MDPI

Article

Salience Bias and Overwork

Fabio Römeis [1], **Fabian Herweg** [2] and **Daniel Müller** [1,*]

1 Chair for Contract Theory and Information Economics, University of Würzburg, Sanderring 2,
D-97070 Würzburg, Germany; fabio.roemeis@uni-wuerzburg.de
2 Faculty of Law, Business and Economics, University of Bayreuth and CESifo, Universitätsstr. 30,
D-95440 Bayreuth, Germany; fabian.herweg@uni-bayreuth.de
* Correspondence: daniel.mueller2@uni-wuerzburg.de

Abstract: In this study, we enrich a standard principal–agent model with hidden action by introducing salience-biased perception on the agent's side. The agent's misguided focus on salient payoffs, which leads the agent's and the principal's probability assessments to diverge, has two effects: First, the agent focuses too much on obtaining a bonus, which facilitates incentive provision. Second, the principal may exploit the diverging probability assessments to relax participation. We show that salience bias can reverse the nature of the inefficiency arising from moral hazard; i.e., the principal does not necessarily provide insufficient incentives that result in inefficiently low effort but instead may well provide excessive incentives that result in inefficiently high effort.

1. Introduction

One of the iconic workhorse models of modern contract theory used to analyze incentive provision and work effort is the moral hazard model with hidden action where both the principal and the wealth-constrained agent have linear utility functions for money.[1] This model, which dates back to work by Innes [2], Baliga and Sjöström [3], and Pitchford [4], has become increasingly popular and was, despite its simplicity, fruitfully applied to a plethora of topics relevant for economics and management science.[2] The key insight of the basic static moral hazard model with hidden action is that the principal provides the agent with inefficiently low incentives, which results in the agent exerting inefficiently low effort, i.e., "underwork" prevails.

The model's singular focus on the agent's work effort being too low, however, seems disconcerting given that several studies point to overwork as a very serious problem. For example, conducting telephone interviews with 1003 wage and salaried employees in the U.S. workforce, the representative study by Galinsky et al. [22] finds that 44% of U.S. employees were overworked or felt overwhelmed by how much work they had to do. Furthermore, Galinsky et al. [22] find that a higher degree of being overworked is associated with the experience of higher levels of stress and of a higher number of symptoms of clinical depression, which other studies have associated with non-negligible personal costs for the overworked employee such as health costs or a strain on their family life.[3] Notably, the causes for employees being overworked seem primarily rooted in employees exerting effort on non-contractable aspects of their respective employment relationship, such as moving quickly from task to task with little time for recovery between tasks, being permanently accessible by cell phone or email outside normal work hours, or continuing work during their holidays in order to keep up with the demands of their job. In order to devise meaningful measures to prevent (or at least reduce) overwork, it seems worthwhile, if not necessary, to first explore what channels potentially give rise to inefficiently high incentive provision and, as a consequence, inefficiently high effort. In order to address this question, we use the iconic moral hazard model with hidden action, where the principal and

a wealth-constrained agent both have linear utility functions for money. This model allows to step away from risk-sharing motives and thus to focus on inefficient effort provision.

To this end, in this paper, we extend the simplest version of said model by allowing the agent's perception of occurrence probabilities to be blurred by the salience of the associated outcomes. According to salience theory as set out by Bordalo et al. [25], a decision maker's attention is unknowingly drawn toward (away from) very salient (rather non-salient) outcomes, which leads to the occurrence probabilities of these outcomes being perceived as inflated (deflated). This implies that the agent inflates the probabilities of those states in which his effort choice has an actual impact on his wage. In other words, the agent may focus too much on receiving a high wage (i.e., obtaining a bonus payment) which facilitates incentive provision. Moreover, the agent's biased probability assessment differs from the objective assessment held by the principal. The principal may exploit these diverging assessments between her and the agent in order to relax the participation constraint. We show that this salience-induced misperception may reverse the nature of the inefficiency compared with the standard model, i.e., that overwork can prevail. Notably, overwork (but not underwork) may occur even in situations where, in the standard model, the first-best effort is always implemented.[4]

At first glance, it might seem at odds with our model that a salience-biased agent feels overworked, as found in the study by Galinsky et al. [22], as he himself considers accepting the principal's contract offer and exerting high effort as the optimal course of action at the moment when he makes these decisions. As we will come to understand, however, if the agent exerts a high level of effort under the optimal contract, then either the participation constraint or the incentive compatibility constraint is binding. A reduction in the agent's salience bias (i.e., a reduction in the agent's focus on receiving the bonus payment) would then lead to either the agent rejecting the contract in question or the agent exerting low rather than high effort under said contract. Thus, if the salience-induced misdirection of the agent's attention to certain aspects of a particular decision fades out once the decision has been taken, the agent, in retrospect, may realize that he has overextended himself in the past by exerting more effort than he now deems worthwhile or by agreeing to accept the principal's contract offer in the first place.

The rest of this paper is structured as follows: After surveying the related literature, we present our model of moral hazard with hidden action and salience-biased perception in Section 2. In Section 3 we analyze the impact of salience bias on incentive provision. We conclude in Section 4. Proofs are deferred to the Appendix A.

Related Literature

Our paper, which incorporates salience theory for choice under risk [25] into an otherwise standard textbook model of static moral hazard with hidden action, adds to two strands of literature. First, our analysis contributes to the vast literature that explores incentive provision in situations with moral hazard. Irrespective of whether the trade-off between efficiency and rent-extraction or the trade-off between incentive provision and risk sharing [28,29] is analyzed, in this literature, underwork is the predominant prediction. In contrast to our paper, the few extant contributions that provide an explanation for overwork rely on a dynamic principal–agent relationship. Analyzing a two-period model, Kräkel and Schöttner [30] find that the agent being replaceable can result in excessive effort (i.e., overwork) in the first period.[5] Englmaier et al. [32] find that excessive effort can also occur in a relational contract between a rational principal and a naïve present-biased agent. In that case, if limited liability imposes a binding restriction, the principal may optimally resort to extracting the agent's rent by inducing the agent to work harder than the efficient effort level (and harder than an agent without a present bias). We complement these contributions by showing that overwork can also prevail in a static one-shot interaction of the principal and the agent.[6]

Second, our paper adds to the growing literature on applied theoretical models that incorporate salience theory. By now, there is a plethora of contributions that apply salience

theory for risk-less consumer choice [34] in models of industrial organization.[7] In particular, regarding applications with an incentive-theoretical flavor, Adrian [36] analyzes the screening problem of a monopolist who deals with a salience-biased consumer in an adverse selection environment. Incentive theoretical contributions of salience theory for choice under risk [25], on the other hand, are rather scarce and so far restricted to law-and-economics literature. Specifically, Friehe and Pham [37] analyze the settling behavior of salience-biased defendants, and Mungan [38] explores whether the certainty of a fine or the severity of a fine has a higher deterrent effect on salience-biased offenders. Our paper complements the aforementioned contributions by providing the first analysis of the impact of salience bias on incentive provision and efficiency in a situation with moral hazard, where the agent makes his choice under risk.

Apart from these two strands of literature, conceptually closest to our paper is the analysis in de la Rosa [39], which incorporates an overconfident agent with biased probability assessment into an otherwise standard static moral hazard model. Specifically, the overconfident agent overestimates the marginal contribution of his effort on the probability of success and thus behaves similar to an agent who has a misguided focus on salient payoffs. As a consequence, according to de la Rosa [39], the agent's overconfidence may help the principal to satisfy incentive compatibility or participation. Importantly, while the diverging probability assessments of the agent and the principal are exogenously given in de la Rosa [39], they are endogenously determined in our setup by the interplay of the agent's salience-biased perception and the specifics of the contract offered by the principal. Moreover, we abstract from risk-sharing motives and assume that the agent is protected by limited liability, whereas de la Rosa [39] assumes that the agent is risk-averse but with deep pockets. Not surprisingly, de la Rosa [39] therefore focuses on how the agent's overconfidence affects the trade-off between risk sharing and incentive provision, whereas our analysis explores how a salience-biased probability assessment affects the agent's limited-liability rent. Finally, and most importantly, de la Rosa [39] defines the first-best outcome as the maximization of the principal's expected profit plus the agent's perceived (i.e., biased) expected utility, whereas we consider salience-induced distortions as a bias that should not obtain any normative weight. Due to these diverging definitions of the first-best outcome, de la Rosa [39] cannot answer the question under which circumstances we observe excessive incentives and overwork.

