Essays in Dynamic Contract Theory

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

vorgelegt von
Sebastian Köhne

Mannheim, Mai 2010

Dekan:	Prof. Tom Krebs, Ph.D.
Referent:	Prof. Dr. Ernst-Ludwig von Thadden
Korreferent:	Prof. Nicola Pavoni, Ph.D.
Tag der mündlichen Prüfung:	28. Juni 2010

Acknowledgements

First of all, I wish to thank my advisor Ernst-Ludwig von Thadden for his invaluable guidance throughout the process of writing this thesis. My research has benefited a lot from his deep knowledge about contract theory, the great research environment at his chair, and his very constructive feedback on my results. I am deeply indebted for the encouragement and support I have received during the last five years.

I also wish to thank Nicola Pavoni for his invaluable advice on my thesis and beyond. His supervision during my visits to the University College London has had a huge impact on my understanding of dynamic public finance and macroeconomics. I am deeply indebted for his continuous help and confidence in me.

Moreover, I am very grateful to Árpád Ábrahám and Nicola Pavoni for the productive and joyful cooperation on Chapter 4 of this thesis. I am looking forward to future collaborations on this topic.

Ian Jewitt provided numerous constructive comments on Chapter 3 of this thesis. I am also very grateful for his help during the job market period.

In addition, I wish to thank my fellow graduate students at the CDSE for the wonderful time I experienced here. I am particularly grateful to all fellows from the class of 2005.

Finally, Jana and my family have been giant sources of support. Thank you!

Contents

Chapter 1

Introduction

This thesis consists of three main chapters, which contribute to the literature on dynamic contract theory and dynamic public finance. The chapters are self-contained and can be read separately. A short summary of each chapter is provided below.

Repeated moral hazard with history-dependent preferences

Chapter 2 introduces history-dependent preferences into the repeated moral hazard framework: the agent's instantaneous consumption utility may depend on both present and past levels of consumption. The paper shows in what sense the Inverse Euler equation generalizes to this environment and derives implications for the intertemporal wedge between the principal's rate of return to saving and the agent's shadow rate of return. While the intertemporal wedge is positive under mild assumptions for all two-period models with history-dependence, the intertemporal wedge is generally indeterminate for models with a longer time-horizon. Finally, the paper contrasts some specific functional forms of history-dependence.

The first-order approach to moral hazard problems with hidden saving

Moral hazard models with hidden saving decisions are useful to study such diverse problems as unemployment insurance, income taxation, executive compensation, or human capital policies. How can we solve such models? In general, this is very difficult. Under the conditions derived in Chapter 3, however, we can replace the incentive constraint with the associated first-order condition. This allows the application of simple Lagrangian methods and yields a precise characterization of optimal contracts. To obtain tractable conditions for the validity of this approach, the paper draws on the concept of log-convexity. Since log-convexity, unlike convexity, is preserved under multiplication, the paper is able to separate the assumptions on the output distribution from the assumptions on the agent's preferences in a specific sense, even though the interaction between these two is important for the agent's incentives.

The first-order approach is valid if the following conditions hold: a) the agent has nonincreasing absolute risk aversion (NIARA) utility, b) the output technology has monotone likelihood ratios (MLR), and c) the distribution function of output is log-convex in effort (LCDF). Finally, the paper shows how the curvature of optimal wage schemes can be used to relax the above conditions.

Optimal income taxation with asset accumulation

Chapter 4 is joint work with Árpád Ábrahám (European University Institute) and Nicola Pavoni (University College London). This paper studies the effect of capital taxation on the shape of labor income taxes. We consider a two-period model of social insurance. By exerting effort, agents can influence their labor income realizations. Moreover, agents can use a risk-free bond for intertemporal self-insurance. We show that the planner's trade-offs in designing optimal effort incentives depend crucially on the tax rate on bond returns. Through this channel, taxing the bond affects the structure of optimal labor income taxes. Specifically, when agents have preferences with convex absolute risk-aversion, we find that taxing the bond leads to more progressivity of optimal marginal tax rates on labor income.

Chapter 2

Repeated moral hazard with history-dependent preferences

2.1 Introduction

The Inverse Euler equation, initially discovered by Diamond and Mirrlees (1978) and Rogerson (1985b), recently generalized by Golosov, Kocherlakota, and Tsyvinski (2003), is arguably one of the most important findings in the dynamic contracting literature. The equation is based on the central observation that the agent's utilities can be intertemporally adjusted without changing the incentives to exert effort. As a consequence, the principal's marginal costs of providing utility today must be equal to the expected marginal costs of providing utility tomorrow. For history-independent (in other words, time-separable) preferences, the marginal costs of providing utility are given by the agent's inverse marginal utility of consumption. The Inverse Euler equation is therefore an immediate consequence of the previous observation.

However, the Inverse Euler equation requires a strong assumption regarding the agent's preferences as the above argument shows. The present paper introduces a more general preference class into the principal-agent framework. Following similar ideas in the macroeconomic literature,[1] the agent's preferences are allowed to be history-dependent. More precisely, the agent's consumption utility may depend on both present and past levels of consumption. In this setup, the simple link between the principal's marginal costs of providing utility and the agent's inverse marginal utility of consumption breaks down and thus the Inverse Euler equation does not longer apply. The goal of this paper is to characterize optimal contracts in this environment and study the robustness of the lessons drawn from the Inverse Euler equation. This task is important not only from a theoretical point of view, but also from a more applied perspective, considering that history-independent preferences might be too narrow to describe and predict economic behavior (Helson 1964, Frederick and Loewenstein 1999).

In the first part of this paper, I study history-dependent preferences in a *two-period* model (with a risk-averse agent and a risk-neutral principal) and identify two key properties of optimal contracts. First, I show that, conditional on present information, the expectation of the agent's marginal rate of intertemporal substitution is equal to the principal's rate of return to saving. Note that the rate of return to saving coincides with the principal's marginal rate of intertemporal substitution. Intuitively, if the expectations of the marginal rates of intertemporal substitution differ between the two parties, then the principal can extract "gains from trade" by rebalancing the agent's utility between the present and the future. If the principal perturbs the agent's future utility in a uniform way across states, then such intertemporal adjustments do not violate the incentive constraint. Therefore, a contract can only be optimal when intertemporal "gains from trade" are impossible.

For history-independent preferences, the above insight is in fact equivalent to the Inverse Euler equation. Hence, we learn in what sense the Inverse Euler equation generalizes to

[1] See Ryder and Heal (1973) for the pioneering contribution.

preferences with history-dependence: The correct interpretation does not consider the agent's marginal utilities, but his marginal rate of intertemporal substitution.

In a second step, I explore how saving distorts the incentive problem. For models with history-independent preferences, the Inverse Euler equation implies that optimal contracts feature a positive *intertemporal wedge*: The principal's rate of return to saving exceeds the shadow rate of return associated with the agent's consumption scheme (Diamond and Mirrlees 1978, Rogerson 1985b, Golosov, Kocherlakota, and Tsyvinski 2003). Therefore, the agent would choose to save if he had access to the same savings technology as the principal. This insight has important consequences for the decentralization of optimal allocations and means that the agent's asset returns must be corrected by a tax system (Kocherlakota 2005, Albanesi and Sleet 2006). In the present setup, by contrast, the intertemporal wedge may be positive or negative. This implies that the agent's access to the savings technology must be distorted as well, but it is possible that *borrowing* becomes the crucial problem.

However, negative intertemporal wedges are not generic (in the *two-period* setup). I show that the intertemporal wedge is positive if the agent's marginal rate of intertemporal substitution increases in future consumption. Put differently, this condition states that the agent's willingness to trade future for present consumption increases with the level of future consumption. This condition is trivially satisfied for preferences that are history-independent. Given its economic meaning, the condition seems also innocuous for the class of history-dependent preferences. Hence, the result is surprisingly strong. Even though history-dependence may change the agent's savings policy quite drastically compared to the standard model, the positivity of the intertemporal wedge is a very robust insight. Moreover, I point out how this result relates to the impact of saving on the effectiveness of the incentive scheme.

In the second part of this paper, I extend the time-horizon to more than two periods. It is straightforward that all results from the two-period model apply to the last two dates of the multi-period model. In general, however, the structure of optimal allocations can be different from the two-period setup. Specifically, I show that the expectation of the agent's marginal rate of intertemporal substitution is *not* equal to the principal's rate of return to saving in general. The intuition for this finding is closely linked to the above discussion. For the multi-period model, it is not generally true that the principal can realize "gains from trade" when the expected marginal rates of intertemporal substitution differ between the two parties. With history-dependent preferences, adjustments in the timing of rewards for periods t and $t + 1$ will change the agent's preferences, and hence the incentives to exert effort, in periods $t + 2$ and later. This effect may be more important than equalizing the expected willingness to trade between periods t and $t + 1$ across the two parties.

The previous finding has direct implications for the intertemporal wedge in the multi-period model: There is no longer such a simple link between monotonicity of the marginal rate of intertemporal substitution and the sign of the intertemporal wedge. More precisely, the intertemporal wedge between periods t and $t + 1$ is now the sum of two terms. The first term is well-known from the two-period model and, ceteris paribus, fosters a positive intertemporal wedge when the agent's willingness to trade period-$(t + 1)$ for period-t consumption increases in the level of period-$(t + 1)$ consumption. From the study of the two-period model, we can relate this term to the impact of saving on the agent's incentives to exert effort in period $t + 1$. The second term captures the difference between the principal's rate of return to saving and the expectation of the agent's marginal rate of intertemporal substitution. Roughly speaking, this term reflects the impact of saving on the incentives to exert effort in periods $t + 2$ and later. Using two simple examples, I show that the sign and the size of this second term are highly sensitive to the exact functional form of history-dependence. Therefore, the sign of the intertemporal wedge in the multi-period model is generally indeterminate.[2]

Overall, this paper bridges the gap between the literature on dynamic moral hazard initiated by Rogerson (1985b) and the literature on history-dependent preferences (or habit formation). In macroeconomic models, history-dependent preferences have been studied by Ryder and Heal (1973), Abel (1990), Constantinides (1990), and Campbell and Cochrane (1999), among others.[3] Frederick and Loewenstein (1999) review the substantial body of empirical research supporting this approach. For instance, workers' self-reported well-being is often strongly related to recent changes in pay, but not so much to absolute levels of pay (Clark 1999, Grund and Sliwka 2007). More generally, the psychological literature documents that the repetition of any stimulus tends to reduce the perception and response to it (Helson 1964). Interpreting consumption as a stimulus, this suggests that, other things being equal, past consumption has an important (negative) influence on current well-being.

So far, however, the impact of time-nonseparable preferences on incentive problems is largely unexplored. The only exception is the work by Grochulski and Kocherlakota (2010), who study a dynamic Mirrleesian taxation model. Their paper provides a first example in which time-nonseparable preferences change the intertemporal wedge. Moreover, the paper finds that the standard way of decentralizing optimal allocations has an important measurability problem when preferences are not time-separable. To remedy this problem, the authors consider a 'social security system' that uses the retirement date to implement the required taxes and transfers. The present paper is complementary to those findings. I focus on how the intertemporal structure of optimal allocations depends on the time-separability of preferences and show under what conditions the

[2]Recall, however, that the intertemporal wedge between the last two periods is positive under mild assumptions; see the discussion of the two-period model.

[3]In particular, models with history-dependent preferences have contributed to the solution of empirical puzzles related to consumption behavior (excess sensitivity puzzle, excess smoothness puzzle), asset pricing (equity premium puzzle) and the relationship between savings and growth; see Messinis (1999) for a review.

intertemporal wedge can be determined.[4]

On a related note, the present paper provides a new perspective on the seminal results by Diamond and Mirrlees (1978) and Rogerson (1985b). By leaving the class of time-separable preferences, the paper identifies the true driving forces for those results. In particular, we see that the positive intertemporal wedge in those models is generated by two factors: First of all, time-separable preferences entail that the agent's willingness to trade present consumption for next period's consumption is increasing in the level of next period's consumption. This property implies that, at the optimal contract, private saving reduces the power of next period's incentives. Secondly, time-separable preferences entail that perturbations of the consumption scheme between two subsequent periods do not affect the agent's preferences (and thus the effort incentives) for any other periods.

In addition, the present paper can be seen as a complement to the literature on effort persistence, which extends the Rogerson (1985b) model by allowing for a *technology* that is not time-separable; see Mukoyama and Sahin (2005), Kwon (2006), Jarque (2010), and Hopenhayn and Jarque (2010). In that setup, the Inverse Euler equation is not affected. Hence, effort persistence and history-dependent preferences ("preference persistence") have fundamentally different effects on the shape of optimal contracts. The reason is that effort persistence alters the information structure of the model, but leaves the preference structure unchanged.

Finally, this paper is related to the discussion on functional forms of history-dependence; compare Rozen (2010). A common approach is to define the agent's utility function on a weighted average of current consumption and current consumption minus the reference level (e.g. Constantinides 1990). Another approach is to aggregate current consumption

[4]Notice that the sign of the intertemporal wedge is also crucial for the shape of the decentralization mechanism. In particular, the 'regressivity' of wealth taxes à la Kocherlakota (2005) is exactly due to the positive sign of the wedge in the time-separable model.

and the reference level in a Cobb-Douglas fashion (e.g. Abel 1990). I contrast these two formulations and show that the latter can have counterintuitive implications for the effect of saving on the agent's incentives to exert effort. Therefore, the weighted average formulation of history-dependence appears much more suitable in the present framework.[5]

The paper proceeds as follows: Section 2.2 sets up the basic two-period model. Section 2.3 generalizes the Inverse Euler equation and characterizes the intertemporal wedge. In Section 2.4, I discuss a few specific functional forms of history-dependence. Section 2.5 extends the model to a longer time-horizon and Section 2.6 concludes.

2.2 Model

I study a two-period moral hazard problem in which the agent's preferences are history-dependent: Second-period consumption utility may depend on consumption in both periods. This setup generalizes the time-separable model introduced by Rogerson (1985b).

2.2.1 Preferences

The relationship between principal (P) and agent (A) lasts for two periods; see Section 2.5 for a longer time-horizon. P maximizes expected profits. She has access to a linear savings technology at the rate R, thus her discount factor equals $1/R$.

A has von-Neumann-Morgenstern preferences and maximizes the expected value of

$$u(c_1) - v(e_1) + \beta(\tilde{u}(c_1, c_2) - v(e_2)),$$

[5] For the multi-period problem, an additional complication arises: one needs to specify a mapping from the consumption history to the reference level. This question is beyond the scope of the present paper, however.

where c_t denotes consumption, e_t represents effort, and $\beta \in (0,1]$ is the discount factor. The functions $u, \tilde{u}, v$ are twice continuously differentiable. The derivatives of u and v are denoted u' and v', while the partial derivatives of $\tilde{u}$ are denoted $\tilde{u}_1$, $\tilde{u}_2$.

Effort disutility $v(e)$ is strictly increasing and weakly convex in e. Lifetime consumption utility is strictly increasing and strictly concave in c_1 and c_2:

$$\forall c_1, c_2: \quad u'(c_1) + \beta\tilde{u}_1(c_1, c_2) > 0, \quad u''(c_1) + \beta\tilde{u}_{11}(c_1, c_2) < 0,$$
$$\forall c_1, c_2: \quad \tilde{u}_2(c_1, c_2) > 0, \quad \tilde{u}_{22}(c_1, c_2) < 0.$$

To rule out boundary solutions, I suppose

$$\forall c_2 > 0: \quad \lim_{c_1 \to 0} \left(u'(c_1) + \beta\tilde{u}_1(c_1, c_2)\right) = \infty,$$
$$\forall c_1 > 0: \quad \lim_{c_2 \to 0} \tilde{u}_2(c_1, c_2) = \infty.$$

I call A's preferences **history-independent** (or time-separable) if $\tilde{u}(c_1, c_2)$ does not depend on c_1. Otherwise, I call A's preferences **history-dependent**. In the latter case, I will refer to c_1 as A's second-period **reference level** of consumption.

2.2.2 Technology

In each period, A exerts a hidden work effort and thereby generates a publicly observable stochastic output. Output realizations for period 1 are denoted x_i, $i = 1, \ldots, N$, with associated probabilities $\pi_i(e_1) > 0$. Second-period output realizations are denoted x_j, $j = 1, \ldots, N$, with associated probabilities $\pi_j(e_2) > 0$.

A cannot save or borrow, hence his consumption equals his wage in each period.

2.2.3 Contracts

A **contract** is a specification (c, e) of wages $c = (c_i, c_{ij})_{i,j}$ and effort levels $e = (e_0, e_i)_i$. Here, e_0 is the recommended effort for period 1, e_i the recommended effort for period 2 given that x_i has realized in period 1. Similarly, c_i is the wage paid in period 1, c_{ij} is the wage paid in period 2, given that outputs (x_i, x_j) have realized.

P offers a contract at the beginning of period 1. She has to respect A's participation constraint

$$\sum_{i,j} \pi_i(e_0)\pi_j(e_i)\Big(u(c_i) - v(e_0) + \beta\left[\tilde{u}(c_i, c_{ij}) - v(e_i)\right]\Big) \geq \underline{U}. \tag{PC}$$

In addition, since effort is not observed by P, contracts must satisfy the incentive compatibility constraint

$$e \in \underset{(e_0', e_i')}{\operatorname{argmax}} \sum_{i,j} \pi_i(e_0')\pi_j(e_i')\Big(u(c_i) - v(e_0') + \beta\left[\tilde{u}(c_i, c_{ij}) - v(e_i')\right]\Big). \tag{IC}$$

2.2.4 Optimal contracts

Contract (c, e) is called **optimal** if it maximizes P's expected profits

$$\sum_{i,j} \pi_i(e_0)\pi_j(e_i)\Big(x_i - c_i + \frac{1}{R}\left[x_j - c_{ij}\right]\Big)$$

subject to the incentive compatibility constraint (IC) and the participation constraint (PC).

2.3 Results on intertemporal optimality

To simplify notation, I denote expectations conditional on first-period information x_i by $E_i[\,\cdot\,]$ in this section. For example, I write

$$\sum_j \frac{\pi_j(e_i)}{\tilde{u}_2(c_i, c_{ij})} = E_i\left[\frac{1}{\tilde{u}_2(c_i, c_{ij})}\right],$$

and so on.

2.3.1 Modification of the Inverse Euler equation

P can divide wages over time in the following way: The total reward for good performance in period 1 need not be given immediately, but may also be included in second-period wages (which makes the second-period wages dependent on the first-period output). If at least one of the parties is risk-averse, as in the model considered here, then it is clearly beneficial to make use of this option.

This reasoning yields the following intertemporal condition.[6]

Proposition 2.1. *Let (c, e) be an optimal contract. Then for all $i = 1, \ldots, N$*

$$\frac{1}{\beta R} E_i\left[\frac{1}{\tilde{u}_2(c_i, c_{ij})}\right] = \frac{1}{u'(c_i)}\left(1 - \frac{1}{R} E_i\left[\frac{\tilde{u}_1(c_i, c_{ij})}{\tilde{u}_2(c_i, c_{ij})}\right]\right). \tag{2.1}$$

Proposition 2.1 states that P's marginal cost of increasing A's utility conditional on output x_i is constant over time: The cost of increasing $\tilde{u}$ by ϵ/β is approximately equal to

$$\frac{\epsilon}{\beta R} E_i\Big[\frac{1}{\tilde{u}_2(c_i, c_{ij})}\Big].$$

[6] All proofs can be found in the appendix.

The cost of increasing u by ϵ while keeping $\tilde{u}$ constant is approximately equal to

$$\frac{\epsilon}{u'(c_i)} - \frac{\epsilon}{Ru'(c_i)} E_i \Big[\frac{\tilde{u}_1(c_i, c_{ij})}{\tilde{u}_2(c_i, c_{ij})} \Big],$$

where the first term denotes the increase of period-1 consumption required to raise u by ϵ, and the second term represents the increase of period-2 consumption needed to keep $\tilde{u}$ at the former level. Equation (2.1) shows that at any optimal contract these two costs are the same.

If preferences are history-independent, with $u = \tilde{u}$, then equation (2.1) collapses to

$$\frac{1}{u'(c_i)} = \frac{1}{\beta R} E_i \left[\frac{1}{u'(c_{ij})} \right]. \tag{2.2}$$

This well-known result from Rogerson (1985b) is often called the Inverse Euler equation. The relation between inverse marginal utilities described in equation (2.1) is slightly different, thus we cannot attach the same label here. Yet, equations (2.1) and (2.2) have exactly the same interpretation: For any optimal contract, P's marginal costs of providing utility conditional on a given output in period 1 must be equal across periods.

Finally, notice that equation (2.1) can equivalently be written as

$$E_i \left[\frac{u'(c_i) + \beta \tilde{u}_1(c_i, c_{ij})}{\beta \tilde{u}_2(c_i, c_{ij})} \right] = R, \tag{2.3}$$

which means that A's conditionally expected marginal rate of intertemporal substitution is equal to P's marginal rate of intertemporal substitution, R. The intuition for equation (2.3) is closely linked to the above discussion: If the conditionally expected marginal rates of intertemporal substitution differ between P and A, then P can realize gains from readjusting the timing of rewards while maintaining incentive compatibility.

