1(n)$ is increasing in n; and ii) $n_{qn} > 0$ while $n_{1n} \leq 0$. Therefore, if at some $\bar{k}$, the total income effect is larger than the substitution effect, and $\Delta(\bar{k}) \equiv \sum_s p'_s(y_s - c_s(\bar{k}))$ is sufficiently large, increasing $\bar{k}$ by some amount $dk > 0$ might reduce the marginal rate of substitution between q and c_1. That is, at the new price $\bar{k} + dk$, since the government's value for perks increases, also the provision of perks does increase.

---

[8]We notice that, though $n_{cn} = 0$ for $\sigma = 1$, a weakly income effect propagates also in this case, by the optimal c.

2.4.3.1 On the scope of providing perks

Let denote by $\rho \equiv \left[(1+\sigma) - \dfrac{np_1'}{p_1} \right]$ a measure of the net efficiency-cost of effort, when taking into account the convexity of the agent's disutility of effort and the effort effectiveness in determining the probability of success, $p_1(n, \theta)$. To simplify notation, let be $\mathcal{Q} \equiv (1 + aq)$. Then, the first-order condition for perks can be written as

$$\mathcal{L}_q: \qquad -\frac{k}{\gamma a} \mathcal{Q}^{\gamma+1} + B(n, \mathbf{c}) \frac{\rho}{\sigma} \mathcal{Q}^\gamma + D(n, \mathbf{c}) \leq 0, \qquad \text{if} <, \ q^\star = 0$$

where $B(n, \mathbf{c}) \equiv \dfrac{n\Delta}{1+\sigma}$ and $D(n, \mathbf{c}) \equiv \dfrac{1}{u_c(c_1)} \dfrac{\delta n^{1+\sigma}}{1+\sigma}$.

Since $q^\star > 0$ if and only if $\mathcal{L}_q|_{q=0} > 0$, we state that $q^\star > 0$ if and only if the following condition is satisfied, namely

$$k < k_2(n_0, \mathbf{c}_{(0)}) \equiv B(n_0, \mathbf{c}_{(0)}) \frac{\rho_0}{\sigma} + D(n_0, \mathbf{c}_{(0)})$$

where the subscripts $_0$ and $_{(0)}$ mean that either n, c_0 and c_1 are evaluated at $q = 0$.

By simple comparison between k_1 of first-best and k_2 of second-best, we obtain

$$\frac{k_1 - k_2}{n_0^{MH}} = \frac{\alpha(y_1 - y_0)\gamma a}{(1+\sigma)} \left[\frac{n_0^{FB}}{n_0^{MH}} - \left[\frac{\rho_0}{\sigma} \left(1 - \frac{c_{1(0)} - c_{0(0)}}{y_1 - y_0} \right) + \frac{\delta n_0^{MH} \sigma}{\alpha u_c(c_{1(0)})(y_1 - y_0)} \right] \right]$$

which is:

1. negative for $\rho_0 > 0$ and n_0^{MH} sufficiently high, such that $n_0^{FB}\sigma \approx n_0^{MH}\rho_0$, and $\dfrac{\delta n_0^{MH} \sigma}{\alpha \rho_0} > u_c(c_{1(0)})(c_{1(0)} - c_{0(0)})$;

2. positive for $\rho_0 > 0$ and $u_c(c_{1(0)})(c_{1(0)} - c_{0(0)}) > \dfrac{\delta n_0^{MH} \sigma}{\alpha \rho_0}$;

3. always positive for $\rho_0 < 0$.

We focus meanly on cases 1 and 2, since - due to our specification form - $\rho > 0$ is satisfied at all n. The economic content of the above expression is the following. For $k_2 > k_1$, the price the benevolent social planner is willing to pay for the first infinitesimal unit of perks is lower in first-best than it is with asymmetric information. Thus, there may be values of k in the positive interval (k_1, k_2) for which perks are provided in the second-best when they are not with perfect information.

First of all, we notice that the technology parameter a is of no consequence in signing $k_1 - k_2$, while this difference is decreasing in either α and ρ_0. Therefore, k_2 is eventually greater than k_1, providing more role for perks in the second-best, the larger is θ (because of the negative effect of higher expected income), and the lower is α (because of the tightness of the moral hazard problem). Moreover, the smaller is $\dfrac{c_{1(0)} - c_{0(0)}}{y_1 - y_0}$, i.e. the more insurance-oriented the tax system is at $q = 0$, the more likely is that $k_2 > k_1$. As far as we can see, in fact, the magnitude of the term $u_c(c_{1(0)})(c_{1(0)} - c_{0(0)})$ is linked to the features of the optimal tax system at $q = 0$, which, of course, does vary with the income gap, $y_1 - y_0$. To see how, let us write

$$\frac{\partial u_c(c_1)(c_1 - c_0)}{\partial y_1} = u_c(c_1)\frac{dc_1}{dy_1}\left[\left[1 - \eta\right] + \frac{c_0}{c_1}\left[1 - \frac{\varepsilon_{c_0,y_1}}{\varepsilon_{c_1,y_1}}\right]\right]$$

where $\eta \equiv \dfrac{|u_{cc}(c_1)|c_1}{u_c(c_1)}$ is the coefficient of relative risk aversion and $\varepsilon_{c_s,y_1} \equiv \dfrac{dc_s}{dy_1}\dfrac{y_1}{c_s}$ is the elasticity of consumption in state s with respect to the level of income in state 1.

Therefore, a progressive tax system (such that at low income levels $\varepsilon_{c_1,y_1} < \varepsilon_{c_0,y_1}$) makes more likely that $k_2 > k_1$. That is, conditional on the income gap being sufficiently low (such that the incentive motive of taxation actually applies) the scope of providing perks with moral hazard exceeds the one of first best the more progressive is the tax system at $q = 0$.

As we already discussed, a progressive tax system imposes larger penalty on low income realizations than it pays rewards at high income realizations. At low consumption level c_0 and low income gaps, when providing incentives is easier, because of the concavity in the consumption utility function, the cut-off price for perks is more likely to be higher under moral hazard then with perfect information. However, we argue that trading-off progressivity/insurance for perks could also be suboptimal. In particular, we expect that if agents have low taste for perks, trading insurance for incentives, may be harmful the higher is their absolute risk aversion.

Finally, we notice that, if $\theta = 0$, then $\rho = \sigma$. The condition stated in case 1, therefore, has the nice interpretation of the marginal disutility of effort evaluated at $q = 0$ being higher than the marginal utility gain from consuming $c_{1(0)} - c_{0(0)}$, an event which only occurs with marginal probability α. The contents of the following propositions are, therefore, quite intuitive.

Proposition7: Let be $\theta = 0$. Then,

$$\text{sign}\left(k_1 - k_2\right) = \text{sign}\left[\frac{n_0^{FB}}{n_0^{MH}} - \left[\left(1 - \frac{c_{1(0)} - c_{0(0)}}{y_1 - y_0}\right) + \frac{\delta n_0^{MH}\,\sigma}{\alpha u_c(c_{1(0)})(y_1 - y_0)}\right]\right]$$

Hence,

- if $n_0^{FB} > n_0^{MH}$, and $\alpha u_c(c_{1(0)})(c_{1(0)} - c_{0(0)}) > g_{nq}(n_0^{MH}, 0)$, then the price the planner is willing to pay for the first infinitesimal unit of perks is higher in FB than it is in SB.

- if $n_0^{FB} \approx n_0^{MH}$, and $g_{nq}(n_0^{MH}, 0) > \alpha u_c(c_{1(0)})(c_{1(0)} - c_{0(0)})$, i.e. if the moral hazard problem is not too much severe but still the marginal disutility of effort is sufficiently high, there is a scope for providing perks even at those values of k for which $q_{FB}^\star = 0$.

Proposition8: Let be $\theta = 0$ and assume (P1) and (P2) are jointly strictly concave. Denote $k_1(n, \mathbf{c}) \equiv \dfrac{\Delta a \gamma n}{1 + \sigma}$ and $k_3(n, c_1) \equiv \dfrac{a \delta n^{1+\sigma}}{(1 + \sigma) u_c(c_1)}$. Then,

a) $q_{FB}^\star > 0$ iff $k < k_1(n_0^{FB}, \bar{c}_0^{FB})$, and $q_{FB}^\star = 0$ otherwise;

b) $q_{SB}^\star > 0$ iff $k < k_1(n_0^{SB}, \mathbf{c}_0^{SB}) + k_3(n_0^{SB}, c_{1(0)}^{SB})$, and $q_{SB}^\star = 0$ otherwise.

Moreover, the following holds,

i) if $k_1 > k > k_2$ for all y_1, then $q_{FB}^\star > q_{SB}^\star = 0$ for all y_1;

ii) if $k_2 > k_1 > k$ for some y_1, there exist two intervals $\mathcal{I}^-$ and $\mathcal{I}^+$ such that $q_{SB}^\star(\hat{y}) > q_{FB}^\star(\hat{y}) = 0$ and $q_{SB}^\star(z) > q_{FB}^\star(z) > 0$, for all $\hat{y} \in \mathcal{I}^-$ and all $z \in \mathcal{I}^+$, with $\hat{y} < y < z$.

iii) Moreover, if $\dfrac{u_c(c_{1(0)}^{SB})}{u_c(\bar{c}_0^{FB})} < \left(\dfrac{n_0^{SB}}{n_0^{FB}}\right)^{1+\sigma}$, it always holds that $k_2 > k_1$.

Though we can derive the cut-off levels for any θ, so that assuming $\theta = 0$ is a bit of loss of generality, to impose $\gamma < \sigma$, which is a sufficient condition for problem (P1), would not be enough to guarantee the strict concavity of problem (P2) (see Appendix 2.C). The second part of Proposition 8 follows from the continuity of the optimal solutions at y_1, which guarantees the continuity of the $k-$cutoff levels. Finally, the last statement is a restriction on the values $k_1^{SB}(n, \mathbf{c})$ and $k_1^{FB}(n, \bar{c})$. In fact, as long as $k_1^{SB} > k_1^{FB}$ (the condition stated in point iii), then $k_2 > k_1$.

In the following section, we explore by a numerical exercise the effect on the progressivity of the optimal tax system of assuming a price for perks above and below the cutoff levels.

2.5 A numerical exercise

This section puts the model at work. The main purpose is the one of assessing the qualitative importance of the channels we have derived analitycally, and through which perks may affect the progressivity of the optimal tax/transfer scheme and, therefore, the social welfare.

It also extend our theoretical results in that that, by allowing to control for the existence of local maxima, it also help us to highlight the mechanism by which an optimal allocation with perks may still be inefficient.

In order to do this, we fix the parameter values in such a way that the predictions derived for our model without perks are consistent with those discussed in the literature. We mainly try to target the optimal tax on state-1 income, when G is fixed at level 0.5 and y_1 is allowed to vary in the range $[1, 3]$. By this, we mean to capture the optimal average tax on labor income when the per-capita public expenditure is between about $15\% - 50\%$ of the (average) top-bracket gross income. We then set $y_0 = 0.3$, $\alpha = 0.5$ and $\theta = 0$. As standard in this class of model, we assume a Frisch labor supply elasticity equal to 0.5, and work with the pair $(\sigma, \gamma) = (2, 1)$. The parameters δ and a are adjusted so that the $k-$cutoff levels range around 1. Since we do not impose any restriction on effort other than the non-negativeness of the probability measures, one can think of our strategy as to target some relative price for perks, and perk-to-consumption ratio, rather than to match some frequency of the data. Finally, we assume preferences such that the coefficient of relative risk aversion takes value $\eta = 2$.

As one can see from Figure 2.1, the following exercise computes the optimal allocation when the cost of perks is 1.2 times the cost of the numeraire and the equilibrium probability of state-1 being realized ranges between 0.3 and 0.7. The model is, therefore, parametrized so that the effort and, thus, the probability of y_1 to be realized decreases in the value of the state. The dynamics for the $k-$cutoffs are the ones we derived. As income increases, both thresholds increases, though at low income level, the willingness to pay for perks is higher with asymmetric information than it is when effort is observable. For some values of y_1, therefore, the actual price k is below k_2 of moral hazard and above $k_1$of first best.

Figures 2.2 and 2.3 report the equilibrium allocation for either state consumption, effort and perks. The former show how our extended model with perks performs when effort is observable and when it is not. For our purposes, it suffices to notice that the profiles of the $k-$cutoff clearly justify the result. With

asymmetric information, perks are provided in positive amounts i) at earlier states and ii) in a higher proportion than they are under perfect information, so triggering a positive response in the agent's effort that offsets the negative effects of increasing income.

When looking at the optimal consumption schemes, perks allow for a higher state-1 consumption level as compared to either the constant consumption level of first best and the consumption obtainable in the perk-less economy when effort is unobservable. However, state-0 consumption falls below its perk-less value. Figures 2.4 and 2.5 shed light on the real effect of this substitution. In fact, either the intrastate and the interstate progressivity indexes, when computed for our model with perks, increase with the value of y_1. Whereas the optimal tax system of the perk-less economy is progressive ($\eta = 2$), the provision of perks makes the optimal tax/transfer scheme being (more) regressive in that that the ex-post tax on state-1 quickly decreases with income, while state-0 transfer becomes a less concave function of y_1.

Though small in its magnitude, the percentage utility gain from a system with perks increases in the state (Figure 2.6). Moreover, those gains are higher with asymmetric information than they are when effort is observable.

To complete our analysis of the welfare implication of providing perks, we repeat the same exercise when controls for the strict concavity of the problem are not imposed. Figures 2.7 and 2.8 prove that optimal (in the sense of been local maxima) but inefficient interior solutions may arise at low income levels. At those income levels, the actual price for perks is above the $k-$cutoff. Nonetheless, as Figure 2.9 shows, the responsiveness of effort to perks is so higher compared to $\dfrac{\partial n}{\partial c_1}$, that the principal may be tempted to provide perks, also on account of their small marginal diminishing returns at those states. The implied fall in the sensitivity of n with respect to c_1, however, results (Figure 2.10) in a less progressive transfer system if we look at the difference in state consumptions, and in a more progressive tax on state-1 income, rather than in a regressive one. Is this profile of the tax system, which trades off insurance for incentives at a too low state of income, what makes the allocation inefficient.
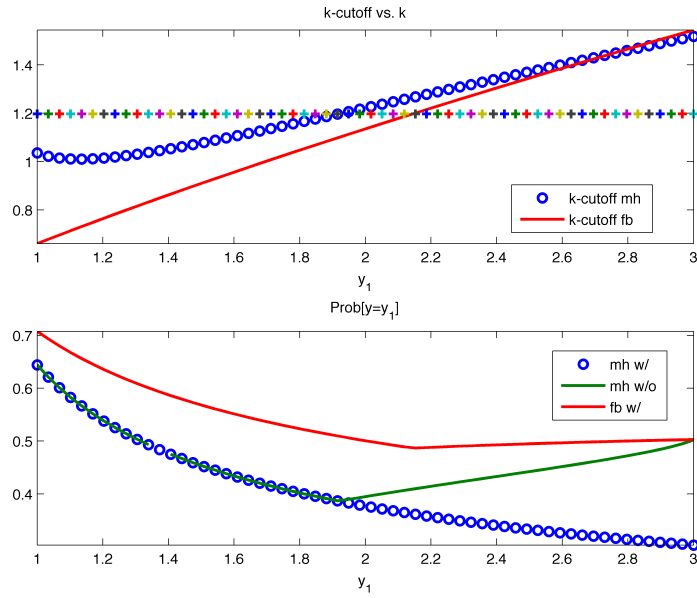
Figure 2.1: Prices and probability

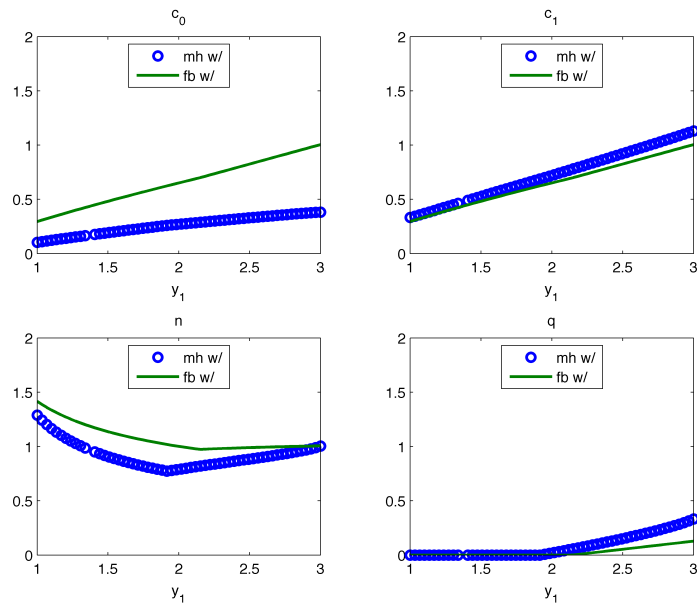Figure 2.2: Optimal allocation with perks: SB vs FB
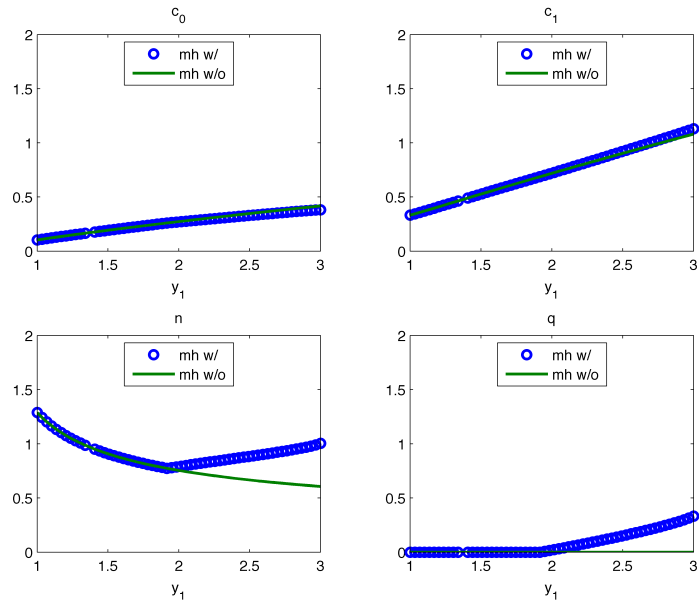
Figure 2.3: Optimal SB allocation: W/ vs W/O

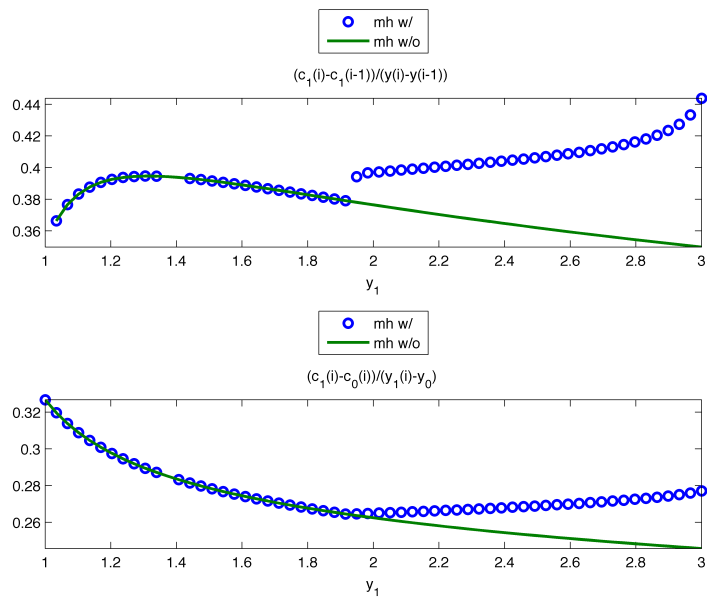Figure 2.4: Progressivity Indexes
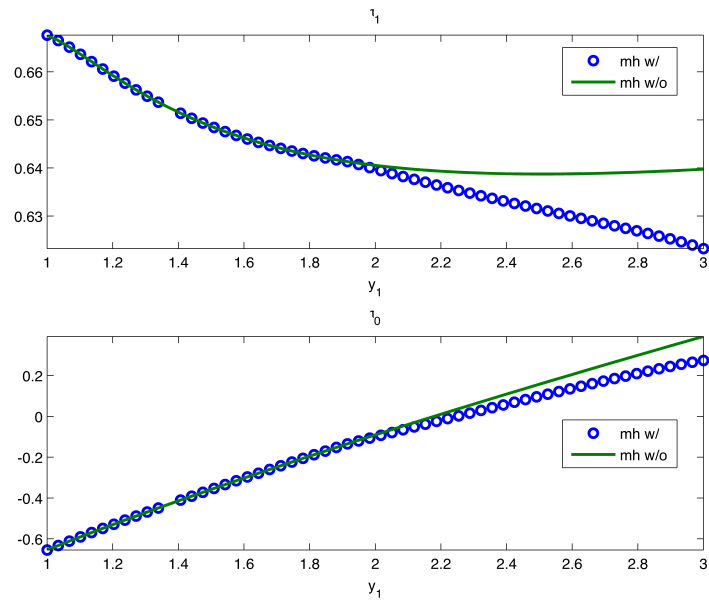
59

Figure 2.5: SB optimal tax and transfer

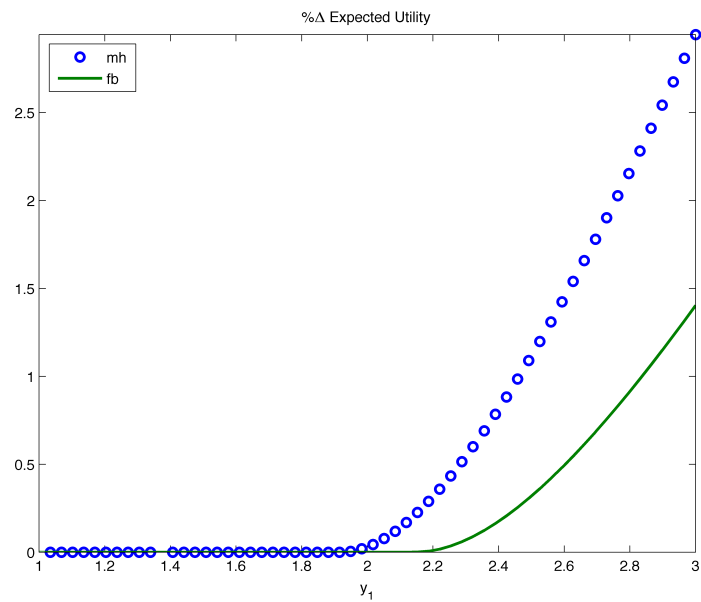Figure 2.6: Utility gain from perks: SB vs FB

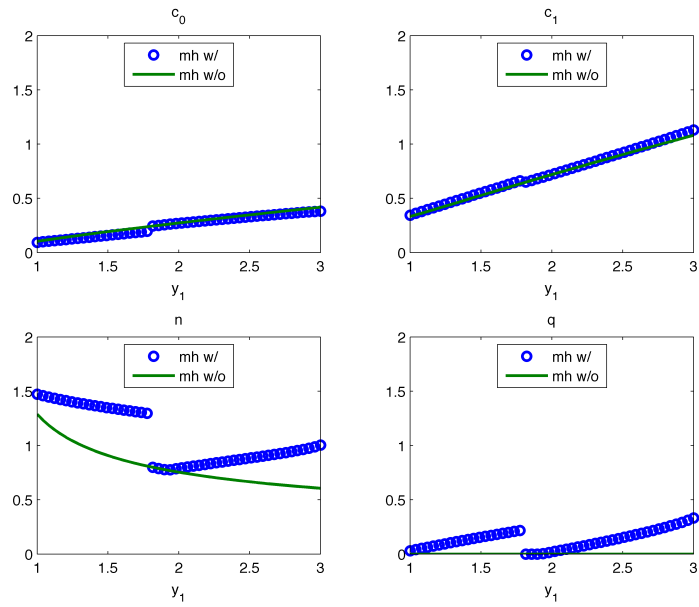Figure 2.7: Inefficient provision of perks at low income states
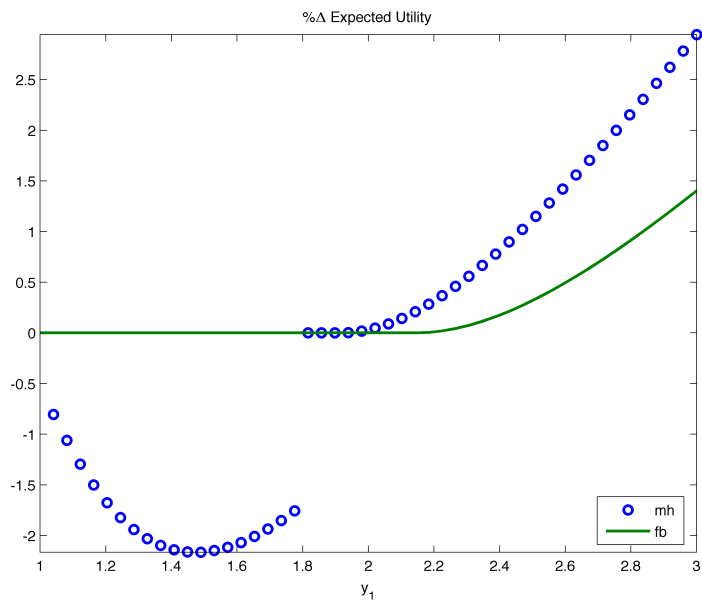

Figure 2.8: Utility loss from perks

61

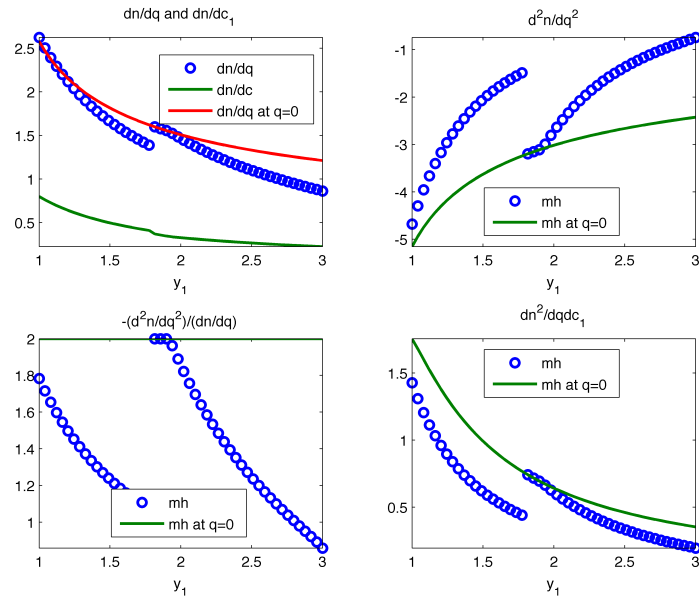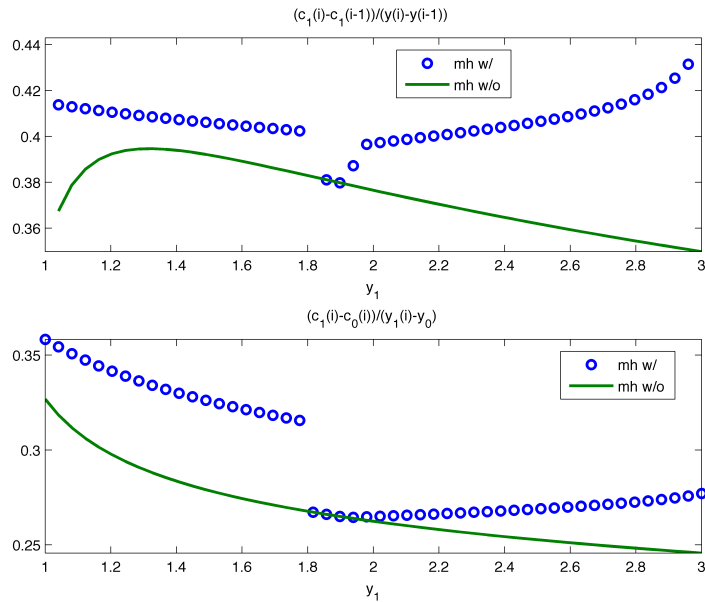Figure 2.9: Responsiveness of effort to perks

Figure 2.10: Progressivity Indexes (Inefficiency)

2.6 Conclusions

This chapter adopts a mechanism design approach to study the optimal labor income tax/transfer system with perks. To the best of our knowledge, it constitutes the first attempt to study analytically the implication of providing perks for optimal taxation. In fact, virtually all the existing papers that analytically study the provision of perks with moral hazard, either focus on their effect on the effort responsiveness to monetary incentives, or they study the cash-perks substitutability in the optimal labor contract, absent any considerations for income taxation.

To deal with the issue at hand, we develop a two-state version of a stochastic Ramsey's problem with a representative, risk-averse agent and an utilitarian, resource-seeking government, who also takes the post of the principal and owner of the only firm in the economy. When designing the optimal rules for cash payments and perk provision, the government, who is constrained by his budget balance, takes into account the agent's unobservable reactions to the tax system. While we assume that perks are not taxable, labor income taxes in our model are set so as to accomplish to either the insurance and the incentive motive.

Methodologically, we rely on the first-order approach (FOA)to replace the set of incentive compatibility constraints with the first order condition derived from the agent's maximization problem, and show that, though we meet the sufficient conditions for the validity of the FOA, a stronger restriction on the convexity of the agent's disutility of effort is needed for the problem to be jointly concave in either effort and perks. To that literature, we also contribute by emphasizing the compliances that naturally arise when a revenue collection problem with wasteful expenditure is taken into account.

Though the complexity of our problem does not allow for a sharp characterization of the solution, we derive simple analytical expressions so as to investigate the main channels through which the provision of perks may affect the optimal tax system. Then, we complete a numerical exercise, in order to test the ability of our model to quantify the progressivity of the optimal tax system, for which we propose two fairly indexes, and prove the robustness of our analysis.

Our results firstly show that there may be a scope for providing perks in the second best that exceeds the one with perfect information, and argue that, when it is the case, the optimal amount of perks in the constrained-efficient allocation is always above its value of first best. Secondly, we prove analytically that, from an ex-post perspective, the equilibrium marginal-rate of substitution be-

tween cash and perks under asymmetric information is always, at all levels of consumption and for all states of the world, greater than its first-best value. In spite of this ex-post inefficiency, however, we also find that the complementarity between perks and effort may be of advantage to the government while counteracting the negative effects on effort that naturally arise at high income levels. Whether or not this feature is beneficial to the agent, it depends on the induced effect of perks on the supporting tax system.