2. Model

2.1. Effort and Production

A principal (P) wants to hire an agent (A) to work on a task on her behalf. The value V generated by A accrues to P and can be either low, $V = V_L$, or high, $V = V_H$, where $0 < V_L < V_H$. If A exerts effort $e \in \{0, 1\}$, he incurs non-monetary cost $\Psi(e) = \psi e$, where $\psi > 0$.[8] As can be seen from Table 1, the actual realization of the value V depends on A's effort choice $e \in \{0, 1\}$ and which state $s \in \mathcal{S} \equiv \{1, 2, 3\}$ of three possible, mutually exclusive states is the true state of the world. Specifically, letting the actual realization of the value being denoted by $V(e, s)$, we obtain $V(e, 1) = V_L$ and $V(e, 3) = V_H$ for all $e \in \{0, 1\}$, i.e., the realized value is invariably low in state 1 and invariably high in state 3. In state 2, on the other hand, the realized value is low if A exerted low effort and high if A exerted high effort, i.e., $V(0, 2) = V_L$ and $V(1, 2) = V_H$. The probability of state $s \in \mathcal{S}$ being the true state of the world is $p_s \in (0, 1)$, where $\sum_{s \in \mathcal{S}} p_s = 1$.[9]

Table 1. Dependence of value V on effort e and state of the world s.

	$s = 1$	$s = 2$	$s = 3$
$e = 0$	V_L	V_L	V_H
$e = 1$	V_L	V_H	V_H

2.2. Information and Contracts

Neither A's effort choice nor the realization of the state of the world is verifiable. The realization of value V is verifiable. Hence, a contract specifies a transfer t from P to A, which is conditional on the realization of the value V. As the value V is a function of e and s, the transfer t is effectively also a function of e and s. Specifically,

$$t(e,s) = \begin{cases} t_L & \text{if} \quad V(e,s) = V_L, \\ t_H & \text{if} \quad V(e,s) = V_H. \end{cases} \tag{1}$$

The agent is protected by limited liability such that t_L and t_H must not be below $\bar{t} \in \mathbb{R}$. In order to ensure that a contract is offered in equilibrium, we assume that $\bar{t} \leq V_L$.

2.3. Sequence of Events

First, P makes a take-it-or-leave-it contract offer to A, which A decides to accept or reject. Second, if A accepts the offer, he decides which effort $e \in \{0,1\}$ to exert. Thereafter, the state of the world $s \in \mathcal{S}$ and, thus, the value $V(e,s)$ is realized, and the transfer $t(e,s)$ is paid. If A rejects the offer, each party obtains her/his outside option that yields a reservation utility equal to zero.

2.4. Preferences and Salience Bias

Both P and A have a linear utility function for monetary outcomes. The principal perceives the occurrence probabilities of the three states correctly, and her expected value from signing a contract that stipulates transfers t_L and t_H and under which A exerts effort $e \in \{0,1\}$ is given by

$$U_P(t_L, t_H, e) = \sum_{s \in \mathcal{S}} p_s [V(e,s) - t(e,s)]. \tag{2}$$

The agent's perception of the occurrence probabilities of the three states is blurred by the salience of the respective outcome combinations in the spirit of salience theory [25]. Specifically, A's expected value from signing a contract that stipulates t_L and t_H and under which he exerts $e \in \{0,1\}$ is given by

$$U_A(t_L, t_H, e) = \sum_{s \in \mathcal{S}} \pi_s [t(e,s) - \Psi(e)], \tag{3}$$

where

$$\pi_s = \frac{\delta^{k_s}}{\sum_{s' \in \mathcal{S}} \delta^{k_{s'}} p_{s'}} \cdot p_s \quad \text{with} \quad \delta \in (0,1). \tag{4}$$

Here, $\pi_s \in (0,1)$ with $\sum_{s \in \mathcal{S}} \pi_s = 1$ is the decision weight that A attaches to state $s \in \mathcal{S}$, $\delta \in (0,1)$ denotes the (inverse) degree of salience bias, and $k_s \in \mathbb{N}_{>0}$ denotes the salience rank of state $s \in \mathcal{S}$. The salience ranking starts at 1, has no jumps, and a lower salience rank denotes higher salience. Hence, above-average (below-average) salience translates into the respective state's objective occurrence probability being perceived as inflated (deflated). Notably, the standard model with A being a risk-neutral expected utility maximizer is captured by $\delta = 1$, in which case A perceives the occurrence probabilities correctly. [10]

We assume that a state's salience is fully determined by the contractually specified transfers that A might receive in that state.[11] Formally, letting $\sigma(\cdot, \cdot)$ denote the salience function, we obtain $k_s \lesseqgtr k_{s'}$ if and only if $\sigma(t(0,s), t(1,s)) \gtreqless \sigma(t(0,s'), t(1,s'))$ for all $s, s' \in \mathcal{S}$. The function $\sigma(\cdot, \cdot)$ is symmetric (i.e., $\sigma(t, t') = \sigma(t', t)$ for all $t, t' \in \mathbb{R}$), assigns identical values to states in which the transfers for both effort levels coincide (i.e., $\sigma(t, t) = \sigma(t', t')$ for all $t, t' \in \mathbb{R}$), and satisfies the following "ordering" property, which captures psychological contrast effects:[12]

(O)   For all $t, t', t'', t''' \in \mathbb{R}$, if $[\min\{t, t'\}, \max\{t, t'\}] \subsetneq [\min\{t'', t'''\}, \max\{t'', t'''\}]$, then $\sigma(t, t') < \sigma(t'', t''')$.

Hence, $k_1 = k_3 \geq k_2$ with the inequality being strict in the case of performance-dependent pay (i.e., for $t_L \neq t_H$). In other words, for performance-independent pay $(t_L = t_H)$, A perceives the probabilities correctly, i.e., $\pi_s = p_s$ for $s = 1, 2, 3$. For performance-dependent pay $(t_L \neq t_H)$, A focuses too much on (over-weighs) state $s = 2$, where payments are different, and puts too little attention on (under-weighs) states $s = 1$ and $s = 3$, where payments are identical. Formally,

$$\pi_2 = \frac{1}{p_2 + \delta(p_1 + p_3)}\, p_2 > p_2 \quad \text{and} \quad \pi_i = \frac{\delta}{p_2 + \delta(p_1 + p_3)}\, p_i < p_i \ \text{ for } i = 1, 3. \tag{5}$$

2.5. Benchmark 1: First-Best Effort

Similar to overconfidence, salience bias represents an unknowing mistake in the assessment of the decision environment. Therefore, in the spirit of the literature that studies transactions in the presence of overconfidence [47–49], our welfare criterion is the unweighted sum of unbiased (expected) utilities, which corresponds to the expected material gains from trade:

$$G(e) = \sum_{s \in \mathcal{S}} p_s \big[V(e, s) - \Psi(e) \big].$$

Observation 1. *The first-best effort level $e^{FB} = \arg\max_{e \in \{0,1\}} G(e)$ is given by*

$$e^{FB} = \begin{cases} 1 & \text{if} \quad \Delta^{FB} \leq V_H - V_L, \\ 0 & \text{if} \quad V_H - V_L \leq \Delta^{FB}, \end{cases} \quad \text{where} \quad \Delta^{FB} := \frac{\psi}{p_2}.$$

Proof. See Laffont and Martimort [41]. □

Intuitively, high effort rather than low effort should be exerted if the associated increase in the expected "extra value", $p_2(V_H - V_L)$, weakly exceeds the associated increase in effort cost, ψ.

3. Analysis

When making her contract offer, P must decide which effort level, $e \in \{0, 1\}$, she wants to induce and what transfers, t_L and t_H, she wants to contractually specify in order to maximize her expected profit. Formally, she solves the following optimization program:

$$\max_{e, t_L, t_H} U_P(t_L, t_H, e)$$

subject to

$$U_A(t_L, t_H, e) \geq U_A(t_L, t_H, e') \ \text{ for } e' \in \{0, 1\}, \ e \neq e' \tag{IC$_e$}$$
$$U_A(t_L, t_H, e) \geq 0 \tag{P$_e$}$$
$$t_L \geq \bar{t}, \ t_H \geq \bar{t}. \tag{LL}$$

The incentive compatibility constraint (IC$_e$), the participation constraint (P$_e$), and the limited liability constraint (LL) bear the usual interpretation.

In order to explore the implications of salience bias for effort provision, we will follow the approach pioneered by Grossman and Hart [29] and decompose P's optimization program into two steps. First, for each effort level, we identify the "cost-minimizing" transfer payments; i.e., those transfer payments that implement the respective effort level at the lowest possible expected transfer payment. Thereafter, in order to determine the overall optimal contract, we compare P's expected utility across effort levels, given that each

effort level is implemented with the respective cost-minimizing transfers. Before launching into this analysis, however, as a second benchmark, we briefly review the outcome of the principal–agent relationship in the absence of salience bias.