2.3.2 The intertemporal wedge

In dynamic models with hidden effort decisions, there is an intertemporal distortion due to the influence of savings on the incentive problem. The intertemporal distortion in models with history-independent preferences is well understood. Diamond and Mirrlees (1978), Rogerson (1985b) and, in a more general framework, Golosov, Kocherlakota, and Tsyvinski (2003) show that there is a positive intertemporal wedge in such models: At optimal contracts, P's rate of return to saving exceeds the shadow rate of return associated with A's consumption scheme.[7]

This subsection characterizes the intertemporal wedge for preferences that are history-dependent. Recall that $E_i[\,\cdot\,]$ denotes expectations conditional on first-period information x_i, i.e., with respect to probability weights $\pi_j(e_i)$. Similarly, $\mathrm{cov}_i(\,\cdot\,,\,\cdot\,)$ represents covariances with respect to $\pi_j(e_i)$.

Consider the difference between P's rate of return to saving and A's shadow rate of return,

$$R - \frac{u'(c_i) + \beta E_i\left[\tilde{u}_1(c_i, c_{ij})\right]}{\beta E_i\left[\tilde{u}_2(c_i, c_{ij})\right]}.$$

It is convenient to divide this difference by P's rate of return, R, and define the **intertemporal wedge** (given output x_i) as

$$W_i := 1 - \frac{u'(c_i) + \beta E_i\left[\tilde{u}_1(c_i, c_{ij})\right]}{R\beta E_i\left[\tilde{u}_2(c_i, c_{ij})\right]}. \tag{2.4}$$

The intertemporal wedge can be characterized as follows.

Proposition 2.2. *Consider an optimal contract. Then, conditional on first-period out-*

[7] The shadow rate of return is defined as the rate at which A does not wish to save or borrow.

put x_i, the intertemporal wedge equals

$$W_i = 1 - \frac{R - \operatorname{cov}_i\left(\tilde{u}_1(c_i, c_{ij}),\ \frac{1}{\tilde{u}_2(c_i, c_{ij})}\right)}{R - R\operatorname{cov}_i\left(\tilde{u}_2(c_i, c_{ij}),\ \frac{1}{\tilde{u}_2(c_i, c_{ij})}\right)}. \tag{2.5}$$

Hence, the intertemporal wedge is positive if and only if

$$\operatorname{cov}_i\left(R\tilde{u}_2(c_i, c_{ij}) - \tilde{u}_1(c_i, c_{ij}),\ \frac{1}{\tilde{u}_2(c_i, c_{ij})}\right) < 0. \tag{2.6}$$

Proposition 2.2 proves that the sign of the intertemporal wedge coincides with the sign of the negated covariance

$$-\operatorname{cov}_i\left(R\tilde{u}_2(c_i, c_{ij}) - \tilde{u}_1(c_i, c_{ij}),\ \frac{1}{\tilde{u}_2(c_i, c_{ij})}\right).$$

Intuitively, this expression captures the hypothetical impact of saving on incentives: The first variable in the covariance, $R\tilde{u}_2 - \tilde{u}_1$, is A's marginal benefit of saving. This benefit consists of the marginal payoff of saving, $R\tilde{u}_2$, plus the marginal utility of reducing the reference level, $-\tilde{u}_1$.[8] The second variable, $1/\tilde{u}_2$, is monotonic in A's second-period wage. If these two variables covary negatively, then saving tends to increase A's utility in the states with low wages relative to the states with high wages. In other words, saving "hedges" against incentives, or "tightens" the incentive compatibility constraint.

Simply put, Proposition 2.2 thus shows that the intertemporal wedge is positive if and only if saving "hedges" against the incentive scheme. This insight extends the intuition from Golosov, Kocherlakota, and Tsyvinski (2003) to the case of history-dependent preferences. The explanation for the emergence of a positive intertemporal wedge, given *history-independent* preferences, is that saving has an adverse effect on incentives. For this reason, the social cost of saving is higher than A's individual cost. It is optimal to

[8] A's marginal cost of saving, $u'(c_i)$, does not show up here, because it is independent of the realization of second-period output.

equate the social marginal cost and benefit of saving, but not A's individual marginal cost and benefit. Therefore, at optimal contracts, A's individual marginal benefit of saving exceeds A's individual marginal cost.

For history-independent preferences, with $u = \tilde{u}$, the above finding takes a particularly simple form and equation (2.5) collapses to

$$W_i = 1 - \frac{1}{1 - \mathrm{cov}_i\left(u'(c_{ij}),\, \frac{1}{u'(c_{ij})}\right)}. \tag{2.7}$$

Since the inverse function is strictly decreasing, the covariance term in (2.7) is nonpositive. Hence, we see that the intertemporal wedge is nonnegative in this case. Moreover, we note that the wedge is positive unless the second-period wage c_{ij} is independent of j.

The intertemporal wedge, as Proposition 2.2 points out, is driven by the impact of saving on A's incentives to exert effort. It is important to notice that this result still involves endogenous variables, since the covariances in (2.5) depend on the contract (c, e). We will see now, however, that the sign of the wedge can be linked to a very simple property of A's preferences.

To simplify notation, write A's **marginal rate of intertemporal substitution** as

$$M(c_1, c_2) := \frac{u'(c_1) + \beta \tilde{u}_1(c_1, c_2)}{\beta \tilde{u}_2(c_1, c_2)}.$$

Proposition 2.3. *Let (c, e) be an optimal contract. Let the first-period output be x_i and suppose $c_{ij} \neq c_{ij'}$ for some j, j'.*[9] *Then the intertemporal wedge is positive (negative) if $M(c_1, c_2)$ is increasing (decreasing) in c_2.*

Proposition 2.3 is a strong result. Assuming that A's willingness to give up future consumption increases in the level of future consumption, i.e., assuming that $M(c_1, c_2)$ is

[9] The case with $c_{ij} = c_{ij'}$ for all j, j' is not interesting. It corresponds to the implementation of minimal effort.

increasing in c_2, we always obtain a positive intertemporal wedge in this model. So, even though history-dependence may change A's savings policy quite drastically compared to the time-separable case, the positivity of the intertemporal wedge is a very robust insight in the present setting. We have to be somewhat careful, however, when extending this finding to models with more than two periods; see Section 2.5.

The next section studies two specific forms of history-dependent preferences. For the *weighted average* formulation, the marginal rate of intertemporal substitution increases in future consumption. For the *Cobb-Douglas* specification, the latter property does not generally hold. As a consequence, the intertemporal wedge can be negative in that setup.

2.4 Functional forms of history-dependence

In this section, I explore the two most common formulations of history-dependent preferences: weighted averages and Cobb-Douglas aggregation. Among many others, Constantinides (1990), Lahiri and Puhakka (1998), and Campbell and Cochrane (1999) follow the first approach. Prominent examples of the Cobb-Douglas approach are Abel (1990) and Carroll, Overland, and Weil (1997, 2000).

Both formulations impose two additional restrictions on A's preferences: Utility in the second period is decreasing in first-period consumption. In other words, A forms *consumption habits*. Secondly, the marginal rate of intertemporal substitution decreases as habits become more important.[10] Notice, however, that the results derived in the previous section did not rely on any of these two properties.

[10] For the Cobb-Douglas model, there is a range of the consumption space in which this property is violated (Wendner 2003). I will only consider cases where the contracts do not fall into this range.

2.4.1 Weighted averages

Let $u : \mathbb{R}_+ \to \mathbb{R}$ be twice continuously differentiable, with $u' > 0$, $u'' < 0$, $\lim_{c \to 0} u'(c) = \infty$. Define period-1 consumption utility as $u(c_1)$ and period-2 consumption utility as

$$\tilde{u}(c_1, c_2) := u(\tilde{c}_2), \quad \tilde{c}_2 := c_2 - \gamma c_1. \tag{2.8}$$

Effective consumption in period 2 can be rewritten

$$\tilde{c}_2 = (1 - \gamma) c_2 + \gamma (c_2 - c_1).$$

That is, effective consumption a *weighted average* of absolute consumption, c_2, and absolute consumption minus the reference level, $c_2 - c_1$. The parameter $\gamma \in [0, 1]$ controls the importance of habits: The higher the value of γ, the more A cares about how period-2 consumption relates to period-1 consumption.

By applying Proposition 2.3, we find that the intertemporal wedge is positive in this setup.

Corollary 2.1. *Let A's preferences be given by* (2.8)*, let* (c, e) *be an optimal contract and suppose that for each* i *there exist* j, j' *such that* $c_{ij} \neq c_{ij'}$*. Then the intertemporal wedge is positive.*

The remainder of this subsection discusses the comparative statics of optimal wage profiles in the habit parameter. For time-separable preferences, the intertemporal slope of wages is determined by the curvature of A's inverse marginal utility, as Rogerson (1985b) has shown. If A's inverse marginal utility is convex, then expected wages are nonincreasing over time. This convexity requirement is satisfied in many generic cases, for instance for CRRA utility with a coefficient of relative risk aversion larger than one.

If A forms habits, then the intertemporal slope of wages does not only depend on the curvature of inverse marginal utility, but also on the importance of habits. This tends to shift wages to the end of the relationship as we will see now.

For logarithmic utility, the functional relation between the slope of wages and the habit parameter γ is particularly simple.

Example 2.1. Let $u(c) = \log(c)$, $\beta R = 1$. Then equation (2.1), the condition for intertemporal optimality of rewards, takes the form

$$E_i\left[\frac{1}{u'(\tilde{c}_{ij})}\right] = \frac{1}{u'(c_i)}(1+\beta\gamma),$$

which implies

$$\frac{E_i[c_{ij}]}{c_i} = 1 + (1+\beta)\gamma.$$

Thus, expected wages are constant over time if A does not form habits ($\gamma = 0$). If he does form habits ($\gamma > 0$), then expected wages are increasing over time. Moreover, the ratio between expected wages paid in period 2 and period 1 is increasing in the habit parameter γ.

In the above example, expected wages are increasing over time for any positive habit parameter. More generally, no matter what utility function is chosen, expected wages will increase over time if the habit parameter is sufficiently large, as the following remark shows.

Remark. Let $\gamma = 1$. If (c, e) is an optimal contract, then $c_{ij} > c_i$ for all i, j.[11]

The ratio between expected wages at date 2 and date 1 is a monotonic function of the habit parameter γ in the log-utility example. This result can be generalized. However, for non-logarithmic utility, the intertemporal condition (2.1) cannot be solved for wage levels. The shape of wages can only be studied if one includes the incentive and

[11] For $\gamma = 1$, second-period utility takes the form $\tilde{u}(c_i, c_{ij}) = u(c_{ij} - c_i)$. Therefore, the remark follows immediately from the assumption $\lim_{c\to 0} u'(c) = \infty$.

participation constraint in the analysis. To keep things tractable, I study the simplest possible case: two effort levels, $\{l, h\}$, and two outputs, $\{x_L, x_H\}$, $x_L < x_H$. Probability distributions satisfy $\pi_H(l) < \pi_H(h)$, and effort costs are $v(l) = 0$, $v(h) = v > 0$. To simplify notation, set $\nu := v/(\pi_H(h) - \pi_H(l))$.

Suppose P wants to implement effort h in both periods (i.e., $x_H - x_L$ is sufficiently large). Then optimal contracts are characterized as follows:

$$\frac{1 + \frac{\gamma}{R}}{u'(c_H)} - \pi_H(h)\frac{1}{\beta R u'(c_{HH} - \gamma c_H)} - (1 - \pi_H(h))\frac{1}{\beta R u'(c_{HL} - \gamma c_H)} = 0 \tag{2.9}$$

$$\frac{1 + \frac{\gamma}{R}}{u'(c_L)} - \pi_H(h)\frac{1}{\beta R u'(c_{LH} - \gamma c_L)} - (1 - \pi_H(h))\frac{1}{\beta R u'(c_{LL} - \gamma c_L)} = 0 \tag{2.10}$$

$$u(c_{HH} - \gamma c_H) - u(c_{HL} - \gamma c_H) - \nu = 0 \tag{2.11}$$

$$u(c_{LH} - \gamma c_L) - u(c_{LL} - \gamma c_L) - \nu = 0 \tag{2.12}$$

$$u(c_H) + u(c_{HL} - \gamma c_H) + (1 + \beta)\nu\pi_H(l) - \nu - \underline{U} = 0 \tag{2.13}$$

$$u(c_L) + u(c_{LL} - \gamma c_L) + (1 + \beta)\nu\pi_H(l) - \underline{U} = 0. \tag{2.14}$$

Here, (2.9),(2.10) are the intertemporal optimality conditions, (2.11),(2.12) ensure incentive compatibility of second-period effort, (2.13),(2.14) ensure incentive compatibility of first-period effort, and together with (2.11),(2.12) also imply that the participation constraint is satisfied.

This system of six equations can be separated into two subsystems with three equations each—one system determining c_H, c_{HL}, c_{HH}, and one determining c_L, c_{LL}, c_{LH}. The comparative statics in the habit parameter γ are as follows.

Proposition 2.4. *Let (c, e) be the optimal contract in the two-effort two-output model described above. Then for each i, j, the ratio c_{ij}/c_i is increasing in γ.*

The intuition underlying Proposition 2.4 is straightforward. If the habit parameter γ increases, A demands a higher compensation for the "comparison effects" generated by date-1 wages. This makes wages paid at date 1 more costly relative to wages paid at date 2, hence P substitutes.

2.4.2 Cobb-Douglas aggregation

Let $u : \mathbb{R}_+ \to \mathbb{R}$ be twice continuously differentiable, with $u' > 0$, $u'' < 0$, $\lim_{c \to 0} u'(c) = \infty$. Let $\gamma \in [0, 1]$, $\epsilon > 0$. Define period-1 consumption utility as $u(c_1)$ and period-2 consumption utility as

$$\tilde{u}(c_1, c_2) := u(\tilde{c}_2), \quad \tilde{c}_2 := \frac{c_2}{(c_1 + \epsilon)^\gamma}. \tag{2.15}$$

Effective consumption in period 2 can be rewritten

$$\tilde{c}_2 = c_2^{1-\gamma} \left(\frac{c_2}{c_1 + \epsilon} \right)^\gamma .$$

That is, effective consumption is a *Cobb-Douglas* aggregate of absolute consumption, c_2, and absolute consumption relative to the reference level, $c_2/(c_1 + \epsilon)$.[12] The parameter γ is the weight attached to the latter term. For an agent with a higher value of γ, the comparison between date-2 consumption and date-1 consumption thus gets more important in this sense.

The next proposition is the main result of this subsection. It shows that the present setup changes the key insight of Rogerson's (1985b) time-separable model: The intertemporal

[12] The introduction of $\epsilon > 0$ in the reference level is a technical necessity for this specification. Note that for $\epsilon = 0$ one obtains

$$u'(c_1) + \beta \tilde{u}_1(c_1, c_2) = u'(c_1) - \beta\gamma c_1^{-\gamma-1} c_2 u'(\tilde{c}_2). \tag{2.16}$$

Hence, in this case, the assumption $u'(0) = \infty$ is no longer sufficient for the condition

$$\forall c_2 > 0 : \quad \lim_{c_1 \to 0} \left(u'(c_1) + \beta \tilde{u}_1(c_1, c_2) \right) = \infty, \tag{2.17}$$

which is needed to make sure that solutions are interior. For $\epsilon > 0$, this problem does not arise.

wedge is not generally positive.

Proposition 2.5. *Let A's preferences be given by* (2.15)*, let* (c, e) *be an optimal contract and suppose that for each* i *there exist* j, j' *such that* $c_{ij} \neq c_{ij'}$*. Then we have the following.*

(i) The intertemporal wedge is not positive in general.

(ii) The intertemporal wedge is positive if A's coefficient of relative risk-aversion, $-\tilde{c}_2 \frac{u''(\tilde{c}_2)}{u'(\tilde{c}_2)}$*, is bounded below by 1.*

The reason why the intertemporal wedge is not generally positive has already been pointed out in Section 2.3.2: While the hypothetical "hedging value" of saving is positive in the time-separable model, this is not in general true for preferences with history-dependence. Notice that A's marginal benefit of saving, $R\tilde{u}_2 - \tilde{u}_1$, consists of two parts. The marginal payoff of saving, $R\tilde{u}_2$, obviously covaries negatively with A's second-period wage. This captures a standard wealth effect of saving. In addition, there is a habit effect of saving, since saving reduces A's reference level. For the Cobb-Douglas formulation of history-dependence, the marginal utility of reducing the reference level, $-\tilde{u}_1$, covaries positively with the second-period wage when the coefficient of relative risk-aversion is smaller than one. In this situation, if the habit effect is sufficiently large, then A's total marginal benefit of saving covaries positively with the second-period wage, hence the "hedging value" of saving is negative.

To provide a simple example where the intertemporal wedge is negative, I explore the following environment.

Example 2.2. Consider the two-effort two-output problem described in the previous subsection. The optimal contract is characterized by a system of equations analogous to the system (2.9)–(2.14) from the weighted average formulation of history-dependence. Assume $R = \beta = 1$, $u(c) = \frac{1}{1-\rho} c^{1-\rho}$, $\rho = 0.4$, $\epsilon = 0.3$, $\upsilon = 0.1$, $\pi_H(l) = 0.25$, $\pi_H(h) = 0.5$, $\underline{U} = 2$.

Table 2.1 depicts optimal wages for different values of the habit parameter γ. Figures 2.1 and 2.2 display wages graphically. We see that first-period wages c_i decrease in the habit parameter γ, whereas second-period effective wages $\tilde{c}_{ij}$ increase in γ.

Figure 2.3 shows the intertemporal wedge for each output realization of period 1, $x_i \in \{x_L, x_H\}$. For $\gamma = 0$, preferences are history-independent and thus the intertemporal wedge is positive, of course. The wedge declines in γ, and eventually becomes negative. Hence, there is a cutoff value of γ, below of which the intertemporal wedge is positive for both first-period output realizations, and above of which the intertemporal wedge is negative for at least one output realization. Define γ^* as this cutoff value. In summary, we have the following.

Observation 2.1. The intertemporal wedge decreases in γ and can become negative (Figure 2.3).

The intuition is as follows. Since $\rho < 1$, the marginal habit effect of saving, $-\tilde{u}_1 = \gamma(c_1 + \epsilon)^{\gamma(\rho-1)-1} c_2^{1-\rho}$, is increasing in second-period consumption c_2. If γ is sufficiently large, this will outweigh the negative relation between the marginal payoff of saving, $R\tilde{u}_2 = (c_1 + \epsilon)^{\gamma(\rho-1)} c_2^{-\rho}$, and second-period consumption. In that case, A's marginal benefit of saving, $R\tilde{u}_2 - \tilde{u}_1$, will covary positively with second-period wages, hence the hypothetical "hedging value" of saving will be negative.

The size of the covariance between the marginal payoff of saving and second-period wages depends crucially on A's risk aversion. The above discussion points out that the habit effect of saving will dominate the wealth effect if the coefficient ρ of relative risk aversion is small. This suggests that for small values of ρ, a relatively small habit parameter γ will suffice to make the intertemporal wedge negative. Figure 2.4 shows the cutoff value γ^*, above of which the intertemporal wedge is negative for at least one realization of first-period output, for varying coefficients ρ of relative risk aversion. Indeed we see the following.