Our analysis suggests that, when perks are efficiently provided, the government trades off progressivity for perk provision. Our main conclusion is that perks make the optimal state-1 consumption be a more convex function of income, while they increase the concavity of the optimal transfer in state-0. Hence, the second-best efficient tax system with perks becomes (more) regressive compared to the case where perks are unavailable. Intuitively, the mechanism designer wants to discourage the agent from exerting less effort. To discourage idleness, not only consumption has to be incentivizing (consumption in higher states needs to increase), but it must also increase more in order to enhance the beneficial effect of perks. Moreover, the lower is the agent's risk aversion, the less-costly is for the principal to trade off insurance for incentives.

As the agent's risk aversion increases, however, there may be an inefficient provision of perks at low income levels. As our quantitative exercise shows, in fact, at lower states, when the revenues collection problem is more severe, the optimal taxation system supporting a positive provision of perks reverses: state-0 transfers become less progressive; state-1 taxes turn to be progressive, while they are regressive in the perk-less scenario. Since these changes occur at those states at which the agent values insurance the most, providing perks is actually inefficient.

Appendix 2.A: The Hessian matrix to (P1-R)

A sufficient condition for $\mathbf{x}^\star \equiv \{n^\star, q^\star\}$ to be a local maximizer for problem (P1-R) is that

1. $(\mathbf{x}^\star, \xi^\star)$ satisfies the first order conditions (with ξ being the Lagrange multiplier on $q \geq 0$), and

2. $|H_2(\mathbf{x}^\star)| > 0$.

We also notice that the constraint qualification condition is always satisfied. By differentiation of the first-order conditions to problem (P1-R), we find

$$|H_2| = \begin{vmatrix} 0 & 0 & 1 \\ 0 & \mathcal{L}_{nn} & \mathcal{L}_{nq} \\ 1 & \mathcal{L}_{nq} & \mathcal{L}_{qq} \end{vmatrix} = -\mathcal{L}_{nn} > 0$$

With $\mathcal{L}_{nn} = -|u_{cc}(\bar{c})|\alpha^2(y_1 - y_0)^2 - \dfrac{\delta \sigma n^{\sigma-1}}{(1+aq)^\gamma} < 0$, for all feasible $\{\bar{c}, n, q\}$, the sufficient condition for a global maximum at $\mathbf{x}^\star$ are satisfied.

2.A.1 Comparative statics for the FB

Let H_2 be the bordered Hessian matrix to problem (P1-R) with positive determinant $H = -\mathcal{L}_{nn}$, and let denote by $g(q)$ the non-negative constraint on q.

1. Whenever $g(q)$ is active, and $q^\star_{FB} = 0$, by applying Cramer's rule we obtain the following comparative statics results.

    (a) The agent's marginal productivity of labor, $\beta \equiv \alpha(y_1 - y_0)$.

    $$\frac{\partial n}{\partial \beta} = \frac{\begin{vmatrix} 0 & -\mathcal{L}_{n\beta} \\ 1 & -\mathcal{L}_{q\beta} \end{vmatrix}}{H} = \frac{\mathcal{L}_{n\beta}}{H} < 0$$

    Proof: $\mathcal{L}_{q\beta} = |u_{cc}|k(n+\theta) > 0$, and

    $\mathcal{L}_{n\beta} = u_c(\bar{c}) - \alpha(y_1 - y_0)(n+\theta)|u_{cc}(\bar{c})| =$

    $$= |u_{cc}(\bar{c})|\left(\frac{u_c - c|u_{cc}|}{|u_{cc}|} - (G + kq^\star - y_0)\right)$$

65

Therefore, $\mathcal{L}_{n\beta} < 0$ if $y_0 < G$ and $u(\cdot)$ is of the class CRRA with $\eta \geq 1$. For all $0 < \eta < 1$ such that $\eta \in \left[\dfrac{1}{1 + (G - y_0)/\bar{c}}, 1\right]$, it also holds that $\mathcal{L}_{n\beta} \leq 0$.

(b) The technological parameter, θ

$$\frac{\partial n}{\partial \theta} = \frac{\begin{vmatrix} 0 & -\mathcal{L}_{n\theta} \\ 1 & -\mathcal{L}_{q\theta} \end{vmatrix}}{H} = \frac{\mathcal{L}_{n\theta}}{H} < 0$$

Proof: $\mathcal{L}_{q\theta} = |u_{cc}|k\alpha(y_1 - y_0) > 0$, and $\mathcal{L}_{n\theta} = -|u_{cc}|\alpha^2(y_1 - y_0)^2 < 0$.

(c) The cost of perks, k

$$\frac{\partial n}{\partial k} = \frac{\begin{vmatrix} 0 & -\mathcal{L}_{nk} \\ 1 & -\mathcal{L}_{qk} \end{vmatrix}}{H} = \frac{\mathcal{L}_{nk}}{H} = 0$$

Proof: $\mathcal{L}_{qk} = -|u_{cc}|\left(kq^\star + \dfrac{u_c}{|u_{cc}|}\right) > 0$, and $\mathcal{L}_{nk} = |u_{cc}|\alpha(y_1 - y_0)q^\star = 0$

2. At an interior solution for q the effect of the same parameters on either n and q is less clear. We can, however, derive the following comparative statics results,

(a) The agent's marginal productivity of labor, $\beta \equiv \alpha(y_1 - y_0)$.
$$\frac{\partial n}{\partial \beta} = \frac{1}{H}\left(\mathcal{L}_{nq}\mathcal{L}_{q\beta} - |\mathcal{L}_{n\beta}||\mathcal{L}_{qq}|\right) > 0 \text{ for } |u_{cc}| \text{ sufficiently high;}$$
$$\frac{\partial q}{\partial \beta} = \frac{1}{H}\left(|\mathcal{L}_{nn}|\mathcal{L}_{q\beta} - |\mathcal{L}_{n\beta}|\mathcal{L}_{nq}\right) > 0 \text{ for } \theta \leq 0.$$

(b) The technological parameter, θ.
$$\frac{\partial q}{\partial \theta} = \frac{1}{H}\left(|\mathcal{L}_{nn}|\mathcal{L}_{q\theta} - |\mathcal{L}_{n\theta}|\mathcal{L}_{nq}\right) = \frac{1}{H}\frac{|u_{cc}|\alpha^2(y_1 - y_0)^2\delta n^\sigma\gamma a}{(1+\sigma)(1+aq)^{\gamma+1}} < 0$$
$$\frac{\partial n}{\partial \theta} = \frac{\partial q}{\partial \theta}\frac{n}{(\frac{1}{a} + q)} < 0 \text{ . That is, } \frac{\partial n}{\partial \theta} < \frac{\partial q}{\partial \theta} \text{ if } q^\star < n^\star - \frac{1}{a}.$$

(c) The cost of perks, k.
$$\frac{\partial n}{\partial k} = \frac{1}{H}\left(|\mathcal{L}_{qq}|\mathcal{L}_{nk} + \mathcal{L}_{qk}\mathcal{L}_{nq}\right) < 0$$
$$\frac{\partial q}{\partial k} = \frac{1}{H}\left(|\mathcal{L}_{nn}|\mathcal{L}_{qk} + \mathcal{L}_{nk}\mathcal{L}_{nq}\right) < 0 \text{ if } \sigma > \gamma.$$

Here, all partial derivatives ($\mathcal{L}_{n\beta}, \mathcal{L}_{q\beta}$, etc.) are as in point 1. Moreover, being

$$\mathcal{L}_{qq} = -k^2|u_{cc}| - \frac{\gamma(\gamma+1)a^2\delta n^{1+\sigma}}{(1+\sigma)(1+aq)^{\gamma+2}} < 0$$

$$\mathcal{L}_{nn} = -\alpha^2(y_1 - y_0)^2|u_{cc}| - \frac{\delta\sigma n^{\sigma-1}}{(1+aq)^\gamma} < 0$$

$$\mathcal{L}_{nq} = \alpha(y_1 - y_0)k|u_{cc}| + \frac{\gamma a\delta n^\sigma}{(1+aq)^{\gamma+1}} > 0,$$

we notice the following,

- if $k > \alpha(y_1 - y_0)$, then $|\mathcal{L}_{nn}| < \mathcal{L}_{nq} < |\mathcal{L}_{qq}|$. That is, the marginal returns from increasing the amount of perks are less persistent;

- if $\dfrac{1+\gamma}{\gamma}k < \alpha(y_1 - y_0)$, then $|\mathcal{L}_{nn}| > \mathcal{L}_{nq} > |\mathcal{L}_{qq}|$, and the marginal returns from increasing effort decrease more quickly.

- If, however, either $|\mathcal{L}_{nn}|$ and $|\mathcal{L}_{qq}|$ are larger than $\mathcal{L}_{nq}$, all comparative statics are such that, at any interior solution for perks, $\dfrac{\partial n}{\partial\beta} - \dfrac{\partial q}{\partial\beta} < 0$, $\dfrac{\partial n}{\partial\theta} - \dfrac{\partial q}{\partial\theta} < 0$ and $\dfrac{\partial n}{\partial k} - \dfrac{\partial q}{\partial k} > 0$.

Appendix 2.B: The agent's optimal response functions

Here $\dfrac{dn}{dx}$ denotes the total derivative of n with respect to x, for $x = \{c_s, q, n\}$, and $\dfrac{\partial n}{\partial x}$ denotes its partial.

$$\frac{dn}{dc_s} = \frac{p_s' u_c(s)}{\sigma\delta n^{\sigma-1}}(1+aq)^\gamma \qquad \text{sign} = +(p_s')$$

$$\frac{dn}{dq} = \frac{\gamma a}{\sigma}\frac{n}{(1+aq)} \qquad > 0$$

It is also convenient to compute the second-order partial and total derivatives. Those are,

$$\frac{d^2n}{dc_s dn} = \begin{cases} -\dfrac{(\sigma-1)}{\sigma}\dfrac{p_s' u_c(s)}{\delta n^\sigma}(1+aq)^\gamma & \text{for } \sigma \geq 2 \\ 0 & \text{for } \sigma = 1 \end{cases} \qquad \frac{d^2n}{dqdn} = \frac{\gamma}{\sigma}\frac{a}{(1+aq)} > 0$$

$$\frac{\partial^2 n}{\partial c_s^2} = \frac{p_s' u_{cc}(s)}{\sigma\delta n^{\sigma-1}}(1+aq)^\gamma \qquad \frac{\partial^2 n}{\partial q^2} = -\frac{\gamma}{\sigma}\frac{na^2}{(1+aq)^2}$$

We conclude, therefore, that,

$$\frac{d^2n}{dc_s^2} = \begin{cases} -\dfrac{p'_s}{\sigma\delta n^{\sigma-1}}(1+aq)^\gamma \left[\dfrac{(\sigma-1)p'_s u_c^2(s)}{\sigma\delta n^\sigma}(1+aq)^\gamma + |u_{cc}(s)|\right] & \text{if } \sigma \geq 2 \\[4mm] \dfrac{p'_s u_{cc}(s)}{\delta}(1+aq)^\gamma & \text{if } \sigma = 1 \end{cases}$$

We notice that for $\sigma = 1$ and $p'_s < 0$, $\dfrac{d^2n}{dc_s^2} > 0$ for all q and γ. For $\sigma \geq 2$, it is always possible to find a value for δ small enough and such that $\dfrac{d^2n}{dc_s^2} < 0$ for all q and $p'_s < 0$.

$$\frac{d^2n}{dq^2} = \frac{\gamma n a^2}{\sigma(1+aq)^2}\left[\frac{\gamma}{\sigma} - 1\right], \text{ which is negative for all } \gamma < \sigma.$$

Finally, the total cross-derivatives are as follows:

$$\frac{d^2n}{dc_s dc_{s'}} = \begin{cases} -\dfrac{\gamma^2(\sigma-1)p'_s p'_{s'} u_c(s) u_c(s')}{\sigma^2\delta^2 n^{2\sigma-1}}(1+aq)^{2\gamma} > 0 & \text{if } \sigma \geq 2 \\[4mm] 0 & \text{if } \sigma = 1 \end{cases}$$

$$\frac{d^2n}{dc_s dq} = \frac{\gamma^2 a p'_s u_c(s)}{\delta\sigma^2 n^{\sigma-1}}(1+aq)^{\gamma-1}$$

Appendix 2.C: The Hessian matrix to (P2)

Each expression in the Hessian matrix results from solving, for all $r = \{c_0, c_1, q, \lambda\}$ and $x = \{c_0, c_1, q, \lambda\}$, the following

$$\mathcal{L}_{rx} = \frac{\partial\mathcal{L}_r}{\partial n}\frac{dn}{dx} + \frac{\partial\mathcal{L}_r}{\partial x}$$

We find,

$$\mathcal{L}_{c_s\lambda} = \sum_s p'_s(y_s - c_s)\frac{dn}{dc_s} - p_s \qquad \text{and} \qquad \mathcal{L}_{q\lambda} = \sum_s p'_s(y_s - c_s)\frac{dn}{dq} - k$$

$$\mathcal{L}_{c_s c'} = \frac{dn}{dc'}\left[\mathcal{L}_n\frac{dn}{dc_s dn} + p'_s\left[u_c(s) - \lambda\right]\right] + \mathcal{L}_{nc'}\frac{dn}{dc_s} + \mathcal{L}_n\frac{\partial^2 n}{\partial c_s \partial c'} + 1_{c_s = c'} p_s u_{cc}(c')$$

$$\mathcal{L}_{c_1 c_0} =$$
$$-\frac{p'_0 p'_1 u_c(c_1)\gamma(1+aq)^\gamma\lambda}{\sigma\delta n^{\sigma-1}}\left[1 + \frac{u_c(c_0)\Delta\gamma(1+aq)^\gamma}{n^\sigma}\left(\sigma - 1 + \frac{n}{n+\theta}\frac{1}{\sigma\delta}\right)\right] > 0$$

Moreover,

$$\mathcal{L}_{c_s c_s} = -\frac{p'^2_s u_c(1+aq)^\gamma}{\sigma\delta n^{\sigma-1}}\lambda\left[\frac{(1+aq)^\gamma\Delta u_c(c_s)}{\sigma\delta n^\sigma}\left(\sigma - 1 + \frac{n}{n+\theta}\right) + \frac{2u_c(1) - u_c(0)}{u_c(1)}\right]$$
$$-p_s|u_{cc}(c_s)|\lambda\left[\frac{1}{u_c(1)} - \frac{\Delta}{\delta\sigma n^{\sigma-1}}\left(\frac{\gamma(1+aq)^\gamma}{(n+\theta)} - \frac{p'_s}{p_s}\right)\right]$$

which is negative either for c_1 and c_0, for δ sufficiently high.

$$\mathcal{L}_{qq} = \frac{dn}{dq}\left[\mathcal{L}_n \frac{dn}{dqdn} + \frac{a\delta n^\sigma}{(t+aq)^2}\right] + \mathcal{L}_n \frac{\partial^2 n}{\partial q \partial q} - \frac{\delta n^{1+\sigma}}{1+\sigma}\frac{\gamma(\gamma+1)a^2}{(1+aq)^{\gamma+2}}$$

$$\mathcal{L}_{qq} = \mathcal{L}_n\left[\frac{\gamma a^2 n}{\sigma(1+aq)^2}\left(\frac{\gamma}{\sigma}-1\right)\right] + \frac{a^2\delta n^{1+\sigma}}{(1+aq)^{\gamma+2}}\left(\frac{\gamma}{\sigma} - \frac{1+\gamma}{1+\sigma}\right)$$

which is negative for all $\gamma < \sigma$. This restriction is necessary for the problem being jointly concave in effort and perks at any interior solution.

Finally, we have,

$$\mathcal{L}_{c_s q} = \frac{p_s' u_c(c_s)a\lambda\Delta\gamma^2(1+aq)^{\gamma-1}}{\delta\sigma^2 n^{\sigma-1}}\left[\frac{\theta}{n+\theta} + \frac{\sigma\delta n^\sigma}{\Delta\gamma(1+aq)^\gamma}\left(\frac{1}{u_c(c_1)} - \frac{1}{u_c(c_s)}\right)\right]$$
$$\text{with sign } = +(p_s')$$

Global concavity of problem (P2), with $\lambda = \{\xi_q, \lambda\}$ and $x_i = \{c_1, c_0, q\}$, requires that the bordered Hessian is such that $|H_3| < 0$, with

$$|H_3(\mathbf{x}; \xi_q, \lambda)| = \begin{vmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & \mathcal{L}_{c_1\lambda} & \mathcal{L}_{c_0\lambda} & \mathcal{L}_{q\lambda} \\ 0 & \mathcal{L}_{c_1\lambda} & \mathcal{L}_{c_1 c_1} & \mathcal{L}_{c_1 c_0} & \mathcal{L}_{c_1 q} \\ 0 & \mathcal{L}_{c_0\lambda} & \mathcal{L}_{c_1 c_0} & \mathcal{L}_{c_0 c_0} & \mathcal{L}_{c_0 q} \\ 1 & \mathcal{L}_{q\lambda} & \mathcal{L}_{c_1 q} & \mathcal{L}_{c_0 q} & \mathcal{L}_{qq} \end{vmatrix} =$$

$$= -\left(2\mathcal{L}_{c1\lambda}\mathcal{L}_{c1c0}\mathcal{L}_{c0\lambda} - \mathcal{L}_{c0\lambda}^2\mathcal{L}_{c1c1} - \mathcal{L}_{c1\lambda}^2\mathcal{L}_{c0c0}\right) < 0$$

Finally, we note that the constraint qualification is met at $\mathbf{x}^\star$ for all sets of effective constraints, $E_1 = \{(GBC)\}$, and $E_2 = \{(GBC), q = 0\}$.

Chapter 3

A Nash non-cooperative game between the fiscal authority and the labor market: the top income marginal tax

This chapter models a two-stage interaction game between the fiscal authority and the labor market, by assuming that the government acts as a Stackelberg leader and the firm as the follower. In fact, it consists of two parts, one for each stage of the game. By a backward induction argument, we first study a two-state version of a static moral hazard problem wherein the independent, self-interest risk-neutral principal (the firm) retains full control of the provision of perks and wage payments to the risk-averse agent (the worker), given the fiscal policy announced by the government. Then, in the second part, we state the government's maximization problem, as restricted by the outcomes generated in the labor market, and solve for the optimal income tax.

To simplify the analysis of the labor market, we abstract from any structural issue (as the market competition, the term structure of the labor contract, the market price of perks, their contribution to productivity, the access to the financial markets, and so on), and concentrate on the effect that perks may have

on (different measures of) the social welfare, when labor income taxes optimally react to their provision.

To deal with the asymmetric nature of our problem, we require that the principal - nor the government - takes into account the agent's effort incentive compatibility constraint, when designing the optimal compensation scheme.

To focus on the main purpose of our research, we also assume that labor income taxes are paid by the firm, so that the worker makes her effort decision only on account of net wages and perk provision. In order to simplify our analysis, however, we fix the low-state income tax to zero, and solve for the top-income tax rate alone. In that respect, we interpret our results in terms of the effect of perks on the marginal tax rate paid by the top-income bracket earners.

We split our analysis in four steps. Firstly, we address our attention to the scope of providing the perquisite good, with and without perfect information. Secondly, we characterize the optimal compensation scheme as the collection of wages and perks that maximizes the firm's expected profits, given the tax rate announced by the government. Thirdly, we enquire on the effect of perks on the rent-extraction problem faced by the principal, and on a measure of the production efficiency. Fourthly, and finally, we state the government's problem and characterize the optimal top-income marginal tax that maximizes the expected utility of an "ideal worker" who does not shrink and is paid her net productivity only in wages.

The rest of the chapter is organized as follows. Section 3.1 briefly reviews some of the more influential works on perk provision. It also emphasizes our contribution to the literature, by pointing out the main dissimilarities and the results we derive from them. The general set-up and the agent's optimal effort policy are described in Section 3.2. Section 3.3 discusses the principal's problem when effort is observable and contractable, and characterizes the first best solution as a reference point. The principal's problem with asymmetric information is then described in Section 3.4. Sections 3.5 and 3.6 characterize the optimal rules for the wage payment and the provision of perks, respectively. The issue on ex-post (in)efficiency is discussed in Section 3.7, while Section 3.8 links it to the agent's information rent and well-being. Section 3.9 derives a closed form solution for the benchmark model, and introduces the concepts of effort and wage elasticity to the tax rate. Section 3.10 proposes and discusses our measure of production efficiency, in order to assess the social desirability of perks. Numerical predictions and comparative statics results are presented in Section 3.11. Section 3.12 states the government's maximization problem and charac-

71

terizes the equilibrium tax rate. All the results in this chapter are summarized in a simple tax-formula, which we propose as a new testable hypothesis. Section 3.13 concludes.

3.1 Related literature

Though there is a long-lasting debate on whether agents are over- or under-provided with perks, and whether perks and cash are complements or substitutes in providing incentives, the theoretical literature on non-monetary benefits does not provide a satisfactory explanation of why, even with perfect information, perks should be provided. Neither it discusses the welfare implications (in terms of excess resources and agent's well-being) of the complementarity/substitutability between perks and wages that, for the most part, mechanically results from those model.

In the financial literature, for example, the standard explanation for perks relies on some form of agency problem (Jensen and Meckling (1976), Grossman and Hart (1980), Bebchuk and Fried (2003), Gabaix and Landier (2008)). Monitoring concerns, which are the key factor of those models, lead managers to use perks in order to appropriate some of the surplus generated by the firm, in a way that is neither approved nor acknowledged by shareholders. This should be beneficial for them, but detrimental for welfare. Unsatisfactory as it might be, in order to be able to apply our analysis to any hierarchical level and any sector, we build instead on the assumption of a principal who has full control of (and, therefore, imposes) the provision of perks to the agents. More generally, we abstract from any issue on the collusion between the worker and the principal, and any misbehavior of the latter.

A second group of researchers consider perks as a "productive" good. This view suggests that perks are useful instruments to align the objectives of principal and agents, and drives the result that the investment in them is always efficient. In Marino and Zabojnik (2008)'s model, for example, perks are either a consumption good and a productivity-enhancement tool. The complementarities between perks and effort, either in the production function and in the consumption utility, have an incentive effect that allows the principal to decrease the pay-performance sensitivity and the uncertainty in the agent's income almost mechanically. To avoid such an entanglements of mechanisms, we focus on a class of pure work-related goods, with no productive attributes.

Among those studies, as ours, that deny fringe benefits any "productive" at-

tribute, there are some that uniquely define perks for their complementarity with the agent's effort, in the usual sense that consuming more of the former good does increase the utility the agent gains from applying the latter. Researchers in this group, however, mainly focus on the issue of ex-post efficiency and agent's utility losses. On that regard, for example, Bennardo, Chiappori and Song (2010) construct a discrete-effort principal-agent model, with no limited liability and binding participation constraint, to find out that whenever the perquisite good is a substitute for leisure, the optimal incentive scheme involves an over-provision of perks, i.e. the agent consumes of perks more than she would have to if given a choice between money and perks at the current market prices. When applied to our setting, however, the analysis of ex-post inefficiency leaves rooms for dynamic improvements as the income gap between states of the world increases, yet providing conditions (on the probability distribution) under which that ex-post inefficiency holds, and it actually holds whenever the agent is enjoying a positive rent. Similarly, in Weinschenk (2013)'s discrete-effort model, the employer may either underinvest or overinvest in perks depending on whether the work-relatedness of perks exceeds or does not a measure of the performance precision. The analysis wherein shows that if perks harm the agent, the principal never underinvests but may overinvest in perks. Though a similar result holds in our model, the derivation of it follows from the extent to which the principal makes use of perks in order to reduce the agent's rent in those scenarios (i.e. technological environments) wherein the wage responsiveness to increasing effort that is needed to preserved incentive compatibility is particularly high.

In the same spirit, though referring to a "motivational action" rather than to a qualified perquisite good, Kvaloy, and Schottner (2012)'s continuous-effort model contributes to a complete discussion of the conditions on the agent's disutility of working under which monetary payments and motivational effort are substitutes in reducing the agent's cost of effort. As we shall see, though we assume, as they do, utility complementarities between perks and effort, the direct or inverse coo-movements in the equilibrium allocations for perks and effort, as long as the complementarity/substitutability between cash and perks, do follow from the interaction between the responsiveness to effort of the probability distribution and that of the agent's marginal disutility of working. Our qualitative result, compared with that in Kvaloy, and Schottner, owes its richness to the different probability measure we assume. Whereas, in their paper, the distribution is defined by just one parameter, normalized to one, we allow the probability distribution to vary with respect to the shape and a location-

parameter, upon which those differences in the comparative statics results turn out to depend.

Moreover, despite the differences between the last three models, they share the same assumption of a profit-maximizing principal, as opposite to the worker's interests. We argue that, though the agents may like perks less than money (as measured by the ex-post inefficiency), yet society (as a whole) may found the perquisite good of some advantage if, by incentivizing effort, perks constitute a resource-improving technology even when agents are compensated (in perks) for their excess of effort.

That is what motivates us to enquire either the real gains from perks in terms of increased capacities, and the agents' utility gains when the tax system itself is allowed to optimize on the tax rates, while taking into account the firm's incentives to provide perks and their effect on production efficiency.

Our results as to regard these questions are completely new to the literature on perks and, therefore, constitute our main contribution to it.

3.2 Preferences, technology and information set

Consider a representative economy, inhabited by an agent (the worker, She) and a firm (the principal, He). The technology of this economy consists in a collection of functions $p^\theta \equiv \langle p_s : \Theta \times \mathcal{N} \to \mathcal{S}(\mathcal{Y}) \rangle$, that, for a given parameter $\theta \in \Theta \equiv (\underline{\theta}, \bar{\theta}) \subset \mathcal{R}$, maps the agent's effort $n \in \mathcal{N} \equiv [0, \bar{n}]$ into a probability measure over the set of income realizations, $y_s \in \mathcal{Y} \equiv \{y_0, y_1\}$, with $0 \leq y_0 < y_1 \leq \bar{y}$ and either $\bar{n}$ and $\bar{y}$ finite. Income can be either consumed by the agent or used by the principal to acquire (at a linear price $k > 0$) any amount of the perk good, $q \in \mathcal{Q} \equiv [0, \bar{q}]$, with $\bar{q} < \bar{y}$. Perks, if paid, are provided to the agent's before effort is taken. We also maintain the assumption that income realizations and the technological parameter θ are both publicly observable and contractable, and call y_s the state of the economy, while the agent's effort may or may not be observable. We refer to the equilibrium allocation arising when effort is observable as the first-best, and define as the second-best (or constrained) optimal allocation, the collection of equilibrium variables when effort is an agent's private information.

A government levies distorting taxes, $T_s(w_s) < y_s$, for $s = \{0, 1\}$, on the state-contingent (gross) wage paid to the worker, $w_s \in \mathcal{R}_+$. It is without loss of generality that we assume that labor income taxes are paid by the principal in

behalf of the worker, so that w_s is the agent's net wage. Restrictions on the functional form of $T_s(\cdot)$ will be imposed as derived by the analysis. Finally, perks are not taxable.

We denote by $p_s(n,\theta)$ the conditional probability that, given θ and effort n, state y_s is realized, i.e. $p_s(n,\theta) \equiv \mathrm{Prob}\{y = y_s|n,\theta\}$. While dealing with a representative economy, in order to simplify notation, we drop θ as an argument of the probability measure, and simply denote by $p_s(n)$ what should be understood as $p_s(n,\theta)$.

We make the following assumption on the discrete probability distribution:

Assumption(PD, Probability Distribution): The density measures $\langle p_s^\theta \rangle$ are such that

- $p_1(n) = \alpha(n+\theta)$, with $\alpha \in (0,1]$
- $0 < p_1(n) < 1$, for all $\theta \in \Theta$ and $n \in \bar{\mathcal{N}} \equiv [0,\bar{n}] \subset \mathcal{N}$, and $p_0(n) = 1 - p_1(n)$

It follows that $p' \equiv \dfrac{\partial p_1(n)}{\partial n} = \alpha > 0$, and $p_0' \equiv \dfrac{\partial p_0(n)}{\partial n} = -\alpha < 0$.

That is, the probability of realizing any y_s is either linearly increasing or linearly decreasing in n. Moreover, for any $\theta, \theta' \in \Theta$ such that $\theta' > \theta$, $\langle p^{\theta'} \rangle$ stochastically dominates $\langle p^\theta \rangle$, in that that for any n, $p_0(n,\theta') < p_0(n,\theta)$.

The factor α, which, of the likelihood ratios, $\dfrac{p_s'}{p_s}$, does directly affect $\dfrac{p_0'}{p_0}$ but not $\dfrac{p'}{p_1}$, measures the effect of changes in n and θ on the probability function p_s. Linearity implies that, once n is fixed, the effect of a marginal increase in n on p_1 , as measured by $\dfrac{p_1}{p'} = n + \theta$, is constant in n.

To respect the full-support assumption, we assume $p_1(\bar{n}) = 1 - \varepsilon$, or, which is equivalent, $p_0(\bar{n}) = \varepsilon$. For the property of a probability measure to be satisfied, we restrict ε to belong to the interval $\mathcal{E} \equiv (1-\alpha, 1)$. That implicitly defines $\bar{n} \equiv \dfrac{1-\varepsilon}{\alpha} - \theta$, for all $\theta \in \Theta \equiv \left(-(1-\dfrac{1-\varepsilon}{\alpha}), \dfrac{1-\varepsilon}{\alpha}\right)$.

We also assume that: i) the agent has an outside option (as an unemployment benefit/research cost) that yields utility $\hat{u} \in \mathcal{R}$, and/or ii) wages are bounded from below by the minimum wage level, $\hat{s} > 0$. Though assumption i) is somehow redundant when the other holds, we maintain both assumptions in our analysis of the first-best, in order to be able to show the link between the agent's bargaining power, i.e. the labor market structure, and the provision of perks.