3.1. Benchmark 2: No Salience Bias

Without salience bias (i.e., for $\delta = 1$), A perceives the probabilities correctly and thus is risk neutral. Letting $\tilde{e}$ denote the effort level that P optimally implements in the absence of salience bias, the outcome of the interaction between P and A takes the following form:

Observation 2. *The principal implements high effort ($\tilde{e} = 1$) if and only if $\tilde{\Delta} \leq V_H - V_L$, where*

$$\tilde{\Delta} := \begin{cases} \frac{p_2 + p_3}{p_2} \Delta^{FB} & if \quad -\frac{p_3}{p_2}\psi < \bar{t}, \\ \Delta^{FB} & if \quad \bar{t} \leq -\frac{p_3}{p_2}\psi. \end{cases} \tag{6}$$

Proof. See Laffont and Martimort [41]. □

Intuitively, if $\bar{t}$ is sufficiently low (i.e., if $\bar{t} \leq -\frac{p_3}{p_2}\psi$), then limited liability does not impose a binding restriction, and, irrespective of which effort level P implements, A's expected utility will equal his reservation utility. In this case, P's problem boils down to maximizing the expected material gains from trade $G(\cdot)$, and, thus, she induces A to exert the materially efficient effort level, i.e., $\tilde{e} = e^{FB}$. Formally, this follows from $\tilde{\Delta}$ being equal to Δ^{FB}.

If, on the other hand, $\bar{t}$ is sufficiently high (i.e., if $\bar{t} > -\frac{p_3}{p_2}\psi$), then limited liability imposes a binding restriction. As a consequence, in case high effort is implemented, the agent does not only receive the minimum wage, $\bar{t}$, or is compensated for his outside option, but he obtains an information rent, which is rooted in the necessity to induce high effort by specifying a sufficiently high bonus payment for good performance ($t_H = \frac{\psi}{p_2} + \bar{t} > \bar{t} = t_L$). Thus, P now trades off maximizing the expected gains from trade and cutting back on A's information rent, which leads to P implementing high effort "less often" than is materially efficient. This well-known result of underwork formally follows from $\tilde{\Delta}$ being strictly smaller than Δ^{FB}. Here, if $\Delta^{FB} \leq V_H - V_L < \tilde{\Delta}$, then $\tilde{e} < e^{FB}$. In summary, in the standard model without salience bias, it is impossible for overwork to prevail.

3.2. Cost-Minimizing Transfers with Salience Bias

The cost-minimization problem for a given effort $e \in \{0, 1\}$ is as follows:

$$\min_{(t_L, t_H) \in \mathbb{R}^2} p_1 t_L + p_2[(1 - e)t_L + e t_H] + p_3 t_H \quad \text{subject to} \quad (\text{IC}_e), (\text{P}_e), (\text{LL}). \tag{7}$$

In the main section of this paper, we are going to provide an intuitive derivation of the cost-minimizing transfer combination based on a graphical analysis. To this end, consider the isocost curve associated with the expected transfer payment T, given effort level e, i.e., the geometric location in (t_L, t_H)−space of all transfer combinations (t_L, t_H) that results in the same expected transfer payment T given A exerts effort e. Formally, this isocost curve is described by

$$\Gamma(t_L | e, T) = \frac{T}{e p_2 + p_3} - \frac{p_1 + (1 - e)p_2}{e p_2 + p_3} t_L. \tag{8}$$

From (8), it follows that the family of isocost curves comprises parallel, negatively sloped straight lines. Importantly, an isocost curve with a smaller vertical intercept is associated with a strictly lower expected transfer payment than an isocost curve with a greater vertical intercept. As a consequence, the solution to the above cost-minimization problem is a transfer combination (t_L^e, t_H^e) that satisfies (IC_e), (P_e), and (LL) such that there

is no other transfer combination $(\hat{t}_L, \hat{t}_H)$ that satisfies (IC$_e$), (P$_e$), and (LL) and lies on a lower isocost curve than (t_L^e, t_H^e).

3.2.1. Implementation of High Effort ($e = 1$)

If P induces $e = 1$, then (IC$_1$) and (P$_1$) take the following form:

$$t_H \geq \frac{\psi}{\pi_2} + t_L =: \Phi_1^{IC}(t_L) \tag{IC$_1$}$$

$$t_H \geq \frac{\psi}{\pi_2 + \pi_3} - \frac{\pi_1}{\pi_2 + \pi_3} t_L =: \Phi_1^{P}(t_L). \tag{P$_1$}$$

As becomes apparent from (IC$_1$), for a contract to be incentive compatible, the transfer t_H must be sufficiently higher than the transfer t_L. Figures 1 and 2 depict the functions $\Phi_1^{IC}(t_L)$ and $\Phi_1^{P}(t_L)$ in (t_L, t_H)-space together with the (LL) constraint. The (LL) constraint requires (t_L, t_H) to be located to the north-east of the orange L-shaped curve, the kink of which lies on the 45 degree line. Notably, $d\Gamma(t_L|1, \cdot)/dt_L < d\Phi_1^{P}(t_L)/dt_L$. The gray-shaded areas in Figures 1 and 2 represent the set of transfer combinations that jointly satisfy (IC$_1$), (P$_1$), and (LL) for $\bar{t} \geq t_L^{ST}$ and $\bar{t} < t_L^{ST}$, respectively, where

$$t_L^{ST} := -\frac{\pi_3}{\pi_2} \psi. \tag{9}$$

Here, t_L^{ST} corresponds to the unique value of transfer t_L, for which $\Phi_1^{IC}(\cdot)$ and $\Phi_1^{P}(\cdot)$ intersect, i.e., denoting the corresponding value of transfer t_H with t_H^{ST}. We can obtain $\Phi_1^{IC}(t_L^{ST}) = \Phi_1^{P}(t_L^{ST}) =: t_H^{ST}$. Thus, the transfer combination $(t_L, t_H) = (t_L^{ST}, t_H^{ST})$ is the only transfer combination for which both the participation constraint and the incentive compatibility constraint are satisfied with equality.
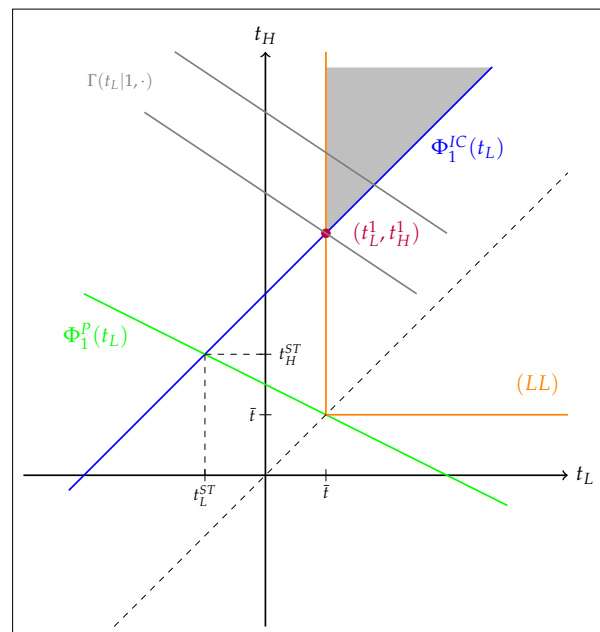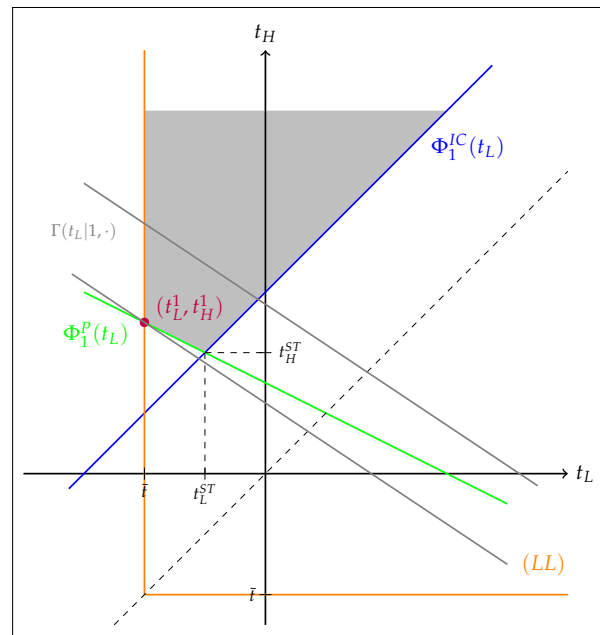

Figure 1. $e = 1$ and $\bar{t} \geq t_L^{ST}$.

Figure 2. $e = 1$ and $\bar{t} < t_L^{ST}$.