γ	c_H	$\tilde{c}_{HH}$	$\tilde{c}_{HL}$	c_L	$\tilde{c}_{LH}$	$\tilde{c}_{LL}$
0	0.57615	0.75031	0.42881	0.42444	0.58109	0.29632
0.1	0.50713	0.83061	0.49353	0.36826	0.63594	0.34152
0.2	0.40427	0.96563	0.60437	0.29368	0.72352	0.41315
0.3	0.24524	1.22542	0.82304	0.19830	0.87988	0.53872
0.4	0.10551	1.54966	1.10302	0.10978	1.09680	0.71313
0.5	0.05256	1.72965	1.26090	0.06041	1.24748	0.83940
0.6	0.03091	1.82740	1.34723	0.03648	1.33626	0.91567
0.7	0.02001	1.88827	1.40119	0.02380	1.39268	0.96471
0.8	0.01377	1.92978	1.43806	0.01638	1.43138	0.99855
0.9	0.00988	1.95983	1.46479	0.01171	1.45945	1.02320
1	0.00730	1.98254	1.48501	0.00861	1.48068	1.04189

Table 2.1: Optimal (effective) wages for different values of the habit parameter γ

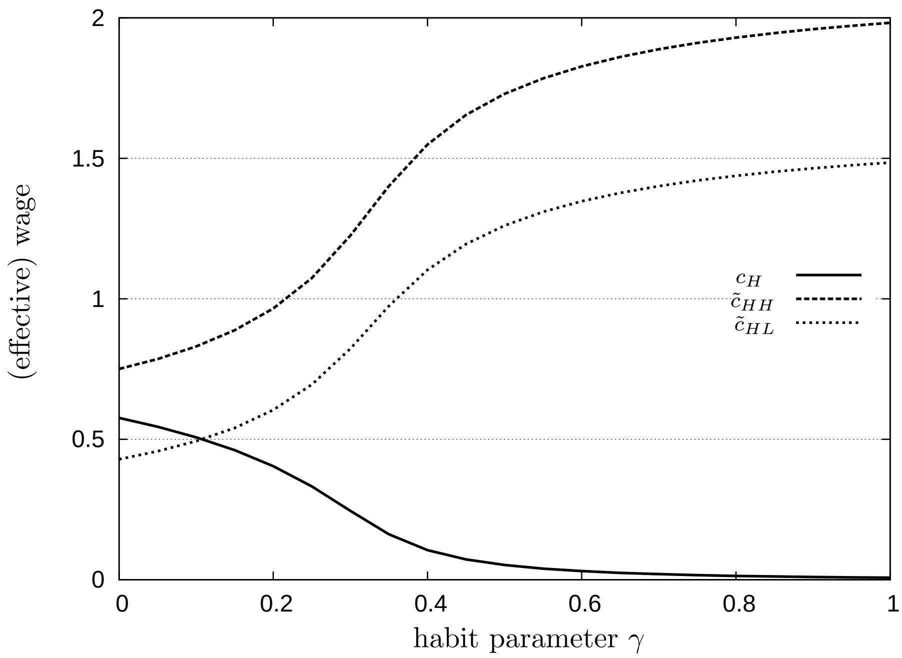

Figure 2.1: (Effective) wages, given high output in period 1

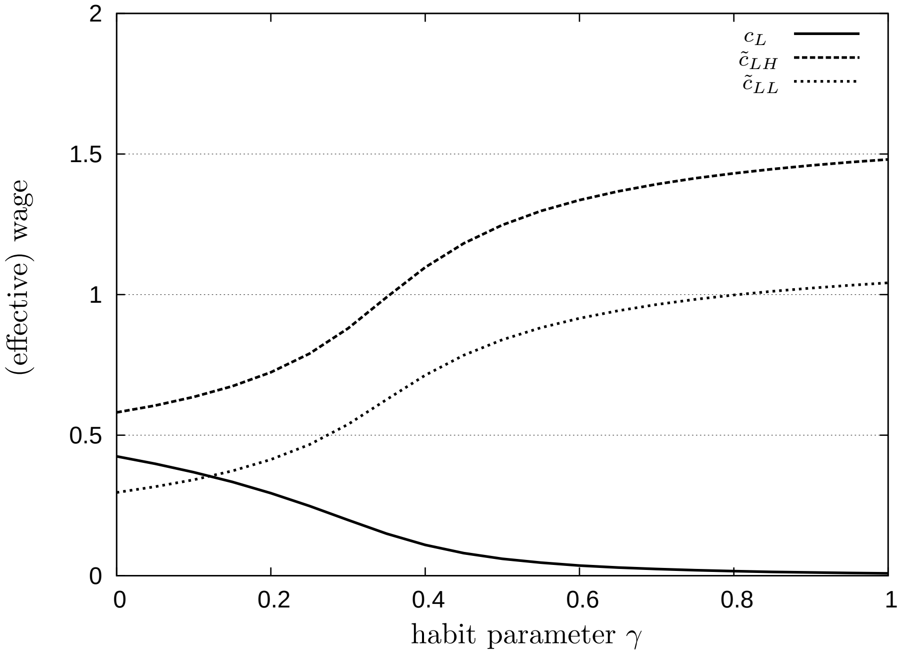

Figure 2.2: (Effective) wages, given low output in period 1

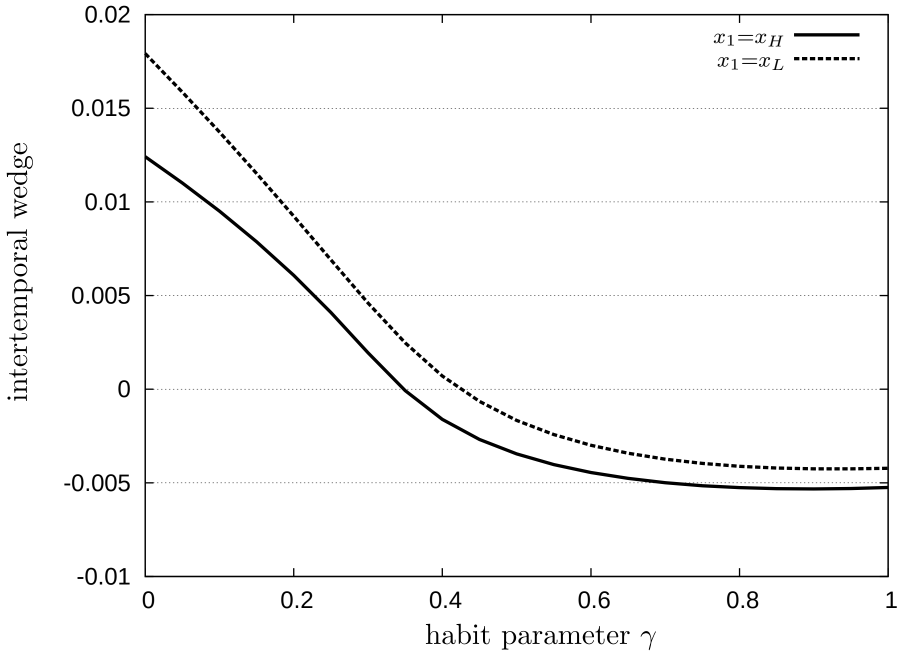

Figure 2.3: Intertemporal wedge for high and low output in period 1. The wedges are negative if the habit parameter is sufficiently large.

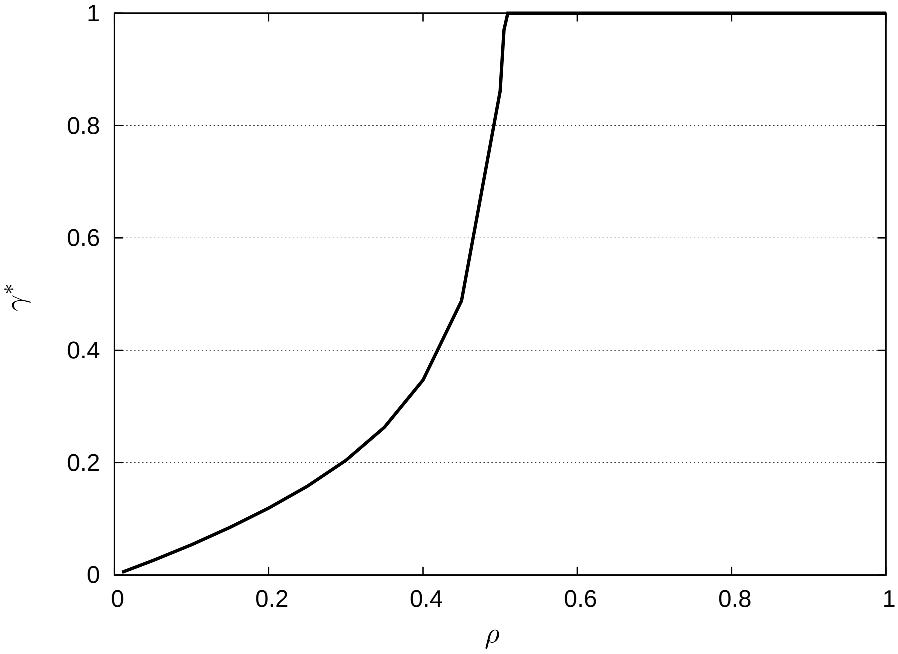

Figure 2.4: Maximum value of the habit parameter for which both intertemporal wedges are positive. This value is increasing in the coefficient of relative risk aversion, ρ.

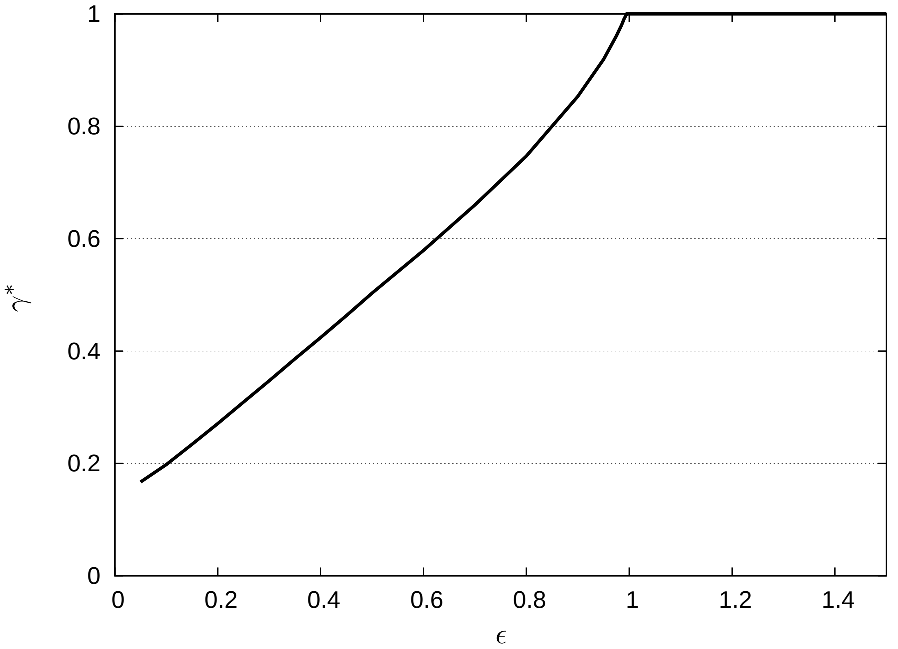

Figure 2.5: Maximum value of the habit parameter for which both intertemporal wedges are positive. This value is increasing in the intercept of the reference level, ϵ.

Observation 2.2. The cutoff value γ^* decreases as ρ gets smaller. Moreover, γ^* approaches zero as ρ goes to zero. (Figure 2.4)

In other words, Figure 2.4 shows that the intertemporal wedge is negative if the habit parameter γ is sufficiently large compared to the coefficient ρ of relative risk aversion. Finally, we note the following.

Observation 2.3. There exist coefficients $\rho < 1$ with $\gamma^* = 1$ (Figure 2.4).

Hence, the highest possible habit parameter ($\gamma = 1$) does not for all risk aversion coefficients $\rho < 1$ lead to negative intertemporal wedges. As Figure 2.5 shows, the set of parameters with negative intertemporal wedges can be increased by decreasing the size of the constant ϵ. This seems due to the fact that ϵ has a dampening effect on changes in the reference level: The larger the size of ϵ, the smaller is the relative change of the reference level $c_1 + \epsilon$ given a reduction of consumption c_1 by one unit.

2.5 Extension to multi-period models

Now I extend the setup introduced in Section 2.2 to 3 periods. It is straightforward to show the respective insights also for models with T periods, with $T > 3$. The choice of $T = 3$, however, allows to maintain much of the previous notation.

2.5.1 Notation

In this section, the relationship between P and A lasts for 3 periods. P maximizes expected profits. A has von-Neumann-Morgenstern preferences and maximizes the expected value of

$$u(c_1) - v(e_1) + \beta(\tilde{u}(c_1, c_2) - v(e_2)) + \beta^2(\hat{u}(c_1, c_2, c_3) - v(e_3)),$$

where c_t denotes consumption, e_t represents effort, and $\beta \in (0,1]$ is A's discount factor. To simplify notation, set

$$U(c_1,c_2,c_3) := u(c_1) + \beta\tilde{u}(c_1,c_2) + \beta^2\hat{u}(c_1,c_2,c_3),$$
$$V(e_1,e_2,e_3) := v(e_1) + \beta v(e_2) + \beta^2 v(e_3).$$

Effort disutility and consumption utility satisfy the usual assumptions: $v(e)$ is strictly increasing and weakly convex in e; consumption utility is strictly increasing and strictly concave in c_1, c_2 and c_3:

$$\forall c_1,c_2,c_3: \quad U_t(c_1,c_2,c_3) > 0, \quad U_{tt}(c_1,c_2,c_3) < 0, \quad t = 1,2,3,$$

where subscripts denote partial derivatives. Moreover, to rule out boundary solutions, I assume

$$\forall c_2,c_3 > 0: \quad \lim_{c_1 \to 0} U_1(c_1,c_2,c_3) = \infty,$$
$$\forall c_1,c_3 > 0: \quad \lim_{c_2 \to 0} U_2(c_1,c_2,c_3) = \infty,$$
$$\forall c_1,c_2 > 0: \quad \lim_{c_3 \to 0} U_3(c_1,c_2,c_3) = \infty.$$

Denote A's **marginal rate of intertemporal substitution** between periods t and s by

$$M_{t,s}(c_1,c_2,c_3) := \frac{U_t(c_1,c_2,c_3)}{U_s(c_1,c_2,c_3)}.$$

In this section, a **contract** (c,e) specifies wages $c = (c_i, c_{ij}, c_{ijk})_{i,j,k}$ and effort levels $e = (e_0, e_i, e_{ij})_{i,j}$. Here, e_{ij} is the recommended effort in period 3 for output history (x_i, x_j), and c_{ijk} is the wage paid in period 3 for output history (x_i, x_j, x_k). The meaning of e_0, e_i and c_i, c_{ij} is as before.

Contract (c, e) is called **optimal** if it maximizes P's expected profits subject to the incentive compatibility constraint (IC) and the participation constraint (PC), i.e., if it solves

$$\max_{c,e} \sum_{i,j,k} \pi_i(e_0)\pi_j(e_i)\pi_k(e_{ij})\Big(x_i - c_i + \frac{1}{R}\left[x_j - c_{ij}\right] + \frac{1}{R^2}\left[x_k - c_{ijk}\right]\Big)$$

s.t.

$$e \in \operatorname*{argmax}_{(e_0', e_i', e_{ij}')} \sum_{i,j,k} \pi_i(e_0')\pi_j(e_i')\pi_k(e_{ij}')\Big(U(c_i, c_{ij}, c_{ijk}) - V(e_0', e_i', e_{ij}')\Big) \tag{IC}$$

$$\sum_{i,j,k} \pi_i(e_0)\pi_j(e_i)\pi_k(e_{ij})\Big(U(c_i, c_{ij}, c_{ijk}) - V(e_0, e_i, e_{ij})\Big) \geq \underline{U}. \tag{PC}$$

2.5.2 Intertemporal optimality

As usual, $E_i[\,\cdot\,]$ and $\text{cov}_i(\,\cdot\,,\,\cdot\,)$ represent expectations and covariances conditional on first-period information x_i. Similarly, $E_{ij}[\,\cdot\,]$ and $\text{cov}_{ij}(\,\cdot\,,\,\cdot\,)$ denote expectations and covariances conditional on second-period information (x_i, x_j).

Analogous to Section 2.3, define the **intertemporal wedge** between periods 1 and 2 as

$$W_i := 1 - \frac{E_i\left[U_1(c_i, c_{ij}, c_{ijk})\right]}{R\, E_i\left[U_2(c_i, c_{ij}, c_{ijk})\right]},$$

and define the **intertemporal wedge** between periods 2 and 3 as

$$W_{ij} := 1 - \frac{E_{ij}\left[U_2(c_i, c_{ij}, c_{ijk})\right]}{R\, E_{ij}\left[U_3(c_i, c_{ij}, c_{ijk})\right]}.$$

Recall that, in the two-period model, the sign of the intertemporal wedge was solely determined by the monotonicity of A's marginal rate of intertemporal substitution; see Proposition 2.3. By following similar steps, we obtain the same result for the intertemporal wedge *between periods 2 and 3* in the present model.

Proposition 2.6. *Let* (c, e) *be an optimal contract and let* (x_i, x_j) *be the output history in period 2. Then* $E_{ij}[M_{2,3}(c_i, c_{ij}, c_{ijk})] = R$*. Moreover, supposing* $c_{ijk} \neq c_{ijk'}$ *for some* k, k'*, the intertemporal wedge between periods 2 and 3 is positive (negative) if* $M_{2,3}(c_1, c_2, c_3)$ *is increasing (decreasing) in* c_3*.*

Proposition 2.6 shows that the intertemporal wedge between periods 2 and 3 is positive whenever A's preferences satisfy the following regularity property: A's willingness to give up future consumption increases in the level of future consumption, i.e., $M_{2,3} = U_2/U_3$ is increasing in c_3. Notice that, for history-*independent* preferences, U_2 does not depend on c_3. Since U_3 decreases in c_3 due to concavity, the regularity property is thus satisfied when preferences are history-independent. Given its sound economic interpretation, the property seems also reasonable when considering history-dependent preferences. Hence, the intertemporal wedge between periods 2 and 3 is positive under rather mild conditions.

The intertemporal wedge between periods 1 and 2 cannot be characterized in the same way, however. First of all, we note the following.

Lemma 2.1. *Let* (c, e) *be an optimal contract. Then we have* $E_i[M_{1,3}(c_i, c_{ij}, c_{ijk})] = R^2$ *for all* i*.*[13]

To gain some intuition, notice that R^2 is P's marginal rate of intertemporal substitution between periods 1 and 3. Intuitively, if the conditionally expected marginal rates of intertemporal substitution differ between P and A, then P can realize gains from readjusting the timing of rewards while maintaining incentive compatibility. Suppose, for instance, that the conditional expectation of A's marginal rate of intertemporal substitution between periods 1 and 3 is higher than R^2. In this case, A's willingness to trade period-3 consumption for period-1 consumption exceeds P's willingness to do so. By

[13] An earlier proof of an analogous result for dynamic Mirrleesian taxation models is provided by Grochulski and Kocherlakota (2010). Instead of the perturbation technique employed in the present paper, they use a Lagrangian approach.

perturbing A's wage scheme as in the above proof (with $\epsilon > 0$), P can extract "gains from trade" without affecting incentive compatibility. At an optimal contract, such gains are impossible. This delivers the result.

Notice, however, that Proposition 2.7 is not informative on A's marginal rate of intertemporal substitution *between periods 1 and 2*. Using the identity $M_{1,2} = M_{1,3}/M_{2,3}$, we have

$$E_i[M_{1,2}] = E_i\left[\frac{M_{1,3}}{M_{2,3}}\right].$$

In general, this expression can be different from

$$\frac{E_i[M_{1,3}]}{E_i[M_{2,3}]} = \frac{E_i[M_{1,3}]}{E_i[E_{ij}[M_{2,3}]]} = \frac{R^2}{R} = R,$$

which suggests that $E_i[M_{1,2}]$ may not be equal to R in general.

As we will see now, $E_i[M_{1,2}]$ can in fact be smaller or larger than R. Moreover, as a direct consequence, the previous link between the monotonicity of the marginal rate of intertemporal substitution and the sign of the intertemporal wedge breaks down.

Proposition 2.7. *Let (c, e) be an optimal contract and let x_i be the output in period 1. Then we have the following.*

(i) $E_i[M_{1,2}(c_i, c_{ij}, c_{ijk})]$ is different from R in general (and can be smaller or larger than R).

(ii) Supposing $c_{ij} \neq c_{ij'}$ for some j, j', the intertemporal wedge between periods 1 and 2 is not positive (negative) in general if $M_{1,2}(c_1, c_2, c_3)$ is increasing (decreasing) in c_2.

The intuition for the first part of Proposition 2.7 is closely related to the above discussion. For periods 1 and 2, it is not necessarily true that P can extract "gains from trade" when the conditionally expected marginal rates of intertemporal substitution differ between P and A. Note that any perturbation of period-1 and period-2 wages will affect A's incentives in the third period if preferences are history-dependent. This effect may offset the efficiency gains from equalizing A's and P's willingness to trade between periods 1 and 2. Hence, a wage scheme may be optimal even when the conditional expectations of A's and P's marginal rates of substitution between periods 1 and 2 do not coincide.

To understand the second part of Proposition 2.7, it is helpful to look into the proof more closely. We find that the intertemporal wedge is positive if and only if

$$\mathrm{cov}_i\left(-U_2,\, M_{1,2}\right) + E_i\left[U_2\right] E_i\left[R - M_{1,2}\right] > 0. \tag{2.18}$$

Suppose for a moment that we have $E_i\left[M_{1,2}\right] = R$, so that the second summand vanishes.[14] In this case, the sign of the intertemporal wedge is fully determined by the sign of the covariance $\mathrm{cov}_i\left(-U_2,\, M_{1,2}\right)$, which is in turn closely related to the monotonicity of $M_{1,2}$ in c_2. For the two-period model, this shows that the intertemporal wedge is positive whenever $M_{1,2}$ is increasing in c_2. Clearly, this insight extends to the last two periods of any multi-period model.

However, as the first part of the proposition points out, $E_i\left[M_{1,2}\right]$ can be different from R in general. Hence, the intertemporal wedge is not solely determined by the aforementioned covariance term. If $E_i\left[M_{1,2}\right]$ is larger than R, then the second summand in (2.18) is negative, hence a positive covariance term does not necessarily imply a positive intertemporal wedge.

It seems rather difficult to conjecture whether, given a specification of A's preferences, $E_i\left[M_{1,2}\right]$ will be larger or smaller than R. The following examples provide concrete setups for each case. Unfortunately, in light of the above discussion, this means that it is generally very difficult to determine the sign of the intertemporal wedge for multi-period models with history-dependent preferences (except for the last two periods).

Example 2.3. (i) In order to study the optimal contract analytically (rather than

[14]Notice that $E_i\left[M_{1,2}\right] = R$ holds true for the two-period model; see Proposition 2.1.

numerically), I suppose there is no incentive problem in the first two periods, i.e., outputs x_i and x_j are non-stochastic. In this case, wages c_i, c_{ij}, c_{ijk} are obviously independent of i, j. Moreover, the conditional expectations $E_i[\,\cdot\,]$ and $E_{ij}[\,\cdot\,]$ coincide.

Let A's consumption preferences be defined as follows:

$$U(c_1, c_2, c_3) := u(c_1) + u(c_2) + u(c_3 - \gamma c_1),$$

where u is a strictly increasing, strictly concave function and $\gamma > 0$.