The agent has preferences over streams of wage levels, $\mathbf{w} = \{w_s\}_{s=0}^1$, with $w_s \in \mathcal{R}_+$, perks $q \in [0, \bar{q}]$, and effort choices, $n \in [0, \bar{n}]$, that are described by an expected utility function given by

$$U^{\langle p \rangle}(\mathbf{w}, q; n) = \sum_{s=0}^1 p_s(n)\, u(w_s) - g(n, q)$$

where $u : \mathcal{R}_+ \to \mathcal{R}$ is a function strictly increasing, strictly concave and twice-continuously differentiable in its argument, i.e. $u_c > 0 > u_{cc}$, for all w_s in $\mathcal{R}_+$; and $g : [0, \bar{n}] \times [0, \bar{q}] \to \mathcal{R}_+$ is a continuous function, twice-continuously differentiable, increasing and convex in n and decreasing and convex in q, i.e. $g_n > 0 < g_{nn}$ and $g_q < 0 < g_{qq}$, for all feasible n and q. Moreover, we assume that,

i) u belongs to the class of NIARA (non-increasing absolute risk aversion) utility functions (either CRRA or CARA);[1]

ii) g is multiplicatively separable in n and q, i.e. $g(n, q) \equiv \tilde{g}(n) h(q)$.

3.2.1 The agent's maximization problem

The agent decides effort n so as to maximize her expected utility, under $\langle p^\theta \rangle$ and given the principal's compensation scheme $\langle \{w\}, q \rangle$. The maximization problem,

$$\max_{n \in \mathcal{N}} U^{\langle p \rangle}(\mathbf{w}, q; n)$$

whose foc with respect to n, namely

$$p'(u(w_1) - u(w_0)) - g_n(n, q) \geq 0, \qquad \text{if } >, \text{ then } n = \bar{n} \qquad \text{(IC)}$$

represents the agent's incentive compatibility constraint that will be taken into account by the principal when designing the compensation package, implies the following optimal response functions (with $s \equiv w_0$ and $b \equiv w_1$, and for all $n \in [0, \bar{n}]$),

$$\frac{\partial n}{\partial s} = -\frac{p' u_c(s)}{\tilde{g}(n) h(q)} \qquad \frac{\partial n}{\partial b} = \frac{p' u_c(b)}{\tilde{g}(n) h(q)} \qquad \frac{\partial n}{\partial q} = \frac{\tilde{g}_n |h_q|}{\tilde{g}_{nn} h(q)}$$

[1]The most straightforward implications of increasing or decreasing absolute risk aversion, is that, if the agent did experience an increase in wealth, she would choose to decrease (or keep unchanged, or increase) the number of dollars spent in the risk-less asset (here, the perk good) if absolute risk aversion is decreasing (or constant, or increasing).

Assumption(N, Disutility of effort): We adopt the following specification form for the agent's disutility of effort,

- $\tilde{g}(n) = \dfrac{\delta n^{1+\sigma}}{1+\sigma}$, with $\delta > 0$ and $\sigma \geq 1$;

- $h(q) = \dfrac{1}{(1+aq)^{\gamma}}$, with $\{\gamma, a\} \gg 0$.

Because of Assumption(N), we have that the agent's responsiveness to the principal's instruments are given by

$$\frac{\partial n}{\partial s} = -\frac{\alpha u_c(s)(1+aq)^{\gamma}}{\delta \sigma n^{\sigma-1}} \qquad \frac{\partial n}{\partial b} = \frac{\alpha u_c(b)(1+aq)^{\gamma}}{\delta \sigma n^{\sigma-1}} \qquad \frac{\partial n}{\partial q} = \frac{\gamma an}{\sigma(1+aq)}$$

Accordingly, n increases either in b and q, and decreases in s. We also note, as it will turn useful later on, that, as long as $\sigma = 1$,

$$\frac{\partial^2 n}{\partial w^2} = \frac{p'_s u_{cc}(w_s)(1+aq)^{\gamma}}{\delta} \qquad \frac{\partial^2 n}{\partial w_s \partial q} = \frac{p'_s u_c(w_s)}{\delta} \frac{\gamma a}{(1+aq)^{1-\gamma}}$$

imply that, for given w_s and q, the responsiveness of the marginal labor supply $\dfrac{\partial n}{\partial w_s}$ to infinitesimal changes in either w_s and q is independent of the effort level that has to be implemented. That is, because of the consumption-leisure separability and linearity in the probability measure, as long as the Frisch elasticity of labor supply - namely, $\dfrac{1}{\sigma}$ - is equal to one, changes in the instruments affect the labor supply only directly (they do not trigger a feedback process).

3.3   The First Best

The characterization of the-first best solution, far from having any policy implications, serves as a reference point for the analysis we conduct in the following section, while dealing with the asymmetric nature of our information structure. It aims at shading light on the dynamics that may lead to a positive provision of perks even with perfect information, and allows us to rule out those scenarios wherein the perquisite good serves as a consumption good, which is not the scope of this research.

Here below is a brief review of our results, so that the reader can easily skip the analysis of the first-best and go directly to Section 3.4.

As our analysis suggests, as far as the agent's effort is either observable and contractable, and taxes on wages are payment-contingent (rather than state-contingent), the principal fully insures the worker against the systemic risk, by

offering her a contract consisting of a fixed wage and, eventually, on a certain amount of the perk good. However, the provision of perks appears to be crucially related to the competitive structure and regulation of the labor market, that, in our model with just one agent and one principal, we model in the form of two fairly assumptions on an agent's reservation utility and the minimal wage requirement.

It turns out, thus, that the highest level of effort may actually be implemented even at $q = 0$, if the restriction on wages guarantees that the worker's participation constraint is satisfied with inequality. On the other hand, however, though the principal's excess of bargaining power pushes the agent's utility from contracting downward to its reservation value, the cost of perks may be sufficiently higher to make the perk good worthless even at the highest effort level. Between these two polar cases, for sufficiently small values of the minimal wage, such that the equilibrium effort in a perk-less economy is interior, the condition for a positive provision of perks may be more or less demanding, depending on the value of the minimal wage. Given the causality between minimal wage and effort, we conclude that, with perfect information, there may be a scope for perks as a consumption good, substitute to leisure, when the labor market structure keeps the fixed wage excessively low, even below the frontier of production possibilities of the firm. When it is the case, the principal reallocates resources to the agent by balancing bonuses and perks according to their relative cost.

3.3.1 The analysis

Under perfect information, the principal's problem consists in maximizing expected profits, net of wages (gross of labor income taxes) and of expenditure on q, subject to: i) the agent's (ex-ante) participation constraint (IR); ii) the minimum-wage requirement, $w_s \geq \hat{s}$, for $s = \{0, 1\}$, and iii) the non-negative constraint on perks. Maximization is taken over $\{w_s\}$, q and n.

$$(\text{P3}) \qquad \max_{\{w_s\}, q, n \in \mathcal{N}} E_s(y_s - w_s - T(w_s)) - kq$$

$$\text{subject to:} \qquad E_s(u(w_s)) - g(n, q) \geq \hat{u} \qquad (\text{IR})$$

$$\forall s, \qquad w_s \geq \hat{s} \qquad (\text{LL})$$

$$q \geq 0$$

The Lagrangian to this problem can be written as follows

$$\mathcal{L} = \max_{n,s,b,q} p_0(n)(y_0 - s - T(s)) + p_1(n)(y_1 - b - T(b)) - kq+$$

$$\lambda \Big[E(u(w_s)) - g(n,q) - \hat{u} \Big] + \textstyle\sum_s \chi_s(w_s - \hat{s}) + \chi_q q + \chi_n(\bar{n} - n)$$

where $\lambda \geq 0$ is the Lagrange multiplier attached to the agent's participation constraint; $\{\chi_s\}$ is the set of Lagrange multipliers, all non-negative, such that $\chi_s(w_s - \hat{s}) = 0$; while the non-negative multiplier χ_q refers to the constraint $q \geq 0$. Finally, $\chi_n \geq 0$ is the Lagrange multiplier corresponding to the feasibility constraint, $n \in \bar{\mathcal{N}}$.

A solution to problem (P3) must satisfy the following first-order necessary conditions:

$\mathcal{L}_n :$ $\qquad p'\Delta + \lambda\Big[p'\Big(u(b) - u(s)\Big) - g_n(n,q)\Big] - \chi_n = 0,$ $\qquad$ if $\chi_n > 0$, then $n = \bar{n}$

with $\Delta \equiv y_1 - y_0 - b + s - T(b) + T(s)$

$\mathcal{L}_b :$ $\qquad -p_1(n)(1 + T'(b)) + \lambda p_1(n)u_c(b) + \chi_b = 0$

$\mathcal{L}_s :$ $\qquad -p_0(n)(1 + T'(s)) + \lambda p_0(n)u_c(s) + \chi_s = 0$

$\mathcal{L}_q :$ $\qquad -k - \lambda g_q(n,q) + \chi_q = 0$

and the agent's participation constraint below,

$\qquad E(u(w_s)) - g(n,q) - \hat{u} \geq 0,$ $\qquad$ if $>$, then $\lambda = 0$

Two different scenarios arise, depending on the bindingness of the participation constraint. Each of them corresponds to a particular structure of the labor market as regard to the agent's bargaining power and implies a different allocation of perks.

1. If $\lambda = 0$, meaning that the agent's participation constraint does not bind, the Lagrangian multipliers χ_b, χ_s, χ_q are all strictly positive. Thus, at an optimum, either $s = b = \hat{s}$ and $q = 0$. Moreover, since $\mathcal{L}_n = p'(y_1 - y_0) > 0$, we have that, for all $\hat{s}$ such that (IR) holds with strictly inequality, $n = \bar{n}$ is implemented at $q = 0$. The principal's expected profits are given by $\Pi = y_0 + \bar{p}_1(y_1 - y_0) - \hat{s} - T(\hat{s})$, where $\bar{p}_1 \equiv p_1(\bar{n})$, while the (certain) agent's utility satisfies, $u(\hat{s}) - g(\bar{n}, 0) > \hat{u}$.

   This scenario, pretty simple, assumes that the agent is able to earn a rent even with perfect information, and that the first best power of monetary

79

incentive is θ-independent. Nonetheless, it allows to exclude a positive provision of perks, independently from the parameter values linked to the perk technology.

Though appealing for its tractability, this scenario looks quite at odds with the reality in the labor market. Since the scope of this work is the one of providing policy recommendations on the provision and on the tax-treatment of perks, which should potentially hold for any hierarchical layer and any sector, we mainly focus on the opposite scenario of a principal that, with perfect information, is able to seize all of the gains from contracting with the agent, i.e. on the case $\lambda > 0$.

2. Assume (IR) is binding, then $\lambda > 0$ and $p_1(n)u(b) + p_0(n)u(s) = \hat{u} + g(n,q)$. Optimality conditions with respect to b and s yield, respectively, to: $\lambda u_c(b) + \tilde{\chi}_b = (1 + T'(b))$ and $\lambda u_c(s) + \tilde{\chi}_s = (1 + T'(s))$, where $\tilde{\chi}_s = \dfrac{\chi_s}{p_s}$ for all $p_s > 0$. Then, $\lambda = \dfrac{1 + T'(w_s) - \tilde{\chi}_s}{u_c(w_s)}$ implies that optimal b and optimal s are allocated as to satisfy $\dfrac{u_c(b)}{u_c(s)} = \dfrac{1 + T'(b) - \tilde{\chi}_b}{1 + T'(s) - \tilde{\chi}_s}$.

   - For all payment-contingent taxes such that $T''(\cdot) \geq 0$, either i) $b = s = \hat{s}$; or ii) $b = s > \hat{s}$. [2]

   - For all state-contingent taxes such that $T_b'(\cdot) > T_s'(\cdot)$, we get the counterfactual result that either: 1) if both limited liability constraints do not bind (or there are no such constraints at all), then $s > b > \hat{s}$, and optimal payments b and s are such that $\dfrac{1 + T_b'(b)}{u_c(b)} = \dfrac{1 + T_s'(s)}{u_c(s)}$; or 2) if only (LL$_b$) binds, and $b = \hat{s}$, then, either $s > b = \hat{s}$, or $s = b = \hat{s}$.

By the preceding analysis, we conclude, therefore, that either $b = s \geq \hat{s}$, or $s > b \geq \hat{s}$. Nonetheless, we can rule out all the equilibria for which $s > b \geq \hat{s}$, by assuming that the income taxation is actually payment-contingent. Under this reasonable assumption, the principal always finds optimal, with perfect information, to equalize wages (and thus taxes) across states. At an optimal solution, therefore, it must be that either $b = s = \hat{s}$, or $b = s > \hat{s}$. The following analysis aims at discussing how the varying of $\hat{s}$, i.e. of the minimum wage requirement, actually varies the scope for perks in the labor market when effort is observable and thus contractable.

---

[2] In both cases, $\chi_b = \chi_s$. However, $\chi_s > 0$ in subcase i), and $\chi_s = 0$ in subcase ii). It is worthy to note that if $s = \hat{s}$, and $\chi_s > 0$, then, neither $b > s = \hat{s}$, because of the decreasing marginal consumption utility, nor $b < s = \hat{s}$, because of the minimum wage requirement.

Given the principal's optimal wage scheme, the binding of (IR) implies $u(b) = \hat{u} + g(n,q)$. Thus, by Assumption(N), we can derive two fairly important conditions:

i) $g_n(n,q) = \dfrac{(1+\sigma)}{n} g(n,q) = \dfrac{1+\sigma}{n}\left[u(b) - \hat{u}\right]$. That is, under perfect information, the agent's consumption utility gain from contracting, namely $u(b) - \hat{u}$, is a fraction $\dfrac{1}{1+\sigma}$ of the marginal cost of effort times the effort itself; and

ii) $n_0 = \min\left\{\bar{n},\ \left[\dfrac{1+\sigma}{\delta}\left(u(b_0) - \hat{u}\right)\right]^{\frac{1}{1+\sigma}}\right\}$. That is, as long as $q = 0$, the first best level of effort at any interior point, is a concave function of the agent's consumption utility gain.

We also notice that, whatever n_0 may be - whether it is at the boundary of the feasible set or not -, q is never part of a the first best solution if the following holds

$$\frac{k}{\gamma a} > \frac{p'(y_1 - y_0)n_0}{1 + \sigma}$$

The utility of the above condition in our model is straightforward. To it we attach the scope of providing perks in an economy where agency problems do not arise, but yet the perk good can serve as a pure consumption good, substitute to leisure. Two scenarios may then arise, depending on whether perks are paid or not at an optimum. We discuss both, though we will focus mainly on the case of a non-positive provision of perks.

- Let assume $b = s = \hat{s}$, and $\hat{s} = \{\tilde{s}:\ u(\tilde{s}) = \hat{u} + g(\bar{n},0)\}$, then $n_0 = \bar{n}$. The binding of (IR) requires that perks are never paid at an optimum (and $n = \bar{n}$), so that all the parameters must satisfy the condition for a non-positive provision of perks, namely $\dfrac{k}{\gamma a} > \dfrac{p'(y_1 - y_0)\bar{n}}{1+\sigma}$. As long as this condition is satisfied, $\{b^\star, n^\star, q^\star\} = \{\hat{s}, \bar{n}, 0\}$ is an equilibrium at which (IR) is satisfied with equality.

- For arbitrarily small $\hat{s} = \{\tilde{s}:\ u(\tilde{s}) > \hat{u}\}$, it may be that $n_0 < \bar{n}$, and the condition for a positive-provision of perks may be more or less demanding depending on n_0's value. However, we notice that if the condition for a non-positive provision of perks is satisfied for $\bar{n}$, it is also satisfied for all $n_0 < \bar{n}$.

  For all $b = s > \hat{s}$, such that $\lambda = \dfrac{1 + T'(b)}{u_c(b)}$, the optimal n_0 is found by solving with respect to $\tilde{n}$

$$\frac{p'(y_1 - y_0)}{1 + T'(s)} - \frac{g_n(\tilde{n})}{u_c(b_0)} \geq 0, \qquad \text{if } >, \text{ then } n_0 = \bar{n},$$

with $b_0 = u^{-1}\left[g(\tilde{n}) + \hat{u}\right]$ by (IR), u^{-1} denoting to the inverse of the agent's consumption utility[3].

If it is optimal to set $q = 0$, then, by substituting $n_0 = \tilde{n}$ and $q = 0$ into (IR), we obtain the equilibrium wage, $b^\star > \hat{s}$. If it is not, the optimal allocation $\{b^\star, q^\star, n^\star\}$ solves $\mathcal{L}_q, \mathcal{L}_n$ and (IR), with $\lambda = \dfrac{1 + T'(b)}{u_c(b)}$. The main lesson we can derive from this result is that, for sufficiently low minimal wage $\hat{s}$, if there is a scope for paying wages above that value, it there may also be a scope for paying perks as an alternative consumption good, substitute to leisure. When it is the case, the principal optimally provides perks depending on their marginal cost and benefit.

3.4 The second-stage of the Stackelberg game: the firm's problem with moral hazard

This section provides a fully characterization of the principal's problem of designing the optimal compensation scheme $\langle\{s, b\}, q\rangle$ when the agent's effort is neither contractable, nor observable. It constitutes the second stage of the strategic game between the fiscal policy and the pair of worker and firm, through which we characterize the equilibrium marginal tax on income. From Sections 3.4 to 3.11, therefore, we will focus on the outcomes of the labor market, when agents take the fiscal policy as given.

---

[3]When $u(c)$ is CRRA, Assumption(N) implies $\dfrac{g_n(\tilde{n})}{u_c(b_0)} = \delta\tilde{n}^\sigma\left[-\left(\dfrac{\eta - 1}{1 + \sigma}\right)\left(\delta\tilde{n}^{1+\sigma} + \hat{u}(1 + \sigma)\right)\right]^{\frac{\eta}{1-\eta}}$.

Therefore, for $Q \equiv \dfrac{1 + T'(b)}{p'(y_1 - y_0)\left(\dfrac{\eta - 1}{1 + \sigma}\right)^{\frac{\eta}{\eta - 1}}}$, n_0 solves the following expression:

$$\mathcal{L}_n : \qquad \frac{\left[-\left(\delta\tilde{n}^{1+\sigma} + \hat{u}(1 + \sigma)\right)\right]^{\frac{\eta}{\eta - 1}}}{\delta\tilde{n}^\sigma} \leq Q, \qquad \text{if } <, \text{ then } n_0 = \bar{n}$$

For $\sigma = 1$, $\eta = 2$ and $\hat{u} < 0$, we get $Q = \dfrac{4(1 + T'(b))}{p'(y_1 - y_0)}$ and, therefore,

$$\mathcal{L}_n : \qquad \frac{\delta n^4}{4} \leq n^2|\hat{u}| + \frac{Q}{4}n - \frac{\hat{u}^2}{\delta}, \qquad \text{if } <, \text{ then } n_0 = \bar{n}$$

Because of the insights gained from the first best solution, we shrink the set of the constraints in the principal's maximization problem, and of the two assumptions on the agent's reservation utility and on the minimal wage requirement, we maintain only the latter. As pointed out in Section 3.8, an adapted version of the last requirement, when combined with the information asymmetry, is enough to guarantee that in our representative economy with just one agent and one principal, the participation constraint of the former does not actually bind (i.e. the worker earns a rent). Furthermore, the limited liability constraint allows us to focus on the role played by the minimal wage requirement (as embodying any given structure and/or regulation of the labor market) on the scope for providing perks under asymmetric information, which, in turn, guarantees more flexibility while parametrizing the model.

As we shall see, the solution to the principal's maximization problem is indexed by the technological parameter θ. At this stage of our work, differences in θ must be understood as differences in the stochastic dominance of the probability distribution $\langle p^\theta \rangle$, i.e. as differences in technology. Since in a representative agent's framework there is no meaning of talking about heterogeneity in workers' productivity or workers' type, by means of θ we aim at accounting for intrinsic differences in the production process, which we believe are tightly linked to the scope of providing perks.

As discussed in Section 3.5, and then in Appendix 3.A, θ is responsible for the complementarity/substitutability of wages and perks in the optimal compensation scheme. Due to the linearity of the probability measure, the second-order effects of increasing (decreasing) the optimal piece-rate at any equilibrium allocation for which the provision of perks is positive depend on the fact that the elasticity with respect to effort of the probability $p_1(n)$ (which is a decreasing function of θ) is either constant or increasing (decreasing) in effort. This sharp result, which also affects the ex-post efficiency of perks, contributes to the existing literature by providing a more general explanation of the scope for perks, based not (or not only) on the way they enter the agent's utility function, but on the relative responsiveness to effort of the probability distribution and the marginal disutility.

Moreover, as long as perks have no productive attributes, and the agent's utility is super modular in effort and perks, the optimal piece-rate, which increases in the likelihood ratio, according to the inference principle, and in the convexity of effort disutility, because of the incentive motive, is only an indirect function of q, through the perks' effect on effort. This allows us to study the positive or

negative correlation between changes in the optimal effort and optimal incentive-rate b, as induced by variations in those parameters that control for either the technology of perks and the stochastic distribution (in particular in y_1).

Section 3.6 states, not very surprisingly, that at any interior solution for perks, the principal's expected gain from saving on perks and paying the agent by cash after state-1 is realized must be zero. That is, since the principal is risk-neutral, substituting an uncertain payment b by a certain payment of q must come at no cost for the firm. Nonetheless, Section 3.7 points out as the agent's marginal rate of substitution between b and q may differ from the relative price of wages and perks, so leading to an (ex-post) inefficiency in the provision of perks, which holds from the agent's point of view. Moreover, depending on the stochastic dominance properties of the probability distribution, the inefficiency the principal imposes on the agent might become larger as income increases. We argue, however, that even an over-provision of perks may make the agent better off, compared to a system without perks, if the excess of perks triggers a sufficiently high excess of effort (Section 3.8), or if it results in a resource-improving technology, that allows for an ex-post compensation of the agent's effort (Section 3.9). Yet, the bottom line of our analysis is that this contingency is ruled out anytime the probability distribution does exhibit a low degree of first-order stochastic dominance. Section 3.9, embodying our welfare analysis based on the concept of production efficiency, makes grounds for farther enquiring the real gains from perks when the tax system itself is allowed to optimize on the tax rate.

To conclude our analysis of the labor market outcomes, Section 3.10 provides a closed form-solution to the principal's maximization problem for the benchmark case with no feedback process, and constant risk-sharing (or independency of b from q). Finally, Section 3.11 puts the model at work and, in a numerical exercise, tests the quantitative importance of our predictions.

3.4.1 The principal's problem: definitions

The principal's problem is the one of maximizing expected profits, net of wages (gross of labor income taxes) and of expenditure on q, subject to the agent's incentive compatibility constraint, the minimum-wage requirements $s, b \geq \hat{s}$, and the non-negative constraint on perks, while taking the tax system, $\{T(s), T(b)\}$, as given.

$$(\text{P4}) \qquad \max_{s,b,q} p_0(n^\star)(y_0 - s - T(s)) + p_1(n^\star)(y_1 - b - T(b)) - kq$$

$$\text{subject to:} \qquad n^\star = \left\{ n \in \bar{\mathcal{N}} \mid p'(u(b) - u(s)) = g(n, q) \right\} \qquad (\text{IC})$$

$$s, b \geq \hat{s} \qquad (\text{LL})$$

$$q \geq 0$$

Because of Assumption(PD), and by substituting $n^\star$ with $n(s, b, q)$ from (IC), the Lagrangian to problem (P4) can be written as

$$\mathcal{L} = y_0 - s - T(s) + \alpha(n+\theta)(y_1 - y_0 - b + s - T(b) + T(s)) - kq + \sum_s \chi_s(w_s - \hat{s}) + \chi_q q$$

where $\{\chi_s\}$ is the set of non-negative Lagrangian multipliers attached to (LL) and χ_q is the multiplier attached to the non-negative constraint on q. Maximization is taken over s, b, q.

Definition1: We refer to the collection $\mathcal{W} \equiv \langle \{s, b\}, q \rangle$ as the contract.

Definition2: A second-best optimal allocation of this economy is a pair $\langle \mathcal{W}, n^\star \rangle$, made of a contract $\mathcal{W}$ and a policy function $n^\star$, such that: i) given $\{T(\cdot)\}$, k, and $\langle p \rangle$, $\langle \{s, b\}, q, n^\star \rangle$ maximizes the principal's expected profits, subject to (LL) and the non-negative constraint on q; ii) the contract $\mathcal{W}$ implements the effort policy $n^\star$ implied by (IC).

In the following, we focus on characterizing an optimal allocation at which $n > 0$. As it is, from the set of the limited liability constraints, we drop the constraint on b, since by strictly monotonicity of u and (IC), for any $s \geq \hat{s}$, and $n > 0$, it must be that $b > \hat{s}$ as well.

The principal's risk neutrality, as compared to the agent's risk aversion, which makes her willing to accept a higher reduction in her expected wage the more disperse the state-contingent payments are, could induce us to believe inefficient any contract such that $s > \hat{s}$, since the firm could, in principal, reduce the agent's expected payment by decreasing either s and b, without violating (LL) and still providing the same incentives.

In our set-up, however, the existence of a tax on wages, along with the chance of substituting out state-0 taxes with state-1 taxes, does not allow to directly apply this reasoning and to fix $s = \hat{s}$, unless we make more specific assumptions on $T(\cdot)$.

So far, we do not assume any restriction on the equilibrium value of the expected profits, since, by (LL), we can always set the value of $\hat{s}$ so as to make the equilibrium expected profits equal to zero.

For convenience, we report below the derivative of the Lagrangian to (P4) with respect to the agent's effort:

$$\mathcal{L}_n: \qquad p'\Big[y_1 - y_0 - b + s - (T(b) - T(s))\Big]$$

The principal's marginal gain from increasing effort is given by the expected marginal revenues he can retain by moving away from state 0 to state 1, i.e. by the extra income, net of wages and income taxes, he will gain with a marginal probability of p'.

By global concavity of (P4) (discussed in Appendix 3.A), either i) $\mathcal{L}_n = 0$ and the implemented effort is the one that maximizes the principal objective; or ii) $\mathcal{L}_n > 0$, and the implemented effort is below its efficient level (from the principal's perspective). If it were $\mathcal{L}_n < 0$, it could be possible to increase the principal's objective by decreasing either b or q, or both, so as to i) discourage the agent from applying a dynamically-inefficient effort; ii) save on payments. It follows, therefore, that any allocation $\langle s, b, q, n \rangle$ such that $\mathcal{L}_n < 0$ does violate optimality.

All this said, and since $\mathcal{L}_n$ can be thought of as a measure of the tightness of the moral hazard problem, i.e. of how difficult is for the principal to implement the optimal effort, we focus on characterizing a solution $\langle s, b, q, n \rangle$ at which $\mathcal{L}_n > 0$.

If we use (IC) to substitute out p' from the expression for $\mathcal{L}_n$, we find that the marginal effect of effort on the principal's expected profit can be measured by the amount of revenues the principal gains by moving from state 0 to state 1, per unit of the payment (in terms of utility) that the principal leaves to the agent to compensate her for increasing the disutility of effort by one unit (when technology is such that the marginal effect of n on p_1 is p'). That is,

$$\mathcal{L}_n: \qquad \Delta \Big/ \frac{\Delta u}{g_n}$$

with $\qquad \Delta \equiv y_1 - y_0 - b + s - (T(b) - T(s))$

and $\qquad \Delta u = u(b) - u(s)$

Definition 3: We refer to Δ as the principal's opportunity-gain from moving to state-1. As one can see, q, which is paid to the agent with certainty before any state is realized, does not enter this identity. As its counterpart, Δu

represents the agent's opportunity-gain of state-1 being realized, given that she spent $g(n, q)$ in providing effort.

The focs to problem (P4), then, solve[4]

$$\mathcal{L}_s : \mathcal{L}_n \frac{\partial n}{\partial s} + \frac{\partial \mathcal{L}}{\partial s} \qquad\qquad \mathcal{L}_b : \mathcal{L}_n \frac{\partial n}{\partial b} + \frac{\partial \mathcal{L}}{\partial b} \qquad\qquad \mathcal{L}_q : \mathcal{L}_n \frac{\partial n}{\partial q} + \frac{\partial \mathcal{L}}{\partial q}$$

3.5 First-order conditions for wages

As to regard to s, the first order condition below

$$\mathcal{L}_s : \qquad -\frac{p'^2 u_c(s)\Delta}{g_{nn}(n, q)} - p_0(n)(1 + T'(s)) + \chi_s \leq 0, \qquad \text{if } <, \text{ then } s = \hat{s}$$

implies that, unless $T'(\hat{s}) < 0$ and such that $|T'(\hat{s})| > 1$, i.e. unless the principal gets a positive transfer and a net gain for any unit paid by s, optimality requires $s = \hat{s}$, regardless of what $q^\star$, $\{y\}$ and $u(\cdot)$ are. If we think of $T(\cdot)$ as a proper tax function and assume $T'(\cdot) \geq 0$, thus, it is always optimum to set $s = \hat{s}$.

More on $\hat{s}$ will be said later on. For the moment being, we assume $s = \hat{s} \geq 0$.

As we are interested in those equilibrium allocations such that $n > 0$, by (IC) it must be $b > \hat{s}$. Optimality, then, requires

$$\mathcal{L}_b : \qquad \frac{p' u_c(b) p' \Delta}{g_{nn}(n, q)} = p_1(n)(1 + T'(b))$$

Since $p'\Delta = \dfrac{\Delta}{\Delta u / g_n(n,q)}$ from (IC), by denoting $p_s \equiv p_s(n)$, we can rewrite it as follows,

$$\mathcal{L}_b : \qquad \frac{\Delta}{\Delta u / u_c(b)} = (1 + T'(b)) \frac{p_1}{p_1'} \frac{\tilde{g}_{nn}}{\tilde{g}_n}$$

That is, combining $\mathcal{L}_b$ and (IC) we find that, at an optimum,

$$\frac{\Delta}{(1 + T'(b)) \frac{\Delta u}{u_c(b)}} = \frac{p_1}{p_1'} \Big/ \frac{\tilde{g}_n}{\tilde{g}_{nn}}$$

Proposition1: The optimal payment b is set by the principal in such a way that the equilibrium revenues he gains by moving from state 0 to state 1, per any dollar of the expenditure in state-1 wage that is needed to

---

[4]Note that, because of the linearity of $\langle p \rangle$, if $\dfrac{\partial^2 n}{\partial w_s \partial n} = 0 = \dfrac{\partial^2 n}{\partial q \partial n}$, as it is in the case where $\sigma = 1$, then, all derivatives $\mathcal{L}_s, \mathcal{L}_b$ and $\mathcal{L}_q$ are linear in effort.

compensate the agent for an exacerbation of her incentive compatibility constraint, equates the ratio of the marginal effects of n on the probability density measure and on the the marginal cost of effort.