In fact, the transfer combination $(t_L, t_H) = (t_L^{ST}, t_H^{ST})$ will be our staring point to intuitively understand the structure of the cost-minimizing contract in order to implement high effort.

Proposition 1. *The cost-minimizing transfers t_L^1 and t_H^1 to implement $e = 1$ satisfy*

$$t_L^1 = \bar{t} \qquad and \qquad t_H^1 = \begin{cases} \bar{t} + \frac{\psi}{\pi_2} & if \quad \bar{t} \geq t_L^{ST}, \\ -\frac{\pi_1}{\pi_2 + \pi_3}\bar{t} + \frac{\psi}{\pi_2 + \pi_3} & if \quad \bar{t} < t_L^{ST}. \end{cases} \tag{10}$$

What is the intuition behind Proposition 1? As usual, the structure of the cost-minimizing contract is shaped by those of the constraints which impose a binding restriction on P's choice of transfers and which constraints do not. So, we consider the transfer combination $(t_L, t_H) = (t_L^{ST}, t_H^{ST})$ for which both the participation constraint and the incentive compatibility constraint are satisfied with equality. If $\bar{t} > t_L^{ST}$, as depicted in Figure 1, this transfer combination violates the limited liability constraint as $t_L = t_L^{ST} < \bar{t}$. Thus, in order to satisfy the limited liability constraint, t_L must be increased. For incentive compatibility to be maintained, however, this requires t_H to be also increased appropriately, and the "cheapest" way to do so is to move to the north-east along the $\Phi_1^{IC}(\cdot)$-curve. Notably, with both transfers increasing compared to the transfer combination that we started out from, the participation constraint is slack under this contractual adjustment. As becomes apparent from Figure 1, the cost-minimizing contract is obtained once the transfer t_L just satisfies the requirement imposed by the limited liability constraint (i.e., $t_L^1 = \bar{t}$), and the transfer t_H is specified such that the incentive compatibility constraint holds with equality (i.e., $t_H^1 = \Phi_1^{IC}(\bar{t})$). Thus, under the cost-minimizing contract it is the limited liability constraint (with regard to transfer t_L) and the incentive compatibility constraint that bind, just as would be the case for an unbiased agent if the minimum transfer $\bar{t}$ were sufficiently high. There is, however, one important difference. Since incentive compatibility requires $t_H > t_L$, a salience-biased agent perceives the occurrence probability of state 2 (i.e., where he receives the "bonus payment" $t_H - t_L$ if and only if he exerted high effort) as inflated (i.e., $\pi_2 > p_2$). On the other hand, the occurrence probabilities of states 1 and 3, where the transfer does not depend on A's effort choice, are perceived as deflated (i.e., $\pi_i < p_i$ for $i = 1, 3$). As a consequence, the minimum bonus payment necessary to induce high effort is lower for an agent whose perception is blurred by salience bias than for an unbiased agent.

Next, suppose that $\bar{t} < t_L^{ST}$, as depicted in Figure 2. Under the transfer combination $(t_L, t_H) = (t_L^{ST}, t_H^{ST})$, for which both the participation constraint and the incentive compatibility constraint are satisfied with equality, the limited liability constraint is slack with regard to both transfers as $\bar{t} < t_L < t_H$. In this case, P can exploit that her own probability assessment and A's probability assessment diverge by decreasing t_L and increasing t_H in a way such that the participation constraint is still satisfied with equality, i.e., by moving to the north-west along the $\Phi_1^P(\cdot)$-curve.[13] Compared with the transfer combination that we started out from, with the bonus payment $t_H - t_L$ being increased by this contractual adjustment, the incentive compatibility constraint becomes slack, and A strictly prefers to exert high effort. Nevertheless, P's expected cost for implementing high effort is strictly reduced. The reason for this is that A, given that he exerts high effort, puts too much weight on the states where he receives the high transfer t_H (i.e., $\pi_2 + \pi_3 > p_2 + p_3$) and too little weight on the state where he receives the low transfer t_L (i.e., $\pi_1 < p_1$). As a consequence, the contractual adjustment strictly decreases the actual, objectively expected transfer payment while, at the same time, leaving the expected transfer payment as perceived by A unchanged. Graphically, this is reflected by the fact that the slopes of the isocost curves are steeper than the slope of the $\Phi_1^P(\cdot)$ curve. As becomes apparent from Figure 2, the cost-minimizing contract is obtained once the transfer t_L just satisfies the requirement imposed by the limited liability constraint (i.e., $t_L^1 = \bar{t}$) and the transfer t_H is specified such that the participation constraint holds with equality (i.e., $t_H^1 = \Phi_1^P(\bar{t})$). The exact size of the high transfer t_H depends on A's degree of salience bias. In particular, the more biased A's perception is, the lower the cost-minimizing specification of transfer t_H.[14]

The principal's expected utility under the cost-minimizing contract to implement $e = 1$ is

$$U_P(t_L^1, t_H^1, 1) = p_1 V_L + (p_2 + p_3) V_H$$
$$- \begin{cases} \left[p_1 \bar{t} + (p_2 + p_3)\left(\bar{t} + \frac{\psi}{\pi_2} \right) \right] & \text{if} \quad \bar{t} \geq t_L^{ST}, \\ \left[p_1 \bar{t} + (p_2 + p_3)\left(-\frac{\pi_1}{\pi_2 + \pi_3} \bar{t} + \frac{\psi}{\pi_2 + \pi_3} \right) \right] & \text{if} \quad \bar{t} < t_L^{ST}. \end{cases} \quad (11)$$

An interesting question, which is answered in the following corollary, is whether P can benefit from A's salience bias.

Corollary 1. *If the degree of salience bias marginally increases (i.e., if δ decreases), then $U_P(t_L^1, t_H^1, 1)$ strictly increases.*

This comparative static result is quite plausible. An increase in salience bias leads to A attaching even more weight to the state where his effort makes an actual difference to his remuneration (i.e., $\frac{d\pi_2}{d\delta} < 0$). In the case where the incentive compatibility constraint imposes a binding restriction (high $\bar{t}$), this allows P to decrease the bonus payment for good performance and, thus, to reduce her expected cost. If the incentive compatibility constraint is slack but the participation constraint is binding (low $\bar{t}$), an increase in salience bias allows P to exploit the divergence in prior beliefs even more to her own advantage. The reason for this is that A then attaches even higher value to an increase of t_H and cares even less about a decrease of t_L (i.e., $\frac{d(\pi_2 + \pi_3)}{d\delta} < 0$).

3.2.2. Implementation of Low Effort ($e = 0$)

If P induces $e = 0$, (IC$_0$) and (P$_0$) take the following form:

$$t_H \leq \frac{\psi}{\pi_2} + t_L =: \Phi_0^{IC}(t_L) \tag{IC$_0$}$$

$$t_H \geq -\frac{\pi_1 + \pi_2}{\pi_3} t_L =: \Phi_0^P(t_L). \tag{P$_0$}$$

As depicted in Figures 3 and 4, which are drawn in analogy to Figures 1 and 2, we now have $d\Phi_0^P(t_L)/dt_L < d\Gamma(t_L|0,\cdot)/dt_L$.
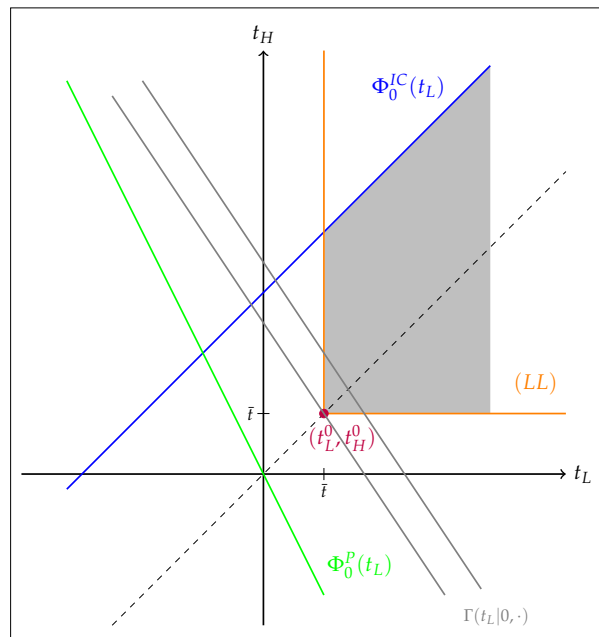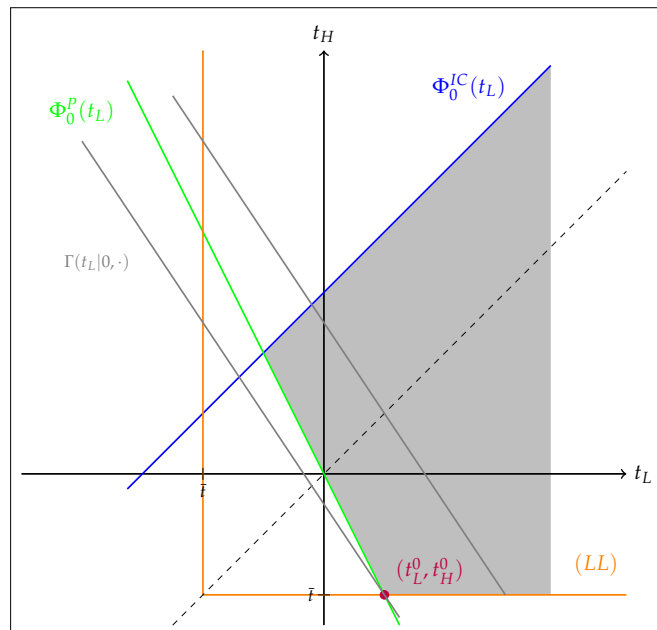

Figure 3. $e = 0$ and $\bar{t} \geq 0$.