By Proposition 2.6, the optimal contract satisfies $R = E_{ij}\left[\frac{U_2}{U_3}\right]$. Since U_2 is constant and the conditional expectations $E_i[\,\cdot\,]$ and $E_{ij}[\,\cdot\,]$ coincide, the previous result is equivalent to $\frac{1}{U_2} = \frac{1}{R} E_i\left[\frac{1}{U_3}\right]$, which implies

$$E_i\left[\frac{U_1}{U_2}\right] = \frac{1}{U_2} E_i\left[U_1\right] = \frac{1}{R} E_i\left[\frac{1}{U_3}\right] E_i\left[U_1\right].$$

Rewriting the right-hand side, we see

$$E_i\left[\frac{U_1}{U_2}\right] = \frac{1}{R} E_i\left[\frac{U_1}{U_3}\right] - \frac{1}{R}\,\mathrm{cov}_i\left(U_1,\, \frac{1}{U_3}\right).$$

Recall that Lemma 2.1 implies $E_i\left[\frac{U_1}{U_3}\right] = R^2$. Hence, the first summand on the right-hand side equals R. The second summand is negative, since $\mathrm{cov}_i\left(U_1,\, \frac{1}{U_3}\right)$ is positive as one easily verifies. Summing up, we have

$$E_i\left[\frac{U_1}{U_2}\right] < R. \tag{2.19}$$

Moreover, we can easily see that W_i, the intertemporal wedge between periods 1 and 2, is positive in this example. Notice that W_i is positive if and only if $E_i[U_1] < R\, E_i[U_2]$. Since U_2 is constant, this condition is in fact equivalent to (2.19).

(ii) In the above environment, change A's preferences to

$$U(c_1, c_2, c_3) := u(c_1) + u(c_2) + u(c_3 - \gamma c_2).$$

Using similar arguments as above, it is not difficult to find[15]

$$E_i \left[\frac{U_1}{U_2} \right] > R. \tag{2.20}$$

On a related note, one can see that the intertemporal wedge between periods 1 and 2 is negative in this setup (cp. Grochulski and Kocherlakota 2010, Section 3.2).

2.6 Concluding remarks

This paper introduces history-dependent preferences into the principal-agent framework. The paper shows how to modify the Inverse Euler equation to account for effects of current consumption on future preferences. Moreover, the paper characterizes the intertemporal wedge. While the results are derived in a moral hazard framework, they can be

[15] We first apply Jensen's inequality and obtain

$$E_i \left[\frac{U_1}{U_2} \right] > \frac{1}{E_i \left[\frac{U_2}{U_1} \right]}.$$

Using Lemma 2.1, the fact that U_1 is constant, and the observation that $E_i[\,\cdot\,]$ and $E_{ij}[\,\cdot\,]$ coincide, this is equivalent to

$$E_i \left[\frac{U_1}{U_2} \right] > \frac{R^2}{E_i \left[\frac{1}{U_3} \right] E_i \left[U_2 \right]}.$$

Rewriting the right-hand side, we have

$$E_i \left[\frac{U_1}{U_2} \right] > \frac{R^2}{E_i \left[\frac{U_2}{U_3} \right] - \operatorname{cov}_i \left(U_2, \frac{1}{U_3} \right)}.$$

By Proposition 2.6, the optimal contract satisfies $E_i \left[\frac{U_2}{U_3} \right] = R$. Now the result follows from the observation that $\operatorname{cov}_i \left(U_2, \frac{1}{U_3} \right)$ is positive.

easily extended to contracting problems with adverse selection. In fact, the driving force for all results is the principal's ability to adjust the timing of rewards. The question whether the incentive problem stems from hidden actions or hidden information is thus secondary.

The results in this paper characterize the efficient way of implementing a given effort profile. For this question, the functional form of the agent's effort preferences is obviously irrelevant. Hence, all results go through if we allow effort costs to depend on past levels of effort. A complementary and also very interesting question is what effort profile the principal wants to implement. In that context, history-dependence of effort costs would make an important difference. However, the latter question cannot be addressed in the general environment studied in this paper. First of all, one needs a better understanding of how past effort choices may influence people's current preferences to work hard. This creates an interesting direction for future research.

2.A Appendix: Proofs

Proof of Proposition 2.1. The proof extends the argument in Rogerson (1985b). Let $h(\cdot)$ be the inverse of $u(\cdot)$ and $\tilde{h}(c_1, \cdot)$ be the inverse of $\tilde{u}(c_1, \cdot)$, c_1 fixed. For $\epsilon \in \mathbb{R}$, construct a wage scheme c^ϵ from c as follows. Fix an arbitrary first-period output x_i. For this output, increase date-1 utility by ϵ and reduce date-2 utility by ϵ/β. Formally, set

$$c_k^\epsilon := \begin{cases} c_k & \text{for } k \neq i, \\ h(u(c_i) + \epsilon) & \text{for } k = i, \end{cases}$$

$$c_{kj}^\epsilon := \begin{cases} c_{kj} & \text{for } k \neq i, \\ \tilde{h}(c_i^\epsilon, \tilde{u}(c_i, c_{ij}) - \frac{\epsilon}{\beta}) & \text{for } k = i. \end{cases}$$

Notice that the incentive and participation constraints under c^ϵ and c coincide. Hence, a necessary condition for optimality of (c, e) is that expected wage payments are minimal at $\epsilon = 0$. That is, $\epsilon = 0$ must minimize

$$h(u(c_i) + \epsilon) + \frac{1}{R} E_i \left[\tilde{h}(c_i^\epsilon, \tilde{u}(c_i, c_{ij}) - \frac{\epsilon}{\beta}) \right],$$

which implies the first-order condition

$$0 = \frac{1}{u'(c_i)} + \frac{1}{R} E_i \left[-\frac{1}{\beta\, \tilde{u}_2(c_i, c_{ij})} - \frac{\tilde{u}_1(c_i, c_{ij})}{u'(c_i)\, \tilde{u}_2(c_i, c_{ij})} \right].$$

Equivalently,

$$\frac{1}{\beta R} E_i \left[\frac{1}{\tilde{u}_2(c_i, c_{ij})} \right] = \frac{1}{u'(c_i)} \left(1 - \frac{1}{R} E_i \left[\frac{\tilde{u}_1(c_i, c_{ij})}{\tilde{u}_2(c_i, c_{ij})} \right] \right).$$

□

Proof of Proposition 2.2. Solve the first-order condition (2.1) for $u'(c_i)$ to get

$$u'(c_i) = \beta \frac{R - E_i \left[\frac{\tilde{u}_1(c_i, c_{ij})}{\tilde{u}_2(c_i, c_{ij})} \right]}{E_i \left[\frac{1}{\tilde{u}_2(c_i, c_{ij})} \right]}.$$

Using

$$E_i \left[\frac{\tilde{u}_1(c_i, c_{ij})}{\tilde{u}_2(c_i, c_{ij})} \right] = \mathrm{cov}_i \left(\tilde{u}_1(c_i, c_{ij}),\ \frac{1}{\tilde{u}_2(c_i, c_{ij})} \right) + E_i \left[\tilde{u}_1(c_i, c_{ij}) \right] E_i \left[\frac{1}{\tilde{u}_2(c_i, c_{ij})} \right],$$

we find

$$u'(c_i) = \beta \frac{R - \mathrm{cov}_i \left(\tilde{u}_1(c_i, c_{ij}),\ \frac{1}{\tilde{u}_2(c_i, c_{ij})} \right)}{E_i \left[\frac{1}{\tilde{u}_2(c_i, c_{ij})} \right]} - \beta E_i[\tilde{u}_1(c_i, c_{ij})].$$

Substitute this into definition (2.4) to obtain

$$W_i = 1 - \frac{R - \mathrm{cov}_i\left(\tilde{u}_1(c_i, c_{ij}),\ \frac{1}{\tilde{u}_2(c_i,c_{ij})}\right)}{R\, E_i\left[\tilde{u}_2(c_i, c_{ij})\right] E_i\left[\frac{1}{\tilde{u}_2(c_i,c_{ij})}\right]}.$$

Now the result follows from

$$E_i[\tilde{u}_2(c_i, c_{ij})]E_i\left[\frac{1}{\tilde{u}_2(c_i, c_{ij})}\right] = 1 - \mathrm{cov}_i\left(\tilde{u}_2(c_i, c_{ij}),\ \frac{1}{\tilde{u}_2(c_i, c_{ij})}\right).$$

□

Proof of Proposition 2.3. Suppose $M(c_1, c_2)$ is increasing in c_2. By definition, the intertemporal wedge is positive if and only if

$$R\beta E_i\left[\tilde{u}_2(c_i, c_{ij})\right] - \left(u'(c_i) + \beta E_i\left[\tilde{u}_1(c_i, c_{ij})\right]\right) > 0.$$

Dividing by β, this condition can be rewritten as

$$E_i\left[\tilde{u}_2(c_i, c_{ij})\left(R - M(c_i, c_{ij})\right)\right] > 0.$$

It is not difficult to see that Proposition 2.1 yields $R - E_i[M(c_i, c_{ij})] = 0$, compare equation (2.3). Hence, the above inequality is equivalent to

$$\mathrm{cov}_i\left(\tilde{u}_2(c_i, c_{ij}),\ R - M(c_i, c_{ij})\right) > 0.$$

Since R is a constant, this holds if and only if

$$\mathrm{cov}_i\left(-\tilde{u}_2(c_i, c_{ij}),\ M(c_i, c_{ij})\right) > 0. \tag{2.21}$$

Since $\tilde{u}(c_i, c_{ij})$ is concave in c_{ij}, the partial derivative $-\tilde{u}_2(c_i, c_{ij})$ is increasing in c_{ij}.

By assumption, $M(c_i, c_{ij})$ is also increasing in c_{ij}. Notice that the covariance of two increasing functions of a random variable is positive. Hence, condition (2.21) is satisfied.

□

Proof of Corollary 2.1. Note $M(c_1, c_2) = \frac{u'(c_1)}{\beta u'(c_2 - \gamma c_1)} - \gamma$. This shows that $M(c_1, c_2)$ is increasing in c_2. Hence, the intertemporal wedge is positive by Proposition 2.3. □

Proof of Proposition 2.4. Let $i = H$. (The case $i = L$ is exactly analogous.) Applying the implicit function theorem to the system of equations (2.9),(2.11),(2.13) yields

$$\begin{aligned}
\frac{\partial c_H}{\partial \gamma} &= -\frac{\beta^2}{D} u'(c_H) u'(\tilde{c}_{HL})^3 u'(\tilde{c}_{HH})^3 \\
\frac{\partial c_{HH}}{\partial \gamma} &= c_H + \frac{\beta}{D} u'(c_H) u'(\tilde{c}_{HL})^3 u'(\tilde{c}_{HH})^2 [u'(c_H) - \gamma \beta u'(\tilde{c}_{HH})] \\
\frac{\partial c_{HL}}{\partial \gamma} &= c_H + \frac{\beta}{D} u'(c_H) u'(\tilde{c}_{HL})^2 u'(\tilde{c}_{HH})^3 [u'(c_H) - \gamma \beta u'(\tilde{c}_{HL})],
\end{aligned}$$

where

$$\begin{aligned}
D = &-(1 - p_H(h)) u'(c_H)^3 u'(\tilde{c}_{HH})^3 u''(\tilde{c}_{HL}) \\
&- u'(\tilde{c}_{HL})^3 [(R + \gamma) \beta^2 u'(\tilde{c}_{HH})^3 u''(c_H) + p_H(h) u'(c_H)^3 u''(\tilde{c}_{HH})] \\
&> 0.
\end{aligned}$$

Hence, the expression $\frac{\partial c_{HH}}{\partial \gamma} c_H - c_{HH} \frac{\partial c_H}{\partial \gamma}$ is equal to

$$c_H^2 + \frac{\beta}{D} u'(c_H) u'(\tilde{c}_{HL})^3 u'(\tilde{c}_{HH})^2 \left[c_H u'(c_H) + \beta \tilde{c}_{HH} u'(\tilde{c}_{HH}) \right],$$

which is positive. This shows that the ratio c_{HH}/c_H is increasing in γ. Monotonicity of c_{HL}/c_H can be seen analogously. □

Proof of Proposition 2.5. (i) See Example 2.2 on page 24.

(ii) The assumption $-\tilde{c}_2 \frac{u''(\tilde{c}_2)}{u'(\tilde{c}_2)} \geq 1$ is equivalent to

$$-u'(c_2(c_1+\epsilon)^{-\gamma}) - c_2(c_1+\epsilon)^{-\gamma} u''(c_2(c_1+\epsilon)^{-\gamma}) \geq 0.$$

Multiplying this by $\gamma(c_1+\epsilon)^{-1-\gamma}$, we have

$$-\gamma(c_1+\epsilon)^{-1-\gamma} u'(c_2(c_1+\epsilon)^{-\gamma}) - \gamma c_2(c_1+\epsilon)^{-1-2\gamma} u''(c_2(c_1+\epsilon)^{-\gamma}) \geq 0,$$

which is equivalent to $\tilde{u}_{12} \geq 0$. Thus, $u'(c_1) + \beta\tilde{u}_1(c_1, c_2)$ is increasing in c_2. Since $\tilde{u}_2(c_1, c_2)$ is decreasing in c_2, we conclude that $M(c_1, c_2)$ is increasing in c_2. Hence, the intertemporal wedge is positive by Proposition 2.3. □

Proof of Proposition 2.6. To verify the first statement, consider the wage scheme c^ϵ that is constructed as follows. Leave period-1 wages unchanged: $c_i^\epsilon := c_i$. Define period-2 and period-3 wages by:

$$\begin{aligned}
\tilde{u}(c_i, c_{ij}^\epsilon) &:= \tilde{u}(c_i, c_{ij}) + \epsilon, \\
\hat{u}(c_i, c_{ij}^\epsilon, c_{ijk}^\epsilon) &:= \hat{u}(c_i, c_{ij}, c_{ijk}) - \frac{\epsilon}{\beta}.
\end{aligned}$$

Now $E_{ij}[M_{2,3}(c_i, c_{ij}, c_{ijk})] = R$ follows from the same steps as the proof of Proposition 2.1. To prove the second statement, use $E_{ij}[M_{2,3}(c_i, c_{ij}, c_{ijk})] = R$ and proceed as in the proof of Proposition 2.3. □

Proof of Lemma 2.1. Consider the wage scheme c^ϵ that is constructed as follows. Leave period-2 wages unchanged: $c_{ij}^\epsilon := c_{ij}$. Define period-1 and period-3 wages by:

$$\begin{aligned}
&u(c_i^\epsilon) := u(c_i) + \epsilon, \\
&\hat{u}(c_i^\epsilon, c_{ij}, c_{ijk}^\epsilon) := \hat{u}(c_i, c_{ij}, c_{ijk}) - \frac{\epsilon}{\beta^2} + \frac{1}{\beta}\left(\tilde{u}(c_i, c_{ij}) - \tilde{u}(c_i^\epsilon, c_{ij})\right).
\end{aligned}$$

Now the result follows from the same steps as the proof of Proposition 2.1. □

Proof of Proposition 2.7. To verify the first statement, see Example 2.3. For the second statement, note that the intertemporal wedge is positive if and only if

$$E_i\left[R\,U_2(c_i, c_{ij}, c_{ijk}) - U_1(c_i, c_{ij}, c_{ijk})\right] > 0,$$

which can be rewritten as

$$E_i\left[U_2(c_i, c_{ij}, c_{ijk})\left(R - M_{1,2}(c_i, c_{ij}, c_{ijk})\right)\right] > 0.$$

Omitting the argument (c_i, c_{ij}, c_{ijk}), this inequality holds if and only if

$$\operatorname{cov}_i\left(U_2,\ R - M_{1,2}\right) + E_i\left[U_2\right] E_i\left[R - M_{1,2}\right] > 0.$$

Since R is a constant, the previous condition is equivalent to

$$\operatorname{cov}_i\left(-U_2,\ M_{1,2}\right) + E_i\left[U_2\right]\left(R - E_i\left[M_{1,2}\right]\right) > 0. \tag{2.22}$$

As the first part of the proposition shows, $R - E_i\left[M_{1,2}\right]$ can be positive or negative in general. Consequently, it is not difficult to construct examples where the left-hand side in (2.22) is negative (positive) even though $\operatorname{cov}_i\left(-U_2,\ M_{1,2}\right)$ is positive (negative); see Example 2.3 (ii), for instance. □

Chapter 3

The first-order approach to moral hazard problems with hidden saving

3.1 Introduction

The study of moral hazard models is enormously simplified if one can use the first-order approach (Mirrlees 1974, Holmström 1979). By replacing the incentive constraint with the associated first-order condition, this approach allows the application of Lagrangian methods. The seminal works of Rogerson (1985a) and Jewitt (1988) validate this procedure for the standard moral hazard problem. Very little is known, however, for more general environments. In particular, the validity of the first-order approach is not well understood for moral hazard problems in which the agent can secretly save (and borrow). This class of problems is particularly important, since observability of the agent's consumption-saving decision appears unrealistic for many common dynamic applications of the moral hazard framework (e.g., employment relationships, insurance problems, income taxation, etc.).

With hidden saving, validating the first-order approach becomes significantly more complex. In addition to making sure that the agent's utility is at a global maximum with respect to the effort decision, one has to ensure the same for the saving decision, and most importantly for *joint* deviations to different effort and saving levels. Typically, the agent would combine a reduction of effort with an increased savings level to insure against the worsened output distribution. Therefore, ruling out joint deviations is the main difficulty in showing that first-order conditions imply incentive compatibility. The assumptions made by Rogerson (1985a) and Jewitt (1988) are not sufficient for such a result, since they apply to the effort dimension only. In fact, it is not difficult to find examples where the first-order approach to the model with hidden saving fails even though Rogerson's or Jewitt's conditions are satisfied (Kocherlakota 2004). Due to the two-dimensional decision space and the complementarity between shirking and saving, there is no obvious way of strengthening those conditions. Therefore, the validity of the first-order approach for this setup has remained unclear.

The present paper validates the first-order approach for two-period moral hazard models with hidden saving. I show that the first-order approach is valid if the agent has non-increasing absolute risk aversion (NIARA) utility, the output technology has monotone likelihood ratios (MLR), and the distribution function of output is log-convex in effort (LCDF). Note that a function is called *log-convex* if the logarithm of that function is convex. Any log-convex function is convex, but not vice versa. Hence, compared to Rogerson's (1985a) assumption that the distribution function is convex in effort, LCDF is a stricter requirement. Under LCDF, the probability that output exceeds a given level is strongly concave in the agent's effort choice. Intuitively, this states that the marginal returns to effort are strongly decreasing in a particular sense.

The link from these conditions to the second-order effects of joint deviations is subtle. Note that by reducing his effort, the agent increases the probability of being punished by a low wage. By increasing his saving at the same time, he alleviates the severity of the punishment, since the utility difference between high and low wages will be reduced. Decreasing marginal returns to effort and a convex marginal utility of consumption limit the gains of such a strategy to some extent. The former implies that the probability of being punished increases more quickly than linearly as effort is reduced; the latter implies that the reduction of the punishment diminishes more quickly than linearly as saving is increased. Since the two effects are *multiplicative*, however, those properties are still too weak to ensure that a joint deviation is not attractive. At this point, the concept of log-convexity proves useful. In contrast to convexity, log-convexity is not only preserved under summation, but also under multiplication. Therefore, log-convexity of the agent's marginal utility of consumption (which is equivalent to NIARA) in combination with log-convexity of the distribution function (LCDF) implies that the 'punishment' described above is *jointly* log-convex (and hence convex) in effort and saving. This makes the agent's optimization problem jointly concave in his decision variables. Thus, first-order conditions yield incentive compatibility and the first-order approach is valid.

I also derive alternative sufficient conditions for the validity of the first-order approach. An important insight of the previous argument is the trade-off between convexity assumptions on the marginal utility of consumption on the one hand and convexity assumptions on the distribution function on the other hand. If one of the two functions is more convex than log-linear, then the assumption on the other function can be weakened. For a large class of utility functions, including CRRA utility, for example, this allows a relaxation of the LCDF condition. Finally, I show how to relax the LCDF condition by exploiting the curvature of the wage scheme. This allows me to validate the first-order approach for some interesting examples in which the LCDF property and even Rogerson's (1985a) CDF condition fail. From a more general point of view, however, the curvature of the contract is not as helpful as in the standard moral hazard problem, since wages tend to become less concave in output under hidden saving.

The present paper is the first to identify the convexity properties of the distribution function and the agent's marginal utility of consumption as the key driving forces for the validity of the first-order approach under hidden saving. Another important contribution is the introduction of log-convexity techniques. This idea gives optimization problems with multiplicatively separable objectives a tractable convex structure and seems to be useful for a much more general class of economic models.[1]

To the best of my knowledge, the only other result on the validity of the first-order approach under hidden saving is the work by Abraham and Pavoni (2009). Like the present paper, they also study a two-period model. However, they impose the 'spanning condition with dominance' from Grossman and Hart (1983), which is a severe restriction to the output technology. In fact, the condition imposes so much structure that moral hazard problems can be characterized quite well even without the first-order approach (Grossman and Hart 1983). Hence, assuming the spanning condition, it appears more natural to look for an extension of the method described in Grossman and Hart (1983), rather than for the validity of the first-order approach. The present paper does not need the spanning condition. The paper obtains the result by Abraham and Pavoni (2009) as a restrictive special case and goes much further.[2]

Establishing the validity of the first-order approach under hidden saving is valuable for a number of reasons. First, it yields a precise characterization of optimal contracts. Questions on the monotonicity of consumption or the value of information can be answered immediately, and one finds many analogies to the model without hidden saving. One also finds an important difference between the two models: Optimal wage schemes tend to be more convex in output under hidden saving. For a detailed discussion of this result, I refer the reader to Chapter 4 of this thesis. Secondly, the first-order approach under

[1] I am not aware of any other work that highlights this idea.