Put it differently,

$$\frac{p_1(1 + T'(b))\frac{\Delta u}{u_c(b)}}{p'\Delta} = \frac{\tilde{g}_n}{\tilde{g}_{nn}}$$

The expected expenditure in b that is needed to compensate the agent for her effort, per unit of the marginal revenues the principal may gain by moving from state 0 to state 1, is adjusted so as to move one-to-one with the responsiveness to effort of the agent's marginal disutility of working.

It is convenient to write the above expression in the following way,

$$\frac{(1 + T'(b))\frac{\Delta u}{u_c(b)}}{\Delta} = \frac{p'}{p_1}\frac{\tilde{g}_n}{\tilde{g}_{nn}} = \frac{\tilde{g}_n}{\tilde{g}_{nn}} \Big/ \frac{p_1}{p'} \qquad (1)$$

In (1), the adjustment in b (on the LH side) is directly linked to changes in n (on the RH side). Expression (1) states that the expenditure in b that is needed to compensate the agent for implementing n, per unit of the revenues the principal gains by moving from state 0 to state 1, is adjusted so as to move one-to-one with the ratio of the effort marginal effects on g_n and p_1.

From (1), we derive the following corollaries.

Corollary1: As long as i) perks have no productive attributes, i.e the density measure $\langle p \rangle$ is not a function of q, and ii) perks and effort are multiplicatively separable in $g(n, q)$, the optimal b is a function of n alone, rather than of n and q.

That is, for $q = 0$, there exists no difference between (1) and the expression we would obtain in the perk-less economy. By this simple assumption, we mean to characterize the effect of perks on optimal compensation only through their direct effect on effort, provided that effort does affect the optimal wage.

Corollary2: By DARA, the LH side of (1) does increase in b and decreases in s, for all tax schemes such that $T''(b) \geq 0$, i.e. for all proportional and progressive tax functions. For those tax schemes such that $T''(b) < 0$, the LH side of (1) does still increase in b for $|T''(b)|$ sufficiently small.

It is, therefore, for seek of simplicity, that we assume taxes are proportional to wages, and constrain the tax function to satisfy the following assumption,

Assumption(TS, Tax System): The tax system is described by a labor-income tax function $T(w_s)$ such that, for $j = \{s, b\}$,

- $T'(j) \equiv \dfrac{dT(j)}{dj} = \tau_j > 0$ and

- $T''(j) \equiv \dfrac{d^2 T(j)}{dj^2} = 0$

Corollary3: Let $g(\cdot)$ and $\langle p^\theta \rangle$ satisfy Assumption(N) and Assumption(PD), respectively. That is, $\tilde{g}(n) = \dfrac{\delta n^{1+\sigma}}{1+\sigma}$ and $p_1(n, \theta) = \alpha(n + \theta)$. Then, at an equilibrium allocation, the optimal payment b is such that the principal's share of the opportunity gains of moving from state 0 to state 1 by incentivizing the agent's through b is either constant, or increasing (decreasing) in effort, depending on whether $\dfrac{p'}{p_1} \dfrac{\tilde{g}_n}{\tilde{g}_{nn}}$ is a constant, or a decreasing (increasing) function of the agent's effort.

By specification of $\tilde{g}(n)$, $\langle p^\theta \rangle$ and $T(\cdot)$, expression (1) becomes

$$\frac{1}{\sigma}\left[\frac{\Delta}{\Delta u / u_c(b)} \right] = (1 + \tau_b)\left(1 + \frac{\theta}{n}\right)$$

Three scenarios are thus possible:

1. If $\theta = 0$, then $\dfrac{\Delta}{(1+\tau_b)\Delta u / u_c(b)} = \sigma$. We refer to this as the perfect risk-sharing benchmark, since, at the equilibrium allocation, the principal's opportunity-gain is proportional to his expenditure in state-1 wage, by a factor which measures the inverse of the Frisch elasticity of the labor supply. That is, the gains for higher y_1 are split between the principal and the agent, in such a way that the ratio of gains $\dfrac{\Delta}{\Delta u / u_c(b)}$ is maintained constant.

   In this scenario, what really matters for the principal is the implementable marginal cost of effort alone, i.e. $g_n(n, q)$, since varying q does not affect the ratio of the marginal cost's and of the marginal cost's responsiveness to effort. For given $g_n(n, q)$, it is possible to define the optimal b, independently of how $g_n(n, q)$ is split between marginal disutility of effort $\tilde{g}_n(n)$ and utility from perks $h(q)$.

2. If $\theta > 0$, $\dfrac{\Delta}{(1+\tau_b)\Delta u / u_c(b)} = \sigma\left(1 + \frac{\theta}{n}\right)$ is decreasing in n. The principal's share of the opportunity gain from contracting is higher than it is in the

perfect risk-sharing benchmark (either because $\theta > 0$, and $n_\theta < 0$), but it is decreasing in the effort level the principal wants to implement.

3. If $\theta < 0$, then $\dfrac{\Delta}{(1+\tau_b)\Delta u / u_c(b)} = \sigma\left(1 - \dfrac{|\theta|}{n}\right)$ is increasing in n. With a negative systemic component θ, for given n, the equilibrium expenditure in b per opportunity-gain is ceteris paribus higher and increasing in n. Since $\langle p^\theta \rangle$ stochastically dominates $\langle p^{-\theta} \rangle$, the provision of incentives when $\theta < 0$ is indeed more costly to the principal.

In the following, we think of the kernel $\langle p^0 \rangle$ as our benchmark. Besides having some economic content, by assuming $\theta > 0$, we mean to characterize those scenarios wherein the production technology first-order stochastically dominates that in the benchmark. By assuming $\theta < 0$, we aim at investigating the optimal allocation when the probability distribution is stochastically dominated by the benchmark.

As it shall be discussed later in Section 3.9, solving for the optimal contract amount to solve a system of non-linear equations. Because of the linearity assumption, and the weak separability of q and n in $g(\cdot)$, a solution to the problem is proved to exist when $\theta = 0$. For all other $\theta \in \Theta$, however, we cannot guarantee the existence of a solution, apart than in a small interval around $\theta = 0$. Since the profit function is continuous, in fact, an application of the Maximum Principle Theorem implies that the solution is continuous, and that it does exist and is also continuous around $\theta = 0$.

3.5.1 Comparative statics for the optimal piece-rate

Because of our specification assumption, $\dfrac{p'}{p_1}\dfrac{\tilde{g}_n}{\tilde{g}_{nn}} = \dfrac{1}{\sigma}\dfrac{n}{n+\theta}$.

Let be $\varepsilon \equiv \dfrac{n}{n+\theta} > 0$.

Therefore, $\qquad \dfrac{\partial \varepsilon}{\partial n} = \dfrac{\theta}{(n+\theta)^2} \qquad \qquad \text{sign}\left(\dfrac{\partial \varepsilon}{\partial n}\right) = \text{sign}\left(\theta\right)$

Let be $\mathcal{A} \equiv y_1 - y_0 + s(1 + \tau_s)$, with the proviso that $\mathcal{A} > b(1 + \tau_b)$ as long as $\mathcal{L}_n > 0$. Then expression (1) becomes

$$\frac{\left[\dfrac{u(b) - u(s)}{u_c(b)}\right]}{\left[\dfrac{\mathcal{A}}{1 + \tau_b} - b\right]} = \frac{\varepsilon}{\sigma} \qquad (2)$$

By differentiating (2), we obtain that, for all parameters j, the following proposition holds,

Proposition2: Let $N \equiv \dfrac{\Delta u}{u_c(b)}$ and $D \equiv \dfrac{\Delta}{1 + \tau_b}$ be, respectively, the agent's and the principal's opportunity gain of moving from state 0 to state 1 per unit of state-1-wage marginal cost. Then,

- for all parameters $j \in \{a, \delta, \gamma, k\}$, and for $j = \alpha$, either

  1. $\dfrac{db}{dj} = 0,$ if $\dfrac{\partial \varepsilon}{\partial n} = 0$, i.e. if $\theta = 0$; or

  2. $\dfrac{db}{dj}$ moves together with $\dfrac{\partial \varepsilon}{\partial n} \dfrac{dn}{dj}$, according to

  $$\frac{db}{dj} = \frac{D^2}{\sigma(D \cdot N_b + N)} \left(\frac{\partial \varepsilon}{\partial n} \cdot \frac{dn}{dj} \right)$$

  When $\theta > 0$, $\text{sign}\left(\dfrac{db}{dj}\right) = \text{sign}\left(\dfrac{dn}{dj}\right)$. When $\theta < 0$, the opposite happens.

- For $j = y_1$,

$$\frac{db}{dy_1} = \frac{1}{D \cdot N_b + N} \left[\frac{D^2}{\sigma} \left(\frac{\partial \varepsilon}{\partial n} \cdot \frac{dn}{dy_1} \right) + \frac{N}{1 + \tau_b} \right]$$

  1. If $\dfrac{dn}{dy_1} \geq 0$, then $\dfrac{db}{dy_1} > 0$ for all $\theta \geq 0$ such that $\dfrac{\partial \varepsilon}{\partial n} \geq 0$;

  2. For $\theta < 0$, the sign of $\dfrac{db}{dy_1}$ is still positive for τ_b sufficiently low.

3.5.1.1   Math Note: The functional forms at work

$\dfrac{\Delta u}{u_c(b)}$ measures the change in the opportunity-gain Δu due to an increase in b when the marginal utility of b is equal to $u_c(b)$. The higher is this index, the easier is for the principal to provide incentives by the payment of b. As one can see below, with a CRRA utility function, we can shape the curvature of $\dfrac{d}{db}\left[\dfrac{\Delta u}{u_c(b)} \right]$ simply by varying the coefficient of the agent's relative risk aversion. We must note, however, that, by increasing/decreasing b over a fixed s, the principal does change the wage dispersion, and by that, the agent's (absolute) risk aversion.

Instead, if $u(c)$ belongs to the class of CARA utility function, not only $\dfrac{d}{db}\left[\dfrac{\Delta u}{u_c(b)}\right]$ is increasing and convex in b for any value of the absolute risk aversion coefficient, but, by varying b the principal does not affect the agent's risk aversion. We expect these differences to play a role in characterizing the welfare implications of the optimal solution, more than the optimal solution itself. Nonetheless, the following analysis suggests that the magnitude of this changes is only of second-order. However, it allows us to focus on a subset of the values for the coefficient of relative risk aversion which is in line with the literature, $\eta \in [1, 2]$.

Lemma1: (See Math Appendix 3.B)

1. If $u(c) = \dfrac{c^{1-\eta}}{1-\eta}$. Then, $N_b \equiv \dfrac{d}{db}\left[\dfrac{\Delta u}{u_c(b)}\right]$ is positive, and i) linearly increasing in b for $\eta = 2$; ii) increasing and concave in b for all $\eta \in (0, 1] \cup (1, 2)$; iii) increasing and convex in b for $\eta > 2$.

2. Assume $u(c) = -\exp(-\eta c)$. Then, $N_b \equiv \dfrac{d}{db}\left[\dfrac{\Delta u}{u_c(b)}\right]$ is positive and increasing and convex in b for all $\eta > 0$.

As far as we can see, CRRA functions with $\eta > 2$ and the $CARA$ utility function behave alike, as to regard the effect of a marginal increase in b on the utility gain the principal must guarantee to the agent in order to ensure incentive compatibility. For $\eta = 2$, the second-order effects of an increase in the piece-rate on the agent's incentive-compatible utility gains are zero; while for $\eta \in (0, 1] \cup (1, 2)$, increases in b have decreasing second-order effects.

3.6   Optimality condition for perks

The foc with respect to q writes

$$\mathcal{L}_q: \qquad \frac{\tilde{g}_n |h_q|}{\tilde{g}_{nn} h(q)} p'\Delta - k \leq 0, \qquad \text{if} <, q^\star = 0$$

Evaluating $\mathcal{L}_q$ at $q = 0$, we get

$$k \geq \frac{\alpha a \gamma n_0}{\sigma}\left(y_1 - y_0 - b_0(1 + \tau_b) + s(1 + \tau_s)\right) \qquad (3)$$

In Appendix 3.A, we discuss the restriction on σ and γ such that $\mathcal{L}_{qq} < 0$ for all $n > 0$. As long as $\mathcal{L}_{qq} < 0$, the condition stated in expression (3), for a non-positive provision of perks, is either necessary and sufficient. We can derive, therefore, the following proposition,

Proposition 3: Fix problem (P4) and assume $\gamma < \sigma$. Then, $q > 0$ if and only if the marginal cost of perk is smaller than the marginal benefit of the first unit of perks when taking into account the effect of n on the principal's expected opportunity-gain, and that of q in the agent's effort policy. That is, if

$$k < \left. \frac{|h_q(0)|}{h(0)} \frac{\tilde{g}_n p' \Delta}{\tilde{g}_{nn}} \right|_{q=0}$$

Because of our specification, with $s_0 = \hat{s}$, we obtain that $q > 0$ if the following holds,

$$k < \bar{k} \equiv \gamma a \frac{\alpha n_0}{\sigma} \left[y_1 - y_0 - b_0(1 + \tau_b) + \hat{s}(1 + \tau_s) \right] \qquad (4)$$

where the subscript $_0$ denotes that the corresponding value is computed for $q = 0$[5]. More conveniently, we can write the above expression as

$$\frac{k}{\gamma a} < \frac{1}{\sigma} \left(n \mathcal{L}_n \right)_{q=0} \qquad (5)$$

Here, $\dfrac{1}{\sigma}$ is the Frisch elasticity of the labor supply. It can be useful to think of $\mathcal{L}_n$ as the firm's labor demand, that is as the price the principal is willing to pay in order to implement (consume) different effort levels. Since $\mathcal{L}_n$ is decreasing in b, as long as b is increasing in n (so as it is under very mild regulatory conditions),

---

[5](See Math Appendix 3.B). When $\sigma = 1$ and $q = 0$, effort must satisfies, from (IC), $n_0 = \dfrac{\alpha \cdot \Delta u|_{q=0}}{\delta}$. Substituting n_0 into $\mathcal{L}_n|_{q=0}$, and assuming $u(\cdot)$ is CRRA with $\eta = 2$, we get the following expressions for b_0 and n_0,

$$b_0 = \left(\frac{\tilde{\mathcal{A}}}{\dfrac{1}{s} + \dfrac{\theta \delta}{\alpha}} \right)^{\frac{1}{2}} \qquad\qquad n_0 = \frac{\alpha}{\delta} \left[\frac{1}{s} - \sqrt{\left(\frac{1}{s} + \frac{\theta \delta}{\alpha} \right) \Big/ \tilde{\mathcal{A}}} \right]$$

where $\tilde{\mathcal{A}} \equiv \dfrac{y_1 - y_0 + s(1 + \tau_s)}{1 + \tau_b}$

For $y_0 = \hat{s} = 1$, $\theta = 0$ and $\tau_s = 0$, our benchmark parametrization, we get $b_0 = \sqrt{y_1/(1 + \tau_b)}$ and $n_0 = \dfrac{\alpha}{\delta} \left[1 - \sqrt{\left(1 + \tau_b \right) \Big/ y_1} \right]$, which is always positive for $y_1 > 1 + \tau_b$.

$\mathcal{L}_n$ is indeed decreasing in n for any variation in all those parameters, but y_1[6] and $\hat{s}$, that are linked to the perks technology and preferences.

Following this interpretation, the area below the $\mathcal{L}_n$ curve and above the $\mathcal{L}_n(n_0)$ line represents the principal's surplus derived from implementing (or purchasing) n_0. The rectangle below $\mathcal{L}_n(n_0)$, i.e. $n_0 \cdot \mathcal{L}_n(n_0)$ represents, therefore, the total amount the principal is actually paying for n_0 units of effort [7].

The above condition can be stated as follows,

Corollary4: Perks are provided in positive amount only if their price, net of the perk's marginal effect on the agent's disutility of effort (namely, $\dfrac{k}{\gamma a}$), is smaller than the expenditure the principal is willing to make in effort at the implementable perk-less effort level n_0, while taking into account the hidden cost of "demanding" n, here represented by the labor supply elasticity, $\dfrac{1}{\sigma}$.

At an interior solution for q, therefore, provided that (IC) binds, we obtain

$$\mathcal{L}_q : \qquad \frac{\Delta}{\Delta u / \frac{\tilde{g}_n}{\tilde{g}_{nn}}} = \frac{k}{|g_{nq}|} = \frac{p'}{\tilde{g}_{nn}} \Delta$$

The first equality says that q is paid till the principal's opportunity-gain Δ, per unit of the utility gap needed to compensate the agent for a unit increase in g_n by n, equals the expenditure in perks that is needed to decrease the agent's marginal cost of effort by one unit.

In more useful terms, the second equality states that the expenditure in perks that is needed to decrease the marginal cost of effort by one unit is a share of the principal's gain of moving away from state 0 to state 1, the factor of proportionality weighting the marginal effect of n on the probability of state 1 to be realized with the responsiveness of the marginal cost of effort to increasing n.

---

[6] In Math Appendix 3.B we prove that, for increasing y_1 and $\dfrac{dn}{dy_1} > 0$, $\mathcal{L}_n$ is increasing in n if and only if $(1 + \tau_b)\dfrac{db}{dy_1} < 1$, which holds whenever either i) $\theta \leq 0$, and $|\theta|$ is such that all variables are well-defined; ii) $\theta > 0$, and either σ is sufficiently high or τ_b is sufficiently low.

[7] For all those changes in the parameters such that $\mathcal{L}_n(n)$ is increasing in n, representing $\mathcal{L}_n$ as a labor demand can be misleading. However, as long as we think of n as the "good" produced by the firm, on the assumption that it is the firm that produces opportunities of work and transforms production capacities (p) into real value (n), we can still represent $\mathcal{L}_n(n)$ as the firm's supply curve. Following this interpretation, condition (5) states that the firm's investment in the intermediate good "perks" is positive, i.e. $q > 0$, if and only if the adjusted price $\dfrac{k}{\gamma a}$ is smaller than the amount for which the principal is willing to offer n_0 units of effort, scaled by σ, that here represents the hidden cost of "producing" n.

By rearranging, so as to get all terms in n on the RH side, we get

$$\mathcal{L}_q: \qquad \frac{k}{|h_q|} = \frac{\tilde{g}_n}{\tilde{g}_{nn}} p' \Delta$$

Since $\frac{\tilde{g}_n}{\tilde{g}_{nn}}$ is increasing in n, the optimal expenditure in q is increasing in n as long as this direct effect overweights the indirect effect of effort on increasing b and, therefore, reducing Δ. However, if b and n change in the opposite direction (for example, for changes in the parameters a, γ, δ, k and α for $\theta \leq 0$), the total expenditure in q and, therefore, q itself, is always increasing in n.

$$\mathcal{L}_q: \qquad \frac{k/|h_q|}{p_1' \Delta} = \frac{\tilde{g}_n}{\tilde{g}_{nn}} \qquad (6)$$

Equation (6) states that perks are paid till the expenditure in perks that is needed to decrease $h(q)$ by one, per unit of the marginal revenues the principal gets by moving from state 0 to state 1, equals the agent's "expenditure" in effort (i.e. amount of effort needed to increase g_n by one unit $(\frac{1}{\tilde{g}_{nn}})$ times the marginal cost of effort $(\tilde{g}_n)$. We can conclude, therefore,

Proposition4: Fix problem (P4) and assume that the condition for a positive provision of perks is satisfied (Proposition 3). Then, the principal sets q such that the following equality holds

$$\frac{k/|h_q|}{p_1 \Delta} = \frac{p'}{p_1} \frac{\tilde{g}_n}{\tilde{g}_{nn}}$$

That is, the share that is invested in perks of the expected gain from moving to state 1 moves one-to-one with the ratio of effort marginal effects on g_n and p_1. This share is constant in n for $\theta = 0$, and it is increasing (decreasing) in n for all $\theta > 0$ ($\theta < 0$).

3.7 A characterization of the MRS between wages and perks

From $\mathcal{L}_q$ and $\mathcal{L}_b$, by substituting out (IC), what we obtain is (See Math Appendix 3.B)

$$\frac{k}{|g_{nq}|} = \frac{p_1(1 + \tau_b)}{p' u_c(b)}$$

The expenditure in perks that is needed to decrease the agent's marginal cost by one unit must be equal to the expected expenditure in state-1 wage that is

needed to repay the agent for increasing that marginal cost by the same amount. Altogether, these variations will keep (IC) binding. We state it in terms of the following proposition,

Proposition5: Fix Problem (P4) and assume Assumptions (N), (PD) and (TS) hold. Thus, whenever the principal wants to induce a one-unit increase in the agent's marginal cost, he may either i) spend in b the amount $\dfrac{(1 + \tau_b)}{p'u_c(b)}$ with probability p_1; or ii) he can reduce his investment in perks by $\dfrac{k}{|g_{nq}|}$. Hence, optimality requires that the expected gains from saving on perks and repaying the agent by cash after state 1 is realized are zero. That is, since the principal is risk-neutral, substituting uncertainty for certainty, and vice-versa, must come at no cost.

By rearranging the above expression, we can write

$$\frac{p'u_c(b)}{|g_{nq}|} = \frac{p_1(1 + \tau_b)}{k} \qquad (7)$$

and state the following,

Corollary5a: From the principal's perspective, the marginal rate of substitution between b and q, in keeping the agent's incentive compatibility constraint binding, must be equal to their relative price, with the proviso that b is paid only with probability p_1, and therefore its expected marginal cost for the principal is $p_1(1 + \tau_b)$, while the marginal cost of perks is k.

By rearranging (7), we also get an expression for the agent's equilibrium marginal rate of substitution.

Corollary5b: At an equilibrium, the agent's (ex-post) marginal rate of substitution between b and q is proportional to the relative price of wages and perks faced by the principal, with a proportional factor that is increasing in the ratio of the effort marginal effects on p_1 and $\tilde{g}$. That is,

$$\frac{u_c(b)}{U_q} = \left(\frac{p_1}{p'} \Big/ \frac{\tilde{g}(n)}{\tilde{g}_n}\right) \frac{1 + \tau_b}{k} \qquad (8)$$

Because of our specification, (8) becomes

$$\frac{u_c(b)}{U_q} = (1 + \sigma)\frac{n + \theta}{n}\frac{(1 + \tau_b)}{k} \qquad (8a)$$

We note that:

1. for $\theta = 0$, the agent's marginal rate of substitution is constant to variations in the income levels, y_0 and y_1, and in α, η, a, δ. It is, then, decreasing in k and increasing in σ and τ_b.

2. For $\theta > 0$, if n increases in y_1 and $q > 0$, the MRS must be decreasing in y_1. As y_1 increases, then, the ex-post inefficiency (associated to larger MRS than the relative price) does decrease, and perks become, for the agent's perspective, less excessively-provided. Moreover, since $u_c(b)$ is decreasing in b and b is increasing in y_1 when $\theta \geq 0$, it must be that U_q decreases slower than the marginal utility of wage.

3. For $\theta < 0$, the equilibrium marginal rate of substitution is, ceteris paribus, lower (for given n), and for higher y_1 it must be that U_q decreases faster than the marginal utility $u_c(b)$. That is, substituting out b with q, the principal imposes on the agent a greater ex-post inefficiency. Since $\langle p^{-\theta} \rangle$ is stochastically dominated by $\langle p^{\theta} \rangle$, the substitution of b with q finds its rationale in the better capacity of perks to provide incentives, which is more valuable the higher is y_1.

Moreover, by differentiation of the LHS of (8a) with respect to y_1, we obtain the following result, stated as a corollary (See Math Appendix 3.B).

Corollary6: The $\text{MRS}_{b,q}$ decreases in y_1 as long as the following inequality is satisfied

$$\frac{|u_{cc}(b)|}{u_c(b)} \frac{|h_q|}{h_{qq}} > \frac{dq}{dy_1} \bigg/ \frac{db}{dy_1}$$

We conclude, therefore, that the principal more heavily substitutes out wage compensation with perks, the higher is the agent's absolute risk aversion and the less diminishing the returns of perks are.

3.8 Rent, $\text{MRS}_{b,q}$ and $\mathbf{E}[U]$

As already discussed, the limited liability constraint and the incentive compatibility constraint univocally determine the wage schedule: that is, $s = \hat{s}$ from (LL) (with $\hat{s} = u^-[\hat{u}]$), and $b = \left\{ \tilde{b} : \; u(\tilde{b}) = \hat{u} + \dfrac{g_n(n,q)}{p'} \right\}$ from (IC).

Nothing has been said, however, on the agent's expected utility. From Microeconomics, we already know that if the (LL) constraint were replaced by an agent's participation constraint of the form $E\left[u(w_s)\right] - g(n, q) = \hat{u}$ (IR), the principal would find optimal to propose a contract $\mathcal{W}$ such that the agent's individual rationality constraint would bind, i.e. such that $E\left[u(w_s)\right] = \hat{u} + g(n, q)$. Therefore, as long as we set $\hat{s}$ greater than the optimal payment (better to say, punishment) the principal would like to impose on the agent at the realization of state 0, and require $s, b \geq \hat{s}$, it may well occur that (IR) does not actually bind, the optimal allocation leaving the agent to enjoy a positive rent.

To our purposes, under Assumption(PD), the expected consumption utility satisfied:

$$E\left[u(w_s)\right] = u(s) + p_1(u(b) - u(s)) =$$

$$= \hat{u} + \frac{p_1}{p'} g_n(n, q) = \hat{u} + g(n, q) + \left[\frac{p_1}{p'} g_n(n, q) - g(n, q)\right]$$

$$= \hat{u} + g(n, q) + R(n, q)$$

$$\geq \hat{u} + g(n, q), \qquad \text{if } \mathcal{R}(n, q) \geq 0$$

Where

$$R(n, q) \equiv g(n, q)\left[\frac{p_1}{p'} \frac{\tilde{g}_n}{\tilde{g}(n)} - 1\right] =$$

$$= g(n, q)\left[\frac{n + \theta}{n}(1 + \sigma) - 1\right] \qquad (9)$$

is the rent the principal has to leave to the agent on top of what is needed to accommodate the (LL) and (IC) constraints. The rent is i) non-negative for all $\theta \geq 0$, and negative for all negative θ such that $|\theta| > \dfrac{n^\star}{2}$; ii) increasing in n for sufficiently small $\theta > 0$ and all $\theta \leq 0$; iii) decreasing in q, at a higher rate the larger is n.

3.8.1   Rent vs. MRS

As we are interested in studying those scenarios wherein the rent is non-negative, we focus on those solutions at which, according to expression (9), $\dfrac{p_1}{p'} \dfrac{\tilde{g}_n}{\tilde{g}(n)} \geq$ 1. By simply inverting it, we obtain that the non-negativeness of the agent's information rent is a restriction on the ratio between the marginal effects of effort

on p_1 and $\tilde{g}(n)$. In fact, the requirement that, at the optimal allocation of effort, $\dfrac{p'}{p_1}\dfrac{\tilde{g}(n)}{\tilde{g}_n} \leq 1$ amounts to say that a positive rent is left to the worker if and only if the convexity of the agent's direct cost function is such that the marginal effect of effort on $\tilde{g}(n)$ is higher than its marginal effect on the probability of state-1 being realized. At the implemented effort, therefore, the (direct) marginal cost of incentivizing effort outweighs its (direct) marginal benefit.

Moreover, by Corollary 5b, we can easily derive the following,

Corollary7: As long as a positive rent is left to the worker, the higher is the rent, the larger is the agent's marginal rate of substitution between b and q, as compared to the relative price paid by the principal. Ex-post, therefore, at those prices, the agent is willing to give up more units of q to obtain one additional unit of b.

Intuitively, this result can be understood as the convexity of the agent's effort cost being so high that, ex-post, the agent had rather received the fair remuneration of her high effort, since it would have amounted in even larger wages.

It is useful to rewrite $\dfrac{p'}{p_1}\dfrac{\tilde{g}(n)}{\tilde{g}_n}$ as the product between two fairly known objects, namely, $\dfrac{p'}{p_1}\dfrac{\tilde{g}_n}{\tilde{g}_{nn}}$, which affects $\mathcal{L}_b$, and $\dfrac{\tilde{g}(n)\cdot\tilde{g}_{nn}}{\tilde{g}_n^2} = \dfrac{\sigma}{1+\sigma}$, where the equality follows from Assumption (N).

As it is, there are rooms for a positive information rent if and only if $\dfrac{p'}{p_1}\dfrac{\tilde{g}_n}{\tilde{g}_{nn}} < \dfrac{1+\sigma}{\sigma}$. Moreover, the lower is σ, the higher is the value of $\dfrac{p'}{p_1}\dfrac{\tilde{g}_n}{\tilde{g}_{nn}}$ at which $R = 0$, with this upper-bound ranging in the interval $(1, 2]$ for $\sigma \in [1, +\infty)$.

Finally, we notice that the higher is $\dfrac{p'}{p_1}\dfrac{\tilde{g}_n}{\tilde{g}_{nn}}$ and, therefore, the smaller (given $g(n,q)$) is the rent, the larger is the responsiveness of b to positive variations in n. Our conclusion is the following.

Whenever an optimal effort n is implemented at which $\dfrac{p'}{p_1}\dfrac{\tilde{g}_n}{\tilde{g}_{nn}}$ is large, the more heavily the principal substitutes perks for wages. In doing so, the principal reduces the agent's rent but increases ex-post efficiency, the agent's marginal rate of substitution being now closer to the relative price of wages and perks.

It is on this ground that our analysis of the efficiency stands. The next two sections are devoted to provide an answer to the issue at hand (in this partial equilibrium framework), looking at it from two different prospectives: the agent's well being (Section 3.8.2), and a measure of production efficiency (Section 3.9).