Figure 4. $e = 0$ and $\bar{t} < 0$.

Proposition 2. *The cost-minimizing transfers t_L^0 and t_H^0 to implement $e = 0$ satisfy*

$$t_L^0 = \begin{cases} \bar{t} & \text{if} \quad \bar{t} \geq 0, \\ -\frac{\pi_3}{\pi_1 + \pi_2}\bar{t} & \text{if} \quad \bar{t} < 0, \end{cases} \qquad \text{and} \qquad t_H^0 = \bar{t}. \tag{12}$$

In order to implement low effort, there is no need for P to reward good performance with a bonus payment. Thus, the natural candidates for an optimal contract are transfer payments t_L and t_H that just satisfy the limited liability requirement, i.e., $t_L = t_H = \bar{t}$. If the transfer combination $(t_L, t_H) = (\bar{t}, \bar{t})$ satisfies the participation constraint, which is the case if the minimum transfer is sufficiently high (i.e., if $\bar{t} \geq 0$), then it is the optimal contract.

If, on the other hand, the minimum transfer is too low (i.e., if $\bar{t} < 0$), then the transfer combination $(t_L, t_H) = (\bar{t}, \bar{t})$ does not satisfy the participation constraint. In this case, the participation constraint imposes a binding restriction, and P can exploit the divergence between her own and A's probability assessments. Specifically, given that A exerts low effort, for $t_L \neq t_H$, he inflates the probability of receiving t_L (i.e., $\pi_1 + \pi_2 > p_1 + p_2$) and deflates the probability of receiving t_H (i.e., $\pi_3 < p_3$). As a consequence, an increase in t_L that is accompanied by a decrease in t_H such that the expected transfer payment as perceived by A is left unchanged strictly reduces the actual objectively expected transfer payment. Therefore, as becomes apparent from Figure 4, the cost-minimizing contract is obtained once the transfer t_H just satisfies the requirement imposed by the limited liability constraint (i.e., $t_H^0 = \bar{t}$), and the transfer t_L is specified such that the participation constraint holds with equality (i.e., $\Phi_0^P(t_L^0) = \bar{t}$).

The principal's expected utility under the cost-minimizing contract to implement $e = 0$ is

$$U_P(t_L^0, t_H^0, 0) = (p_1 + p_2)V_L + p_3 V_H - \begin{cases} \bar{t} & \text{if} \quad \bar{t} \geq 0, \\ \left[(p_1 + p_2)\left(-\frac{\pi_3}{\pi_1 + \pi_2} \bar{t} \right) + p_3 \bar{t} \right] & \text{if} \quad \bar{t} < 0, \end{cases} \quad (13)$$

and once again, an increase in the degree of A's salience bias makes P (weakly) better off.

Corollary 2. *If the degree of the agent's salience bias marginally increases (i.e., if δ decreases),* $U_P(t_L^0, t_H^0, 0)$ *remains unchanged if $\bar{t} \geq 0$ and strictly increases if $\bar{t} < 0$.*

Clearly, if $\bar{t} \geq 0$, the cost-minimizing transfers are determined by the limited liability constraint (LL) alone, i.e., $t_L^0 = t_H^0 = \bar{t}$. As a consequence, as the liability threshold $\bar{t}$ remains unchanged, a change in the degree of the agent's salience bias leaves the cost-minimizing transfers and, thus, the principal's expected utility unaffected. If $\bar{t} < 0$, on the other hand, an increase in the degree of the agent's salience bias strictly increases the principal's maximum utility from implementing low effort. In this case the participation constraint is binding, and the contract offered by the principal is a gamble that exploits the diverging probability assessments. An increase in the degree of the agent's salience bias increases the divergence in prior beliefs and thus allows the principal to exploit this divergence more effectively.

3.3. Second-Best Effort with Salience Bias

Having analyzed the second-best optimal contracts for given effort levels, we can now investigate which effort level is optimally induced by the principal. Let $e^* = \arg\max_{e \in \{0,1\}} U_P(t_L^e, t_H^e, e)$ denote this second-best effort level. As a tie-breaking rule, we assume that P induces high effort when being indifferent, i.e., $e^* = 1$ if and only if $U_P(t_L^1, t_H^1, 1) \geq U_P(t_L^0, t_H^0, 0)$. Intuitively, P will induce high effort rather than low effort if and only if the associated increase in value is sufficiently high.

Proposition 3. *The principal implements high effort, i.e., $e^* = 1$, if and only if $\Delta(\bar{t}) \leq V_H - V_L$, where*

$$\Delta(\bar{t}) := \begin{cases} \frac{p_2 + p_3}{\pi_2} \Delta^{FB} & \text{if} \quad 0 \leq \bar{t}, \\ \frac{p_2 + p_3}{\pi_2} \Delta^{FB} + \frac{p_1 + p_2}{p_2(\pi_1 + \pi_2)} \bar{t} & \text{if} \quad t_L^{ST} \leq \bar{t} < 0, \\ \frac{p_2 + p_3}{\pi_2 + \pi_3} \Delta^{FB} - \frac{\bar{t}}{p_2}\left(\frac{1 - p_1}{1 - \pi_1} - \frac{1 - p_3}{1 - \pi_3} \right) & \text{if} \quad \bar{t} < t_L^{ST}. \end{cases} \quad (14)$$

We are interested in the distortion that arises under the second-best optimal contract, and whether this is caused by the (degree of) salience bias. According to Proposition 3, how $\Delta(\bar{t})$ relates to Δ^{FB} depends on the liability threshold ($\bar{t}$), the objective occurrence probabilities (p_1, p_2, and p_3), and the degree of salience bias (δ), which enters into the decision weights (π_1, π_2, and π_3).

Recall that, for $\bar{t} \geq 0$, the second-best optimal contract that implements $e = 0$ is independent of the salience bias. On the other hand, implementing $e = 1$ becomes cheaper, the stronger the salience bias, and thus the excess weight on receiving the bonus payment, is. Not surprisingly, if the agent's probability assessment is distorted, high effort is implemented "more often" compared to the standard case without salience bias ($\delta = 1$). Interestingly, the salience bias can make the implementation of high effort so cheap, that P induces high effort even in cases where low effort maximizes the undistorted gains from trade. To see this formally, note that for $\bar{t} \geq 0$, $\Delta(\bar{t}) \lesseqgtr \Delta^{FB}$ if and only if $\frac{p_2+p_3}{\pi_2} \lesseqgtr 1$, which is equivalent to

$$\delta \lesseqgtr \frac{p_1 p_2}{(1-p_1)(1-p_2)} \in (0,1). \tag{15}$$

This observation already shows how salience bias can fundamentally alter the outcome of the principal–agent relationship in comparison with the standard model without salience bias. In the standard model with an unbiased agent (i.e., $\delta = 1$), we have $\Delta(\bar{t}) = \tilde{\Delta} = (1 + \frac{p_3}{p_2})\Delta^{FB} > \Delta^{FB}$, such that any distortion in effort must take the form of underwork (i.e., $\tilde{e} < e^{FB}$). In contrast, if A's perception is sufficiently blurred by salience, then P induces high effort even though the gains from trade are maximized by low effort; i.e., if $\delta < \frac{p_1 p_2}{(1-p_1)(1-p_2)}$ and $\Delta(\bar{t}) \leq V_H - V_L < \Delta^{FB}$, then the distortion in effort takes the form of overwork (i.e., $e^* > e^{FB}$).