[2] Example 3.4 in Section 3.3 shows that the distribution function is log-convex in effort given the assumptions made by Abraham and Pavoni (2009). Moreover, Section 3.4 contains important relaxations of the log-convexity assumption.

hidden saving is helpful for a large range of applied questions. For instance, Bertola and Koeniger (2009) use this approach to develop a theoretical model on cross-country differences between public and private insurance. Gottardi and Pavoni (2010) build on the first-order approach to address optimal capital taxation. Chade (2009) uses the first-order approach to study efficient compensation contracts. Finally, the first-order approach is important because it gives multi-period moral hazard problems with hidden saving a tractable recursive form, as discussed by Werning (2001), Werning (2002), Kocherlakota (2004), and others. Analytical results for the validity of the first-order approach provide a theoretical foundation for this procedure. The present paper marks an important first step towards this goal. However, the extension from the present two-period results to the multi-period problem remains a task for future research.[3]

The paper proceeds as follows: Section 3.2 describes the setup of the model. Section 3.3 validates the first-order approach given NIARA, MLR and LCDF. Section 3.4 shows how to relax the latter assumption. Section 3.5 concludes. All proofs are presented in the appendix.

3.2 Model

I study a two-period principal-agent problem. In the first period, the agent makes a hidden saving decision. In the second period, the agent exerts a hidden work effort. The contract is signed at the beginning of the first period and there is no renegotiation.

[3]For the multi-period problem, it will be crucial to understand how the *value function* changes with the agent's savings level. A characterization of the value function is beyond the scope of the present paper, however. Note that conventional macroeconomic techniques are of limited help here, because the model involves a hidden state in combination with an *endogenously* determined probability distribution.

3.2.1 Preferences

The Principal (P) maximizes expected profits. P's discount factor is $1/R$, with $R > 0$. The Agent (A) has von-Neumann-Morgenstern preferences and maximizes the expected value of

$$u(c_1) + \beta\left(u(c_2) - v(e)\right),$$

where c_t denotes consumption and e represents effort. Consumption utility u is twice continuously differentiable and satisfies $u' > 0$, $u'' < 0$. Effort disutility v is twice continuously differentiable and satisfies $v' > 0$, $v'' \geq 0$.

3.2.2 Technology

In the first period, A is endowed with w_0 units of the consumption good and can save at the rate $R > 0$. Negative saving, i.e., borrowing, is allowed. The set of feasible saving choices s is the real interval J, which may be bounded or unbounded.[4] A's saving decision is not observable.

In the second period, A exerts an unobservable work effort $e \in I$, where I is a real interval. This generates a publicly observable stochastic output $x \in [\underline{x}, \overline{x}]$. (All results go through for discrete output spaces as well.) The output is distributed according to the probability density $f(x, e)$, which is continuously differentiable and has full support for all $e \in I$.

3.2.3 Contracts

At the beginning of the first period, P proposes a **contract** $(w(\cdot), e, s)$ consisting of an output-contingent wage scheme $w(\cdot)$ and recommended choices (e, s). A's utility from

[4] The interval J may be bounded below due to a borrowing constraint and bounded above due to a nonnegativity constraint.

rejecting the contract is $\underline{U}$. The contract is called **optimal** if it maximizes expected profits subject to the incentive compatibility constraint and the participation constraint, i.e., if it solves the following problem:

$$\max_{w(\cdot),e,s} \frac{1}{R}\int_{\underline{x}}^{\overline{x}} (x - w(x))f(x,e)\,dx \tag{P1}$$

s.t.

$$(e,s) \in \operatorname*{argmax}_{(e',s')\in I\times J} u(w_0 - \frac{s'}{R}) + \beta\int_{\underline{x}}^{\overline{x}} u(w(x)+s')f(x,e')\,dx - \beta v(e') \tag{IC}$$

$$u(w_0 - \frac{s}{R}) + \beta\int_{\underline{x}}^{\overline{x}} u(w(x)+s)f(x,e)\,dx - \beta v(e) \geq \underline{U} \tag{PC}$$

3.2.4 First-order approach

Problem (P1) is extremely intricate. The incentive constraint (IC) consists of a two-dimensional continuum of inequalities. For all $e' \in I, s' \in J$, it requires

$$\begin{aligned} &u(w_0 - \frac{s}{R}) + \beta\int_{\underline{x}}^{\overline{x}} u(w(x)+s)f(x,e)\,dx - \beta v(e) \\ \geq\ &u(w_0 - \frac{s'}{R}) + \beta\int_{\underline{x}}^{\overline{x}} u(w(x)+s')f(x,e')\,dx - \beta v(e'). \end{aligned} \tag{3.1}$$

To obtain a problem that can be solved by standard methods, one replaces the incentive constraint by the agent's first-order necessary conditions. This gives rise to the following problem:

$$\max_{w(\cdot),e,s} \frac{1}{R}\int_{\underline{x}}^{\overline{x}} (x - w(x))f(x,e)\,dx \tag{P2}$$

s.t.

$$\beta \int_{\underline{x}}^{\overline{x}} u(w(x)+s)f_e(x,e)\,dx - \beta v'(e) = 0 \tag{FOCe}$$

$$\frac{1}{R}u'(w_0 - \frac{s}{R}) - \beta \int_{\underline{x}}^{\overline{x}} u'(w(x)+s)f(x,e)\,dx = 0 \tag{FOCs}$$

$$u(w_0 - \frac{s}{R}) + \beta \int_{\underline{x}}^{\overline{x}} u(w(x)+s)f(x,e)\,dx - \beta v(e) \geq \underline{U} \tag{PC}$$

Solutions to (P2) are denoted by $(w^*(\cdot), e^*, s^*)$. The associated consumption levels are denoted by $c_0^* = w_0 - s^*/R$ and $c^*(x) = w^*(x) + s^*$.

Replacing the true problem (P1) by the first-order problem (P2) is a valid procedure only if their solutions coincide. Assuming that the solutions to (P1) are interior with respect to effort and saving, this will be the case if and only if the contracts solving (P2) are incentive compatible. A sufficient condition for incentive compatibility is that the agent's decision problem is concave at those contracts. The remainder of this paper will identify conditions under which this is the case.

3.3 A sufficient condition for concavity of the agent's problem

In this section, I validate the first-order approach using nonincreasing absolute risk aversion, monotonicity of the wage scheme, and an assumption on the curvature of the output distribution function. This procedure strengthens the classic approach of Mirrlees (1979) and Rogerson (1985a).

Using λ, μ and ξ as the Lagrange multipliers associated with the constraints (PC), (FOCe), (FOCs), respectively, the first-order condition of the Lagrangian of problem (P2) with respect to wages, for $x \in [\underline{x}, \overline{x}]$, is

$$0 = -\frac{1}{R} f(x, e^*) + \mu\beta u'(c^*(x)) f_e(x, e^*) - \xi\beta u''(c^*(x)) f(x, e^*) + \lambda\beta u'(c^*(x)) f(x, e^*). \tag{3.2}$$

Equivalently, as shown by Abraham and Pavoni (2009),

$$\frac{1}{R\beta u'(c^*(x))} = \lambda + \mu \frac{f_e(x, e^*)}{f(x, e^*)} + \xi a(c^*(x)), \quad x \in [\underline{x}, \overline{x}], \tag{3.3}$$

where $a(c) = -u''(c)/u'(c)$ is A's coefficient of absolute risk-aversion.

Expression (3.3) equates the principal's costs and benefits of marginally increasing the agent's utility at output x, normalized by the probability density. Compared to the standard moral hazard problem, there is now the additional term $\xi a(c^*(x))$, because an increase of $u(c^*(x))$ relaxes the agent's Euler equation.[5]

I will often use the following two assumptions to give equation (3.3) more structure.

MLR. The likelihood ratio function, $f_e(x, e)/f(x, e)$, is continuously differentiable and nondecreasing in output x for all effort levels e.

NIARA. The agent's coefficient of absolute risk aversion, $a(c) = -u''(c)/u'(c)$, is continuously differentiable and nonincreasing in consumption c.

MLR is standard and simply means that more output is indicative of higher effort. NIARA is also unproblematic, since it is satisfied by most common utility functions.

[5]Note that an increase of $\beta u(c^*(x))$ by one marginal unit costs the principal $1/(R\beta u'(c^*(x)))$ units of consumption. On the other hand, it generates a benefit of λ because the participation constraint is relaxed and a benefit (or cost) of $\mu f_e/f$ because the first-order incentive constraint is relaxed (or tightened). In addition, there is a benefit of $\xi a(c^*(x))$ because an increase of $\beta u(c^*(x))$ mitigates the agent's wish to save (Abraham and Pavoni 2009).

NIARA implies that the multipliers λ, μ, ξ in the Kuhn-Tucker condition (3.3) are positive: $\lambda > 0$, $\mu > 0$, $\xi > 0$ (Abraham and Pavoni 2009). Moreover, MLR plus NIARA is sufficient for the wage scheme $w^*(x)$, with $w^*(x) = c^*(x) - s^*$, to be continuously differentiable and nondecreasing in output x; see equation (3.3).[6]

As noted before, the first-order approach is valid if A's objective function

$$(e, s) \mapsto u(w_0 - \frac{s}{R}) + \beta \int_{\underline{x}}^{\overline{x}} u(w^*(x) + s) f(x, e)\, dx - \beta v(e) \tag{3.4}$$

is concave in (e, s) for the wage scheme $w^*(\cdot)$ that solves (P2). One can restrict attention to A's second-period consumption utility as the next result shows.

Lemma 3.1. *A's decision problem is concave in (e, s) if A's second-period consumption utility*

$$(e, s) \mapsto \int_{\underline{x}}^{\overline{x}} u(w^*(x) + s) f(x, e)\, dx \tag{3.5}$$

is concave in (e, s).

By focusing on A's second-period consumption utility, I ignore the curvature generated by the effort disutility function v and by the effect of saving on first-period utility. In principle, one could obtain more general results by including these two effects. It is not apparent, however, how far this would relax the curvature requirement imposed on the agent's second-period utility. Note, besides, that the role of the effort disutility function is limited anyway, since effort units can always be normalized such that this function is linear.

The following lemma identifies a sufficient condition for concavity of (3.5).

[6] NIARA can be relaxed. Equation (3.3) implies that $w^*(\cdot)$ is nondecreasing under MLR if $-(u'''u' - (u'')^2) \leq -u''(R\beta\xi)^{-1}$. This requires that the coefficient of absolute risk aversion does not increase too quickly.

Lemma 3.2. *Suppose $w^*(\cdot)$ is continuously differentiable and nondecreasing. Suppose the distribution function of output, $F(x,e) = \int_{\underline{x}}^{x} f(z,e)\,dz$, is convex in e and for all $x \in [\underline{x}, \overline{x}]$, $e \in I$, $s \in J$, we have*

$$\frac{F_{ee}(x,e)F(x,e)}{(F_e(x,e))^2}\frac{u'''(w^*(x)+s)u'(w^*(x)+s)}{(u''(w^*(x)+s))^2} \geq 1. \tag{3.6}$$

Then A's second-period consumption utility is concave in (e,s).

To understand condition (3.6), note that $F_{ee}F/(F_e)^2$ is nonnegative if and only if F is convex in e, and at least 1 if and only F is log-convex in e.[7] Hence, $F_{ee}F/(F_e)^2$ measures the convexity of the distribution function F as a function effort. This motivates the following concept.

LCDF. The distribution function of output, $F(x,e) = \int_{\underline{x}}^{x} f(z,e)\,dz$, is log-convex in effort e for all output levels x.

A necessary but not sufficient condition for LCDF is that the distribution function is convex in effort. Hence, LCDF tightens the CDF condition from Mirrlees (1979) and Rogerson (1985a). To interpret LCDF, note that $F(x',e)$ equals $1 - P(x > x'|e)$. Therefore, stating that $F(x',e)$ is log-convex in effort implies that the probability $P(x > x'|e)$ is 'strongly' concave in effort. For this reason, LCDF requires that the (stochastic) returns to effort are strongly decreasing: The probability $P(x > x'|e)$ that output is larger than some level x' is strongly concave in the agent's effort choice e for all values of x'.

Analogous to the interpretation of $F_{ee}F/(F_e)^2$, note that $u'''u'/(u'')^2$ is a measure of convexity of A's marginal utility of consumption. This measure is nonnegative if and only if u' is convex, and at least 1 if and only if u' is log-convex. Log-convexity of u' is

[7] A function $f : \mathbb{R} \to \mathbb{R}_{++}$ is called *log-convex* if the logarithm of that function is convex. Assuming differentiability, f is log-convex if and only if $f''f/(f')^2 \geq 1$. Any log-convex function is convex, but not vice versa.

equivalent to

$$\frac{u'''u' - (u'')^2}{(u')^2} \geq 0. \tag{3.7}$$

This is the case if and only if

$$\frac{d}{dc}\left(-\frac{u''(c)}{u'(c)}\right) \leq 0. \tag{3.8}$$

Hence, log-convexity of u' is equivalent to NIARA.

The main result is now a direct consequence of these observations: MLR, NIARA and LCDF validate the first-order approach.

Theorem 3.1. *Let* $(w^*(\cdot), e^*, s^*)$ *be a solution to the first-order problem* (P2)*. Suppose MLR, NIARA and LCDF. Then, given this contract, the agent's decision problem is concave. Hence, the contract solves the original problem* (P1)*.*

Compared to the model without hidden saving, Theorem 3.1 additionally requires NIARA and LCDF instead of Rogerson's (1985a) CDF condition. NIARA is unproblematic, because it is satisfied by most common utility functions and confirmed by many empirical and experimental studies. As argued above, LCDF states that the returns to effort are strongly decreasing in a particular sense. The following examples, in particular the case with only two possible output levels, clarify this property.

Example 3.1 (Rogerson 1985)**.** Rogerson's paper contains the following distribution function that is convex in effort and satisfies MLR:

$$F(x, e) = \left(\frac{x}{\overline{x}}\right)^{e-\underline{e}}, \quad x \in [0, \overline{x}],\ e \in (\underline{e}, \infty). \tag{3.9}$$

This distribution function is not only convex in e, but even satisfies LCDF. Note

$$\log(F(x, e)) = (e - \underline{e}) \log\left(\frac{x}{\overline{x}}\right), \tag{3.10}$$

which shows that $F(x, e)$ is log-linear in e for all i.

Example 3.2 (Log-logistic distribution)**.** Let $0 < b \leq 1$. Consider the following distribution function:

$$F(x,e) = \frac{1}{1+(e/x)^b}, \quad x \in [0,\infty),\ e \in (0,\infty). \tag{3.11}$$

It is not difficult to see that MLR is satisfied. Moreover, note

$$\log(F(x,e)) = -\log\left(1+(e/x)^b\right). \tag{3.12}$$

Since $b \leq 1$, the expression $(e/x)^b$ is concave in e. Since the logarithm is increasing and concave, equation (3.12) shows that $\log(F(x,e))$ is convex in e. Thus, LCDF is satisfied.

The next two examples apply to discrete output spaces $X = \{x_1, \ldots, x_n\}$, $x_i < x_j$ for $i < j$. In this setup, wages are vectors $(w_1, \ldots, w_n) \in \mathbb{R}^n$, and probability weights $(p_1(e), \ldots, p_n(e))$ replace the density function $f(x,e)$. The previous results extend to the discrete setup without difficulty.

Example 3.3 (Two outputs)**.** Consider the case with two possible outputs, $x_L < x_H$, and associated probabilities $p_L(e) = 1 - p(e)$, $p_H(e) = p(e)$, for some increasing function p with $0 \leq p(e) \leq 1$. Since p is increasing, MLR is satisfied. LCDF is equivalent to the log-convexity of $1 - p(e)$, which holds if and only if

$$\frac{-p''(e)(1-p(e))}{(p'(e))^2} \geq 1 \text{ for all } e \in I. \tag{3.13}$$

To yield LCDF, the probability $p(e)$ of the high output level thus has to be sufficiently concave in effort. One example that satisfies this condition is the function $p(e) = 1 - \exp(-f(e))$, where $f : I \to (0,\infty)$ is increasing and concave.

Example 3.4 (Spanning condition)**.** Let $(\pi_{1h}, \ldots, \pi_{nh})$, $(\pi_{1l}, \ldots, \pi_{nl})$ be two probability distributions on $\{x_1, \ldots, x_n\}$ such that π_{ih}/π_{il} is nondecreasing in i. (This implies that

π_h first-order stochastically dominates π_l.) Let

$$p_i(e) = \Gamma(e)\pi_{ih} + (1 - \Gamma(e))\pi_{il} \tag{3.14}$$

for some increasing function Γ, with $0 \leq \Gamma(e) \leq 1$. Monotonicity of Γ, combined with the fact that π_{ih}/π_{il} is nondecreasing, yields MLR. Note

$$F_i(e) = F(x_i, e) = \sum_{j=1}^{i} p_j(e) = (1 - \Gamma(e)) \sum_{j=1}^{i} (\pi_{il} - \pi_{ih}) + \sum_{j=1}^{i} \pi_{ih}. \tag{3.15}$$

First-order stochastic dominance implies $\sum_{j=1}^{i}(\pi_{il} - \pi_{ih}) \geq 0$. Therefore, LCDF holds if $1 - \Gamma(e)$ is log-convex. This requirement is equivalent to

$$\frac{(\Gamma'(e))^2}{-\Gamma''(e)(1 - \Gamma(e))} \leq 1, \tag{3.16}$$

which is exactly the condition under which Abraham and Pavoni (2009) validate the first-order approach for the spanning condition and NIARA utility. Their proof relies heavily on the spanning condition and there is no obvious way how it might generalize to the setting considered in this paper. Moreover, Abraham and Pavoni's reading of the property in (3.16) is that the Frisch elasticity of leisure must not be larger than one (Abraham and Pavoni 2009, p. 16). This does not capture the true sense in which (3.16) tightens the CDF condition from Mirrlees (1979) and Rogerson (1985a), in contrast to the argument provided in the present paper.

3.4 Alternative sufficient conditions for concavity

In this section, I discuss a few important relaxations of the assumptions made in Theorem 3.1. First, I study the case of CRRA utility functions. Then, I exploit curvature properties of the wage scheme. Note that, so far, the results have only used monotonicity of the wage scheme.

3.4.1 CRRA utility

Recall that the two crucial assumptions from Theorem 3.1, LCDF and NIARA, are strong convexity conditions for the distribution function and the agent's marginal utility of consumption, respectively. As Lemma 3.2 highlights, each of the conditions can be relaxed by strengthening the other. This insight is useful, because in many cases we do not only have log-convexity of the agent's marginal utility of consumption (NIARA), but stronger results.

The following proposition is a formal statement of this idea. It relaxes the LCDF property by restricting the class of preferences.

Proposition 3.1. *Let* $(w^*(\cdot), e^*, s^*)$ *be a solution to the first-order problem* (P2). *Suppose MLR and NIARA. Suppose there exists a number* $\eta > 1$ *such that for all* c

$$\frac{u'''(c)u'(c)}{(u''(c))^2} \geq \eta, \tag{3.17}$$

and for all $e \in I$, $x \in [\underline{x}, \overline{x}]$,

$$\frac{F_{ee}(x,e)F(x,e)}{(F_e(x,e))^2} \geq \frac{1}{\eta}. \tag{3.18}$$

Then, given this contract, the agent's decision problem is concave. Hence, the contract solves the original problem (P1).

Note that the right-hand side of (3.18) is a number between 0 and 1. Thus, the convexity requirement for the distribution function is somewhere between Rogerson's (1985a) CDF condition and the LCDF property introduced in Theorem 3.1.[8]

[8]Clearly, we cannot expect to obtain a better condition than CDF in the present setup, given the counterexample by Kocherlakota (2004).

As an important application of Proposition 3.1, consider CRRA utility: $u(c) = c^{1-\gamma}/(1-\gamma)$. Then we have

$$\frac{u'''(c)u'(c)}{(u''(c))^2} = 1 + \frac{1}{\gamma}. \tag{3.19}$$

Hence, using Proposition 3.1, we conclude that the first-order approach is valid if for all $e \in I$, $x \in [\underline{x}, \overline{x}]$,

$$\frac{F_{ee}(x,e)F(x,e)}{(F_e(x,e))^2} \geq \frac{\gamma}{1+\gamma}. \tag{3.20}$$

Under the spanning condition from Example 3.4, for instance, this property is equivalent to

$$\frac{(\Gamma'(e))^2}{-\Gamma''(e)(1-\Gamma(e))} \leq 1 + \frac{1}{\gamma} \quad \text{for all } e \in I. \tag{3.21}$$

This relaxes condition (3.16).

3.4.2 Exploiting the curvature of the contract

To validate the first-order approach, I have previously derived conditions on the output distribution and the agent's preferences under which, given that the contract is monotonic in output, the agent's decision problem is concave. Finding such conditions will be much simpler if the contract is not only monotonic, but also exhibits some form of concavity. For the moral hazard problem without hidden saving, Jewitt (1988) provides a general analysis of this idea. He shows that, under quite general conditions, the agent's utility changes with output in a concave way, and therefore the convexity conditions on the distribution function can be substantially relaxed.