3.8.2 Rent vs. the agent's well-being

As for the model with perks, the agent's expected utility from working in a perk-less economy is given by $E\left[u(w_s)\right] - g(n) = \hat{u} + \mathcal{R}(n)$. By simple comparison, therefore, we conclude that the expected utility of the agent leaving in the perk-less economy is either greater or lower than the expected utility of the agent leaving in the more sophisticate, perk-allowed economy, depending on whether $\mathcal{R}(n)$ is larger or lower than $\mathcal{R}(n,q)$. Formally,

Proposition6: For all $\langle p^\theta \rangle$ and σ, the agent's equilibrium expected utility with perks, $U^{\langle p \rangle}(\mathbf{w}, q, n)$, is larger than the agent's equilibrium expected utility without perks, $U^{\langle p \rangle}(\mathbf{w}, n)$, if and only if the equilibrium values of effort and perks satisfy the following condition,

$$g(n,q) > \Phi g(n) \;\;\rightarrow\;\; (1+aq)^\gamma < \frac{\tilde{g}(n^q)}{\tilde{g}(n)} \frac{1}{\Phi}$$

with $\Phi \equiv \dfrac{n^q}{n}\left[\dfrac{\sigma n + \theta(1+\sigma)}{\sigma n^q + \theta(1+\sigma)}\right]$, where n^q is the solution to problem P(4) when $q \geq 0$, and n is the effort equilibrium solution in the perk-less economy.

A simple way to interpret the above inequality is that, as long as there is an over-excess of perks, a system with perks might make the agent worse off if their provision fails to trigger a sufficiently high excess of effort. In fact, it is straightforward to see that, for $\sigma = 1$ and $\theta = 0$, the condition for a non-harmful provision of perks results in the following,

$$\frac{\tilde{g}(x,q)}{(1+aq)^\gamma} \equiv g(n,q) > g(n)$$

Besides this limiting scenario, the problem also displays an interesting feature as long as $\theta < 0$ and n^q and n are such that $n < \dfrac{|\theta|(1+\sigma)}{\sigma} < n^q$. In this case, in fact, $U^{\langle p \rangle}(\mathbf{w}, q, n) < U^{\langle p \rangle}(\mathbf{w}, n)$ for all $q \geq 0$. The conclusion is that providing perks to the workers reduces their well-being if it occurs when the technology exhibits a low degree of stochastic dominance (with respect to the benchmark). Finally, it is worth noticing that, if the effort levels satisfy $n = n^q = \bar{n}$, a system with perks does never make the agent's strictly better-off, while it makes her strictly worse off at any equilibrium allocation for which perks are paid in positive amounts. Intuitively, a sufficient condition for the perks being harmful for the agents is that their effort is already at its maximum level.

3.9 A closed-form solution to (P4)

The main purpose of this section is to bring to unity the partial analysis of the optimal allocation as it has been discussed so far. Since finding a solution to the principal's problem comes down to solve a system of non-linear equations, for which the existence of a solution itself is not guaranteed, we derive a specific closed-form solution to problem (P4) only for a benchmark case, in which any complementarity/substitutability between wages and perks is ruled out.

The focs to problem P(4), left out the one on s, are the following,

$$\mathcal{L}_b : \qquad \frac{p' u_c(b)}{g_{nn}(n,q)} \frac{p'}{p_1} \Delta = (1 + \tau_b)$$

$$\mathcal{L}_q : \qquad \frac{|h_q| \tilde{g}_n}{\tilde{g}_{nn}} p' \Delta = k$$

$$IC : \qquad p'(u(b) - u(s)) = g_n(n,q)$$

Each of them simply states that, at the equilibrium allocation, the marginal benefit of setting any instrument to the corresponding optimal amount equals its marginal cost - for the principal as well as for the agent. Because of the specific functional forms we have assumed, we obtain

$$\mathcal{L}_b : \qquad \frac{(1+aq)^\gamma}{\delta n^\sigma} = \frac{n+\theta}{n} \frac{\sigma(1+\tau_b)}{\alpha \Delta \cdot u_c(b)}$$

$$\mathcal{L}_q : \qquad \frac{(1+aq)^\gamma}{\delta n^\sigma} = \frac{\gamma a \alpha \Delta}{k(1+aq)\delta \sigma n^{\sigma-1}}$$

$$IC : \qquad \frac{(1+aq)^\gamma}{\delta n^\sigma} = \frac{1}{\alpha(u(b) - u(s))}$$

The LHS of each of these equations represents the (inverse of) the agent's marginal cost of effort. As already discussed, if we look at $\mathcal{L}_b$, its equilibrium value is a function of n and b alone. However, as long as, the probability distribution is such that the effect of effort on p_1 equals that on $\tilde{g}_n$, for any given value of the parameters, the principal sets b regardless of which n has to be implemented. For $\sigma = 1$, moreover, optimality with respect to q imposes to fix the agent's marginal cost of effort as a function of q and b alone. It is so because, for $\sigma = 1$, the responsiveness of n with respect to b and that of $\dfrac{\partial n}{\partial b}$ with respect to q are independent of n.

Dividing the two focs for IC, we obtain

$$\mathcal{L}_q/IC : \qquad 1 = \frac{\gamma a \alpha \Delta}{k(1+aq)\sigma} \frac{\alpha(u(b) - u(s))}{\delta n^{\sigma-1}}$$

$$\mathcal{L}_b/IC: \qquad 1 = \left(\frac{n+\theta}{n}\right)\frac{\sigma\delta(1+\tau_b)}{\alpha\Delta \cdot u_c(b)}\frac{\alpha(u(b)-u(s))}{\delta}$$

and thus,

$$\left(\frac{n+\theta}{n}\cdot n^{\sigma-1}\right)\frac{k(1+aq)}{\gamma a} = \left(\frac{\alpha\Delta}{\sigma}\right)^2\frac{u_c(b)}{\delta(1+\tau_b)}$$

Therefore, if $\sigma = 1$, i.e. there is no feedback process through n on the labor optimal policies, and $\theta = 0$, i.e. the marginal effect of n on $\tilde{g}_n$ equals its effect on p_1, either i) b is not a function of n; ii) given b, g_n is fixed; iii) the substitutability between b and q does not depend on n.

For $\sigma = 1$, however, concavity on q demands $0 < \gamma < 1$. For the most part we can neglect what γ is, and look only to the relative cost $\tilde{k} \equiv \dfrac{k}{a\gamma}$ when computing q. Then, for higher γ, (IC) implies higher n.

With $\sigma = 1$ and $\theta = 0$, $\mathcal{L}_b$ becomes

$$\frac{b^\eta}{s^{\eta-1}} - (2-\eta)b - (\eta-1)\mathcal{A} = 0$$

with $\mathcal{A} \equiv \dfrac{y_1 - y_0 + s(1+\tau_s)}{1+\tau_b}$.

By differentiating this equation with respect to the parameter η, we obtain an expression for the sensitivity of the optimal b to the agent's relative risk aversion, that will be useful later on, when discussing the dynamics that lead to an optimal marginal tax rate. In fact,

$$\frac{\partial b}{\partial \eta} = \frac{(\mathcal{A}-b) - s^{1-\eta}b^\eta \log b}{\eta\left[\left(\dfrac{b}{s}\right)^{\eta-1}+1\right]-2}$$

which is negative for sufficiently low $0 < \eta < 1$ and it is negative for $\eta \geq 2$, provided that $\dfrac{\mathcal{A}}{b} < 1 + \left(\dfrac{b}{s}\right)^{\eta-1}\log b$. For $\eta = 2$, moreover, the optimal b is set as to satisfy

$$\frac{b^2}{s} - \mathcal{A} = 0$$

That is, $b = \sqrt{\mathcal{A}s}$. When $y_0 = s = \hat{s}$, and $\tau_s = 0$, the optimal rule for the bonus states $b = \sqrt{\dfrac{\hat{s}}{1+\tau_b}}y_1$, which clearly relates the equilibrium wage to i) a measure of the agent's effort productivity, y_1; ii) the distorting marginal tax

rate (where the marginality of τ_b comes from the fact that we fixed $\tau_s = 0$); iii) the labor market structure (through the minimal wage requirement, $\hat{s}$); iv) the agent's risk aversion (from which the square rout has been derived).

From $\mathcal{L}_q/\mathcal{L}_b$, we know that $\dfrac{k(1+aq)}{\gamma a} = \left(\dfrac{\alpha \Delta}{\sigma}\right)^2 \dfrac{u_c(b)}{\delta(1+\tau_b)}$.

As long as $\eta = 2$, therefore, $(1+aq) = \dfrac{\gamma a}{k} \dfrac{\alpha^2}{\delta(1+\tau_b)} \dfrac{\Delta^2}{b^2}$ implies

$$q = \max\left\{\frac{\gamma \alpha^2}{\delta} \frac{(1+\tau_b)}{k} \left[\sqrt{\frac{\mathcal{A}}{s}} - 1\right]^2 - \frac{1}{a}, \ 0\right\}$$

The main result as to regard q is that its optimal amount is not monotonic in τ_b. As we can see, indeed, the effect of the top-income tax on perks can be decomposed into a (positive) substitution effect, that operates through the relative price of wages to perks, and a (negative) income effect derived from the fact that an higher tax reduces the principal's expected gain from incentivizing a movement from state 0 to state 1. Moreover, to this direct effect of the labor income tax on Δ, we have to add the indirect effect that changes on n, which occur because of the direct substitution effect of τ_b on b, have on the probability that Δ is gained. Whenever the complementarities between n and q are strong, and the agent's effort is sufficiently sensitive to variations of b, the negative income effect is more likely to outweigh the substitution effect.

In fact, whenever $y_0 = s = \hat{s}$ and $\tau_s = 0$ hold, $\dfrac{\partial q}{\partial \tau_b} < 0$ for all y_1 such that the following condition is satisfied, namely $\dfrac{y_1}{\hat{s}} > 4(1+\tau_b)$.

Finally, from the agent's incentive compatibility constraint, $n = \dfrac{\alpha}{\delta}(1+aq)^\gamma \left[u(b) - u(s)\right]$. For $\eta = 2$, therefore, the equilibrium effort is

$$n = \frac{\alpha^{1+2\gamma}}{\delta^{1+\gamma}} \left[\frac{\gamma a}{k}(1+\tau_b)\right]^\gamma \left[\sqrt{\frac{\mathcal{A}}{s}} - 1\right]^{2\gamma} \left[\frac{\sqrt{\mathcal{A}} - \sqrt{s}}{s\sqrt{\mathcal{A}}}\right]$$

By letting be $\gamma = \dfrac{1}{2}$, we obtain that

$$n = \frac{\alpha^2}{\delta^{\frac{3}{2}}} \sqrt{\frac{a}{2k}(1+\tau_b)} \frac{\left(\sqrt{\mathcal{A}} - \sqrt{s}\right)^2}{\sqrt{\mathcal{A}s^3}}$$

3.9.1 Elasticities w.r.t. the tax rate

To our main purpose, we compute and briefly discuss the elasticities with respect to the tax rate of either the wage-bonus and the agent's effort. It is, unfortunately, with loss of generality that we focus on the case where $\theta = 0$, $\eta = 2$, $y_0 = \hat{s} = 1$, and $\tau_s = 0$. Then,

$$\epsilon_{b,\tau} \equiv \frac{\partial b}{\partial \tau_b} \frac{\tau_b}{b} = -\frac{1}{2} \frac{\tau_b}{1 + \tau_b},$$

which is negative, decreasing and convex. Moreover, we note that $\epsilon_{b,\tau} \to 0$ as $\tau_b \to 0$.

Under our parametrization, $n^\star = \phi \frac{1 + \tau_b}{\sqrt{y_1}} \left(\sqrt{\frac{y_1}{1 + \tau_b}} - 1 \right)^2$, where $\phi = \frac{\alpha^2}{\delta^{\frac{3}{2}}} \sqrt{\frac{a}{2k}}$. Therefore,

$$\epsilon_{n,\tau} \equiv \frac{\partial n}{\partial \tau_b} \frac{\tau_b}{p_1} = -\frac{\tau_b}{1 + \tau_b} \bigg/ \left(\sqrt{\frac{y_1}{1 + \tau_b}} - 1 \right)$$

Unsatisfactory as it may be, that the elasticity of the effort (and therefore of the equilibrium probability, p_1) with respect to the top-income marginal tax rate is not affected by the preferences for perks (being somehow less unsatisfactory its unresponsiveness to their price), we stress on the result that, for any $\tau_b > 0$, $\epsilon_{n,\tau} \to 0$ if and only if $y_1 \to \infty$. Nonetheless, we must notice that, for $y_1 \to \infty$, also $\mathcal{L}_n \to \infty$, so that $n = \bar{n}$.

3.10 A measure of efficiency for perks

In this section, we propose a simple measure of efficiency so as to assess the (market) value of an equilibrium allocation with perks and compare it to the value of the equilibrium allocation we would obtain in a perk-less economy.

For our purposes, we denote by "efficiency" the capacity of either technology system (the first one applying both effort and perks; the second one applying only effort) to realize an higher expected income (i.e. real resources), net of i) the agent's work cost, and ii) all the expenses, in perks (if any) and taxes.

Definition(E, Efficiency): We say that the equilibrium allocation with perks is efficient, or that the perk system is a resource-improving technology, if and only if the equilibrium expected value of production, net of taxes, expenditure in perks and agents' cost of effort, is larger in the economy with perks than in the perk-less economy. That is,

$$E^{(q)}\Big[y_s - T(w_s)\Big] - kq - g(n,q) > E\Big[y_s - T(w_s)\Big] - g(n) \qquad (10)$$

By doing so, we are interested in understanding whether the payment by perks, whose amount is decided by the principal in order to maximize his own profits, can make the equilibrium allocation with perks (socially) inefficient as compared to the equilibrium allocation we would obtain in a perk-less economy. If that is not the case, then, it must be that the second-order effects of perks on either the implementable effort (which is positive) and taxes (which is ambiguous) does balance the direct negative effect of their cost.

Up to this point, the analysis assumes that the tax system is the same for either technology. Nonetheless, it makes grounds for the enquiring of the general gains from perks when the tax system itself is allowed to optimize on the tax rates, when taking into account the firm's incentives to provide perks, and their effect on efficiency/redistribution. We will attempt to give an answer to this question in Section 3.11, by solving the first-stage of the interaction game between the government and the labor market. Since that problem requires the definition of a sufficiently tractable social welfare function, however, we find convenient and yet interesting to adapt the framework discussed so far, in order to assess the (monetized value of the) efficiency of providing perks.
The next section goes through all the details of our efficiency indexes.

3.10.1 The normalization procedure

Because of the separability between wages and effort in the agent's utility function, it seems unreasonable to directly refer to $g(n,q)$ as the monetary equivalent of the agent's effort. The drawback of this is that, since $g(\cdot)$ is expressed in utility terms whereas monetary values are expressed in units of the numeraire, our efficiency index, and the comparative statics results for it, may be affected by the scale of the $g(\cdot)$ function.
In order to get rid of this effect, we apply a simple compensation principle and compute the principal's expenditure in perks that, given the optimal allocations with and without perks, is needed in order to impose on the agent the same cost of effort across models. Computed the required differential in q, we value it at the market price k [8]. The rest of this section provides a detailed explanation of

---

[8]Alternative to this pricing would be to apply to the differential in perks their shadow price, i.e. the marginal price the principal is willing to pay for a relaxation of the constraint $q = 0$, namely χ_q.

how the compensation is carried over. Results, based on numerical simulation, are discussed in Section 3.11.

Let denote by $x \in [0, \bar{n}]$ the optimal effort, solution to (P4) when $q \geq 0$; and by $n \in [0, \bar{n}]$, the optimal effort in the restricted model with $q = 0$. Then, we require, either

1. $g(x, q + dq) = g(n)$, or

2. $g(x, q) = g(n, dq)$

where dq is the differential change in q that, starting from the equilibrium, is required in order to equate the cost of effort across the two models.

In case 1, dq is computed so as to bring the agent's cost of effort that prevails in the perk economy to the value it takes in the perk-less economy (i.e. he agent's cost of effort of the perk-less economy is the one that has to be implemented). Therefore, either $dq > 0$ if $g(n) < g(x, q)$ or $dq \leq 0$, otherwise.

The history behind this adjustment can be easily told as follows. Let assume that the contracting game is such that the principal only commits to provide q, and that, on that account, the agent takes effort x. The die is cast and either state of the world is realized. Then, a social planner call on the principal and ask him to adjust the provision of perks such that the end-of-game agent's disutility of effort is $g(x, q + dq) = g(n)$. The question is therefore the following: from an ex-ante perspective, whenever perks are provided, do they guarantee that there are sufficiently high expected resources in the economy for the principal being able, if he were asked to do so, to buy extra dq units of perks, at the market price k , in order to leave the agent indifferent (as to regard her cost of effort) between the two systems?

By solving the following equation,

$$\frac{\delta x^{1+\sigma}}{1+\sigma} \frac{1}{(1 + a(q + dq))^\gamma} = \frac{\delta n^{1+\sigma}}{1+\sigma}$$

we obtain that the dq required to implement $g(n)$ amounts to

$$dq = \frac{1}{a}\left(\left[\frac{x}{n}\right]^{\frac{1+\sigma}{\gamma}} - 1\right) - q \qquad (11)$$

When, in the model w/o perks, we impose $q = 0$, the Lagrangian associated to that constraint, can be either positive or negative, depending on which direction the principal would like to violate the constraint. (See Math Appendix 3.B)

For case 1, therefore, the compensated efficiency index for the perk economy is as follows,

$$E^{(q)}\Big[y_s - T(w_s)\Big] - k \cdot (q + dq) - g(n)$$

In case 2, instead, the effort cost that has to be implemented is the one with perks. If it is therefore the case that $g(x, q) > g(n)$, a negative dq is needed to bring $g(n, dq)$ at the level $g(x, q)$. Alike to the previous case, we can think of the social planner buying from the principal in the perk-less economy, at price k, dq units of his plant, so as to impose on the agent an end-of-the-game disutility of effort equal to $g(x, q)$. The compensated efficiency index, for the perk-less economy, is given by

$$E\Big[y_s - T(w_s)\Big] + k \cdot dq - g(x, q)$$

That is, in the scenario without perks, we compute dq such that

$$\frac{\delta x^{1+\sigma}}{1+\sigma} \frac{1}{(1 + aq)^\gamma} = \frac{\delta n^{1+\sigma}}{1+\sigma} \frac{1}{(1 - adq)^\gamma}$$

and obtain

$$dq = \frac{1}{a}\left(1 - \left[\frac{n}{x}\right]^{\frac{1+\sigma}{\gamma}}\right) - q \cdot \left[\frac{n}{x}\right]^{\frac{1+\sigma}{\gamma}} \qquad (12)$$

Since it can be argued against the planner eating some of the "intangible" structure of the perk-less firm, we propose an alternative compensation (named case 2b) that, as for case 1, implements $g(n)$ and requires the planner of the perks economy to provide the agent with extra units of perks.

The difference between case 1 and case 2b relies on the marginal effect of the extra units being different across types of compensation. While, in fact, dq is weighted according to the parameter a in case 1), in case 2b) the sensitivity of $h(q)$ to dq is measured by $a \cdot \left[\frac{x}{n}\right]^{\frac{1+\sigma}{\gamma}}$. We can, therefore, think of $\left[\frac{x}{n}\right]^{\frac{1+\sigma}{\gamma}}$ as the personal benefit the agent derives by his position in the firm. For given x, the larger is that responsiveness, the less units of dq are required in order to bring the agent's work disutility to its value $g(n)$. Formally, we solve the following identity,

$$\frac{\delta x^{1+\sigma}}{1+\sigma} \frac{1}{\left(1 + a(q + dq \cdot \left[\frac{x}{n}\right]^{\frac{1+\sigma}{\gamma}})\right)^{\gamma}} = \frac{\delta n^{1+\sigma}}{1+\sigma}$$

It is straightforward to see that case 1) and case 2b) are such that $dq^{1)} = dq^{2b)} \left[\frac{x}{n}\right]^{\frac{1+\sigma}{\gamma}}$, with $dq^{2b)} < dq^{1)}$. The difference is due to the fact that in case 1) $g(n)$ is implemented, whereas in case 2b) the responsiveness to the extra units of perks is adjusted and dq is computed so as to implement $g(n,q)$. To conclude, the efficiency index that is relevant in case 2b), which modifies the expression for the perks economy, is the following,

$$E^{(q)}\left[y_s - T(w_s)\right] - k \cdot (q + dq^{2b)}) - g(n)$$

3.11   A numerical exercise

This simulation exercise serves two purposes. Firstly, it extends our theoretical results and comparative statics. The reference is precisely to the discussion on the second-order effects associated with the curvature of the utility function, and on the complementarity/substitutability of perks and cash at the equilibrium allocation, which unambiguously requires to control for the strict concavity of the problem. The second target is to evaluate quantitatively our efficiency indexes and draw some conclusions on the convergent or divergent incentives between firms and fiscal authority as to regard the payment by perks.

In order to proceed, we postulate the main parameters according to either our feasibility analysis and the values that appear typical in the literature on optimal taxation. We use the values $\eta = \{0.5, 1, 2\}$, and $\sigma = \{1, 2\}$. In fact, according to Bernasconi (1998), the coefficient of relative risk aversion should lie in the range 1-2, whereas many empirical studies seem to confirm a Frisch elasticity that is less than one. As we need $\gamma < \sigma$, we set $\gamma = \sigma/2$ for convenience. Since we do not know the appropriate values of α, ε and θ, which determine the stochastic distribution of income, we let them range around the values $0.8, 0.05$ and 0, respectively. We set $\tau_b = 0.25$ if $\eta \leq 1$, and $\tau_b = 0.45$ for $\eta = 2$, or let it vary in the interval $[0.2, 0.6]$. We fix $\tau_s = 0$, and $y_0 = \hat{s} = 1$, while we allow y_1 to vary in the interval $(2, 4)$, or we set it at 3.5, when varying some other parameter. The values of $\delta = \{0.25, 0.5, 1\}$ and $a = \{2, 4.5, 5\}$, are adjusted so as to maintain the cost of perks k around 1.

The effect of a change in the agent's relative risk aversion, η. We first notice that an increase in η leads to a drop in the optimal effort which, everything else equal, reduces the principal's incentives to provide perks. As it is, either $\frac{q}{y_1}$ and $\frac{q}{b}$ fall down as η increases. The intuition behind the result is the following. The more relatively risk-averse the agent is, the lower is the risk-premium that the principal has to pay in order to incentivize the agent's effort. Nonetheless, the level of effort that can actually be implemented decreases in η, so reducing the scope for providing perks. As an overall result, the principal expected profits decrease, but, because of the provision of perks, they fall less than they do in the perk-less economy. The agent, however, is made worse off from the provision of the fringe benefit at those η's for which it is still convenient for the principal to pay her in work-related goods and incentivize effort above the optimal level it takes in the perk-less economy. When it is the case, in fact, the information rent paid to the agent in the perk system is higher (though still being decreasing in η) than it is in the perk-less economy, because of the positive effect of perks on effort. Nonetheless, the utility derived from the lower wages and perks that are needed to induced the more risk-averse worker to apply the desired effort is not enough to outweigh her disutility from working. The mismatch between the linearity of the principal's payoff and the concavity of the agent's utility gives an incentive to the principal to use perks in order to distort the agent's working decision.

When looking at the behavior of our efficiency indexes as η increases, we see that, besides an overall tendency to state the inefficiency of perks when compensation is performed so as to implement the perk-less economy's effort (case 1), the gain from providing perks is U-shaped when efficiency is measured as in case 2. That is, the system with perk performs better than the perk-less economy at sufficiently low values of the coefficient of relative risk aversion. Nonetheless, as η increases, the index decreases, so making perks a less-efficient technology for a bunch of relatively higher coefficients, with a change in tendency at some point and a progressively decreasing inefficiency as we approach the level of η for which $q = 0$.

The inefficiency that arise in case 1 can be easily understood if one looks at the principal's willingness to pay for a relaxation of the constraint on a non-positive provision of perks. Given our simple parametrization, that willingness is very low compared to the market price of perks, k. Since, because of the combined effect of relatively low risk-aversion and perk provision, the amount

of extra units of perks that the principal has to pay to the agent, in order to bring her disutility to its perk-less counterpart level, is very large (higher than the amounts required in case 2 and 2b), at a price k there may be not sufficient resources left for that compensation to take place.

The effect of a change in the stochastic parameter, θ. Here, the most striking result, somehow related to the second-order gains that the principal can potentially obtain at $\theta \neq 0$, by deviating from the optimal contract, is that, for all η, as θ varies in a small interval centered at zero, either $\frac{q}{y_1}$ and $\frac{q}{b}$ exhibit an inverse U-shape pattern, with a maximum at $\theta = 0$. Though this behavior reflects the increased concavity of the ratio $\frac{n}{\bar{n}}$, both being decreasing and concave functions of θ, it does not affect the trend of $\frac{b}{y_1}$, which strictly decreases as θ increases in its interval. When compared with its value in the perk-less economy, the strictly positive (respectively, strictly negative) variation in the optimal compensation b for $\theta \in (0, \bar{\theta}]$ (respectively, $\theta \in [\underline{\theta}, 0)$) is, however, too small (in quantitative terms) to explain the worker's excess of utility derived from the provision of perks. In fact, that gain chiefly results from the higher rent the agent is let free to enjoy along the interval, in the perk economy. Moreover, since i) because of our parametrization, the rent equals the disutility of effort, and ii) the effect on the optimal wage is very small, the percentage increment in the principal's expected profits, achieved by providing perks, is quantitatively unimportant.

As θ increases, either the model with perks and the perk-less economy perform better in terms of net expected resources, after compensation has taken place. However, all our efficiency indexes show the relative inefficiency of providing perks, as compared to a system that does not allow for them, with a trough near $\theta = 0$. The result being robust to changes in σ and η, our analysis suggests that, by focusing on $\theta = 0$, we are actually overestimating the effect of perks in the economy.

The effect of increasing y_1. For given prices, τ_b and k, and fixed $y_1 = \hat{s}$, the overall effect of increasing y_1 amounts to a steady increase in the ratio $\frac{q}{y_1}$ (as soon as the income gap is such that it becomes optimal for the principal to pay by perks also) and, therefore, to a higher ratio $\frac{n}{\bar{n}}$ in that region, as compared to a system without perks. In spite of the negative trend of $\frac{b}{y_1}$, which is robust

to changes in η, θ and σ, the perk-to-wage ratio increases in the income level, and reaches its maximum in correspondence of the y_1's value at which $n = \bar{n}$. Further increments in y_1 are then associated with lower $\dfrac{q}{y_1}$ and $\dfrac{b}{y_1}$, so as to keep the ratio perk-to wage constant.

The net effect of the agent's higher productivity, if positive on the principal's expected profits, is ambiguous from the agent's prospective. Indeed, either the information rent and the agent's expected utility may fall below the values they take in a perk-less economy. Moreover, differences in the relative risk aversion may lead to different dynamics as to regard the worker's well-being. In fact, though the agent's expected utility with perks is initially decreasing in y_1, as the principal starts using them, reaches a minimum and then increases above its perk-less level, for $0 < \eta < 1$; for $\eta \geq 1$, this dynamics is reverted, with the agent's utility gains from perks being firstly positive and increasing, as the rent increases in her effort, and then starting to fall, as soon as the implementable effort hits its maximum. We will refer to these differences later on in this chapter, when discussing the implications of a lower relative risk-aversion on the designing of the optimal top-income tax.

Finally, let notice that, whenever the focus is the production efficiency, the agent's relative risk aversion does not change the behavior of any of the efficiency indexes as the income level increases. In fact, other than a difference in quantitative terms, which is larger the higher is η, all indexes split the interval of income gaps into two regions. At first, as soon as perks are provided, the less advanced system performs, in expected term, better than our model with perks. Then, after a trough, the latter system decreases its inefficiency, until it overcomes the perk-less economy in exploiting the higher potential of the agents, as embodied in larger y_1's. Furthermore, it is worth mentioning that, because of the linearity with respect to effort and perks of our indexes, for all y_1 such that $n = \bar{n}$, though the perk-to-wage ratio and the effort keep constant as income increases, the efficiency from providing perks is characterized by a less steep, but still increasing, trend.

The effect of increasing τ_b. When designed so as to measure the impact of an increase in the tax rate (but the same hold for an increase in k), our numerical solutions corroborate the analytical result of the complementarities between wages and effort and those between effort and perks being the driving forces in explaining the reduction in q as the income tax increases.

In fact, as soon as the model's parametrization is such that incentives are so strong that effort hits its maximum in a range of positive values for τ_b, the substitution effect alone (those complementarities have being ruled out) leads to an increase in the optimal provision of perks and to a drop in wages. Nonetheless, as soon as the agent's effort is pushed down by the reduction in wages and by the decreasing marginal returns of perks, higher tax rates are associated with less perks and a fall (at a decreasing speed) in wages, as compared to the value they take in the perk-less economy. In the broader interval, therefore, the net effect on the perk-to-wage ratio is no longer monotonic in τ_b, being initially increasing, and then sharply decreasing in the income tax, until it goes to zero. Though higher distorting taxes generally imply lower expected utilities and smaller information rents, the gain/loss from being paid with perks, as compared to a perk-less scenario, depends on the agent's risk aversion. For $\eta = 0.5$, the worker is made worse off by the (temporary) provision of perks, whereas, for $\eta = 2$, the agent is better off when provided by perks, for at least some values of τ_b in our interval. Moreover, in either case, optimal responses are such that they change their trend just once, i.e. the agent's expected utility is (weakly) concave in the tax rate. As we already pointed out, these are the dynamics we will be concerned about when characterizing the equilibrium top-income marginal tax rate.

Finally, strong of the intuitions on how variations in the principal's expected profit's, through their effect on the willingness to incentivize effort, spread over the optimal policy functions, it is not surprising that the efficiency indexes' response to changes in the tax-rate shears the same pattern, but reflected, than the response to changes in y_1. In fact, our system with perks shows an index of relative efficiency that: i) is strictly positive and firstly increasing (if $n = \bar{n}$) and then decreasing (as effort decreases) for some τ_b sufficiently low; ii) becomes negative, with a global minimum, over a range of intermediate values; iii) slightly reduces its inefficiency, until the provision of perks turns to zero for sufficiently high τ_b.