For $\bar{t} < 0$, P benefits from A's distorted probability assessments irrespective of whether she induces high or low effort. This makes the comparison of $\Delta(\bar{t})$ and Δ^{FB} more complicated and significantly less tractable. To retain tractability and to streamline the exposition, we impose the following simplifying assumption:

Assumption 1. $p_2 = p \in (0,1)$ and $p_1 = p_3 = \frac{1-p}{2}$.

According to Assumption 1, states 1 and 3, where the agent's effort choice has no impact on the value that is generated for the principal, are equiprobable. If Assumption 1 holds, how $\Delta(\bar{t})$ compares to Δ^{FB} can be related to the degree of salience bias as follows:

Lemma 1. *Suppose Assumption 1 holds. Then* $\Delta(\bar{t}) \lesseqgtr \Delta^{FB}$ *if and only if* $\delta \lesseqgtr \hat{\delta}(\bar{t})$*, where*

$$\hat{\delta}(\bar{t}) := \begin{cases} \frac{p}{1+p} & \text{if} \quad 0 \leq \bar{t}, \\ \frac{p\left\{-[\psi(1+3p)+2\bar{t}(1+p)]+\sqrt{\psi^2(p+3)^2-4\bar{t}\psi(1-p^2)+4\bar{t}^2(1+p)^2}\right\}}{2\psi(1-p^2)} & \text{if} \quad t_L^{ST} \leq \bar{t} < 0, \\ 1 & \text{if} \quad \bar{t} < t_L^{ST} \end{cases} \tag{16}$$

is a continuous and weakly decreasing function.

The function $\hat{\delta}(\cdot)$ is depicted in Figure 5. Given liability level $\bar{t}$, whenever $\delta \neq \hat{\delta}(\bar{t})$, there is scope for the second-best effort level to diverge from the first-best effort level.

Proposition 4. *Suppose Assumption 1 holds. If* $\delta < \hat{\delta}(\bar{t})$ *and* $\Delta(\bar{t}) \leq V_H - V_L < \Delta^{FB}$*, then the principal induces overwork (i.e., $e^* > e^{FB}$). If $\delta > \hat{\delta}(\bar{t})$ and $\Delta^{FB} \leq V_H - V_L < \Delta(\bar{t})$, then the principal induces underwork (i.e., $e^* < e^{FB}$). Otherwise the principal induces the first-best effort (i.e., $e^* = e^{FB}$).*

If A's salience bias is sufficiently weak, $\delta > \hat{\delta}(\bar{t})$, then, as in the standard model, any distortion away from the first-best effort results in underwork. In contrast to the standard model, however, if A's salience bias is sufficiently strong, $\delta > \hat{\delta}(\bar{t})$, then the inefficient effort provision takes the form of overwork. Notably, if $\bar{t} \leq -\frac{p_3}{p_2}\psi = -\frac{1-p}{2p}\psi$, then $\hat{\delta}(\bar{t}) = 1$ such that P always induces the first-best effort in the standard model, even a small degree of salience bias opens the door for overwork to prevail.

These observations are summarized in Figure 5. Note that the results of the standard model are also contained in this figure, which can be seen by looking at the dashed horizontal line at $\delta = 1$.
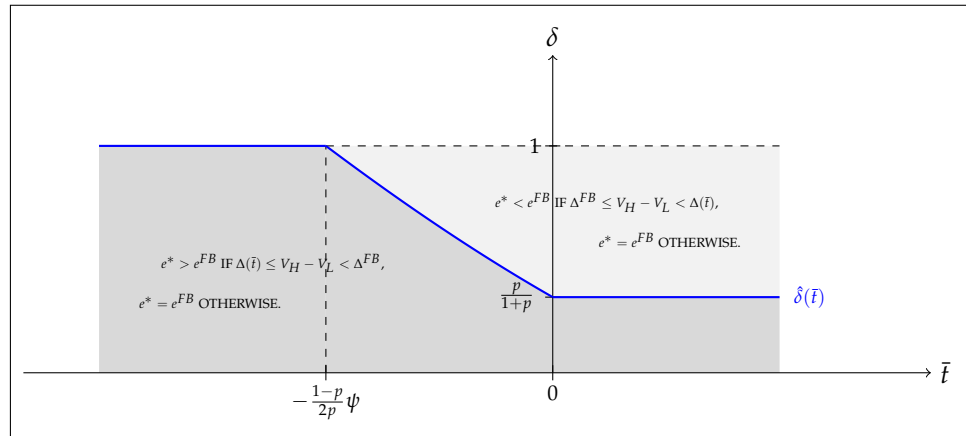

Figure 5. Second-best effort.

Proposition 4 relates the effort level, e^*, induced under the optimal contract for a salience-biased agent at Benchmark 1, i.e., to the materially efficient effort level e^{FB}. Next, we want to compare e^* to Benchmark 2, i.e., to the effort level, $\tilde{e}$, induced under the optimal contract for an unbiased agent. One can verify that $\frac{d\Delta(\bar{t})}{d\delta} > 0$, which implies that $\tilde{\Delta} > \Delta(\bar{t})$ for all $\delta \in (0,1)$ and $\bar{t} \in \mathbb{R}$. As a consequence, $\tilde{e} \leq e^*$ for all $\delta \in (0,1)$ and $\bar{t} \in \mathbb{R}$. Hence, if the principal faces an agent whose perception is blurred by salience bias, she always induces weakly, and in some circumstances strictly, higher effort than in the case where she faces an unbiased agent who perceives the occurrence probabilities correctly. Specifically, if $\tilde{e} = e^{FB} < e^*$, then salience bias is detrimental for material efficiency as it induces overwork, which is never the result for an unbiased agent. If, in contrast, $\tilde{e} < e^{FB} = e^*$, then salience bias fosters material efficiency as underwork would prevail in the case of an unbiased agent.

To conclude, let us entertain a thought experiment that might help to understand why employees may actually feel overworked, as found by Galinsky et al. [22]. At first glance, this observation seems at odds with our model, given that a salience-biased agent also considers his decision to accept a contract and his decision to exert high effort as utility maximizing at the moment when he makes the respective decision. Imagine that a salience-biased agent accepted a contract under which he exerts high effort, i.e., under which $e^* = 1$. As we know from Proposition 1, under this contract either the incentive compatibility constraint binds and the participation constraint is slack (if $\bar{t} > t_L^{ST}$), or the participation constraint binds and the incentive compatibility constraint is slack (if $\bar{t} < t_L^{ST}$). Suppose that we can slightly reduce the degree of the agent's salience bias right before these decisions. Irrespective of which constraint was binding in the presence of salience bias, the respective constraint is violated for the lower degree of salience bias. That is to say, if the incentive compatibility constraint is binding in the presence of salience bias, then, once we weaken the bias, the agent would agree that he had almost accidentally exerted more effort than is good for him. Likewise, if the participation constraint is binding in the presence of salience bias, then, once we weaken the bias, the agent would agree that he almost accidentally entered into a work relationship even though he is better off rejecting the principal's contract offer. While it may not be possible for us to reduce the degree of the agent's salience bias at will, one might imagine that the salience-blurred perception, which shapes a particular decision will fade out once the decision has been taken and the agent's attention is shifted elsewhere. In this case, it stands to reason that the agent, in retrospect, will realize that he has overextended himself in the past by exerting more effort on a task than he now deems worthwhile or by agreeing to take on the task in the first place.

4. Discussion

In this paper, we enriched an otherwise standard principal–agent model of moral hazard with hidden action by salience-biased perception [25] on the agent's side and analyzed its implications for incentive provision. In comparison to the standard model, we found that salience bias can reverse the nature of the inefficiency that arises from the agent being privately informed about his own work effort, i.e., the principal does not necessarily provide insufficient incentives that result in inefficiently low effort, but instead may provide excessive incentives that result in inefficiently high effort. The reason overwork may prevail is that the agent's misguided focus on salient payoffs facilitates incentive provision and allows the principal to exploit diverging prior beliefs to her advantage.

The first immediately related question is to what extent this finding is specific to the model of salience theory for choice under risk outlined by Bordalo et al. [25]. As the result is rooted in the excess weight or excess attention that is placed on states with rather different payoffs, any state-space based theory that embodies the psychological principle of similarity judgments [52,53] in one way or the other should yield qualitatively similar results.[15] The advantage of salience theory is its tractability, because, even if the agent's remuneration is contingent on his performance (i.e., even if the transfer paid in case of generation of high value is strictly different from the transfer paid in case of generation of low value), the distortion in the decision weights does not depend on the specifics of the contractually specified transfer payments, and the contracting problem remains a simple linear programming problem.