In the present subsection, I try to exploit the curvature of the contract in the context of hidden saving.

Integration of the distribution function

For the standard moral hazard problem, Jewitt (1988, Theorem 1) validates the first-order approach when the agent's (ex-post) utility $u(w^*(x))$ is concave in output x and the integral of the distribution function satisfies a convexity property. He then relates the latter property to the concept of Total Positivity (Karlin 1968) and shows that it is satisfied for a rather general class of probability distributions. In the setup with hidden saving, such a result is more difficult to obtain. Since the agent's consumption level can be changed by saving, one needs to specify not only the curvature between utility and output, but also how this curvature changes with the savings level s. I obtain the following result.[9]

Proposition 3.2. *Let* $(w^*(\cdot), e^*, s^*)$ *be a solution to the first-order problem* (P2). *Suppose MLR and NIARA. Suppose that for all output levels* x

$$-\frac{d^2(u(w^*(x)+s))}{dx^2} \quad \text{is positive and log-convex in saving } s, \tag{C1}$$

$$\widetilde{F}(x,e) = \int_{\underline{x}}^{x} F(z,e)\,dz \quad \text{is log-convex in effort } e. \tag{LCI}$$

Then, given this contract, the agent's decision problem is concave. Hence, the contract solves the original problem (P1).

Unfortunately, there is no simple way of expressing condition (C1) in terms of the fundamentals of the model. However, it is easy to see that (C1) is a concavity property of the contract: A's (ex-post) consumption utility, $u(w^*(x)+s))$, depends on output x in a concave way. In addition, the curvature between utility and output changes with saving s in a log-convex way.

The next result shows that (C1) is satisfied if the wage scheme $w^*(x)$ is concave in output

[9] To facilitate the argument, I suppose that the wage scheme is twice continuously differentiable. As the Kuhn-Tucker condition (3.3) shows, the wage scheme $w^*(x) = c^*(x) - s^*$ will be C^2 in x if $f_e(x,e)/f(x,e)$ is C^2 in x and $u'(c), a(c)$ are C^2 in c.

x. Hence, (C1) is guaranteed under an appropriate concavity property of the likelihood ratio function $f_e(x,e)/f(x,e)$; see the appendix for details.

Lemma 3.3. *Suppose NIARA and suppose $-u'''(c)/u''(c)$ is nonincreasing in c. Then condition* (C1) *is necessary but not sufficient for $w^*(x)$ to be concave in x.*

The assumption that $-u'''(c)/u''(c)$ is nonincreasing in c (nonincreasing absolute prudence) is innocuous. For instance, it is satisfied for all utility functions with hyperbolic absolute risk aversion (HARA).

To capture the second condition in Proposition 3.2, it is important to note that log-convexity is preserved under integration (Boyd and Vandenberghe 2004, p. 106). Therefore, log-convexity of the integral $\widetilde{F}(x,e) = \int_{\underline{x}}^{x} F(z,e)\,dz$ is a weaker assumption than log-convexity of $F(x,e)$ (LCDF). Intuitively, the integral $\widetilde{F}(x,e)$ will be log-convex in e if the distribution function $F(x,e)$ is log-convex in e for small values of x and "not too misbehaved" for large values of x. In fact, $F(x,e)$ does not even have to be convex in e as the following example shows.

Example 3.5 (Beta Prime distribution)**.** Consider the Beta Prime distribution with parameter $b = 2$:

$$f(x,e) = \frac{x^{e-1}(1+x)^{-e-2}}{B(e,2)}, \quad x \in [0,\infty),\ e \in (0,\infty), \tag{3.22}$$

where $B(e,b)$ represents the Beta function. The likelihood ratio function $f_e(x,e)/f(x,e)$ is nondecreasing concave in x, hence the class of preferences satisfying (C1) is nonempty. The distribution function is

$$F(x,e) = (1+e+x)x^e(1+x)^{-e-1}. \tag{3.23}$$

It is easy to see that $F(x,e)$ is not convex in e for all x. However, the integral of the distribution function,

$$\widetilde{F}(x,e) = x\left(\frac{x}{1+x}\right)^{e}, \tag{3.24}$$

is log-linear in e. Therefore, LCI is satisfied.

Given this example, Proposition 3.2 even validates the first-order approach for a class of setups where the distribution function is not convex in effort. However, the gain of Proposition 3.2 in terms of relaxing the LCDF condition is limited. It is difficult to find many other examples that satisfy LCI without satisfying LCDF. In addition, condition (C1) is a stronger requirement than concavity of the agent's utility in output. Hence, compared to Jewitt's (1988) Theorem 1, it is much more difficult to relax the convexity properties of the distribution function in the present framework.

Quasiconvex distribution functions

To exploit the curvature of contracts, it is often helpful to study distribution functions that are (jointly) quasiconvex in output and effort, because, roughly speaking, such distributions are equivalent to production functions with nonincreasing returns to scale. For instance, Jewitt (1988, Theorem 3) and Conlon (2009) use this property to validate the first-order approach for multi-signal moral hazard problems. In the present section, I show that quasiconvex distribution functions also have attractive properties for the model with hidden saving.

Recall that, to establish concavity of the agent's decision problem, it is sufficient to consider the agent's utility in the second period (Lemma 3.1). Moreover, note that the distribution function $F(x,e)$ is quasiconvex in (x,e) if and only if output can be represented by a 'production function' $x = \varphi(e,z)$ that is nondecreasing concave in e and nondecreasing in the stochastic state of nature z (Jewitt 1988, Lemma 2). Using the production function, we can write the agent's second-period utility as

$$\int_{\underline{x}}^{\overline{x}} u(w^*(x)+s)f(x,e)\,dx = \mathbb{E}[u(w^*(\varphi(e,z))+s)], \tag{3.25}$$

where $\mathbb{E}[\cdot]$ denotes expectations with respect to the state of nature z. Now, notice that concavity is preserved under summation and under nondecreasing concave transformations. Hence, since u is nondecreasing concave, the agent's decision problem will be concave in (e, b) if $w^*(\varphi(e,z))$ is concave in e for all z.[10] This insight is key for the following result.

Proposition 3.3. *Let* $(w^*(\cdot), e^*, s^*)$ *be a solution to the first-order problem* (P2). *Suppose that the following conditions hold:*

$$F(x,e) \text{ is quasiconvex in } (x,e), \tag{3.26}$$

$$f_e(x,e)/f(x,e) \text{ is nondecreasing and concave in } x \text{ for all } e, \tag{3.27}$$

$$g(c) := \left(\frac{1}{R\beta u'(c)} - \xi a(c)\right) \text{ is increasing and convex in } c. \tag{3.28}$$

Then, given this contract, the agent's decision problem is concave. Hence, the contract solves the original problem (P1).

The function g defined in (3.28) links the likelihood ratio function $f_e(y,e)/f(y,e)$ to the shape of the wage scheme: Recall the Kuhn-Tucker condition (3.3),

$$\frac{1}{R\beta u'(c^*(x))} = \lambda + \mu\frac{f_e(x,e^*)}{f(x,e^*)} + \xi a(c^*(x)), \quad x \in [\underline{x}, \overline{x}]. \tag{3.29}$$

[10] At first glance, this insight seems to suggest that the validity of the first-order approach is not affected by the introduction of hidden saving. Notice, however, that the shape of the wage scheme is crucially influenced by the agent's ability to save; see equation (3.3).

Hence, the wage scheme $w^*(x) = c^*(x) - s^*$ is characterized by

$$w^*(x) = g^{-1}\left(\lambda + \mu \frac{f_e(x, e^*)}{f(x, e^*)}\right) - s^*, \quad x \in [\underline{x}, \overline{x}]. \tag{3.30}$$

Thus, since g^{-1} is nondecreasing concave under assumption (3.28), concave likelihood ratio functions will generate concave wage schemes $w^*(x)$ in this case.

As an important application of Proposition 3.3, consider CARA utility:

$$u(c) = -\exp(-\alpha c)/\alpha.$$

Then we have $g(c) = (R\beta)^{-1} \exp(\alpha c) - \xi\alpha$. Obviously, this function is increasing and convex in c. Therefore, Proposition 3.3 validates the first-order approach for CARA utility when the distribution function $F(x, e)$ is quasiconvex in (x, e) and the likelihood ratio function $f_e(x, e)/f(x, e)$ is nondecreasing concave in x.[11]

There are other examples, such as CRRA utility, for which g is not convex, however. In that case, the concavity property of the likelihood ratio function formulated in (3.27) has to be strengthened to obtain a concave wage scheme; the details can be found in the appendix. Essentially, the difficulty in obtaining concave wage schemes is driven by the coefficient of absolute risk aversion, which tends to make the right-hand side of the optimality condition (3.29) less concave in x. This result is also important for the characterization of optimal contracts and hints that wages become a more convex function of output under hidden saving.[12]

[11] In the present paper, preferences over consumption and effort are additively separable. For CARA utility and *multiplicatively separable* preferences, by contrast, the validation of the first-order approach becomes much simpler, since the agent's effort choice will be independent of his wealth level. For most applications, this does not seem to be a useful approximation, however.

[12] See Chapter 4 for a more detailed discussion of this insight.

3.5 Concluding remarks

This paper validates the first-order approach for two-period moral hazard problems with hidden saving. Compared to the model without hidden saving, I additionally impose an assumption on the convexity of the agent's marginal utility of consumption and a restriction of Rogerson's (1985a) CDF condition. I obtain alternative sets of sufficient conditions by relaxing the latter property and including conditions on the curvature of the wage scheme.

These results show under what conditions the first-order approach can be safely applied. Besides, they indicate that the approach is slightly less general than in the standard moral hazard problem. Therefore, understanding how to characterize optimal contracts without applying the first-order approach could be a useful complement to the present work. Given the findings by Grossman and Hart (1983), however, such an approach will also involve strong structural assumptions in general.

3.A Appendix: Proofs

Proof of Lemma 3.1. A's objective function is

$$(e, s) \mapsto u(w_0 - \frac{s}{R}) + \beta \int_{\underline{x}}^{\overline{x}} u(w^*(x) + s) f(x, e)\, dx - \beta v(e). \tag{3.31}$$

Since u is concave, the first summand is concave in (e, s). Since v is convex, the third summand is concave in (e, s). □

Proof of Lemma 3.2. Using partial integration, A's second-period consumption utility can be rewritten as

$$\int_{\underline{x}}^{\overline{x}} u(w^*(x)+s)f(x,e)\,dx = u(w^*(\overline{x})+s) - \int_{\underline{x}}^{\overline{x}} (w^*)'(x)u'(w^*(x)+s)F(x,e)\,dx. \tag{3.32}$$

Hence, A's second-period consumption utility is concave in (e,s) if the function

$$(e,s) \mapsto -\int_{\underline{x}}^{\overline{x}} (w^*)'(x)u'(w^*(x)+s)F(x,e)\,dx \tag{3.33}$$

is concave, or equivalently if the function

$$(e,s) \mapsto \int_{\underline{x}}^{\overline{x}} (w^*)'(x)u'(w^*(x)+s)F(x,e)\,dx \tag{3.34}$$

is convex. I want to show that

$$g(e,s;x) = u'(w^*(x)+s)F(x,e) \tag{3.35}$$

is convex in (e,s) for all x. Since $(w^*)'(x) \geq 0$ by assumption, and since convexity is preserved under integration, this will imply convexity of (3.34).

The function $g(e,s;x)$ is convex in (e,s) if and only if its Hessian has a nonnegative diagonal and a nonnegative determinant. Omitting all arguments, the Hessian equals

$$H = \begin{pmatrix} F_{ee}u' & F_e u'' \\ F_e u'' & F u''' \end{pmatrix}. \tag{3.36}$$

The first diagonal entry is nonnegative by assumption. Condition (3.6) is equivalent to the statement that the determinant of H is nonnegative. In that case, the second diagonal entry of H must also be nonnegative. □

Proof of Theorem 3.1. By Lemma 3.1, it is sufficient to establish concavity of A's second-period consumption utility. Due to MLR and NIARA, the Kuhn-Tucker condition (3.3) implies that the wage scheme $w^*(x) = c^*(x) - s^*$ is continuously differentiable and nondecreasing in output x. Moreover, LCDF and NIARA imply that condition (3.6) from Lemma 3.2 is satisfied. Hence, A's second-period consumption utility is concave. □

Proof of Proposition 3.1. Using similar steps as in the proof of Theorem 3.1, the result follows from Lemma 3.1 and Lemma 3.2. □

Proof of Proposition 3.2. As Lemma 3.1 shows, it is sufficient to establish concavity of

$$(e, s) \mapsto \int_{\underline{x}}^{\overline{x}} u(w^*(x) + s) f(x, e)\, dx. \tag{3.37}$$

This is equivalent to establishing convexity of

$$(e, s) \mapsto -\int_{\underline{x}}^{\overline{x}} u(w^*(x) + s) f(x, e)\, dx. \tag{3.38}$$

Using two steps of partial integration, the latter function can be rewritten as

$$-u(w^*(\overline{x}) + s) + (w^*)'(\overline{x}) u'(w^*(\overline{x}) + s)) \widetilde{F}(\overline{x}, e) + \int_{\underline{x}}^{\overline{x}} \left(-\frac{d^2(u(w^*(x) + s))}{dx^2} \right) \widetilde{F}(x, e)\, dx. \tag{3.39}$$

First, note that the expression $-u(w^*(\overline{x}) + s)$ is convex in (e, s) due to the concavity of u. Moreover, the expression

$$(w^*)'(\overline{x}) u'(w^*(\overline{x}) + s)) \widetilde{F}(\overline{x}, e) \tag{3.40}$$

is convex in (e, s) by an argument similar to Lemma 3.2. For the third term in (3.39), note that

$$-\frac{d^2(u(w^*(x)+s))}{dx^2}\widetilde{F}(x,e) \tag{3.41}$$

is the product of a function that is log-convex in s and a function that is log-convex in e. Such products are convex in (e,s) as one easily verifies. Since convexity is preserved under integration, the third term in (3.39) is thus convex as well. This completes the proof. □

Proof of Lemma 3.3. Suppose $w^*(x)$ is concave in x. The function studied in (C1) can be represented as

$$-\frac{d^2(u(w^*(x)+s))}{dx^2} = -(w^*)''(x)u'(w^*(x)+s) + ((w^*)'(x))^2\left(-u''(w^*(x)+s)\right). \tag{3.42}$$

The first summand in (3.42) is log-convex in s, since $-(w^*)''(x) \geq 0$ and since u' is log-convex due to NIARA. The second summand is log-convex in s, since $((w^*)'(x))^2 \geq 0$ and since $-u''$ is log-convex when $-u'''/u''$ is nonincreasing. Since log-convexity is preserved under summation (Boyd and Vandenberghe 2004, p. 105), the function studied in (C1) is therefore log-convex in the variable s.

On the other hand, suppose that the function studied in (C1) is log-convex in s. As (3.42) shows, this does not imply that $(w^*)''(x)$ is nonpositive in general. □

Proof of Proposition 3.3. By Lemma 3.1, it is sufficient to consider A's second-period consumption utility. Moreover, due to quasiconvexity of the distribution function, the output technology can be represented by a production function $x = \varphi(e,z)$, with $\varphi(e,z)$ nondecreasing concave in effort e and nondecreasing in the stochastic state of nature z (Jewitt 1988, Lemma 2). Using this representation, we can write A's second-period consumption utility as

$$\int_{\underline{x}}^{\overline{x}} u(w^*(x) + s) f(x, e)\, dx = \mathbb{E}[u(w^*(\varphi(e, z)) + s)], \tag{3.43}$$

where $\mathbb{E}[\cdot]$ denotes expectations with respect to the state of nature z.

I claim that w^* is nondecreasing concave. Recall from the Kuhn-Tucker condition (3.3) that solutions to (P2) are characterized by

$$w^*(x) = g^{-1}\left(\lambda + \mu \frac{f_e(x, e^*)}{f(x, e^*)}\right) - s^*, \tag{3.44}$$

with $g(c) = 1/(R\beta u'(c)) - \xi a(c)$. By assumption, g is increasing and convex. Equivalently, g^{-1} is increasing and concave. Since $f_e(x, e^*)/f(x, e^*)$ is nondecreasing concave in x by assumption, this implies that $w^*(x)$ is nondecreasing concave in x.

Now, since $\varphi(e, z)$ is concave in e and $w^*(x)$ is nondecreasing concave in x, the composition $w^*(\varphi(e, z))$ is concave in e. Hence, the function $w^*(\varphi(e, z)) + s$ is concave in (e, s). Since u is nondecreasing concave, and since concavity is preserved under taking expectations, this completes the proof. □

3.B Appendix: Concave wage schemes

This section characterizes when the consumption scheme $c^*(x)$ solving the first-order problem (P2) is concave in output x. Since $w^*(x) = c^*(x) - s^*$, this property is equivalent to the wage scheme $w^*(x)$ being concave in x.

Due to equation (3.3), the consumption scheme is characterized by

$$c^*(x) = g^{-1}\left(\lambda + \mu L(x)\right), \tag{3.45}$$

with $g(c) = 1/(R\beta u'(c)) - \xi a(c)$, $L(x) = f_e(x, e^*)/f(x, e^*)$. The first derivative of g equals

$$g'(c) = \frac{\xi u'''(c)u'(c) - ((R\beta)^{-1} + \xi u''(c))u''(c)}{u'(c)^2}, \tag{3.46}$$

which yields

$$(c^*)'(x) = \frac{\mu L'(x)u'(c^*(x))^2}{\xi u'''(c^*(x))u'(c^*(x)) - ((R\beta)^{-1} + \xi u''(c^*(x)))u''(c^*(x))}. \tag{3.47}$$

Omitting the arguments x and $c^*(x)$, the latter implies

$$\begin{aligned}(c^*)''(x) =& \frac{\mu}{(\dots)^2}\Big[\left(L''(u')^2 + 2L'u''u'(c^*)'\right)\left(\xi u'''u' - ((R\beta)^{-1} + \xi u'')u''\right) \\ & - L'(u')^2(c^*)'\left(\xi u^{(4)}u' - ((R\beta)^{-1} + \xi u'')u'''\right)\Big].\end{aligned} \tag{3.48}$$

Hence, given the assumption

$$(u')^2[\xi(u'''u' - (u'')^2) - u''(R\beta)^{-1}] > 0, \tag{3.49}$$

which is true under NIARA, $c^*(x)$ is concave in x if and only if the likelihood ratio function satisfies the following concavity condition:

$$\begin{aligned}L''(x) \leq \frac{L'(c^*)'}{(u')^2[\xi(u'''u' - (u'')^2) - u''(R\beta)^{-1}]}\Big[&2\xi u'(-u'')(u'''u' - (u'')^2) + 2u'(u'')^2(R\beta)^{-1} \\ &+ \xi(u')^2(-u'')u''' + \xi(u')^3u^{(4)} - (u')^2u'''(R\beta)^{-1}\Big].\end{aligned} \tag{3.50}$$

Chapter 4

Optimal income taxation with asset accumulation

4.1 Introduction

The progressivity of the income tax code is a central question in the economic literature and the public debate. While we observe progressive tax systems in most developed countries nowadays, theoretical insights on progressivity are rather limited. Previous research has examined how the optimal degree of progressivity depends on the skill distribution (Mirrlees 1971), the welfare criterion (Sadka 1976), and earnings elasticities (Saez 2001).

The existing approaches to income tax progressivity have largely focused on models where labor is the only source of income. However, we argue in the present paper that the optimal shape of labor income taxes cannot be determined in isolation from the tax code on capital income. Specifically, we show that taxing capital has an important effect on the optimal progressivity of labor income taxes: When capital is taxed, the optimal tax on labor income becomes more progressive.

We derive this result in a two-period model of social insurance. A continuum of ex-ante identical agents influence their labor incomes by exerting effort. Labor income realizations are not perfectly controllable, which creates a moral hazard problem. Thus, the social planner faces a trade-off between insuring the agents against idiosyncratic income uncertainty on the one hand and the associated disincentive effects on the other hand. In addition, agents have access to a risk-free bond, which gives them a limited means for self-insurance.

There is an efficiency reason for taxing bond returns in our model: The bond provides insurance against the realization of labor income and thereby reduces the incentives to exert effort. Consequently, the planner faces an additional constraint when determining the optimal social insurance scheme, since she has to satisfy the agent's Euler equation. If the bond is appropriately taxed, however, the constraint becomes non-binding. Through this channel, taxing the bond affects the structure of optimal labor income taxes. More precisely, if the bond is *not* taxed, the planner needs to consider how the distribution of after-tax income changes the agent's saving motive in the previous period. We show that, for a given income level, a local reduction in the tax rate affects the agent's saving motive according to his coefficient of absolute risk aversion. Since absolute risk aversion is typically a convex function, we find that optimal after-tax income changes in a more convex way with labor income when bond returns are not taxed. Equivalently, optimal after-tax income changes in a more concave way when the bond is taxed, which means that the tax on labor income becomes more progressive.