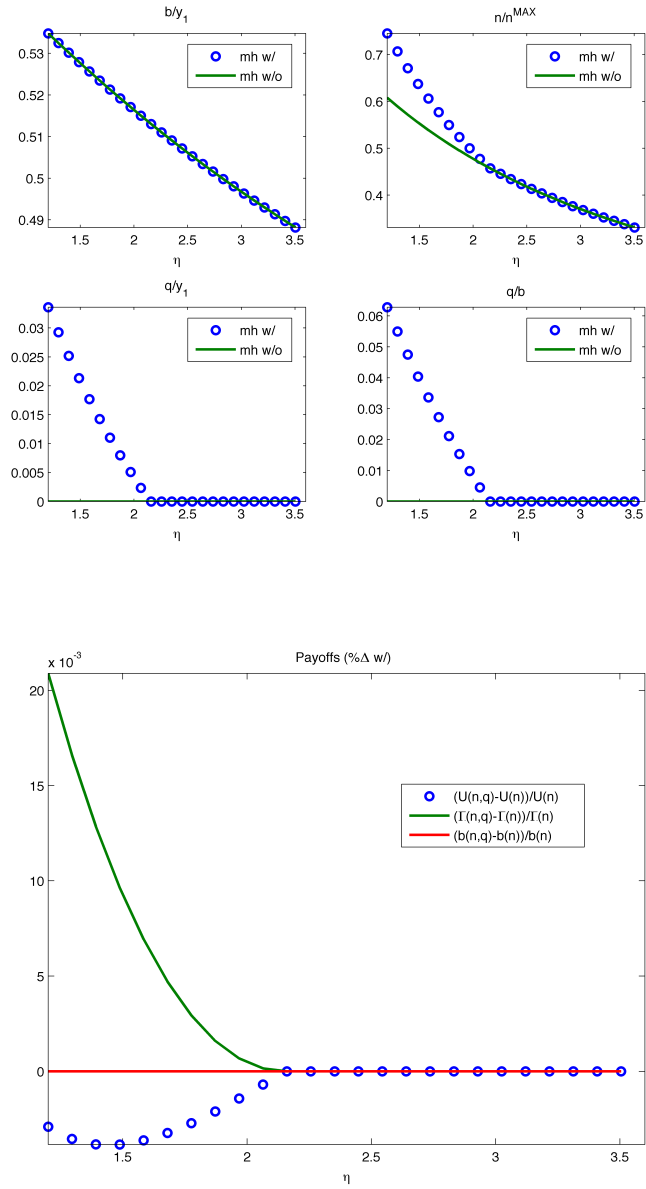Figure 3.1: For η: Optimal allocation ($\theta = 0$)

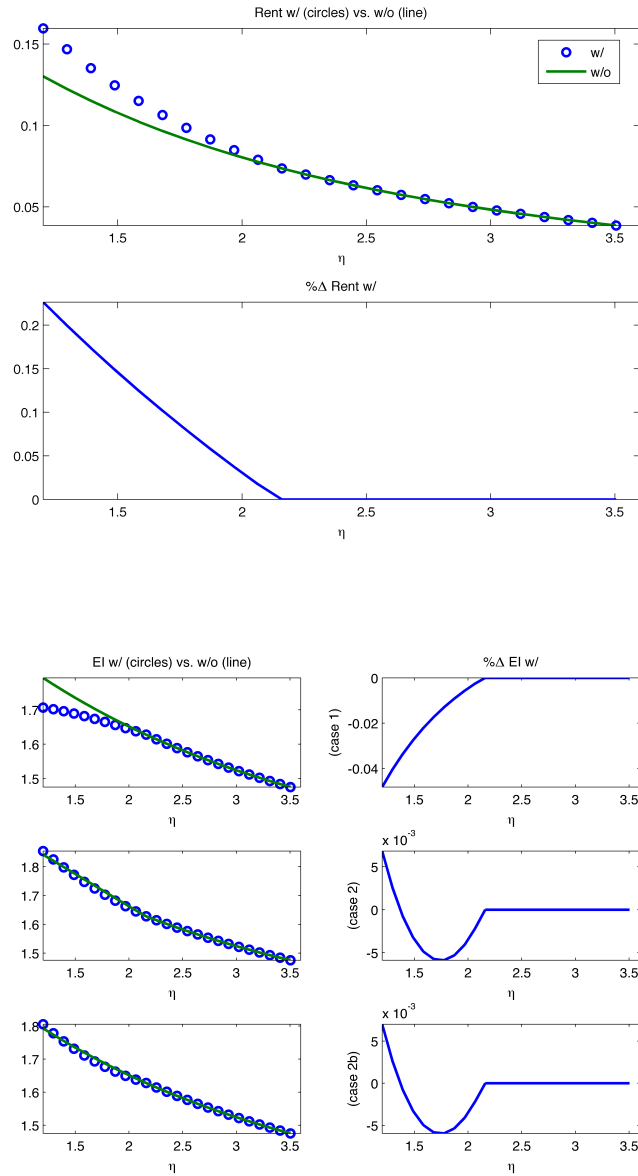Figure 3.2: For η: Information rent and efficiency $(\theta = 0)$

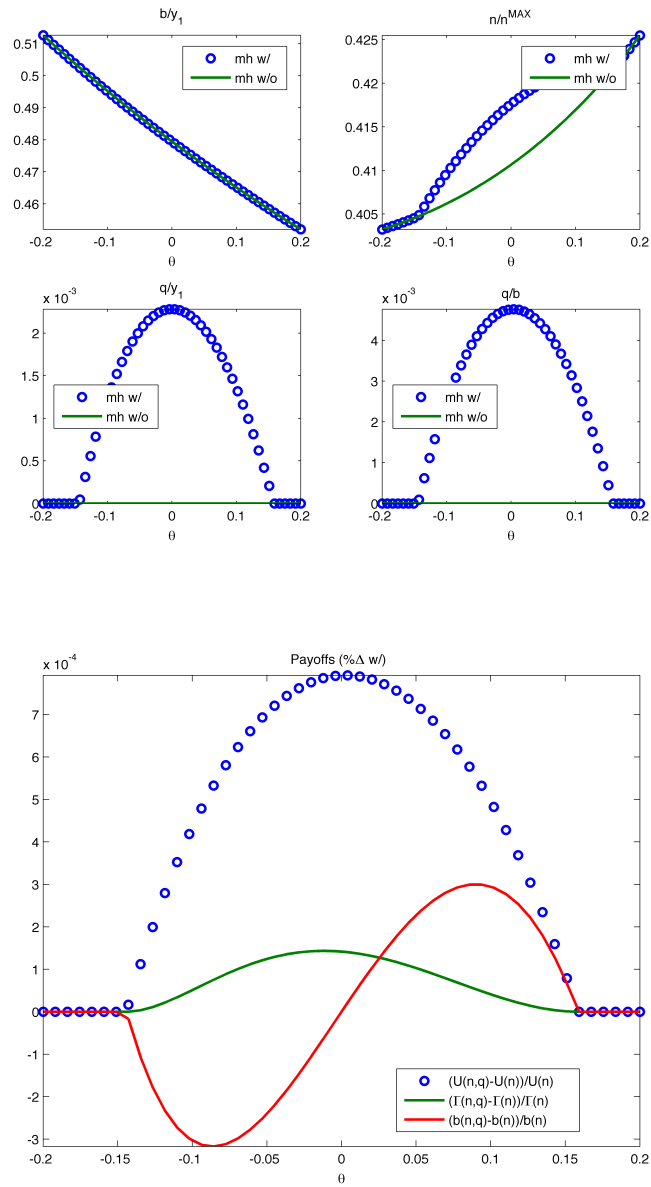Figure 3.3: For θ: Optimal allocation ($\eta = 2$)

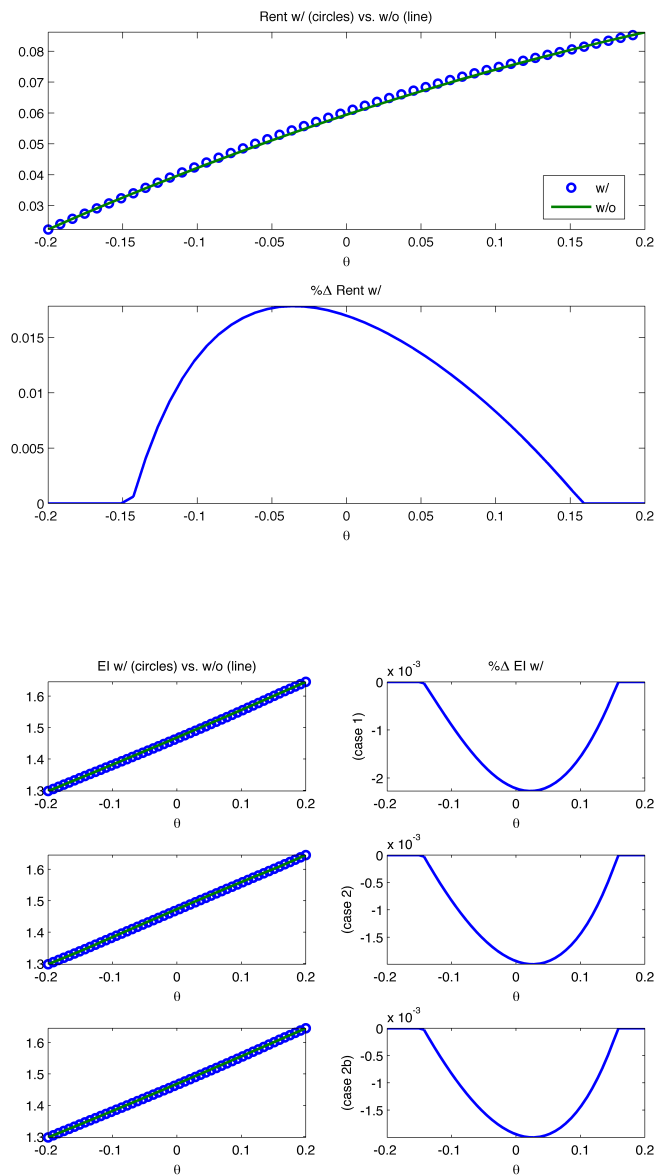Figure 3.4: For θ: Information rent and efficiency ($\eta = 2$)

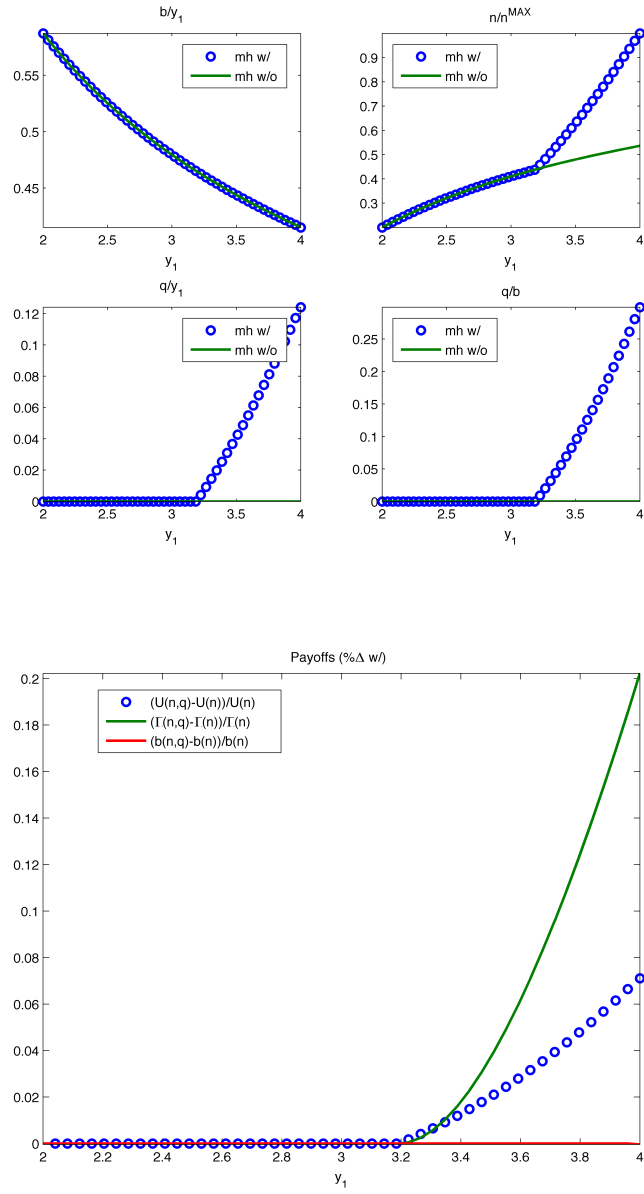Figure 3.5: For y_1: Optimal allocation ($\eta = 2$)

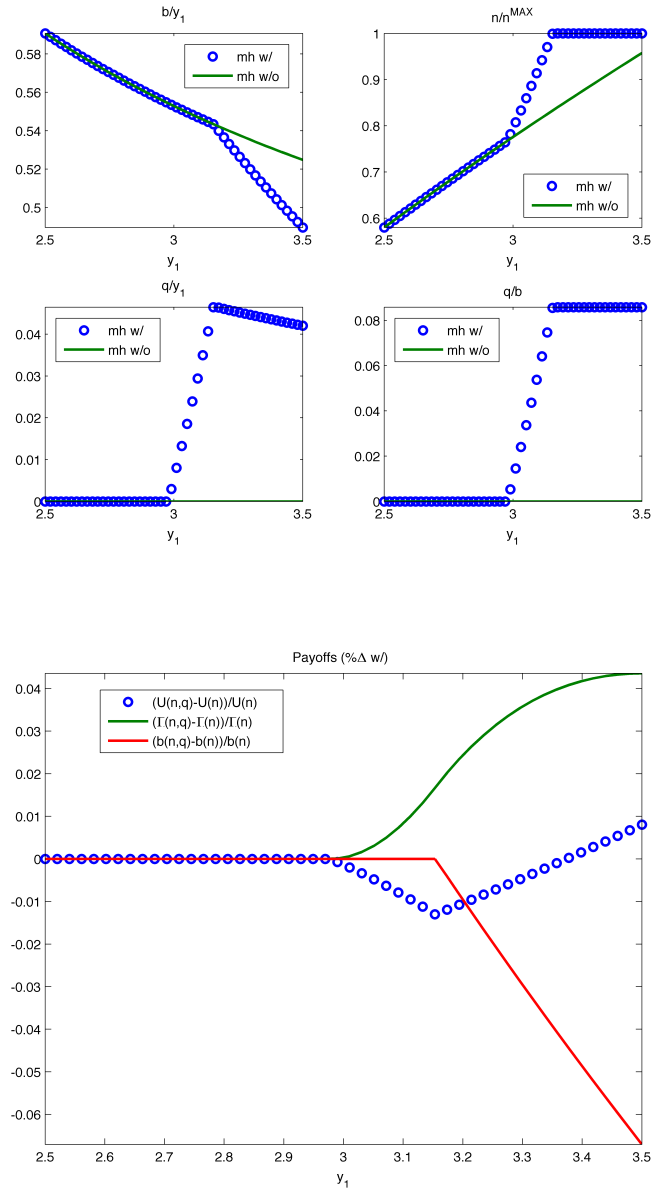Figure 3.6: For y_1: Optimal allocation ($\eta = 0.5$)

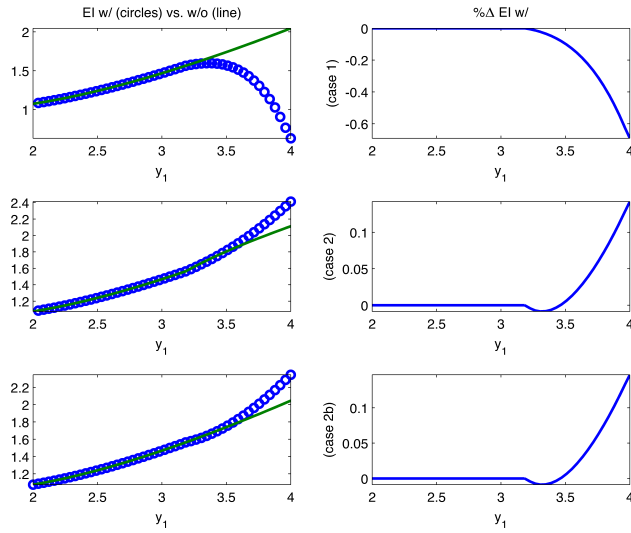Figure 3.7: For y_1: Efficiency ($\eta = 2$)
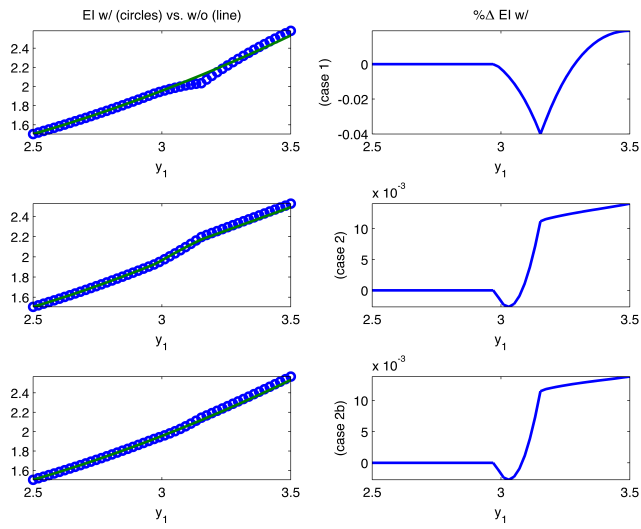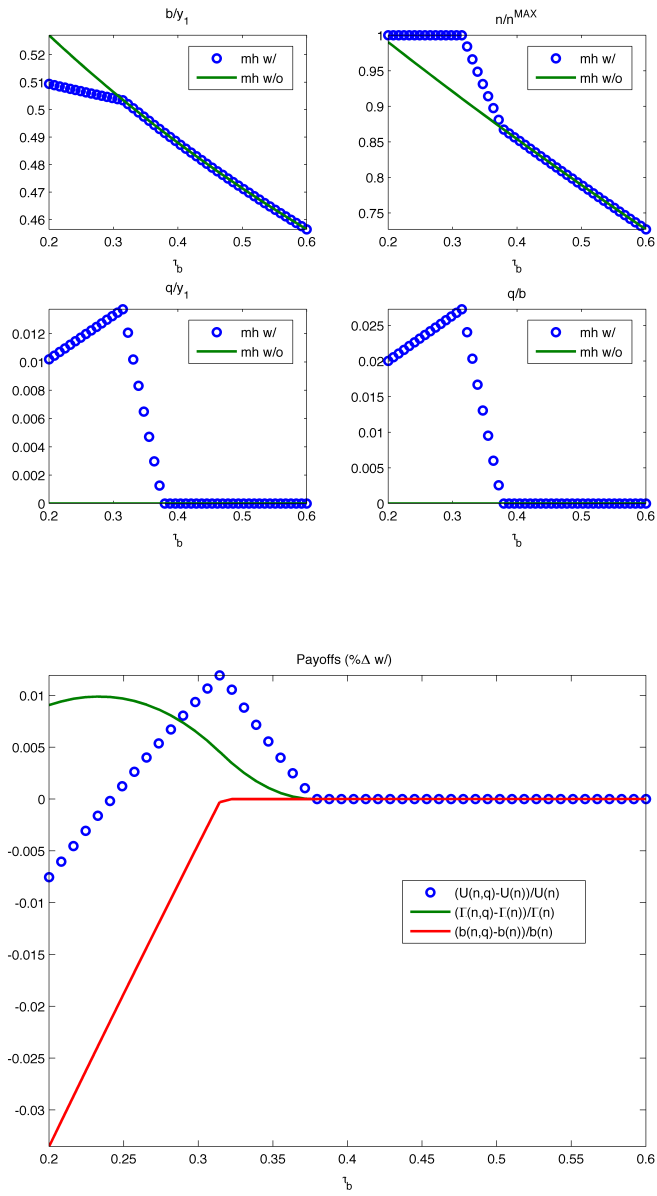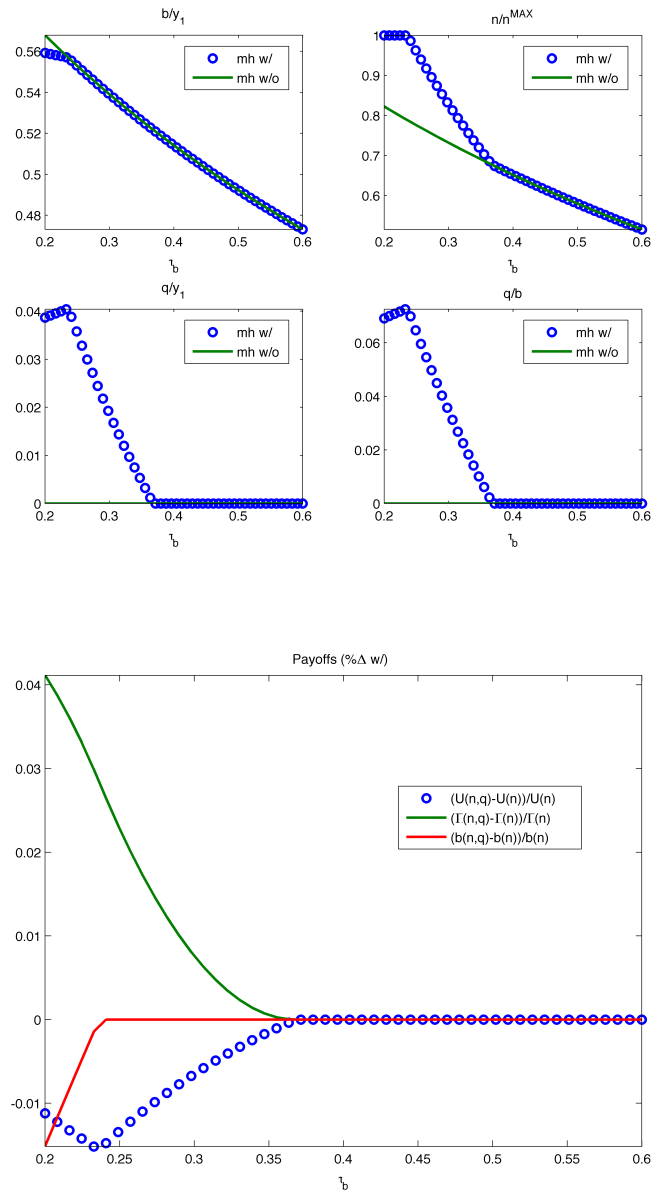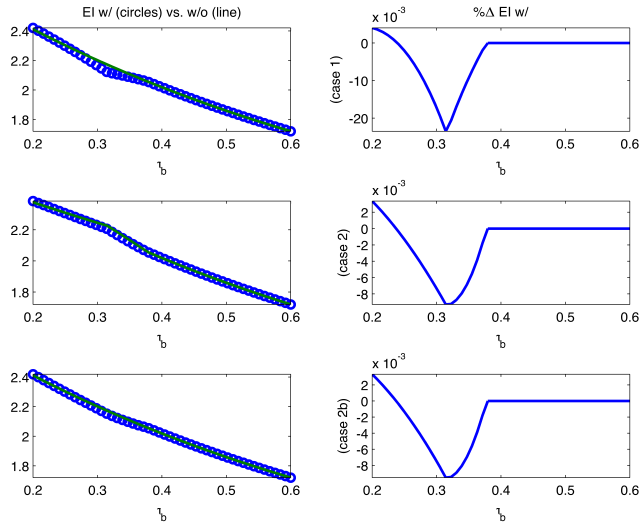

Figure 3.8: For y_1: Efficiency ($\eta = 0.5$)

Figure 3.9: For τ_b: Optimal allocation ($\eta = 2$)

Figure 3.10: For τ_b: Optimal allocation ($\eta = 0.5$)

121

Figure 3.11: For τ_b: Efficiency $(\eta = 2)$

Figure 3.12: For τ_b: Efficiency $(\eta = 0.5)$
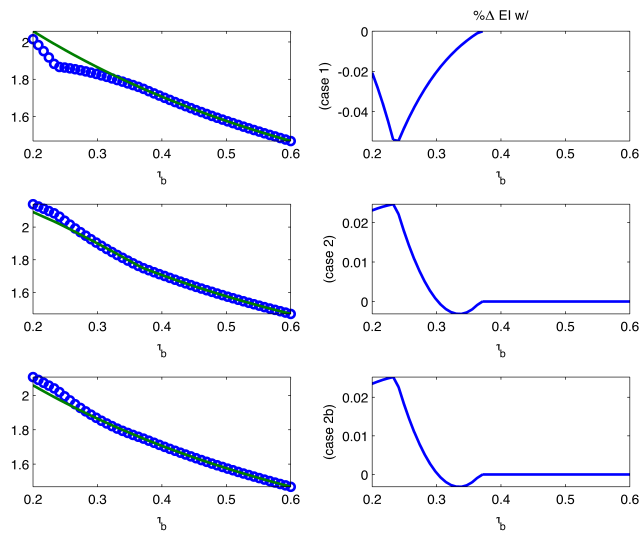
3.12 The first-stage of the Stackelberg game: the optimal top income marginal tax

So far, we have discussed and characterized the outcomes of the second of the two stages in the Nash non-cooperative game between the fiscal authority and the labor market. The sequence of the moves for the complete game can now be stated more formally as follows:

1. The fiscal authority announces the tax scheme and the monetary value of the public expenditure, $\langle \tau_s, \tau_b, G \rangle$.

2. Given the fiscal policy $\langle \tau_s, \tau_b, G \rangle$, the principal designs the optimal compensation scheme $\mathcal{W} \equiv \langle s, b, q \rangle$, and provides the agent with perk q.

3. Given $\langle \tau_s, \tau_b, G \rangle$ and $\mathcal{W}$, the worker decides how much effort n to apply.

4. Income is realized, net wages are paid to the worker, while taxes are collected by the government.

In this section, by solving the first stage of the game we aim at endogenizing the income tax rate that arises in the decentralized economy, when the fiscal authority allows for the provision of the non-taxable perquisite good. Either for tractability purposes and because the part of the income distribution that is more concerned about the provision of perks is the upper tail, we focus on the equilibrium top-income marginal tax rate, and assume that G, which was mostly excluded from the analysis in the second stage, does not now directly affect the agent's incentive compatibility constraint.

Technically, in fact, the problem at hand requires the specification of an objective function for the government (and, therefore, for society) that, though unable to accurately describe the behavior of the fiscal authority, builds at least on some simple normative criterium. On that regard, we develop an alternative approach to optimal taxation based on the notion, not new to other branches of the literature on fiscal policies, of an "ideal" agent.

We then discuss how our results have a bearing on the debate on i) labor supply elasticities, ii) zero top-income tax rate, iii) incentive versus insurance motive of taxation. In doing so, we clearly mean to explore the ability of our theory to explain, jointly, the facts on wages and perk provision, the dynamics of tax revenues and the level of government's expenditure. Our results are in fact in the form of a simple tax-formula that, though in line with those in the literature,

builds on a new theoretical explanation which accounts for the provision of perks.

The next of the section proceeds as follows. The nature of the externality problem and the social welfare function are described in Section 3.12.1. Section 3.12.2 state the maximization problem formally, and derives our tax formula. The main results of the chapter are combined together and discussed in Section 3.12.3. Finally, Section 3.12.4 proposes a numerical solution for the equilibrium marginal tax rate which is compatible with perks being either provided and not.

3.12.1 Set-up and information structure

As before, we consider a model where the government levies taxes to finance the public provision of goods and services. Now, however, to deal with the issue at hand, we explicitly assume that the agent derives utility from both private consumption (i.e. the allocation of wages, perks and effort) and the public expenditure.

The government, thus, chooses the optimal tax rate, given social preferences and the behavior of the labor market. That is, although the government recognize correctly that perks are used to incentivize workers' effort and that perks might substitute out wages, and he derives his budget constraint on that account, in designing the optimal tax rate, he seeks to maximize a properly defined social welfare function.

Since the "ideal", both for the society and the government, is for the worker not to shrink, we assume that the government takes as his objective the expected utility of a representative agent who always applies the maximum effort, $\bar{n}$. We may call such an agent the "ideal" agent. She takes no decision, hence, has no incentive compatibility conditions to satisfy. Decisions, however, are taken in her behalf by the government who, as we said, derives his budget constraint from the behavior of self-interested agents in the labor market.

Characterized the revenues that can be collected when agents take the fiscal policy as given, we derive the optimal marginal tax by equating G to the actual expected revenues and then searching for the value of the tax rate that maximizes the social welfare function.

Throughout all this section we assume $y_0 = \hat{s}$, and $\tau_s = 0$. Then, τ_b can be thought of as the top-income marginal tax rate of this economy.

The extended information set. The nature of the externality problem is as follows. Workers derive utility from labor market outcomes ($\langle s, b, q, n \rangle$) and the publicly provided goods and services (G). Nonetheless, when deciding how much to work and how much to pay, agents take as given the tax system and the public provision of goods, namely $\langle \tau_b, G \rangle$. The government, however, spends the actual expected tax revenues, $\tilde{T}(\tau_b)$, under the balanced budget constraint, $G = \tilde{T}(\tau_b)$. Since the firm may have an interest to pay by perks in addition to wages, by substituting out wages with perks, it can cause a reduction in the amount of taxes, though expected taxes, and therefore, the public expenditure G, may either increase or decrease by the effect of perks on effort and, thus, on probabilities.

3.12.1.1 A refinement of the agent's preferences and SWF

Workers' preferences over public expenditure and private outcomes are given by

$$V^p(\mathcal{W}, G; n) = \mu u(G) + (1 - \mu) \left[p_1 u(b) + p_0 u(s) - g(n, q) \right]$$

The term into square brackets is what we have defined in Section 3.4 as $U^p(\mathcal{W}; n)$. We notice that, whenever the agent takes the public expenditure G as given, the incentive compatibility constraint derived from maximizing either $U^p(\mathcal{W}; n)$ and $V^p(\mathcal{W}, G; n)$ with respect to n is the same, so that our derivation of the optimal allocation $\langle s, b, q, n \rangle$ still applies here[9].

The government uses $V^p(\mathcal{W}, G; n)$ to maximize over τ_b the expected utility of an "ideal" worker for whom $n = \bar{n}$ and $q = 0$. In fact, we are implicitly assuming that the government behaves "as if" he does not know what the provision of perks is. To keep things as simple as possible, we also assume that the government is indifferent about who pays the tax, if the principal or the worker, so that, in designing the optimal tax, he maximizes the expected utility of the ideal worker whose net wages are $b = y_1(1 - \tau_b)$ and $s = y_s(1 - \tau_s) = y_s$. Actually, we can think of this limiting worker as being the entrepreneur of the firm, who retains for himself his productivity (namely, y_1) and pays $\tau_b y_1$ in taxes.