A second related question is whether our study allows us to identify ways to deal with overwork as a social problem. An obvious answer would be to reduce salience bias by training the agent (e.g., a worker or a manager) in his assessment of stochastic processes. The adequacy of this solution is, however, open to question. First, inferential errors are commonly regarded as errors of application [57], i.e., individuals alerted to an error in judgment in one instance seem unable to avoid making the very same error in subsequent judgments. Thus, awareness training is arguably not a fail-safe device to cope with overwork. Second, the principal has no incentive to incur the cost for implementing such awareness training for the agent, because her expected utility increases with the severity of salience bias (cf. Corollaries 1 and 2). Hence, coping with overwork seems to call for regulatory intervention. A practical approach suggested by our analysis is the imposition of a sufficiently high minimum wage. As can be seen from Figure 5, the severity of salience bias necessary for overwork to possibly prevail is highest if transfers from the principal to the agent cannot be negative. In fact, with $\hat{\delta}(\bar{t}) < \frac{1}{2}$ for $\bar{t} > 0$, the estimation in Königsheim et al. [45], who estimate the salience parameter δ to equal 0.7–0.8, suggests that overwork should not arise in the presence of a non-negative minimum wage.[16]

Throughout our analysis, we assumed that the agent's outside option has no effect on his salience-blurred perception of the occurrence probabilities of the different monetary outcomes feasible under the principal's contract offer. As noted before, this assumption seems reasonable if the reservation utility associated with *A*'s outside option has a non-monetary origin such as the stigma or the shame of being unemployed. However, one might also imagine that, if the agent's outside option is unemployment, his reservation utility is primarily tied to monetary unemployment benefits. As monetary unemployment benefits can rather easily be compared to the monetary transfer payments specified in the principal's contract offer, they might affect salience. Extending our model in this regard would allow us to address the heretofore unexplored participation and incentive effects of unemployment benefits, which, we believe, is a highly interesting and policy-relevant venue for future research.

Author Contributions: Conceptualization: F.R., F.H., and D.M.; formal analysis: F.R.; writing—original draft preparation: F.R.; writing—review and editing: F.R., F.H., and D.M.; funding acquisition: F.H., and D.M. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Throughout this appendix, as opposed to the main text, π_s always denotes the salience-biased probabilities in case of performance-dependent pay (i.e., $t_H \neq t_L$) as stated in (5). For performance-independent pay (i.e., $t_H = t_L$) we will explicitly write down the objective probabilities, p_s, as there is no salience bias.

Proof of Proposition 1. In the case that she wants to implement high effort $e = 1$, the principal's cost-minimization problem takes the following form:

$$\max_{(t_L, t_H) \in \mathbb{R}^2} p_1 t_L + (p_2 + p_3) t_H$$

subject to

$$\pi_2(t_H - t_L) \geq \psi \tag{IC$_1$}$$

$$\pi_1 t_L + (\pi_2 + \pi_3) t_H \geq 0 \tag{P$_1$}$$

$$t_L \geq \bar{t}, \ t_H \geq \bar{t}. \tag{LL}$$

Strictly speaking, if $t_H = t_L$, then (IC$_1$) and (P$_1$) take the following respective form: $p_2(t_H - t_L) \geq \psi$ and $p_1 t_L + (p_2 + p_3) t_H \geq 0$. Note that $\lim_{t_L \to t_H} \pi_2(t_H - t_L) = \lim_{t_L \to t_H} p_2(t_H - t_L) = 0$ and $\lim_{t_L \to t_H} \pi_1 t_L + (\pi_2 + \pi_3) t_H = \lim_{t_L \to t_H} p_1 t_L + (p_2 + p_3) t_H = t_H$. Therefore, in the case where $t_H = t_L$, we can w.l.o.g. replace the objective probabilities with the distorted probabilities in (IC$_1$) and (P$_1$).

In order to determine the cost-minimizing transfers, we distinguish two cases, namely $t_L^{ST} \leq \bar{t}$ and $\bar{t} < t_L^{ST}$.

Case 1: $t_L^{ST} \leq \bar{t}$

First, we have

$$t_H \geq \frac{\psi}{\pi_2} + t_L \geq \frac{\psi}{\pi_2} + t_L^{ST} = \frac{\psi}{\pi_2 + \pi_3} - \frac{\pi_1}{\pi_2 + \pi_3} t_L \geq \frac{\psi}{\pi_2 + \pi_3} - \frac{\pi_1}{\pi_2 + \pi_3} t_L,$$

where the first inequality holds by (IC$_1$), and the remaining inequalities hold by (LL). The last inequality implies that (P$_1$) is automatically satisfied (and thus can be ignored) if (IC$_1$) and (LL) are jointly satisfied.

Next, note that we must have $t_L = \bar{t}$ under the cost-minimizing contract. To understand this, we proceed by proof by contradiction, i.e., suppose that (t_L', t_H') is a solution to the cost-minimization problem with $t_L' > \bar{t}$. Consider an alternative contract with transfers $\tilde{t}_H = t_H'$ and $\tilde{t}_L = t_L' - \varepsilon$, where $\varepsilon > 0$. As $\tilde{t}_H - \tilde{t}_L > t_H' - t_L'$, this alternative contract satisfies (IC$_1$). Moreover, as long as ε is sufficiently small, the alternative contract also satisfies (LL). With this contractual adjustment resulting in a strictly lower expected transfer payment than the original contract, (t_L', t_H'), without violating any constraint,

the original contract (t'_L, t'_H) cannot be a solution to the cost-minimization problem, i.e., it is a contradiction.

With $t_L = \bar{t}$, (IC$_1$) requires that $t_H \geq \frac{\psi}{\pi_2} + \bar{t} > \bar{t}$, such that (LL) is automatically satisfied (and thus can be ignored) if (IC$_1$) holds. As a consequence, with the expected transfer payment being strictly increasing in transfer t_H, (IC$_1$) must be satisfied with equality under the cost-minimizing contract; i.e., we must have $t_H = \frac{\psi}{\pi_2} + \bar{t}$.

Case 2: $\bar{t} < t_L^{ST}$

First, we establish that $t_L = \bar{t}$ under the cost-minimizing contract. Proceeding by proof by contradiction, suppose (t'_L, t'_H) is a solution to the cost-minimization problem with $t_L > \bar{t}$. Now, consider an alternative contract with transfers $\tilde{t}_H = t'_H + \frac{\varepsilon \pi_1}{\pi_2 + \pi_3}$ and $\tilde{t}_L = t'_L - \varepsilon$, where $\varepsilon > 0$. As $\tilde{t}_H - \tilde{t}_L > t'_H - t'_L$, (IC$_1$) is satisfied under this alternative contract. Furthermore, as $\pi_1 \tilde{t}_L + (\pi_2 + \pi_3)\tilde{t}_H = \pi_1 t'_L + (\pi_2 + \pi_3)t'_H$ by construction of the transfers $\tilde{t}_L$ and $\tilde{t}_H$, (P$_1$) is unaffected by this contractual adjustment. Finally, as long as ε is sufficiently small, the transfers $\tilde{t}_L$ and $\tilde{t}_H$ also satisfy (LL). Regarding the expected transfer payment from the principal, we obtain $p_1 \tilde{t}_L + (p_2 + p_3)\tilde{t}_H = p_1 t'_L + (p_2 + p_3)t'_H - \varepsilon \frac{p_1 - \pi_1}{\pi_2 + \pi_3}$. As $\pi_1 < p_1$, the alternative contract $(\tilde{t}_L, \tilde{t}_H)$, which does not violate any constraint, results in a strictly lower expected transfer payment than the original contract (t'_L, t'_H), such that the latter cannot be a solution to the cost-minimization problem, i.e., it is a contradiction.

Next, we show that (P$_1$) has to be satisfied with equality under the cost-minimizing contract. Again, we proceed by proof by contradiction, i.e., supposing that (t'_L, t'_H) is a solution to the cost-minimization problem with $\pi_1 \bar{t} + (1 - \pi_1)t'_H > \psi$. As $t_L = \bar{t}$ under the cost-minimizing contract, we have

$$t'_H > \frac{\psi}{1 - \pi_1} - \frac{\pi_1}{1 - \pi_1}\bar{t} > \frac{\psi}{1 - \pi_1} - \frac{\pi_1}{1 - \pi_1}t_L^{ST} = \frac{\psi}{\pi_2} + t_L^{ST} > \frac{\psi}{\pi_2} + \bar{t} > \bar{t},$$

where the first inequality holds by (P$_1$) being satisfied with strict inequality under the contract (t'_L, t'_H), and the second and the third inequalities hold by $\bar{t} < t_L^{ST} = -\frac{\pi_3}{\pi_2}\psi$. Importantly, with $t_L = \bar{t}$, the third and the fourth inequalities imply that the lower bounds on t_H imposed by (IC$_1$) and (LL) are strictly less restrictive than the lower bound on t_H imposed by (P$_1$), i.e., given $t_L = \bar{t}$, any transfer t_H that satisfies (P$_1$) automatically satisfies both (IC$_1$) and (LL). Now consider an alternative contract with transfers $\tilde{t}_H = t'_H - \varepsilon$ and $\tilde{t}_L = t'_L = \bar{t}$, where $\varepsilon > 0$. As long as ε is sufficiently small, the alternative contract satisfies (P$_1$) and thus also (IC$_1$) and (LL). Since the contractual adjustment results in a strictly lower expected transfer payment than the original contract (t'_L, t'_H) without violating any constraint, the original contract (t'_L, t'_H) cannot be a solution to the cost-minimization contract, i.e., it is a contradiction.