To the best of our knowledge, this is the first paper that examines how capital taxation affects the optimal progressivity of the labor income tax code. Recent work on dynamic Mirrleesian taxation models has highlighted a somewhat complementary question. Kocherlakota (2005), Albanesi and Sleet (2006), and Golosov and Tsyvinski (2006) focus on the optimal taxation of capital, given nonlinear taxation of income. In that literature, the reason for capital taxation is similar to our model and stems from disincentive effects associated with the accumulation of wealth.[1]

While the Mirrlees (1971) framework focuses on redistribution in a population with heterogeneous skills, our approach highlights the social insurance (or ex-post redistribution) aspect of income taxation. In spirit, our model is therefore closer to the works by Varian (1980) and Eaton and Rosen (1980). The main reason for excluding ex-ante redistribution in the present paper is tractability. Conesa and Krueger (2006) study optimal taxation numerically when both redistribution and social insurance motives are present. However, they consider a restricted class of tax schemes. In particular, they do not distinguish between the taxation of labor income and capital. The present model is simpler in some dimensions, but allows us to analytically explore the fully optimal tax scheme.

The paper proceeds as follows: Section 2 describes the setup of the model. Section 3 presents the main result of the paper: Capital taxation increases the progressivity of the optimal income tax code. In Section 4, we explore alternative concepts of concavity/progressivity. Section 5 concludes and discusses some directions for future research. All proofs are provided in the appendix.

4.2 Model

Consider a benevolent social planner (the principal) whose objective is to maximize the welfare of the citizens. The (small open) economy contains a continuum of ex-ante identical agents who live for two periods, $t = 0, 1$, and can influence their date-1 income realizations by working hard or shirking. The planner offers a tax/transfer system

[1] However, the tax on capital takes a much simpler form in our model: a linear tax on the aggregate trade of the bond is sufficient to implement the second best; compare Gottardi and Pavoni (2010). Notice, in particular, that this tax can be implemented in an anonymous asset market where the planner only observes the aggregate trading volume.

to insure them against idiosyncratic risk and provide them appropriate incentives for working hard. The planner's budget must be (intertemporally) balanced.

4.2.1 Preferences

The agent derives utility from consumption $c_t \geq \underline{c} \geq -\infty$ and effort $\infty \geq \bar{e} \geq e_t \geq 0$ according to: $u(c_t) - v(e_t)$, where both u and v are strictly increasing and twice continuously differentiable functions, and u is strictly concave whereas v is convex. We normalize $v(0) = 0$. The agent's discount factor is denoted by $\beta > 0$.

4.2.2 Production and endowments

At date $t = 0$, the agent has a *fixed* endowment y_0. At date $t = 1$, the agent has a stochastic income $y \in Y := [\underline{y}, \overline{y}]$. The realization of y is publicly observable, while the probability distribution over Y is affected by the agent's unobservable effort level e_0 that is exerted at $t = 0$. The probability density of this distribution is given by the smooth function $f(y, e_0)$. As in most of the the optimal contracting literature, we assume *full support*, that is $f(y, e_0) > 0$ for all y, e_0. There is no production or any other action at $t \geq 2$.

4.2.3 Markets

At each date, the agent can buy or (short)-sell a risk-free bond b_t which costs $q \geq 0$ consumption units today and pays one unit of consumption tomorrow. The agent has no access to any insurance market other than that delivered by the planner. The planner can impose a linear tax τ^k on the price of the bond.[2] Therefore, the net price of the bond is $\tilde{q} = (1 + \tau^k)q$.

[2] A tax on the bond price is equivalent to a tax on the return in our model.

There are two ways of motivating the linearity assumption of the tax τ^k. First, the planner only needs to observe the aggregate trade of the bond in order to implement such a tax. Therefore, the tax is feasible even in an anonymous market where individual asset decisions and consumption levels are private information. Second, linearity is in fact without loss of generality in the present model, since the planner is able to generate the second best with such a tax (see Proposition 4.1).

Given the structure of the problem, the agent will never be able to borrow at $t = 1$, hence we have $b_1 \geq 0$. Monotonicity of preferences guarantees that the agent will not want to leave any positive amount of assets at date 1 either. So, $b_1 = 0$ for all states y. Similarly, since v is strictly increasing, $e_1 = 0$ for all states y.

4.2.4 Contracts

A **contract** $\mathcal{W} := (\boldsymbol{T}, \tau^k, e_0, b_0)$ consists of a tax/transfer scheme $\boldsymbol{T}$, a tax rate on the bond τ^k, and choices (e_0, b_0). The tax/transfer scheme $\boldsymbol{T} := (T_0, T(\cdot))$ has two components: T_0 denotes the transfer the individual receives in period $t = 0$, and $T(y)$, $y \in Y$, denotes the transfer the individual receives in period $t = 1$ conditional on income realization y.

Given a contract $\mathcal{W}$, the agent's utility is

$$U(e_0, b_0; \boldsymbol{T}, \tau^k) := u(y_0 + T_0 - (1 + \tau^k)qb_0) - v(e_0) + \beta \int_{\underline{y}}^{\overline{y}} u(y + T(y) + b_0) f(y, e_0)\, dy.$$

To guarantee solvency of the agent for every contingency, we impose the 'natural' borrowing limit: $b_0 \geq \underline{c} - \inf_y \{y + T(y)\}$.

The social planner faces the same credit market as the agent, therefore her discount rate is q. The planner's expenditures are

$$T_0 + q \int_{\underline{y}}^{\overline{y}} T(y) f(y, e_0)\, dy - \tau^k q b_0 + G,$$

where G denotes government consumption.

4.2.5 Efficiency

An **optimal contract** is a contract that maximizes ex-ante welfare

$$\max_{\mathcal{W}} \quad U(e_0, b_0; \boldsymbol{T}, \tau^k) \tag{4.1}$$

subject to the planner's budget constraint

$$-T_0 - q \int_{\underline{y}}^{\overline{y}} T(y) f(y, e_0)\, dy + \tau^k q b_0 - G \geq 0 \tag{4.2}$$

and the incentive compatibility constraint

$$(e_0, b_0) \in \arg \left\{ \max_{e,b} \ U(e, b; \boldsymbol{T}, \tau^k) \text{ s.t. } e \geq 0,\ y_0 + T_0 - \underline{c} \geq \tilde{q} b \geq -\tilde{q} \inf_y \{ y + T(y) - \underline{c} \} \right\}. \tag{4.3}$$

Note that there is indeterminacy in the contract between T_0 and b_0. The planner can implement the same allocation with a contract

$$\left(T_0, T(\cdot), \tau^k, e_0, b_0\right)$$

and with a contract

$$\left(T_0 - \tilde{q}\varepsilon, T(\cdot) + \varepsilon, \tau^k, e_0, b_0 - \varepsilon\right).$$

In other words, since the planner and the agent face the same credit market, there is a continuum of optimal contracts. Throughout this paper, without loss of generality, we will study the one specific optimal contract that implements $b_0 = 0$. Because of these observations, we will sometimes refer to the combination of e_0, τ^k, and $\boldsymbol{c} = (c_0, c(\cdot))$, with $c_0 := y_0 + T_0$, $c(y) := y + T(y)$, $y \in Y$, as a contract.

4.2.6 First-order approach

Throughout this paper, we assume that the first-order approach (FOA) is justified. Hence, we can replace the incentive constraint (4.3) by the first-order conditions of the agent's maximization problem with respect to e_0 and b_0. Sufficient conditions for the validity of the FOA in this setup are given in Chapter 3 of this thesis. Specifically, the FOA is valid if the agent has nonincreasing absolute risk aversion and the cumulative distribution function of income is log-convex in effort.[3]

Using the normalization $b_0 = 0$ and the notation $\tilde{q} = (1 + \tau^k)q$, $c_0 = y_0 + T_0$, $c(y) = y + T(y)$, $y \in Y$, we can thus rewrite the planner's problem as

$$\max_{\mathbf{c}, \tilde{q}, e_0} \; u(c_0) - v(e_0) + \beta \int_{\underline{y}}^{\overline{y}} u(c(y)) f(y, e_0)\, dy \tag{4.4}$$

subject to $c_0 \geq \underline{c}$, $c(y) \geq \underline{c}$, $e_0 \geq 0$, the planner's budget constraint

$$y_0 - c_0 + q \int_{\underline{y}}^{\overline{y}} (y - c(y)) f(y, e_0)\, dy - G \geq 0 \tag{4.5}$$

[3] As argued by Abraham, Koehne, and Pavoni (2010), both conditions have quite a broad empirical support. First, virtually all estimations for u reveal NIARA; see Guiso and Paiella (2008) for example. Second, while the condition on the distribution function cannot be taken to the data directly, the authors show that it can be interpreted as a restriction on the agent's Frisch elasticity of labor supply. This restriction is satisfied as long as the Frisch elasticity is smaller than unity. In fact, most empirical studies find values of this elasticity between 0 and 0.5; see Domeij and Floden (2006), for instance.

and the first-order incentive conditions

$$
\begin{aligned}
-v'(e_0) + \beta \int_{\underline{y}}^{\overline{y}} u(c(y)) f_e(y, e_0)\, dy &\geq 0 & (4.6)\\
\tilde{q} u'(c_0) - \beta \int_{\underline{y}}^{\overline{y}} u'(c(y)) f(y, e_0)\, dy &\geq 0. & (4.7)
\end{aligned}
$$

Under mild assumptions, the optimal contract is interior: $c_0 > \underline{c}$, $c(y) > \underline{c}$, $e_0 > 0$.[4] In this case, using λ, μ and ξ as the (nonnegative) Lagrange multipliers associated with the constraints (4.5), (4.6), (4.7), respectively, the first-order conditions of the Lagrangian with respect to consumption are

$$
\begin{aligned}
\frac{\lambda q}{\beta u'(c(y))} &= 1 + \mu \frac{f_e(y, e_0)}{f(y, e_0)} + \xi a(c(y)), \quad y \in [\underline{y}, \overline{y}], & (4.8)\\
\frac{\lambda}{u'(c_0)} &= 1 - \xi \tilde{q} a(c_0), & (4.9)
\end{aligned}
$$

where $a(c) := -u''(c)/u'(c)$ denotes the agent's absolute risk aversion.

4.2.7 Preliminary characterization of optimal contracts

Because of the incentive problem, it is efficient to impose a positive tax on savings in our model; compare Gottardi and Pavoni (2010).

Proposition 4.1. *Assume that the FOA is justified and that the optimal contract is interior. Then the tax on savings is positive: $\tau^k > 0$. Moreover, equations (4.8) and (4.9) characterizing the consumption scheme are satisfied with $\xi = 0$.*

The above result is intuitive. It is efficient to tax the bond, because saving provides insurance against the incentive scheme. By appropriately reducing the rate of return, however, the planner can control the agent's intertemporal decision and therefore cir-

[4] A sufficient condition for interiority is $\lim_{c \to \underline{c}} u'(c) = \infty$, $v'(0) = 0$.

cumvent the (first-order) incentive constraint for saving.[5] Consequently, when savings can be taxed, condition (4.8) takes the form that is familiar from dynamic moral hazard models without asset accumulation:

$$\frac{\lambda q}{\beta u'(c(y))} = 1 + \mu \frac{f_e(y, e_0)}{f(y, e_0)}, \quad y \in [\underline{y}, \overline{y}].$$

Since we are interested in understanding how the shape of optimal consumption depends on the possibility of taxing savings, we contrast the optimal contract when τ^k is a choice variable for the planner with the optimal contract when τ^k is restricted to zero (i.e., $\tilde{q} = q$).

Proposition 4.2. *Consider the above problem with τ^k restricted to zero. Assume that the FOA is justified and that the optimal contract is interior. Then equations (4.8) and (4.9) characterizing the consumption scheme are satisfied with $\xi > 0$.*

4.3 Simple results

We are interested in the shape of the optimal tax/transfer scheme $\boldsymbol{T}$. Clearly, this shape is closely related to the curvature of consumption $c(y) = y + T(y)$. Recall that we can fix $b_0 = 0$ without loss of generality.

Definition 4.1. We say that the transfer scheme $\boldsymbol{T}$ is **progressive (regressive)** if $c'(y)$ is decreasing (increasing) in y. We call $\boldsymbol{T}$ **proportional** if $c'(y)$ is constant in y.

This definition implies that whenever consumption is a concave (convex) function of income we have a progressive (regressive) tax system supporting it. In terms of the taxes and transfers $T(y)$, in a progressive system taxes ($T(y) < 0$) are increasing faster than

[5] The validity of the first-order approach is crucial here, since it allows to characterize the agent's saving decision exclusively based on the rate of return.

income does. At the same time, for the states when the agent is receiving a transfer ($T(y) > 0$), transfers are increasing slower than income is decreasing. The opposite happens when we have a regressive scheme. Intuitively, if the scheme is progressive, incentives are provided more by imposing 'large penalties' for low income realizations, since consumption decreases relatively quickly when income decreases. Regressive schemes, by contrast, put more emphasis on rewards for high income levels than punishments for low income levels. If the scheme is proportional, these rewards and punishments are in some sense balanced. The next proposition provides conditions for progressivity and regressivity of the optimal scheme.

Proposition 4.3 (Sufficient conditions for progressivity/regressivity)**.** *Assume that the FOA is justified and that the optimal contract is interior.*
(i) If the likelihood ratio function $l(y,e) := \frac{f_e(y,e)}{f(y,e)}$ *is concave in* y *and* $\frac{1}{u'(c)}$ *is convex in* c*, then* $\boldsymbol{T}$ *is progressive. If, in addition, absolute risk aversion* $a(c)$ *is decreasing and concave,*[6] *then this result continues to hold when* τ^k *is restricted to zero.*
(ii) On the other hand, if $l(y,e)$ *is convex in* y *and* $\frac{1}{u'(c)}$ *is concave in* c*, then* $\boldsymbol{T}$ *is regressive. If, in addition, absolute risk aversion* $a(c)$ *is decreasing and convex, then this result continues to hold when* τ^k *is restricted to zero.*

Proposition 4.3 implies that CARA utilities with concave likelihood ratios lead to progressive schemes, no matter whether savings are taxed or not.[7] When savings are taxed, progressive schemes are also induced by concave likelihood ratios and CRRA utilities with $\sigma \geq 1$, since $\frac{1}{u'(c)} = c^\sigma$ is convex in this case. For logarithmic utility with linear likelihood ratios we obtain a scheme that is *proportional*, since $\frac{1}{u'(c)} = c$ is both concave and convex. Interestingly, when savings are not taxed, the scheme becomes *regressive*

[6]Notice that absolute risk aversion is bounded below by zero. Therefore, the function $a(\cdot)$ can only be decreasing and concave over $[0,\infty)$ if it is constant.

[7]Other cases where the tax on savings does not affect regressivity/progressivity are when a has the same shape as $\frac{1}{u'}$ (quadratic utility) and when a is linear (and hence increasing).

in this case (since absolute risk aversion $a(c) = \frac{1}{c}$ is convex).[8] This particular finding sheds light on a more general pattern under convex absolute risk aversion: when savings are taxed, the allocation has a 'more concave' relationship between income and consumption. In other words, taxing savings calls for more progressivity in the income tax/transfer system. The next result formalizes this insight.

Proposition 4.4 (Concavity)**.** *Assume that the FOA is justified. Let $\mathbf{c}$ be an interior, monotonic optimal consumption scheme for the general model and let $\hat{\mathbf{c}}$ be an interior, monotonic optimal consumption scheme for the model when τ^k is restricted to zero, both implementing effort level e_0. Moreover, assume that u has convex absolute risk aversion and that the likelihood ratio $l(y, e_0)$ is linear in y. Under these conditions, if $\hat{\mathbf{c}}$ changes with y in a concave way, then $\mathbf{c}$ does as well.*

In order to obtain a clearer intuition of this result, we further examine condition (4.8), namely

$$\frac{\lambda q}{\beta u'(c(y))} = 1 + \mu \frac{f_e(y, e_0)}{f(y, e_0)} + \xi a(c(y)).$$

This expression equates the discounted present value (normalized by $f(y, e_0)$) of the costs and benefits of increasing the agent's utility by one unit in state y. The increase in utility costs $\frac{q}{\beta u'(c(y))}$ units in consumption terms. Multiplied by the shadow price of resources λ, we obtain the left-hand side of the above expression. In terms of benefits, first of all, since the agent's utility is increased by one unit, there is a return of 1. Furthermore, increasing the agent's utility also relaxes the incentive constraint for effort, generating a return of $\mu \frac{f_e(y, e_0)}{f(y, e_0)}$.[9] Finally, by increasing $u(c(y))$ the planner alleviates the saving motives of the agent. This gain, measured by $\xi a(c(y))$, depends crucially on whether savings are taxed or not. When savings are appropriately taxed, we have $\xi = 0$ and this

[8] More precisely, consumption is characterized by $\frac{\lambda q}{\beta} c(y) - \xi \frac{1}{c(y)} = 1 + \mu l(y, e)$ in this case. Since the left-hand side is concave in c and the right-hand side is linear in y, the consumption scheme $c(y)$ must be convex in y.

[9] Of course, if the increase in consumption is done in a state with a negative likelihood ratio, this represents a cost since the incentive constraint is in fact tightened.

gain vanishes. Intuitively, by controlling the net price of the bond, the planner is able to circumvent the incentive constraint for saving. However, when a tax on savings is ruled out, this constraint is binding and we have $\xi > 0$. Under convex absolute risk aversion, the gain $\xi a(c(y))$ is convex. This implies that, ceteris paribus, the benefits of increasing the agent's utility change in a more convex way with income. As a consequence, at the optimal contract the costs of increasing the agent's utility must also change in a more convex way with income, hence consumption becomes more convex in y in this case.

4.4 More elaborate results

Since at least Holmström (1979), it is well understood that consumption patterns under moral hazard are crucially influenced by the shape of the likelihood ratio function $l(\cdot, e)$. Stated in more negative terms, one can always find functions $l(\cdot, e)$ so that the shape of consumption is almost arbitrary. To make the impact of the savings tax on the shape of optimal consumption easier to observe, we have therefore normalized the curvature of the likelihood ratio by assuming linearity in Proposition 4.4.

In this section, we study how the savings tax changes the curvature of the consumption scheme for arbitrary likelihood ratio functions. As usual, we assume that the FOA is justified and that $\boldsymbol{c}$ and $\hat{\boldsymbol{c}}$ are interior, monotonic optimal contracts for the general model and the model without the savings tax, respectively, implementing the same effort level e_0.

Probably the most well known ranking in terms of concavity in economics is that dictated by concave transformations (e.g. Gollier 2001).

Definition 4.2. We say that f_1 is a concave (convex) transformation of f_2 if there is an increasing and concave (convex) function v such that $f_1 = v \circ f_2$.

Proposition 4.5. *Assume that u has convex absolute risk aversion. Then, if $\hat{c}$ is a* *concave transformation of l, then c is a concave transformation of l. Conversely, if c is a convex transformation of l, then $\hat{c}$ has the same property.*

The previous finding induces an ordering that has the flavor of c being 'more concave' than $\hat{c}$. Note that this result generalizes Proposition 4.4 to arbitrary shapes of the likelihood ratio function l. As a drawback, we can rank the curvature of c and $\hat{c}$ only when, for example, c is more concave than l. We will now reduce the set of possible utility functions to facilitate such comparisons.

Let us consider the class of HARA (or linear risk tolerance) utility functions, namely

$$u(c) = \rho \left(\eta + \frac{c}{\gamma} \right)^{1-\gamma}$$

$$\text{with } \rho \frac{1-\gamma}{\gamma} > 0, \text{ and } \eta + \frac{c}{\gamma} > 0.$$

For this class, we have $a(c) = \left(\eta + \frac{c}{\gamma} \right)^{-1}$. Hence, absolute risk aversion is convex. Special cases of the HARA class are CRRA, CARA, and quadratic utility (e.g. Gollier 2001).

Lemma 4.1. *Given a utility function $u : C \to \mathbb{R}$, consider the two functions defined as follows:*

$$\begin{aligned} g_\lambda(c) &:= \frac{\lambda q}{\beta u'(c)}, \\ \hat{g}_{\hat{\lambda},\hat{\xi}}(c) &:= \frac{\hat{\lambda} q}{\beta u'(c)} - \hat{\xi} a(c). \end{aligned}$$

Then, if u belongs to the HARA class with $\gamma \geq -1$, then $\hat{g}_{\hat{\lambda},\hat{\xi}}$ is a concave transformation of g_λ for all $\hat{\lambda}, \hat{\xi} \geq 0$, $\lambda > 0$.

The restriction $\gamma \geq -1$ in the above result is innocuous and allows for all HARA functions with nonincreasing absolute risk aversion as well as quadratic utility, for instance. To state the consequences of this Lemma, we introduce the concept of G-convexity (e.g.

Avriel, Diewert, Schaible, and Zang 1988), which is widely used in optimization. A function f is G-convex if once we transform f with G we get a convex function. More formally:

Definition 4.3. Let f be a function and G an increasing function mapping from the image of f to the real numbers. The function f is called **G-convex (G-concave)** if $G \circ f$ is a convex (concave) function.

This concept generalizes the standard notion of convexity. It is easy to see that a function f is convex if and only if it is G-convex for any increasing affine function G. Moreover, it can be shown that if G is concave and f is G-convex then f must be convex, but the converse is false.[10]

Proposition 4.6. *Assume u belongs to the HARA class with $\gamma \geq -1$. Then c is g_λ-convex (g_λ-concave) if and only if $\hat{c}$ is $\hat{g}_{\hat{\lambda},\hat{\xi}}$-convex ($\hat{g}_{\hat{\lambda},\hat{\xi}}$-concave).*[11]

Corollary 4.1. *If $\hat{c}$ is g_λ-concave then c is g_λ-concave. Conversely, if c is g_λ-convex then $\hat{c}$ is g_λ-convex.*

The corollary shows that whenever $\hat{c}$ satisfies the g_λ-concavity property, then c satisfies this property. In this sense, we note again that c is 'more concave' than $\hat{c}$.