We, therefore, define the following

Definition(IEU, Ideal agent's expected utility): Let denote by

---

[9]The key assumption is, in fact, that though derived from expected revenues, the public provision of good is a certain outcome. In that, our searching for the optimal tax wants to shed light on the desirability of some crowding-out effect between perks and public services.

$$V^p(\tau_b, G) \equiv \mu u(G) + (1 - \mu) \left[\bar{p}_1 u(y_1(1 - \tau_b)) + \bar{p}_0 u(y_s) - g(\bar{n}, 0) \right]$$

with

$$u(c) = \begin{cases} \dfrac{c^{1-\eta}}{1 - \eta} & \forall \eta \neq 1 \\ \log c & \text{if } \eta = 1 \end{cases} \qquad g(\bar{n}, 0) = \dfrac{\delta \bar{n}^{1+\sigma}}{1 + \sigma}$$

the expected utility of an "ideal" worker, whose effort is $n = \bar{n}$ and the wages are her net-of-tax productivity.

Assumption(SWF, Social welfare function): Let $V^p(\tau_b, G)$ give the social preferences between public expenditure and private labor market outcomes in this economy.

Definition(NNE, Nash non-cooperative equilibrium): Each τ_b chosen by the government in the first stage determines a (second-stage) sub-game. The profile of strategies $\langle b^\star(\tau_b), q^\star(\tau_b), n^\star(\tau_b) \rangle$ forms a Nash non-cooperative equilibrium in the sub-game. A sub-game perfect equilibrium of the game is then the tuple $\langle b^\star, q^\star, n^\star, \tau_b^\star \rangle$, where $\tau_b^\star$ maximizes the government's social welfare function subject to the budget constraint, $G = \tilde{T}(\tau_b^\star)$.

The next subsection states the problem formally and characterize our solution.

3.12.2 The government's problem

The government's problem can be stated as follows

$$\text{(P5)} \qquad \max_{\tau_b} V^p(\tau_b, G) \qquad \text{subject to:}$$

$$G = \tau_b p_1(n^\star) b^\star$$

$$0 \leq \tau_b \leq 1$$

$$b^\star = \arg\max_b \Pi^p(n; \mathcal{W})$$

$$n^\star = \arg\max_n U^p(\mathcal{W}; n)$$

The existence of a solution to (P5) depends, of course, on a multiplicity of parameters: directly on $y_1, \langle p(\bar{n}) \rangle, \mu$ and η, and indirectly on all the other parameters which affect $b^\star$ and $n^\star$ [10] . One of the most important parameters, however, is μ, which determines society's preferences for public services compared to the consumption of the private outcomes.

Since we do not know the right value of μ, but - due to our tractable form for the agent's utility function - expected tax revenues (and, therefore, G) do not depend on μ, a calibration strategy should consist in picking a value for $\mu \in (0,1)$ so as to meet a target value for G. We leave it to future works. For the moment, we limit ourselves to pick the value of μ, such that at the equilibrium marginal tax-rate either perks are optimally provided or they are prevented, and test the ability of our framework to account for the differences in the two cases.

3.12.2.1 A formula for the top-income marginal tax

Whenever an interior solution does exist, the first-order necessary condition for an optimum requires

$$\mathcal{L}_\tau : \qquad \left[p_1 b + \tau_b \left(b \frac{\partial p_1}{\partial n} \frac{\partial n}{\partial \tau_b} + p_1 \frac{\partial b}{\partial \tau_b} \right) \right] G^{-\eta} = \frac{1-\mu}{\mu} \left[\bar{p}_1 y_1 (y_1 (1-\tau_b))^{-\eta} \right]$$

Simple rearrangement of the above expression leads to the following tax formula,

$$\frac{\tau_b}{1-\tau_b} = \begin{cases} \left(\dfrac{\mu}{1-\mu} \dfrac{1}{\bar{p}_1} \right)^{\frac{1}{\eta}} (1 + \alpha \epsilon_{n,\tau} + \epsilon_{b,\tau})^{\frac{1}{\eta}} \left(\dfrac{p_1 b}{y_1} \right)^{-\frac{\eta-1}{\eta}} & \forall \eta \neq 1 \\[4mm] \left(\dfrac{\mu}{1-\mu} \dfrac{1}{\bar{p}_1} \right) (1 + \alpha \epsilon_{n,\tau} + \epsilon_{b,\tau}) & \text{if } \eta = 1 \end{cases}$$

where $\epsilon_{n,\tau}$ and $\epsilon_{b,\tau}$ are, respectively, the elasticity of effort and wages with respect to the marginal tax rate, and $p_1 b$ can be thought of as the product of the (average) wage rate in the top-income bracket times its density.

While the first term on the RHS accounts for the social marginal weight on public expenditure, so that the optimal marginal tax is increasing in μ, the second term directly reflects the efficiency cost of increasing taxes, due to the behavioral responses of either the worker and the firm. It is an efficiency cost because such behavioral responses have no first-order positive welfare effects on the agent (not just on the ideal agent) and the firm, but have a second-order negative effect on revenues.

---

[10] As we shall show, however, when the agent's utility takes the logarithmic form, the existence of the solution and, therefore, its equilibrium value, depend only on μ and $p_1(\bar{n})$.

Since the ϵ's are both non-positive for y_1 sufficiently high, this expression requires for an optimum to exist, that the elasticities with respect to τ_b of the probability measure and the optimal wage sum up together to a number less the one (in absolute value), i.e. either effort and wages must be not too much elastic.

It is also worth stressing that ϵ is an uncompensated elasticity concept. It differs slightly from the literature on optimal taxation, which focuses, instead, on the elasticity of reported income with respect to the participation/post-tax wage rate, $(1 - \tau)$.

Furthermore, since all the variables involved are endogenous functions of τ_b and, thus, they do not allow for a closed form solution of the equilibrium tax, we propose the above formula as a testable hypothesis. According to it, our framework gives leave to i) replace the elasticity of effort supply, $\epsilon_{n\tau}$, with the elasticity of the density measure, $\epsilon_{p\tau} = p'\epsilon_{n\tau}$, and ii) handle with the wage distribution instead of the one for the reported income. In that, the first difference seems to account for the extensive margin of the top-earners' labor supply, while the second captures its intensive margin, net, however, the portion of it that is repaid by perks.

3.12.3   Comparative statics results

In the standard optimal taxation theory, the optimal top tax rate increases in the density of the top-income agents (the revenue-collection motive). However, our model predicts that, whenever that density increases, y_1 being fixed, the optimal tax rate should decrease if the economy is sufficiently high relative risk averse, i.e. if $\eta > 1$. The marginal tax rate, however, becomes more progressive if the agent's is relatively low-risk averse. In this last scenario, in which the provision of incentives through wages is more costly for the principal (risk-premia are in fact higher) and firms more easily substitute wages with perks, our model predicts higher optimal marginal taxes.

The economic content behind the result can be explained as follows. Since at $n = \bar{n}$, which is the desirable level of effort for the government's perspective, a drop in η is associated with lower either $\dfrac{b}{y_1}$ and $\dfrac{q}{y_1}$, the government finds optimal to counteract the leading forces in the labor market, and acts so as to substitute the low private consumption with a "forced" increment in the public expenditure, since for those values of the agent's risk aversion, the ideal agent's expected loss from a system with perks (in terms of her private consumption)

is positive. In fact, when society is low-risk averse, the social value of public expenditure is high and persistent, so that decreasing the marginal consumption gain of higher b by increasing the tax rate, increases the utility of the ideal agent in its public-good-consumption component.

On the other hand, when agents are sufficiently high risk-averse, there are two possible effects: firstly, since guaranteeing incentive compatibility is easier, higher tax rates at lower wages are consistent with the optimizing behavior of the firm and the workers. At those states, in fact, the outcome implementable in the labor market is less likely to make the agent worse off. Secondly, since the ideal agent is more risk-averse, the government can find optimal to substitute labor-outcome uncertainty with the certainty of providing the public-good.

Intuitively, one can find values for μ and η for which either effect dominates the other. Nonetheless, for μ sufficiently small and fixed, we expect the decentralized labor market outcome to converge to the centralized efficient allocation as the agent's risk aversion increases over a small range.

According to Bernasconi (1998), the coefficient of relative risk aversion lies indeed in the interval $[1, 2]$, so that our theoretical result is in favor of a more regressive tax rate. As we discussed in Chapter 2, this is indeed what an utilitarian social planner would have liked to do in the centralized economy.

In the next section, we provide a graphical representation of our results, based on the numerical exercise discussed in Section 3.11.

The following proposition summarizes the main results of this chapter.

Proposition7: Fix problem (P4) and (P5) and let Assumptions (PD) and (N) hold. Then, for fixed y_1, the equilibrium marginal tax is:

1. decreasing (respectively, increasing) in the expected wage, $p_1 b$, if $\eta > 1$ (respectively, if $0 < \eta < 1$), and does not depend on it if the agent's utility function is logarithmic;

2. increasing in the social weight to public expenditure, μ and decreasing in both elasticities, $\epsilon_{n,\tau}$ and $\epsilon_{w,\tau}$.

   Moreover,

3. When society is low risk-averse, the equilibrium marginal tax rate is consistent with a positive provision of perks, even for large values of μ.

4. When agents are sufficiently high risk-averse, and the provision of incentives is less-costly for the principal, either the equilibrium tax rate

discourages the provision of perks, for relatively high μ and η, or it is consistent with a positive provision of the perquisite good. Furthermore, the latter case is more likely to occur if the outcome implementable in the labor market, as compared to the perk-less economy, does not make the agent worse off.

3.12.4 A numerical exercise for the equilibrium marginal tax rate

So far we have proposed an argument for why, for $\eta \in (0, 1]$, the optimal tax rate is increasing in wages; whereas, for $\eta > 1$, the optimal tax is decreasing in them. Here, we test the ability of our model to pin down the equilibrium tax rate that is compatible with perks being supplied in the economy.

As we already discussed, we should be able to find a set of parameters for which, at the equilibrium tax rate: i) perks are provided even when social preferences attach a large weight to the public expenditure, if the agent is sufficiently low risk averse, and ii) perks are not provided in positive amount as the agent's risk aversion increases.

Because of the continuity of our solution, however, there must be some parameter values for which, given $\eta = 2$, which is our upper-bound for the agent's relative risk aversion, the equilibrium tax rate takes two values, one of which allows for a positive provision of perks in the economy. Intuitively, then, the complementarity/substitutability between cash and perks plays now a key role in determining the optimal marginal tax rate.

The following figures assume $\mu = 0.2$, $(\sigma, \delta) = (1, 0.5)$ and $y_1 = 3$. However, for case 1, which is the case of an equilibrium with perks, we assume $a = 5$; while, for case 2, for which at the equilibrium tax rate perks are not provided by the principal, $a = 4.5$. In either case the price of perks is set to $k = 1$.

As we can see form the Figure 3.13 and 3.14, our model predicts a top-income marginal tax rate of 30% in case 1. It implies a public expenditure, $G = 0.20$, which accounts for about 7% of the state, y_1, and 13% of the equilibrium wages $b(\tau_b^\star)$. Moreover, at their equilibrium value, perks amount to 1.6% of the gross income and 3.2% of the wage.

However, in case 2 (Figures 3.15, 3.16) the optimal marginal tax rate is found to be 45%. At those value, it is not optimal for the firm to pay perks. The actual expected revenues imply $G = 0.22$, which account for 7.5% of the gross income and 15% of the wage.

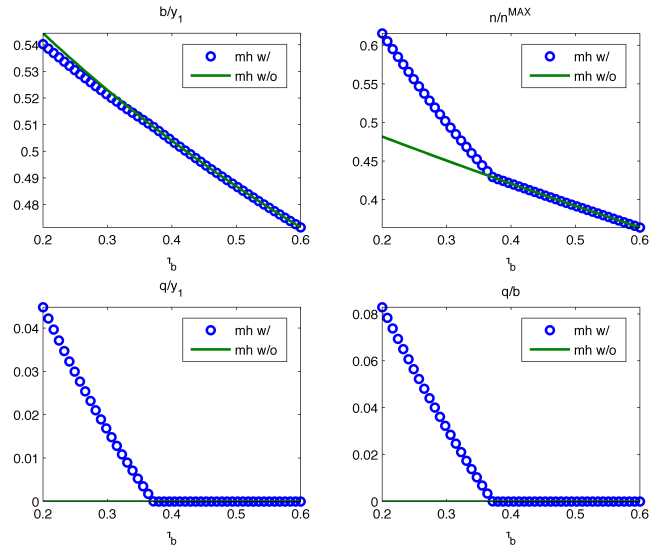Figure 3.13: Stage two: The wages and perk strategies (case 1)

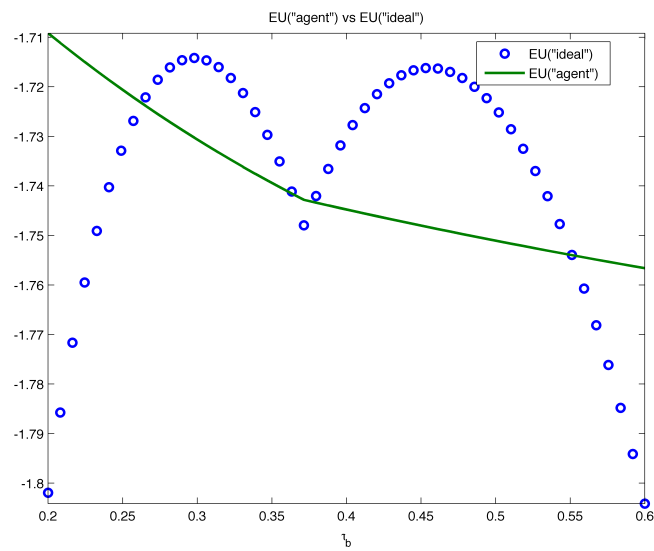Figure 3.14: The ideal agent's expected utility (Equilibrium w/ perks)

Figure 3.15: Stage two: The wages and perk strategies (case2)
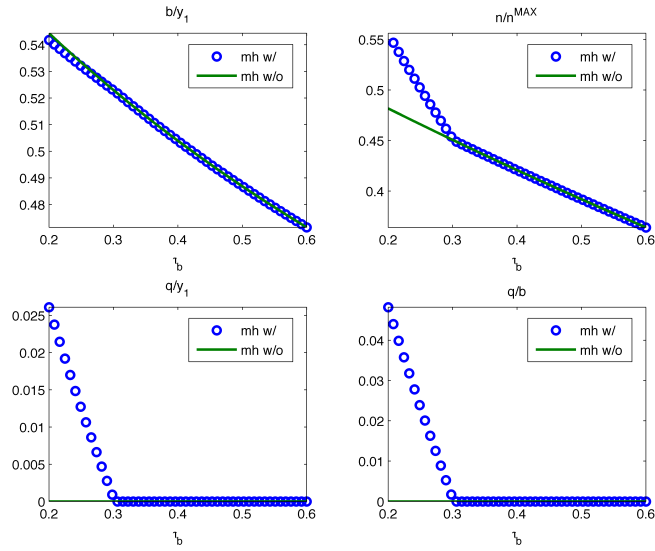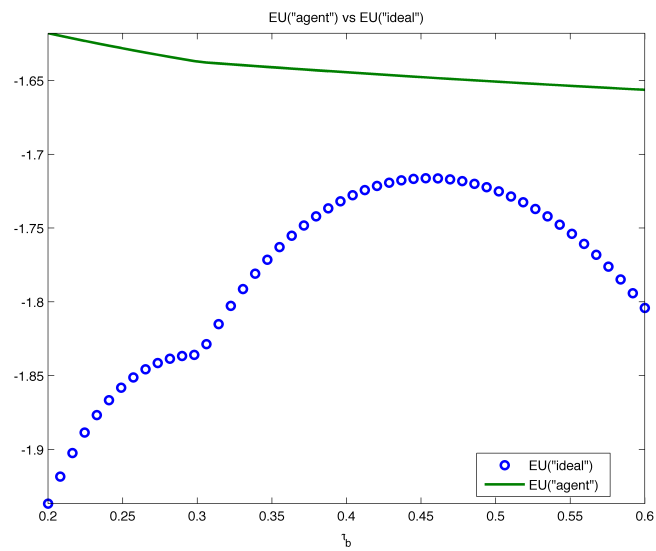

Figure 3.16: The ideal agent's expected utility (Equilibrium w/o perks)

3.13   Conclusions

This chapter develops a two-stage Nash non-cooperative game between a fiscal authority and the agents in the labor market, in order to pin down the equilibrium top-income marginal tax rate that results when firms are allowed to repay workers by perks.

The nature of the externality problem implied by our setting is the following. Workers derive utility from either the labor market outcomes and the goods and services publicly provided by the government. Nonetheless, when deciding how much to work and how much to pay, both agents take as given the tax system and the public expenditure announced by the fiscal authority. The government, however, sets his current purchases equal to the value of the expected tax revenues he can actually collect, given the behavior of the labor market. Since the firm may have an interest to pay by perks in addition to wages, the government may also have an interest to either discourage or support the provision of perks. Technically, the problem at hand requires the definition of a social welfare function that is either tractable and simple to justify. On that regard, we adopt an utilitarian approach based on the concept, not new to the literature on tax evasion, of an ideal agent, who never shrinks nor receives perks.

We split our analysis into two parts. In the first part we focus on the labor market and, given the fiscal policy, we characterize the optimal allocation of perks, wages and effort. Our analysis suggests that the agent's marginal rate of substitution between perks and wages may differ from their relative price, so leading to an (ex-post) inefficiency in the provision of perks. Nonetheless, we argue that even an over-provision of perks may make the agent better off, compared to a system without perks, if the excess of perks triggers a sufficiently high excess of effort, or if it results in a resource-improving technology that allows for an ex-post compensation of the agent's effort.

However, either theoretically and through numerical simulations, we show how these results are sensitive to the stochastic dominance properties of the probability distribution and the value of the state in the economy. If, in the former case, deviations from the benchmark are qualitatively important but quantitatively small, in the latter case, however, the quantitative implications of different income levels are quite significant.

In the second part of the chapter, we formally state the government's maximization problem over the optimal marginal tax rate, given society's preferences and the outcomes generated in the labor market. Our results are in terms of a simple

tax-formula, that we propose as a new testable hypothesis, which builds on a theoretical explanation that explicitly accounts for the provision of perks. Its novelty should be found on its replacing the elasticity of the reported income with the combination of the elasticity to the tax rate of effort (which account for the extensive margin of the top-earners' labor supply) and wages (which captures the intensive margin of the labor supply, net, however, the portion of it that is repaid by perks).

Our main contribution consists in showing that a regressive marginal tax rate is consistent with a positive provision of perks. Though we argue how this result is sensitive to the match between the agent's relative risk aversion and the level of preferences for public expenditure, we also justify it according to our theoretical predictions.

Finally, in a numerical exercise, we test the ability of our model to quantify the equilibrium top-income marginal tax and find that a tax rate about 30% is consistent with a positive provision of perks when they account for 1.6% of the gross income (3.2% of the agent's wage), and the implementable per-capita public expenditure is about 7% of the top-brackets taxpayers' gross income (15% of their earnings). An equilibrium allocation without perks is also found for a 45% marginal tax rate.

A more precise calibration of the model is left for future works, along with a better characterization of the results in terms of the degree of stochastic dominance implied by the income distribution.

Appendix 3.A: The Hessian Matrix

Remember Assumption (N),

Assumption(N): $g(n,q) = \tilde{g}(n)h(q)$,

$$\tilde{g}(n) = \frac{\delta n^{1+\sigma}}{1+\sigma} \qquad h(q) = \frac{1}{(1+aq)^\gamma}.$$

By differentiating $\mathcal{L}_b$ with respect to b, we obtain

$$\mathcal{L}_{bb}: \qquad \frac{dn}{db}\left[\mathcal{L}_n \frac{d^2n}{dbdn} - p_1'(1+\tau_b)\right] + \mathcal{L}_{nb}\frac{dn}{db} + \mathcal{L}_n \frac{d^2n}{db^2}$$

where

$$\frac{d^2n}{dbdn} = -\frac{p_1' u_c(b)}{g_{nn^2}}g_{nnn} \qquad \mathcal{L}_{nb} = -p_1'(1+\tau_b) \qquad \frac{d^2n}{db^2} = -\frac{p_1'|u_{cc}|}{g_{nn}}$$

Therefore,

$$\mathcal{L}_{bb}: \qquad -\frac{p_1' u_c(b)}{g_{nn}}\left[2p_1'(1+\tau_b) + \mathcal{L}_n p_1' u_c(b)\frac{g_{nnn}}{g_{nn}^2}\right] - \frac{\mathcal{L}_n p_1'|u_{cc}(b)|}{g_{nn}} < 0$$

The derivative with respect to q of $\mathcal{L}_q$ is given by

$$\mathcal{L}_{qq}: \qquad \left[\mathcal{L}_n \frac{d^2n}{dqdn}\right]\frac{dn}{dq} + \mathcal{L}_n \frac{d^2n}{dq^2}$$

with

$$\frac{d^2n}{dqdn} = \frac{|g_{nnq}|g_{nn}}{g_{nn^2}} - \frac{|g_{nq}|g_{nnn}}{g_{nn}^2} =$$

$$= \frac{|h_q|}{h(q)}\left[1 - \frac{g_n g_{nnn}}{g_{nn}^2}\right] = \frac{|h_q|}{h(q)}\left[1 - \frac{\sigma-1}{\sigma}\right] = \frac{|h_q|}{h(q)}\left[\frac{1}{\sigma}\right]$$

and

$$\frac{d^2n}{dq^2} = \frac{g_{nn}g_{nqq}}{g_{nn}^2} - \frac{g_{nnq}|g_{nq}|}{g_{nn}^2} =$$

$$= \frac{\tilde{g}_n}{\tilde{g}_{nn}h(q)}\left[-|h_{qq}| + \frac{|h_q|^2}{h(q)}\right]$$

Therefore,

$$\mathcal{L}_{qq}: \qquad \mathcal{L}_n \frac{h_q^2}{h(q)^2}\frac{\tilde{g}_n}{\tilde{g}_{nn}}\frac{1}{\sigma} - \mathcal{L}_n \frac{h_q^2}{h(q)^2}\frac{\tilde{g}_n}{\tilde{g}_{nn}}\left[\frac{|h_{qq}|h(q)}{h_q^2} - 1\right] =$$

$$= \mathcal{L}_n \frac{h_q^2}{h(q)^2}\frac{\tilde{g}_n}{\tilde{g}_{nn}}\left[\frac{1}{\sigma} + 1 - \frac{\gamma+1}{\gamma}\right] =$$

135

$$= \mathcal{L}_n \frac{h_q^2}{h(q)^2} \frac{\tilde{g}_n}{\tilde{g}_{nn}} \left[\frac{1}{\sigma} - \frac{1}{\gamma} \right] < 0$$

if and only if $\sigma > \gamma$. Because of our specification, thus,

$$\mathcal{L}_{qq} : \qquad \frac{\alpha \gamma a^2 n \Delta}{\sigma (1 + aq)^2} \left[\frac{\gamma}{\sigma} - 1 \right] < 0 \quad \forall \gamma < \sigma$$

The cross-derivative $\mathcal{L}_{qb}$ is,

$$\mathcal{L}_{qb} : \qquad \mathcal{L}_n \frac{d^2 n}{db dq} + \frac{dn}{dq} \left[\mathcal{L}_n \frac{d^2 n}{db dn} - p_1'(1 + \tau_b) \right]$$

$$\text{with} \qquad \frac{d^2 n}{db dq} = \frac{p_1' u_c(b) |h_q|}{\tilde{g}_{nn} h(q)^2} \qquad \frac{d^2 n}{db dn} = -\frac{p_1' u_c(b) \tilde{g}_{nnn}}{h(q) \tilde{g}_{nn}^2}$$

$$\mathcal{L}_{qb} : \qquad \mathcal{L}_n p_1' \frac{u_c(b)}{\tilde{g}_{nn}} \frac{|h_q|}{h(q)^2} \left[1 - \frac{g_{nnn} g_n}{g_{nn^2}} \right] - p_1'(1 + \tau_b) \frac{\tilde{g}_n}{\tilde{g}_{nn}} \frac{|h_q|}{h(q)} =$$

$$= \frac{p_1' |h_q| \tilde{g}_n}{h(q) \tilde{g}_{nn}} \left[\left(1 - \frac{g_{nnn} g_n}{g_{nn^2}} \right) \frac{\mathcal{L}_n u_c(b)}{\tilde{g}_n h(q)} - (1 + \tau_b) \right] \lesseqgtr 0$$

We remember that optimality condition for b states that $\frac{\mathcal{L}_n u_c(b)}{\tilde{g}_n h(q)} = (1 + \tau_b) \frac{p_1}{n p_1'} \frac{\tilde{g}_{nn} n}{\tilde{g}_n}$.

Therefore, as long as the following holds, $\frac{\tilde{g}_n}{\tilde{g}_{nn} n} = 1 - \frac{g_{nnn} g_n}{g_{nn^2}}$, and it holds for our iso-elastic disutility function, we obtain

$$\mathcal{L}_{qb} : \qquad \frac{p_1' |h_q| \tilde{g}_n}{h(q) \tilde{g}_{nn}} (1 + \tau_b) \left[\frac{p_1}{n p_1'} - 1 \right]$$

By Assumption(SP),

$$\mathcal{L}_{qb} : \qquad \frac{p_1' |h_q| \tilde{g}_n}{h(q) \tilde{g}_{nn}} (1 + \tau_b) \left[\frac{\theta}{n} \right]$$

Then,

- if $\theta = 0$, $\mathcal{L}_{qb} = 0$, meaning that there is no gain from joint deviation of increasing b and reducing q at the equilibrium allocation $\langle s, b, n, q \rangle$.

- if $\theta > (<) 0$, $\mathcal{L}_{qq} > (<) 0$, there may be a profitable deviation from increasing (decreasing) b and increasing q at the equilibrium allocation $\langle s, b, n, q \rangle$.

Anyway, for the problem being jointly concave in q and b (when s is exogenously fixed), those second-order gains must be outweighed by the losses the principal would face by moving away from the optimal levels of b, q (as measured by the main entrees in the Hessian matrix).

- At any interior solution for q, joint concavity requires

$$|H_2| = \mathcal{L}_{bb} \cdot \mathcal{L}_{qq} - \left(\mathcal{L}_{qb}\right)^2 > 0$$

which, for $\theta = 0$, requires $\sigma > \gamma$.

- When $q = 0$, that is, when the non-negative constraint on q is binding, joint concavity demands

$$|H_2| = \begin{vmatrix} \mathcal{L}_{bb} & \mathcal{L}_{qb} & 0 \\ \mathcal{L}_{qb} & \mathcal{L}_{qq} & 1 \\ 0 & 1 & 0 \end{vmatrix} = -\mathcal{L}_{bb} > 0$$

Appendix 3.B: Math notes

Proof of Proposition 2:

Let denote $N \equiv \dfrac{u(b) - u(s)}{u_c(b)}$, and $D \equiv \dfrac{\mathcal{A}}{1 + \tau_b} - b$, and let be $N_b \equiv \dfrac{\partial N}{\partial b} > 0$ (see Lemma1 below). By differentiating (2), we get that, for all parameters j:

$$\frac{N_b}{D}\frac{db}{dj} + \frac{1}{D}\frac{\partial N}{\partial j} - \frac{N}{D^2}\left(\frac{\partial D}{\partial j} - \frac{db}{dj}\right) = \frac{1}{\sigma}\left(\frac{\partial \varepsilon}{\partial n} \cdot \frac{dn}{dj}\right) - 1_{j=\sigma}\left[\frac{\varepsilon}{\sigma^2}\right]$$

$$\frac{db}{dj}\left(\frac{D \cdot N_b + N}{D^2}\right) + \frac{1}{D^2}\left(D\frac{\partial N}{\partial j} - N\frac{\partial D}{\partial j}\right) = \frac{1}{\sigma(1-\mu)}\left(\frac{\partial \varepsilon}{\partial n} \cdot \frac{dn}{dj}\right) - 1_{j=\sigma}\left[\frac{\varepsilon}{\sigma^2}\right]$$

from which we derive our results.

Proof of Lemma1:

- Assume $u(c) = \dfrac{c^{1-\eta}}{1 - \eta}$. Then,

$$\frac{u(b) - u(s)}{u_c(b)} = \begin{cases} \dfrac{1}{1-\eta}\left[\dfrac{b^{1-\eta} - s^{1-\eta}}{b^{-\eta}}\right] = \dfrac{1}{1-\eta}\left[b - \dfrac{s^{1-\eta}}{b^{-\eta}}\right] & \text{if } \eta \neq 1 \\ b\left(\log b - \log s\right) & \text{if } \eta = 1 \end{cases}$$

Therefore, its derivatives (total and partial coincides) with respect to b are equal to

$$N_b \equiv \frac{d}{db}\left[\frac{u(b) - u(s)}{u_c(b)}\right] = \begin{cases} \dfrac{1}{1-\eta}\left[1 - \eta\left(\dfrac{s}{b}\right)^{1-\eta}\right] > 0 & \text{if } \eta \neq 1 \\ 1 + \log\left(\dfrac{b}{s}\right) > 0 & \text{if } \eta = 1 \end{cases}$$

1. $\dfrac{d^2}{db^2}\left[\dfrac{\Delta u}{u_c(b)}\right] = \begin{cases} \dfrac{\eta b^{\eta-2}}{s^{\eta-1}} > 0 & \text{if } \eta \in (0,1) \cup (1,\infty) \\[2mm] \dfrac{1}{b} > 0 & \text{if } \eta{=}1 \end{cases}$

2. $\dfrac{d^3}{db^3}\left[\dfrac{\Delta u}{u_c(b)}\right] = \begin{cases} \dfrac{\eta(\eta-2)b^{\eta-3}}{s^{\eta-1}} & \text{if } \eta \in (0,1) \cup (1,\infty) \\[2mm] -\dfrac{1}{b^2} < 0 & \text{if } \eta{=}1 \end{cases}$

That is, N_b is i) linearly increasing in b for $\eta = 2$; ii) increasing and concave in b for all $\eta \in (0,1] \cup (1,2)$; iii) increasing and convex in b for $\eta > 2$.