Finally, it appears natural to ask whether the concavity of c and $\hat{c}$ can also be ranked according to the concavity notion of Definition 4.2. In other words, can we conclude that c is a concave transformation of $\hat{c}$ for HARA utility with $\gamma \geq -1$? In general, the answer is negative. After a small modification of the above lemma, it can be shown that there exists a concave function $\tilde{h}$ such that c and $\hat{c}$ are related as follows:

$$c(y) = \tilde{g}^{-1} \circ \tilde{h} \circ \tilde{g}\left(\hat{c}(y)\right),$$

[10] For example, suppose $f(x) = x^2$ and $G(\cdot) = \log(\cdot)$, then $G(f(x)) = 2\log(x)$, which is obviously not convex.

[11] In fact, this statement is not only true for concavity and convexity, but more generally for any property defined with respect to the transformations g_λ and $\hat{g}_{\hat{\lambda},\hat{\xi}}$.

where $\tilde{g}(c) = \frac{1}{\mu}\left(\frac{\lambda q}{u'(c)} - 1\right)$ is increasing. If $\tilde{g}$ is an affine function (u is logarithmic utility), then one can easily verify that the composition $\tilde{g}^{-1} \circ \tilde{h} \circ \tilde{g}$ is concave whenever $\tilde{h}$ is concave. For the logarithmic case, c is hence a concave transformation of $\hat{c}$. In general, however, we cannot be sure that the composition $\tilde{g}^{-1} \circ \tilde{h} \circ \tilde{g}$ is concave when $\tilde{h}$ is concave.[12]

4.5 Conclusion and outlook

This paper analyzes how capital taxation changes the optimal tax code on labor income. Whenever preferences exhibit convex absolute risk aversion, we find that optimal consumption changes in a 'more concave' way with labor income when the bond is taxed. In this sense, labor income taxes become more progressive when capital is taxed.

The theoretical results derived above are rigorous, but have a few drawbacks. For future research, we thus plan to perform some quantitative exercises to provide a more precise understanding of the above findings. In particular, we plan to estimate the likelihood ratio function from the model. This will show whether the assumptions made, for instance, in Propositions 4.4 and 4.5 are likely to be satisfied in the data. Moreover, we want to explore within a concrete counterfactual how much capital taxation changes the curvature of labor income taxes from a *quantitative* point of view. Specifically, we aim at comparing the shape of labor income taxes under the existing capital tax code with the shape under the optimal code.

In addition, quantitative explorations will allow for the study of alternative progressivity measures. While the theoretical part of this paper has focused on the *marginal* tax rate, progressivity can also be defined with respect to the *average* tax rate as in Sadka (1976). It may also be interesting to consider some alternative measures of progressivity, such

[12]Consider the following example: $\tilde{g}(c) = \exp(c)$ for $c > 0$, $\tilde{h}(x) = (x+d)^{\alpha}$ for $x > 0$, with $0 < \alpha < 1$, $d > 0$. Then $\tilde{h}(x)$ is concave in x, but $\tilde{g}^{-1} \circ \tilde{h} \circ \tilde{g}\,(c) = \alpha \log(\exp(c) + d)$ is convex in c.

as the progressivity wedge discussed in Guvenen, Kuruscu, and Ozkan (2009).[13] Finally, quantitative explorations will highlight the welfare implications of capital taxation in our model. Obviously, capital taxation is welfare-increasing in the present framework, since it reduces the insurance value of asset accumulation. The magnitude of this effect is not evident, however.

On a more general level, some further extensions of the model could be useful for future research. First of all, the study of assets different from the bond might highlight the boundaries of the present findings. Secondly, it might be interesting to see how the impact of capital taxation changes when there is a desire for ex-ante redistribution.

4.A Appendix: Proofs

Proof of Proposition 4.1. See Gottardi and Pavoni (2010). □

Proof of Proposition 4.2. If $\xi > 0$, we are done. If $\xi = 0$, then the first-order conditions of the Lagrangian read

$$\begin{aligned} \frac{\lambda}{u'(c_0)} &= 1, \\ \frac{\lambda q}{\beta u'(c(y))} &= 1 + \mu \frac{f_e(y, e_0)}{f(y, e_0)}, \quad y \in [\underline{y}, \overline{y}]. \end{aligned}$$

[13] Guvenen, Kuruscu, and Ozkan (2009) define the progressivity wedge as follows:

$$-\frac{\frac{c_{s+1}}{y_{s+1}} - \frac{c_s}{y_s}}{\frac{c_s}{y_s}}.$$

When income is continuous, the corresponding expression is:

$$-\frac{d}{dy} \log \frac{c(y)}{y} = -\frac{d}{dy} \left[\log c(y) - \log y\right] = \frac{1}{y} - \frac{c'(y)}{c(y)}.$$

Since $f(y,e)$ is a density, integration of the last line yields

$$\int_{\underline{y}}^{\overline{y}} \frac{\lambda q}{\beta u'(c(y))} f(y,e_0)\,dy = 1.$$

As a consequence, we obtain

$$\frac{\lambda}{u'(c_0)} = \int_{\underline{y}}^{\overline{y}} \frac{\lambda q}{\beta u'(c(y))} f(y,e_0)\,dy \geq \frac{\lambda q}{\beta \int_{\underline{y}}^{\overline{y}} u'(c(y)) f(y,e_0)\,dy},$$

where the inequality follows from Jensen's inequality. The inequality is in fact strict, since the agent cannot be fully insured when effort is interior. Hence, we conclude

$$\lambda\beta \int_{\underline{y}}^{\overline{y}} u'(c(y)) f(y,e_0)\,dy > \lambda q u'(c_0).$$

For $\tau^k = 0$, we have $q = \tilde{q}$, however. Therefore, the above inequality is incompatible with the agent's Euler equation (4.7). This shows that ξ cannot be zero. □

Proof of Proposition 4.3. We only show (i), since statement (ii) can be seen analogously. Define

$$g(c) := \frac{\lambda q}{\beta u'(c)} - \xi a(c).$$

Notice that $\frac{1}{u'(\cdot)}$ is always increasing. Therefore, if $\frac{1}{u'(\cdot)}$ is convex and $\xi = 0$ (or $\xi > 0$ and $a(\cdot)$ decreasing and concave), then $g(\cdot)$ is increasing and convex. Given the validity of the FOA, Proposition 4.1 (Proposition 4.2) shows that optimal consumption is defined as follows:

$$g\left(c(y)\right) = 1 + \mu\, l\left(y, e_0\right),$$

where, by assumption, the right-hand side is a positive affine transformation of a concave function. By applying the inverse function of $g(\cdot)$ to both sides, we see that $c\left(\cdot\right)$ is concave since it is an increasing and concave transformation of a concave function. □

Proof of Proposition 4.4. Given validity of the FOA, $c(y)$ and $\hat{c}(y)$ are defined as follows (see Propositions 4.1 and 4.2):

$$g_{\lambda}\left(c(y)\right) = 1 + \mu\, l\left(y, e_0\right), \text{ where } g_{\lambda}\left(c\right) := \frac{\lambda q}{\beta u'(c)}, \tag{4.10}$$

$$\hat{g}_{\hat{\lambda},\hat{\xi}}\left(\hat{c}(y)\right) = 1 + \hat{\mu}\, l\left(y, e_0\right), \text{ where } \hat{g}_{\hat{\lambda},\hat{\xi}}\left(c\right) := \frac{\hat{\lambda} q}{\beta u'(c)} - \hat{\xi} a(c), \text{ with } \hat{\xi} > 0. \tag{4.11}$$

Since $l\left(y, e\right)$ is linear in y by assumption, concavity of $\hat{c}$ is equivalent to convexity of $\hat{g}_{\hat{\lambda},\hat{\xi}}$. Moreover, since $a(c)$ is convex in c by assumption, convexity of $\hat{g}_{\hat{\lambda},\hat{\xi}}$ implies convexity of $g_{\lambda} = \frac{\lambda}{\hat{\lambda}}\left(\hat{g}_{\hat{\lambda},\hat{\xi}} + \hat{\xi} a\right)$. Finally, notice that convexity of g_{λ} is equivalent to concavity of c, since $l\left(y, e\right)$ is linear in y. □

Proof of Proposition 4.5. Recall that we have

$$g_{\lambda}\left(c(y)\right) = 1 + \mu\, l\left(y, e_0\right), \tag{4.12}$$

$$\hat{g}_{\hat{\lambda},\hat{\xi}}\left(\hat{c}(y)\right) = 1 + \hat{\mu}\, l\left(y, e_0\right), \tag{4.13}$$

where the functions g_{λ} and $\hat{g}_{\hat{\lambda},\hat{\xi}}$ are defined as in (4.10) and (4.11), respectively. First, suppose that $\hat{c}$ is a concave transformation of l. Since the right-hand side of (4.13) is a positive affine transformation of l, this implies that $\hat{g}_{\hat{\lambda},\hat{\xi}}$ is convex. Now, notice that convexity of $\hat{g}_{\hat{\lambda},\hat{\xi}}$ implies that $g_{\lambda}\left(c\right) = \frac{\lambda}{\hat{\lambda}}\left(\hat{g}_{\hat{\lambda},\hat{\xi}}(c) + \hat{\xi} a(c)\right)$ is convex as well (since $a(c)$ is convex by assumption). Hence, using (4.12), we see that c is a concave transformation of l.

Conversely, suppose that c is a convex transformation of l. Using (4.12), we see that g_{λ} is then concave. Convexity of $a(c)$ implies that $\hat{g}_{\hat{\lambda},\hat{\xi}}$ is then also concave, which shows that $\hat{c}$ is a convex transformation of l. □

Proof of Lemma 4.1. If u belongs to the HARA class, we obtain

$$\hat{g}_{\hat{\lambda},\hat{\xi}}(c) = \frac{\hat{\lambda}}{\lambda} g_\lambda(c) - \hat{\xi} a(c) = \frac{\hat{\lambda}}{\lambda} g_\lambda(c) - \hat{\xi} \lambda^{\frac{1}{\gamma}} \kappa \left(g_\lambda(c)\right)^{-\frac{1}{\gamma}}, \text{ with } \kappa = \left[\frac{\gamma q}{\beta \rho (1-\gamma)}\right]^{\frac{1}{\gamma}} > 0.$$

In other words, we have

$$\hat{g}_{\hat{\lambda},\hat{\xi}}(c) = h\left(g_\lambda(c)\right), \text{ where } h\left(g\right) = \frac{\hat{\lambda}}{\lambda} g - \hat{\xi} \lambda^{\frac{1}{\gamma}} \kappa\, g^{-\frac{1}{\gamma}}.$$

The second derivative of h with respect to g is $-\frac{\hat{\xi}\lambda^{\frac{1}{\gamma}}\kappa}{\gamma}\left(\frac{1}{\gamma}+1\right) g^{-\frac{1}{\gamma}-2}$, which is negative whenever $\gamma \geq -1$. □

Proof of Proposition 4.6. Recall that consumption is determined as follows:

$$\begin{aligned} g_\lambda\left(c(y)\right) &= 1 + \mu\, l\left(y, e_0\right), \\ \hat{g}_{\hat{\lambda},\hat{\xi}}\left(\hat{c}(y)\right) &= 1 + \hat{\mu}\, l\left(y, e_0\right). \end{aligned}$$

As a consequence, we can relate the two consumption functions as follows:

$$\frac{1}{\mu}\Big(g_\lambda\left(c(y)\right) - 1\Big) = \frac{1}{\hat{\mu}}\Big(\hat{g}_{\hat{\lambda},\hat{\xi}}\left(\hat{c}(y)\right) - 1\Big). \tag{4.14}$$

Now the result follows from the simple fact that convexity/concavity is preserved under positive affine transformations. □

Proof of Corollary 4.1. Let $\hat{c}$ be g_λ-concave. By Lemma 4.1, we have $\hat{g}_{\hat{\lambda},\hat{\xi}} = h \circ g_\lambda$ for some increasing and concave function h. Hence, when $\hat{c}$ is g_λ-concave, then $\hat{c}$ must also be $\hat{g}_{\hat{\lambda},\hat{\xi}}$-concave. Now Proposition 4.6 implies that c is g_λ-concave

To verify the second statement, let c be g_λ-convex. From Proposition 4.6, we see that $\hat{c}$ is $\hat{g}_{\hat{\lambda},\hat{\xi}}$-convex, i.e., $\hat{g}_{\hat{\lambda},\hat{\xi}} \circ \hat{c}$ is convex. By Lemma 4.1, we have $\hat{g}_{\hat{\lambda},\hat{\xi}} = h \circ g_\lambda$ for some increasing and concave function h. Since the inverse of h must be convex, we conclude that $g_\lambda \circ \hat{c} = h^{-1} \circ \hat{g}_{\hat{\lambda},\hat{\xi}} \circ \hat{c}$ is convex. □

Bibliography

ABEL, A. B. (1990): "Asset Prices under Habit Formation and Catching up with the Joneses," *American Economic Review*, 80(2), 38–42.

ABRAHAM, A., S. KOEHNE, AND N. PAVONI (2010): "On the First-Order Approach in Principal-Agent Models with Hidden Borrowing and Lending," University College London. Mimeo.

ABRAHAM, A., AND N. PAVONI (2009): "Principal-Agent Relationships with Hidden Borrowing and Lending: The First-Order Approach in Two Periods," University College London, January 2009. Mimeo. http://www.ucl.ac.uk/∼uctpnpa/FOC.pdf.

ALBANESI, S., AND C. SLEET (2006): "Dynamic Optimal Taxation with Private Information," *Review of Economic Studies*, 73(1), 1–30.

AVRIEL, M., W. E. DIEWERT, S. SCHAIBLE, AND I. ZANG (1988): *Generalized Concavity.* Plenum Publishing Corporation, New York.

BERTOLA, G., AND W. KOENIGER (2009): "Public and Private Insurance: Theory and Cross-Country Facts," Queen Mary University and Collegio Carlo Alberto. Mimeo.

BOYD, S., AND L. VANDENBERGHE (2004): *Convex optimization.* Cambridge University Press.

Campbell, J. Y., and J. H. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107(2), 205–251.

Carroll, C. D., J. Overland, and D. N. Weil (1997): "Comparison Utility in a Growth Model," *Journal of Economic Growth*, 2(4), 339–67.

——— (2000): "Saving and Growth with Habit Formation," *American Economic Review*, 90(3), 341–355.

Chade, H. (2009): "Moral Hazard with Unobservable Consumption," Arizona State University. Mimeo.

Clark, A. E. (1999): "Are wages habit-forming? evidence from micro data," *Journal of Economic Behavior & Organization*, 39(2), 179–200.

Conesa, J. C., and D. Krueger (2006): "On the optimal progressivity of the income tax code," *Journal of Monetary Economics*, 53(7), 1425–1450.

Conlon, J. R. (2009): "Two New Conditions Supporting the First-Order Approach to Multisignal Principal-Agent Problems," *Econometrica*, 77(1), 249–278.

Constantinides, G. M. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98(3), 519–43.

Diamond, P. A., and J. A. Mirrlees (1978): "A model of social insurance with variable retirement," *Journal of Public Economics*, 10(3), 295 – 336.

Domeij, D., and M. Floden (2006): "The Labor-Supply Elasticity and Borrowing Constraints: Why Estimates are Biased," *Review of Economic Dynamics*, 9(2), 242–262.

Eaton, J., and H. S. Rosen (1980): "Optimal Redistributive Taxation and Uncertainty," *The Quarterly Journal of Economics*, 95(2), 357–64.

Frederick, S., and G. Loewenstein (1999): "Hedonic Adaptation," in *Well-being: The foundations of hedonic psychology*, ed. by D. Kahneman, E. Diener, and N. Schwarz, pp. 302–329. Russell Sage Foundation Press.

Gollier, C. (2001): *The Economics of Risk and Time*. MIT Press.

Golosov, M., N. Kocherlakota, and A. Tsyvinski (2003): "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, 70(3), 569–587.

Golosov, M., and A. Tsyvinski (2006): "Designing Optimal Disability Insurance: A Case for Asset Testing," *Journal of Political Economy*, 114(2), 257–279.

Gottardi, P., and N. Pavoni (2010): "Ramsey Asset Taxation under Asymmetric Information," University College London. Mimeo.

Grochulski, B., and N. Kocherlakota (2010): "Nonseparable preferences and optimal social security systems," *Journal of Economic Theory*, forthcoming.

Grossman, S. J., and O. D. Hart (1983): "An Analysis of the Principal-Agent Problem," *Econometrica*, 51(1), 7–45.

Grund, C., and D. Sliwka (2007): "Reference-Dependent Preferences and the Impact of Wage Increases on Job Satisfaction: Theory and Evidence," *Journal of Institutional and Theoretical Economics (JITE)*, 163(2), 313–335.

Guiso, L., and M. Paiella (2008): "Risk Aversion, Wealth, and Background Risk," *Journal of the European Economic Association*, 6(6), 1109–1150.

Guvenen, F., B. Kuruscu, and S. Ozkan (2009): "Taxation of Human Capital and Wage Inequality: A Cross-Country Analysis," NBER Working Papers 15526, National Bureau of Economic Research, Inc.

Helson, H. (1964): *Adaptation-Level Theory.* Harper & Row New York.

Holmström, B. (1979): "Moral hazard and observability," *Bell Journal of Economics*, 10(1), 74–91.

Hopenhayn, H., and A. Jarque (2010): "Unobservable Persistent Productivity and Long Term Contracts," *Review of Economic Dynamics*, 13(2), 333–349.

Jarque, A. (2010): "Repeated Moral Hazard with Effort Persistence," *Journal of Economic Theory*, forthcoming.

Jewitt, I. (1988): "Justifying the First-Order Approach to Principal-Agent Problems," *Econometrica*, 56(5), 1177–1190.

Karlin, S. (1968): *Total Positivity. Volume I.* Stanford, California: Stanford University Press.

Kocherlakota, N. R. (2004): "Figuring out the impact of hidden savings on optimal unemployment insurance," *Review of Economic Dynamics*, 7(3), 541–554.

Kocherlakota, N. R. (2005): "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, 73(5), 1587–1621.

Kwon, I. (2006): "Incentives, Wages, and Promotions: Theory and Evidence," *RAND Journal of Economics*, 37(1), 100–120.

Lahiri, A., and M. Puhakka (1998): "Habit Persistence in Overlapping Generations Economies under Pure Exchange," *Journal of Economic Theory*, 78(1), 176–186.

Messinis, G. (1999): "Habit Formation and the Theory of Addiction," *Journal of Economic Surveys*, 13(4), 417–42.

Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38(114), 175–208.

——— (1974): "Notes on Welfare Economics, Information and Uncertainty," in *Essays in Economic Behavior Under Uncertainty*, ed. by M. Balch, D. McFadden, and S. Wu. North-Holland, Amsterdam.

——— (1979): "The Implications of Moral Hazard for Optimal Insurance," Seminar given at Conference held in honor of Karl Borch, Bergen, Norway. Mimeo.

Mukoyama, T., and A. Sahin (2005): "Repeated moral hazard with persistence," *Economic Theory*, 25(4), 831–854.

Rogerson, W. P. (1985a): "The First-Order Approach to Principal-Agent Problems," *Econometrica*, 53(6), 1357–1367.

——— (1985b): "Repeated Moral Hazard," *Econometrica*, 53(1), 69–76.

Rozen, K. (2010): "Foundations of Intrinsic Habit Formation," *Econometrica*, forthcoming.

Ryder, Jr., H. E., and G. M. Heal (1973): "Optimum Growth with Intertemporally Dependent Preferences," *Review of Economic Studies*, 40(1), 1–33.

Sadka, E. (1976): "On Progressive Income Taxation," *American Economic Review*, 66(5), 931–35.

Saez, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68(1), 205–29.

Varian, H. R. (1980): "Redistributive taxation as social insurance," *Journal of Public Economics*, 14(1), 49–68.

Wendner, R. (2003): "Do habits raise consumption growth?," *Research in Economics*, 57(2), 151–163.

Werning, I. (2001): "Repeated Moral-Hazard with Unmonitored Wealth: A Recursive First-Order Approach," MIT. Mimeo. http://econ-www.mit.edu/files/1264.

——— (2002): "Optimal Unemployment Insurance with Unobservable Savings," MIT. Mimeo. http://econ-www.mit.edu/files/1267.

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig angefertigt und mich keiner anderen als der in ihr angegebenen Hilfsmittel bedient zu haben. Insbesondere sind sämtliche Zitate aus anderen Quellen als solche gekennzeichnet und mit Quellenangaben versehen.

Mannheim, 6. Mai 2010

Sebastian Köhne

Lebenslauf – Sebastian Köhne

10/2005 - 06/2010	Promotionsstudium im Fach Volkswirtschaftslehre an der Universität Mannheim, Center for Doctoral Studies in Economics and Management (CDSE)
07/2005	Diplom-Wirtschaftsmathematiker, Universität Bielefeld
02/2004 - 08/2004	Auslandsstudium an der University of New South Wales in Sydney, Australien
10/2000 - 07/2005	Studium im Fach Wirtschaftsmathematik an der Universität Bielefeld
05/1999	Abitur, Städtisches Gymnasium Steinheim