- Assume $u(c) = -\exp(-\eta c)$. Then,

$$\frac{u(b) - u(s)}{u_c(b)} = \frac{-\exp(-\eta b) + \exp(-\eta s)}{\eta \exp(-\eta b)}$$
$$= -\frac{1}{\eta} + \frac{1}{\eta}\exp(\eta(b-s))$$

whose derivative with respect to b is given by

$$N_b \equiv \frac{d}{db}\left[\frac{1}{\eta}\Big(\exp(\eta(b-s)) - 1\Big)\right] = \exp(\eta(b-s)) > 0$$

1. $\dfrac{d^2}{db^2}\left[\dfrac{1}{\eta}\Big(\exp(\eta(b-s)) - 1\Big)\right] = \eta \exp(\eta(b-s)) > 0$

2. $\dfrac{d^3}{db^3}\left[\dfrac{1}{\eta}\Big(\exp(\eta(b-s)) - 1\Big)\right] = \eta^2 \exp(\eta(b-s)) > 0$

The perk-less values of effort and wage

In our most preferred specification, we assume $\sigma = 1$ and $u(c) = \dfrac{c^{1-\eta}}{1-\eta}$. From expression (IC), therefore, as long as $q = 0$, effort must satisfy $n_0 = \dfrac{p' \cdot \Delta u}{\delta}$. Substituting n_0 into $\mathcal{L}_n$ and evaluating it at $q = 0$, we obtain

$$\frac{p' u_c(b) p' \Delta}{g_{nn}(n, q)(1 + \tau_b)} = p_1 \text{ and, therefore, } \frac{u_c(b) p' \Delta}{\delta(1 + \tau_b)} = n + \theta$$

Let $\tilde{\mathcal{A}} \equiv \dfrac{y_1 - y_0 + s(1 + \tau_s)}{1 + \tau_b}$. The optimal b_0 solves

$$\tilde{\mathcal{A}} - b = \frac{1}{1-\eta}b - b^{\eta}\Big(\frac{s^{1-\eta}}{1-\eta} - \frac{\theta\delta}{p'}\Big)$$
$$\tilde{\mathcal{A}} - \Big(\frac{2-\eta}{1-\eta}\Big)b = -b^{\eta}\Big(\frac{s^{1-\eta}}{1-\eta} - \frac{\theta\delta}{p'}\Big)$$

$$\left(\frac{2-\eta}{1-\eta}\right)b - \tilde{\mathcal{A}} = b^\eta\left(\frac{s^{1-\eta}}{1-\eta} - \frac{\theta\delta}{p'}\right)$$

For $\eta = 2$, $b_0 = \left(\dfrac{\tilde{\mathcal{A}}}{\dfrac{1}{s} + \dfrac{\theta\delta}{p'}}\right)^{\frac{1}{2}}$. Substituting it into n_0 above, we find

$$n_0 = \frac{p'}{\delta}\left[\frac{1}{s} - \sqrt{\left(\frac{1}{s} + \frac{\theta\delta}{p'}\right)\Big/\tilde{\mathcal{A}}}\right]$$

For $y_0 = \hat{s} = 1$, $\theta = 0$ and $\tau_s = 0$, we obtain our result.

Proof of $\mathcal{L}_n$ being increasing in n

For changes in y_1, $\mathcal{L}_n$ is increasing in n if and only if $(1+\tau_b)\dfrac{db}{dy_1} < 1$, provided that $\dfrac{dn}{dy_1} > 0$. Differentiating $\mathcal{L}_n(n)$ with respect to y_1 we obtain:

$$\frac{d\mathcal{L}_n}{dn}\frac{dn}{dy_1} = p'\left(1 - (1+\tau_b)\frac{db}{dy_1}\right)$$

Since, from $\mathcal{L}_b$ we know that,

$$\frac{db}{dy_1} = \frac{1}{D\cdot N_b + N}\left[\frac{D^2}{\sigma}\left(\frac{\partial\varepsilon}{\partial n}\cdot\frac{dn}{dy_1}\right) + \frac{N}{1+\tau_b}\right]$$

- if $\theta = 0$, $(1+\tau_b)\dfrac{db}{dy_1} = \dfrac{N}{D\cdot N_b + N} < 1$;

- if $\theta < 0$, $(1+\tau_b)\dfrac{db}{dy_1} = \dfrac{N}{D\cdot N_b + N} - \dfrac{\Delta^2\big/(1+\tau_b)}{\sigma(D\cdot N_b + N)}\left|\dfrac{\partial\varepsilon}{\partial n}\right|\dfrac{dn}{dy_1} < 1$;

- If $\theta > 0$, with $\dfrac{d\varepsilon}{dn} > 0$, condition $(1+\tau_b)\dfrac{db}{dy_1} < 1$ is more demanding. However, if second-order effects are less important than direct effects, it is always satisfied for sufficiently high σ and low τ_b.

The MRS

From $\mathcal{L}_q$ and $\mathcal{L}_b$ we know, respectively, that

$$p'\Delta = k\frac{\tilde{g}_{nn}h(q)}{\tilde{g}_n|h_q|} \qquad p'\Delta = \frac{p_1(1+\tau_b)}{p'u_c(b)}g_{nn}(n,q)$$

Substituting out $p'\Delta$ we find

$$\frac{k\tilde{g}_{nn}h(q)}{\tilde{g}_n|h_q|} = \frac{(1+\tau_b)}{u_c(b)}\frac{p_1}{p'}\tilde{g}_{nn}h(q)$$

$$\frac{k}{\tilde{g}_n|h_q|} = \frac{(1+\tau_b)}{u_c(b)}\frac{p_1}{p'}$$

$$\frac{k}{|h_q|} = \frac{(1+\tau_b)}{u_c(b)}\frac{p_1\tilde{g}_n}{p'}$$

$$\frac{k/|h_q|}{(1+\tau_b)/u_c(b)} = \frac{\tilde{g}_n}{p'/p_1}$$

$$\frac{k}{|g_{nq}|} = \frac{p_1(1+\tau_b)}{p'u_c(b)}$$

Corollary5b

From (7) , by noticing that $\tilde{g}(n)|h_q| = U_q$, we obtain (8) in the following steps

$$\frac{p'u_c(b)}{\tilde{g}_n|h_q|} = \frac{p_1(1+\tau_b)}{k}$$

$$\frac{p'u_c(b)}{\frac{\tilde{g}_n}{\tilde{g}(n)}\tilde{g}(n)|h_q|} = \frac{p_1(1+\tau_b)}{k}$$

$$\frac{u_c(b)}{U_q} = \left(\frac{p_1}{p'}\bigg/\frac{\tilde{g}(n)}{\tilde{g}_n}\right)\frac{1+\tau_b}{k} \qquad (8)$$

Corollary6

By differentiation of the LHS of (8a) w.r.t. y_1 we obtain,

$$\frac{u_{cc}(b)\dfrac{db}{dy_1}}{U_q} - \frac{u_c(b)U_{qq}\dfrac{dq}{dy_1}}{U_q^2} =$$

$$= \frac{u_c(b)}{U_q}\left[-\frac{|u_{cc}(b)|}{u_c(b)}\frac{db}{dy_1} + \frac{|h_{qq}|}{|h_q|}\frac{dq}{dy_1}\right] =$$

$$= \frac{u_c(b)}{U_q}\frac{|h_{qq}|}{|h_q|}\left[-\frac{|u_{cc}(b)|}{u_c(b)}\frac{|h_q|}{|h_{qq}|}\frac{db}{dy_1} + \frac{dq}{dy_1}\right]$$

from which we derive our conclusions.

χ_q

From the foc w.r.t. q, we can conclude the following:

- If $\chi_q < 0$, then $\left.\dfrac{\tilde{g}_n|h_q|}{\tilde{g}_{nn}h(q)}p'\Delta\right|_{q=0} > k$, and the principal would have liked to increase q above zero.

  A negative χ_q stands for the reduction in the maximum value of the objective function the principal is willing to accept for each increase in q.

- If $\chi_q > 0$, then $\left.\dfrac{\tilde{g}_n|h_q|}{\tilde{g}_{nn}h(q)}p'\Delta\right|_{q=0} < k$, and the principal would have liked to reduce q below zero. That is, χ_q is the price the principal would have asked for himself, to be compensate for paying $q=0$ to the agent, instead of paying her a negative amount of perks.

Normalization procedure (case 1)

$$\frac{\delta x^{1+\sigma}}{1+\sigma}\frac{1}{(1+a(q+dq))^\gamma} = \frac{\delta n^{1+\sigma}}{1+\sigma}$$

$$(1+a(q+dq))^\gamma = \frac{x^{1+\sigma}}{n^{1+\sigma}}$$

$$dq = \frac{1}{a}\left[\frac{x}{n}\right]^{\frac{1+\sigma}{\gamma}} - \frac{1}{a} - q$$

$$dq = \frac{1}{a}\left(\left[\frac{x}{n}\right]^{\frac{1+\sigma}{\gamma}} - 1\right) - q \qquad (11)$$

Normalization procedure (case 2)

$$\frac{\delta x^{1+\sigma}}{1+\sigma}\frac{1}{(1+aq)^\gamma} = \frac{\delta n^{1+\sigma}}{1+\sigma}\frac{1}{(1-adq)^\gamma}$$

$$(1-adq)^\gamma = \frac{n^{1+\sigma}}{x^{1+\sigma}}(1+aq)^\gamma$$

$$(1-adq) = \left[\frac{n}{x}\right]^{\frac{1+\sigma}{\gamma}}(1+aq)$$

$$dq = \frac{1}{a} - \left[\frac{n}{x}\right]^{\frac{1+\sigma}{\gamma}}\frac{(1+aq)}{a}$$

$$dq = \frac{1}{a}\left(1 - \left[\frac{n}{x}\right]^{\frac{1+\sigma}{\gamma}}\right) - q\cdot\left[\frac{n}{x}\right]^{\frac{1+\sigma}{\gamma}} \qquad (12)$$

Closed-form solution for perks and effort

$$(1 + aq) = \frac{\gamma a}{k} \frac{\alpha^2}{\delta(1 + \tau_b)} \frac{\Delta^2}{b^2}$$

$$= \frac{\gamma a}{k} \frac{\alpha^2}{\delta(1 + \tau_b)} \frac{\left[\mathcal{A}(1 + \tau_b) - b(1 + \tau_b)\right]^2}{b^2}$$

$$= \frac{\gamma a}{k} \frac{\alpha^2 (1 + \tau_b)}{\delta} \frac{b^2 \left[\frac{\mathcal{A}}{b} - 1\right]^2}{b^2}$$

$$= \frac{\gamma a}{k} \frac{\alpha^2 (1 + \tau_b)}{\delta} \left[\frac{\mathcal{A}}{b} - 1\right]^2$$

$$= \frac{\gamma a}{k} \frac{\alpha^2 (1 + \tau_b)}{\delta} \left[\frac{\mathcal{A}}{\sqrt{\mathcal{A}s}} - 1\right]^2$$

Finally, we obtain

$$q = \frac{\gamma}{k} \frac{\alpha^2 (1 + \tau_b)}{\delta} \left[\sqrt{\frac{\mathcal{A}}{s}} - 1\right]^2 - \frac{1}{a}$$

- Derivative w.r.t. τ_b of q

$$\frac{\partial q}{\partial \tau_b} = \frac{\gamma}{k} \frac{\alpha^2}{\delta} \left\{ \left[\frac{\mathcal{A}}{s} - \sqrt{A/s} + 1\right] + (1 + \tau_b)\left[-\frac{y}{s(1 + \tau_b)^2} + \frac{1}{2}\frac{\sqrt{y}}{\sqrt{s(1 + \tau_b)^{3/2}}}\right] \right\}$$

$$\frac{\partial q}{\partial \tau_b} = \frac{\gamma}{k} \frac{\alpha^2}{\delta} \left\{ 1 - \frac{1}{2}\frac{\sqrt{y}}{\sqrt{s(1 + \tau_b)}} \right\}$$

$$\frac{\partial q}{\partial \tau_b} = \frac{\gamma}{k} \frac{\alpha^2}{\delta} \left\{ 1 - \frac{1}{2}\sqrt{\frac{y}{s(1 + \tau_b)}} \right\}$$

which is negative if $y > 4s(1 + \tau_b)$.

- Derivatives w.r.t. τ_b for n

$$n = \frac{\alpha}{\delta} \left[\frac{\gamma a}{k} \frac{\alpha^2}{\delta}(1 + \tau_b)\right]^{\gamma} \left[\sqrt{\frac{\mathcal{A}}{s}} - 1\right]^{2\gamma} \left[\frac{1}{s} - \frac{1}{\sqrt{\mathcal{A}s}}\right]$$

Therefore,

$$\frac{\partial n}{\partial \tau_b} = \frac{\phi}{\sqrt{y_1}} \left[\left(\sqrt{\frac{y_1}{1 + \tau_b}} - 1\right)^2 - \sqrt{\frac{y_1}{1 + \tau_b}}\left(\sqrt{\frac{y_1}{1 + \tau_b}} - 1\right)\right] =$$

$$= \frac{\phi}{\sqrt{y_1}} \left(\sqrt{\frac{y_1}{1 + \tau_b}} - 1\right)\left[\sqrt{\frac{y_1}{1 + \tau_b}} - 1 - \sqrt{\frac{y_1}{1 + \tau_b}}\right] =$$

$$= -\frac{\phi}{\sqrt{y_1}} \left(\sqrt{\frac{y_1}{1 + \tau_b}} - 1\right)$$

which is negative for all $y_1 > 1 + \tau_b$, but is positive for any τ_b such that $1 < y_1 < 1 + \tau_b$.

The effort and wage elasticities w.r.t. τ_b

- The elasticity of p_1 with respect to τ_b

$$\epsilon_{p,\tau} = \frac{\partial p_1}{\partial n}\frac{\partial n}{\partial \tau_b}\frac{\tau_b}{p_1} = -\frac{\phi}{\sqrt{y_1}}\left(\sqrt{\frac{y_1}{1+\tau_b}}-1\right)\frac{\tau_b}{n} =$$

$$= -\frac{\phi\tau_b}{\sqrt{y_1}}\left(\sqrt{\frac{y_1}{1+\tau_b}}-1\right)\bigg/\left[\phi\frac{1+\tau_b}{\sqrt{y_1}}\left(\sqrt{\frac{y_1}{1+\tau_b}}-1\right)^2\right] =$$

$$= -\frac{\tau_b}{1+\tau_b}\bigg/\left(\sqrt{\frac{y_1}{1+\tau_b}}-1\right)$$

- The elasticity of b with respect to τ_b is derived as follows,

$$\frac{\partial b}{\partial \tau_b} = -\frac{1}{2}\sqrt{\frac{y_1}{(1+\tau_b)^3}} \qquad \epsilon_{b,\tau} = \frac{\partial b}{\partial \tau_b}\frac{\tau_b}{b} = -\frac{1}{2}\frac{\tau_b}{1+\tau_b}.$$

The tax-formula

From the foc w.r.t. τ_b in (P5), we obtain,

$$\left[p_1 b + \tau_b\left(b\frac{\partial p_1}{\partial n}\frac{\partial n}{\partial \tau_b} + p_1\frac{\partial b}{\partial \tau_b}\right)\right]G^{-\eta} = \frac{1-\mu}{\mu}\left[\bar{p}_1 y_1 (y_1(1-\tau_b))^{-\eta}\right]$$

$$\left[p_1 b\left(1 + \left(\frac{\partial p_1}{\partial n}\frac{\partial n}{\partial \tau_b}\frac{\tau_b}{p_1} + \frac{\partial b}{\partial \tau_b}\frac{\tau_b}{b}\right)\right)\right]^{\frac{1}{\eta}}\bigg/G = \left(\frac{1-\mu}{\mu}\bar{p}_1\right)^{\frac{1}{\eta}}y_1^{\frac{1-\eta}{\eta}}\bigg/(1-\tau_b)$$

$$\frac{\left(p_1 b\right)^{\frac{1}{\eta}}}{G}(1+\epsilon_{p,\tau}+\epsilon_{b,\tau})^{\frac{1}{\eta}} = \left(\frac{1-\mu}{\mu}\bar{p}_1\right)^{\frac{1}{\eta}}y_1^{\frac{1-\eta}{\eta}}\bigg/(1-\tau_b)$$

$$\left(\frac{p_1 b}{y_1}\right)^{\frac{1-\eta}{\eta}}(1+\epsilon_{p,\tau}+\epsilon_{b,\tau})^{\frac{1}{\eta}}\frac{1-\tau_b}{\tau_b} = \left(\frac{1-\mu}{\mu}\bar{p}_1\right)^{\frac{1}{\eta}}$$

$$\frac{1-\tau_b}{\tau_b} = \left(\frac{1-\mu}{\mu}\bar{p}_1\right)^{\frac{1}{\eta}}(1+\alpha\epsilon_{n,\tau}+\epsilon_{b,\tau})^{-\frac{1}{\eta}}\left(\frac{p_1 b}{y_1}\right)^{\frac{\eta-1}{\eta}}$$

$$\frac{\tau_b}{1-\tau_b} = \left(\frac{\mu}{1-\mu}\frac{1}{\bar{p}_1}\right)^{\frac{1}{\eta}}(1+\alpha\epsilon_{n,\tau}+\epsilon_{b,\tau})^{\frac{1}{\eta}}\left(\frac{p_1 b}{y_1}\right)^{-\frac{\eta-1}{\eta}}$$

- In the log case, it simplifies to:

$$\frac{1-\tau_b}{\tau_b} = \left(\frac{1-\mu}{\mu}\bar{p}_1\right)(1+\alpha\epsilon_{n,\tau}+\epsilon_{b,\tau})^{-1}$$

References

1. Abraham A. and Pavoni, N. 2008. Optimal Income Taxation and Hidden Borrowing and Lending: The First-Order Approach in Two Periods. Carlo Alberto Notebooks 102, Collegio Carlo Alberto.

2. Abraham, A. and Koehne, S. and Pavoni, N. 2011. On the first-order approach in principal-agent models with hidden borrowing and lending, Journal of Economic Theory, Elsevier, vol. 146(4), 1331-1361.

3. Abraham, A. and Koehne, S. and Pavoni, N. 2014. Optimal Income Taxation with Asset Accumulation, Economics Working Papers ECO2014/14, European University Institute.

4. Abraham, A. and Pavoni, N. 2008. Efficient Allocations with Moral Hazard and Hidden Borrowing and Lending: A Recursive Formulation, Review of Economic Dynamics, in press.

5. Akerlof, G. 1982. Labour Contracts as Partial Gift Exchange. Quarterly Journal of Economics, 97:4, 543-63. Akerlof, G. 1984. Gift Exchange and Efficiency Wage Theory: Four Views. American Economic Review 74, 79-83.

6. Albanesi, S. and Sleet, C. 2006. Dynamic Optimal Taxation with Private Information, Review of Economic Studies, Vol. 73(1), 1-30.

7. Bebchuk, L.A. and Fried, J. M. 2003. Executive compensation as an agency problem. The Journal of Economic Perspectives 17, 71-92.

8. Bennardo, A. and Chiappori, P.A. and Song, J. 2010. Perks as Second Best Optimal Compensations. CSEF Working Papers 244, Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy.

9. Bernasconi, M. 1998. Tax evasion and orders of risk aversion. Journal of Public Economics 67, 123-134.

10. Bernasconi, M. and Zanardi, A. 2004. Tax evasion, tax rates and reference dependence. FinanzArchiv 60, 422-445

11. Bolton, P. and Scheinkman, J. and Wei, X. 2006. Executive Compensation and Short-termist behaviour in speculative markets. Review of Economic Studies vol. 73, 577-610.

12. Carrothers, A. and Han, S. and Qiu, J. 2012. CEO Pay with Perks, Department of Economics Working Papers 2012-05, McMaster University.

13. Chien, Y. and Song, J. 2011. Excessive perks to restrain the hidden saving problem. Mimeo .

14. Chien, Y. and Song, J. 2014. A Note On Using Excessive Perks To Restrain The Hidden Saving Problem, Macroeconomic Dynamics, Cambridge University Press, vol. 18(02), 480-496, March.

15. Cole, H. L. and Kocherlakota, N. R. 2001. Efficient Allocations with Hidden Income and Hidden Storage, Review of Economic Studies, 68(3), 523-542.

16. Cremer, H. and Gahvari, F. 1996. Tax evasion and optimal general income tax. Journal of Public Economics 60, 235-249.

17. D'Aspremont, C. and Gevers, L. 1977. Equity and the informational basis of social choice. Review of Economic Studies 46, 199-210.

18. Del Negro, M. and Perri, F. and Schivardi, F. 2010. Tax buyouts, Journal of Monetary Economics, Elsevier, vol. 57(5), 576-595.

19. Dhami, S. and al-Nowaihi, A. 2007. Why do people pay taxes? Prospect theory versus expected utility theory. Journal of Economic Behavior and Organization 64, 171-192.

20. Dhami, S. and al-Nowaihi, A. 2010. Optimal taxation in the presence of tax evasion: Expected utility versus prospect theory. Journal of Economic Behavior and Organization 75, 313-337

21. Diamond, P. 1975. A Many-Person Ramsey Tax Rule, Journal of Public Economics 4(4), 335-342.

145

22. Diamond, P. 1980. Income Taxation with Fixed Hours of Work, Journal of Public Economics 13, February, 101-110.

23. Diamond, P. 1998. Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates, American Economic Review 88, 83-95.

24. Diamond, P. and Mirrlees, J. 1971. Optimal Taxation and Public Production I: Production Efficiency and II: Tax Rules. American Economic Review, 61, 8-27 and 261-278.

25. Diamond, P. and Mirrlees, J. 2002. Optimal Taxation and the Le Chatelier Principle, unpublished MIT working paper.

26. Diamond, P. and Saez, E. 2011. The Case for a Progressive Tax: From Basic Research to Policy Recommendations, Journal of Economic Perspectives 25(4), 165-190.

27. Edgeworth, F. Y. 1897. The Pure Theory of Taxation, Economic Journal 7, 46-70, 226-238, and 550-571.

28. Edmans, A. and Gabaix, X. and Landier, A. 2008. A Multiplicative Model of Optimal CEO Incentives in Market Equilibrium, Journal of Financial Economics.

29. Fama, E. F. 1980. Agency Problems and the Theory of the Firm, Journal of Political Economy, University of Chicago Press, vol. 88(2), 288-307.

30. Feldstein, M. 1995. The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act. Journal of Political Economy, 103(3), 551-572.

31. Feldstein, M. 1999. Tax Avoidance and the Deadweight Loss of the Income Tax. Review of Economics and Statistics, 81(4), 674-680.

32. Feldstein, M. 2012. The Mirrlees Review, Journal of Economic Literature 50(3), 781-790.

33. Gabaix, X. and Landier, A. 2008. Why Has CEO Pay Increased So Much?. The Quarterly Journal of Economics, MIT Press, vol. 123(1), 49-100.

34. Golosov, M. and Tsyvinski, A. 2007. Optimal Taxation with Endogenous Insurance Markets, Quarterly Journal of Economic, 122(2), 487-534.

35. Golosov, M. and Tsyvinski, A. and Werning, I. 2006. New Dynamic Public Finance: A User's Guide, NBER Macroeconomics Annual.

36. Goolsbee, A. 2000. What Happens When You Tax the Rich? Evidence from Executive Compensation. Journal of Political Economy, 108(2), 352-378.

37. Grinstein, Y. and Weinbaum, D. and Yehuda ,N. 2011. The economic consequences of perks disclosure. Johnson School Research Paper Series n° 06-2011

38. Grossman, S. J. and Hart, O.D., 1980. Takeover bids, the free-rider problem, and the theory of the corporation. Bell Journal of Economics 11, 42–64.

39. Guesnerie, R. 1995. A Contribution to the Pure Theory of Taxation, Cambridge University Press: Cambridge.

40. Internal Revenue Service United Stated, Department of the Treasury. Publication 15-B(2015): Employer's Tax Guide to Fringe Benefits.

41. Jensen, M.C. and Meckling, W.H. 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure, Journal of Financial Economics, Elsevier, vol. 3(4), 305-360.

42. Kocherlakota, N.R. 2004. Figuring out the Impact of Hidden Savings on Optimal Unemployment Insurance. Review of Economic Dynamics,vol. 7(3), 541-554.

43. Kocherlakota, N.R. 2005. Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation, Econometrica, 74. Kocherlakota, N.R. 2010. The New Dynamic Public Finance, Princeton, Princeton University Press.

44. Kvaløy, O. and Schöttner, A. 2015. Incentives to motivate, Journal of Economic Behavior and Organization, Elsevier, vol. 116(C), 26-42.

45. Laroque, G.R. 2005. Indirect Taxation is Superfluous under Separability and Taste Homogeneity: A Simple Proof. Economics Letters, 87(1), 141-4.

46. Lee, L. and Saez, E. 2012. Optimal Minimum Wage in Competitive Labor Markets, Journal of Public Economics 96(9-10), 739-749.

47. Marino, A.M. and Zábojník J. 2008. Work-related perks, agency problems, and optimal incentive contracts. RAND Journal of Economics, RAND Corporation, vol. 39(2), 565-585.

48. Mirrlees, J.A. 1986. The Theory of Optimal Taxation,in : K. J. Arrow and M.D. Intriligator (ed.), Handbook of Mathematical Economics volume 3, chapter 24, 1197-1249.

49. Mirrlees, J.A. 2010. Dimension of Tax Design: The Mirrlees Review, Institute for Fiscal Studies, Oxford: Oxford University Press.

50. Mirrlees, J.A. 2011. Tax By Design: The Mirrlees Review, Institute for Fiscal Studies, Oxford: Oxford University Press.

51. Mirrlees, J.A. 1971. An Exploration in the Theory of Optimal Income Taxation. Review of Economic Studies, 38, 175-208.

52. Mirrlees, J.A. 1976. Optimal Tax Theory: A Synthesis, Journal of Public Economics 6, 327-358.

53. Mirrlees, J.A. 1982. Migration and Optimal Income Taxes. Journal of Public Economics 18, 319-41.

54. Piketty, T. and Saez, E. and Stantcheva, S. 2011. Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities, NBER Working Paper No. 17616.

55. Piketty, T. and Saez. E. 2007. How Progressive is the U.S. Federal Tax System? A Historical and International Perspective, Journal of Economic Perspectives, 21(1), 3-24.

56. Rajan, R.G. and Wulf, J., 2006. Are perks purely managerial excess? Journal of Financial Economics, Elsevier, vol. 79(1), 1-33.

57. Rogerson, W. 1985b. The First Order Approach to Principal-Agent Problems, Econometrica, 53, 1357-1368.

58. Rosen, S. 2000. Does the Composition of Pay Matter?, in William T. Alpert and Stephen A. Woodbury, editors, Employee Benefits and Labor Markets in Canada and the United States, W. E. Upjohn Institute For Employment Research, Kalamazoo, Michigan, 2000, 13-30.

59. Rothschild, C. and Scheuer, F. 2011. Optimal Taxation with Rent-Seeking"
    NBER working paper No. 17035. Sadka, E. 1976. On Income Distribu-
    tion, Incentive Effects and Optimal Income Taxation, Review of Economic
    Studies, 43(1), 261-268.

60. Saez, E. 2001. Using Elasticities to Derive Optimal Income Tax Rates,
    Review of Economic Studies 68, 205-229.

61. Saez, E. 2002a. Optimal Income Transfer Programs: Intensive Versus
    Extensive Labour Supply Responses. Quarterly Journal of Economics,
    117(2), 1039-73.

62. Saez, E. 2002b. The Desirability of Commodity Taxation under Non-linear
    Income Taxation and Heterogeneous Tastes. Journal of Public Economics,
    83(2), 217-230.

63. Saez, E. 2004a. The Optimal Treatment of Tax Expenditures, Journal of
    Public Economics, 88(12), 2657-2684.

64. Saez, E. 2004b. Direct or Indirect Tax Instruments for Redistribution:
    Short-Run versus Long-Run, Journal of Public Economics, 88(3-4), 503-
    518.

65. Saez, E. 2004c. Reported Incomes and Marginal Tax Rates, 1960-2000:
    Evidence and Policy Implications. in James Poterba, ed., Tax Policy and
    the Economy, 18: 117-174.

66. Saez, E. and Slemrod, J. and Giertz, S. 2012. The Elasticity of Taxable
    Income with Respect to Marginal Tax Rates: A Critical Review, Journal
    of Economic Literature 50(1), 3-50.

67. Saez, E. and Stantcheva., S. 2013. Generalized Social Marginal Welfare
    Weights for Optimal Tax Theory, NBER Working Paper.

68. Seade, J.K. 1977. On the Shape of Optimal Tax Schedules, Journal of
    Public Economics, 7(1), 203-236.

69. Seade, J.K. 1982. On the Sign of the Optimum Marginal Income Tax,
    Review of Economic Studies, 49, 637-643.

70. Security Exchange Commission, 2006. Executive Compensation and Re-
    lated Person Disclosure, SEC Release No. 33-8732A

71. Slemrod, J. 1996. High Income Families and the Tax Changes of the 1980s: The Anatomy of Behavioral Response. In Empirical Foundations of Household Taxation, eds. Martin Feldstein and James Poterba, 169-192. Chicago: University of Chicago Press.

72. Slemrod, J. and Kopczuk, W. 2002. The Optimal Elasticity of Taxable Income. Journal of Public Economics, 84(1), 91-112.

73. Slemrod, J. and Yitzhaki, S. 2002. Tax Avoidance, Evasion and Administration. In Handbook of Public Economics, 1st edition, Volume 3, eds. Alan Auerbach and Martin Feldstein, 1423-1470.

74. Weinschenk, P. 2013. Compensation, perks, and welfare. Economics Letters, Elsevier, vol. 120(1), 67-70.

75. Yermack, D. 2006. Flights of fancy: corporate jets, CEO perquisites, and inferior shareholder returns. Journal of Financial Economics 80, 211–242.

76. Zabojnik, J. 1996. Pay-Performance Sensitivity and Production Uncertainty, Economics Letters 53, 291-96.

77. Zabojnik, J. 2002. Centralized and Decentralized Decision-Making in Organizations, Journal of Labor Economics 20, January, 1-22.