Finally, inserting $t_L = \bar{t}$ into the binding participation constraint (P$_1$) yields $t_H = \frac{\psi}{\pi_2 + \pi_3} - \frac{\pi_1}{\pi_2 + \pi_3}\bar{t}$. □

Proof of Corollary 1. Differentiation of (11) with respect to δ yields

$$\frac{dU_P(t_L^1, t_H^1, 1)}{d\delta} = \begin{cases} -\frac{(p_1 + p_3)(p_2 + p_3)}{p_2}\psi & \text{if } \bar{t} > t_L^{ST}, \\ -\frac{p_1 p_2 (p_2 + p_3)}{(p_2 + \delta p_3)^2}(\psi - \bar{t}) & \text{if } \bar{t} < t_L^{ST}. \end{cases} \tag{A1}$$

The statement follows from $t_L^{ST} < 0$ and $U_P(t_L^1, t_H^1, 1)$ being continuous at $\bar{t} = t_L^{ST}$. □

Proof of Proposition 2. In case that she wants to implement low effort $e = 0$, the principal's cost-minimization problem takes the following form:

$$\max_{(t_L, t_H) \in \mathbb{R}^2} (p_1 + p_2)t_L + p_3 t_H$$

subject to

$$\pi_2(t_H - t_L) \leq \psi \tag{IC_0}$$

$$(\pi_1 + \pi_2)t_L + \pi_3 t_H \geq 0 \tag{P_0}$$

$$t_L \geq \bar{t}, \ t_H \geq \bar{t}. \tag{LL}$$

Strictly speaking, if $t_H = t_L$, then (IC_0) and (P_0) take the following respective form: $p_2(t_H - t_L) \leq \psi$ and $(p_1 + p_2)t_L + p_3 t_H \geq 0$. Note that $\lim_{t_L \to t_H} \pi_2(t_H - t_L) = \lim_{t_L \to t_H} p_2(t_H - t_L) = 0$ and $\lim_{t_L \to t_H} (\pi_1 + \pi_2)t_L + \pi_3 t_H = \lim_{t_L \to t_H} (p_1 + p_2)t_L + p_3 t_H = t_H$. Therefore, in the case where $t_H = t_L$, we can w.l.o.g. replace the objective probabilities with the distorted probabilities in (IC_0) and (P_0).

In order to determine the cost-minimizing transfers, we distinguish two cases, namely $0 \leq \bar{t}$ and $\bar{t} < 0$.

Case 1: $0 \leq \bar{t}$

With the expected transfer payment being strictly increasing in both t_L and t_H, the best that the principal can hope for is to set both transfers as low as possible, i.e., to set $t_L = t_H = \bar{t}$. Clearly, this specification satisfies both (IC_0) and (LL). Furthermore, with $\bar{t} \geq 0$, the expected transfer payment from the principal to the agent is non-negative, such that (P_0) is also satisfied. As a consequence, setting $t_L = t_H = \bar{t}$ must be a solution to the principal's cost-minimization problem.

Case 2: $\bar{t} < 0$

First, we establish that $t_H = \bar{t}$ under the cost-minimizing contract. Proceeding by proof by contradiction, suppose (t'_L, t'_H) is a solution to the cost-minimization problem with $t_H > \bar{t}$. Now, consider an alternative contract with transfers $\tilde{t}_H = t'_H - \varepsilon$ and $\tilde{t}_L = t'_L + \frac{\varepsilon \pi_3}{\pi_1 + \pi_2}$, where $\varepsilon > 0$. As $\tilde{t}_H - \tilde{t}_L < t'_H - t'_L$, this contractual adjustment satisfies (IC_0). Furthermore, as $(\pi_1 + \pi_2)\tilde{t}_L + \pi_3 \tilde{t}_H = (\pi_1 + \pi_2)t'_L + \pi_3 t'_H$ by construction of the transfers $\tilde{t}_L$ and $\tilde{t}_H$, (P_0) is unaffected by this contractual adjustment. Finally, as long as ε is sufficiently small, the transfers $\tilde{t}_L$ and $\tilde{t}_H$ also satisfy (LL). Regarding the expected transfer payment from the principal to the agent, we have $(p_1 + p_2)\tilde{t}_L + p_3 \tilde{t}_H = (p_1 + p_2)t'_L + p_3 t'_H - \varepsilon \frac{p_3 - \pi_3}{\pi_1 + \pi_2}$. As $\pi_3 < p_3$, the alternative contract $(\tilde{t}_L, \tilde{t}_H)$, which does not violate any constraint, results in a strictly lower expected transfer payment than the original contract (t'_L, t'_H), such that the latter cannot be a solution to the cost-minimization problem, i.e., it is a contradiction.

Next, we establish that (P_0) must be satisfied with equality under the cost-minimizing contract. To understand this, we proceed by proof by contradiction, i.e., supposing that (t'_L, t'_H) is a solution to the cost-minimization problem with $(\pi_1 + \pi_2)t'_L + \pi_3 t'_H > 0$. With $t'_H = \bar{t} < 0$, (P_0) requires $t_L > 0$, which in turn implies that $t_L > \bar{t}$ and that, under contract (t'_L, t'_H), not only (P_0) but also (IC_0) is satisfied with strict inequality. Consider an alternative contract with $\tilde{t}_H = t'_H$ and $\tilde{t}_L = t'_L - \varepsilon$, where $\varepsilon > 0$. As long as ε is sufficiently small, (IC_0), (P_0), and (LL) are all satisfied under this alternative contract. With this contractual adjustment resulting in a strictly lower expected transfer payment than the original contract (t'_H, t'_L) without violating any constraint, the original contract (t'_H, t'_L) cannot be a solution to the cost-minimization problem, i.e., it is a contradiction.

Finally, inserting $t_H = \bar{t}$ into the binding participation constraint (P_0) yields $t_L = -\frac{\pi_3}{\pi_1 + \pi_2}\bar{t}$. As $\bar{t} < 0 < -\frac{\pi_3}{\pi_1 + \pi_2}\bar{t}$, it follows that these transfers also satisfy (IC_0) and (LL). $\quad\square$

Proof of Corollary 2. Differentiation of (13) with respect to δ yields

$$\frac{dU_P(t_L^0, t_H^0, 0)}{d\delta} = \begin{cases} 0 & \text{if } \bar{t} > 0, \\ \frac{p_2 p_3 (p_1 + p_2)}{(\delta p_1 + p_2)^2} \bar{t} & \text{if } \bar{t} < 0. \end{cases} \tag{A2}$$

The statement follows from $U_P(t_L^0, t_H^0, 0)$ being continuous at $\bar{t} = 0$. $\quad\square$

Proof of Proposition 3. The principal implements high effort $e = 1$ if and only if $U_P(t_L^1, t_H^1, 1) \geq U_P(t_L^0, t_H^0, 0)$. Given the contractual specifications identified in Propositions 1 and 2, this comparison of expected utilities depends on whether $0 \leq \bar{t}$, $t_L^{ST} \leq \bar{t} < 0$, or $\bar{t} < t_L^{ST}$. We consider each of these cases in turn.

Case 1: $0 \leq \bar{t}$

From (13) and (11), it follows that the principal's expected utility from inducing low effort is given by

$$U_P(t_L^0, t_H^0, 0) = (p_1 + p_2)V_L + p_3 V_H - \bar{t}, \tag{A3}$$

and her expected utility from inducing high effort is given by

$$U_P(t_L^1, t_H^1, 1) = p_1 V_L + (p_2 + p_3)V_H - \bar{t} - \frac{p_2 + p_3}{\pi_2}\psi. \tag{A4}$$

Comparison of (A3) and (A4) reveals that $U_P(t_L^1, t_H^1, 1) \geq U_P(t_L^0, t_H^0, 0)$ if and only if

$$V_H - V_L \geq \frac{p_2 + p_3}{\pi_2}\Delta^{FB}. \tag{A5}$$

Case 2: $t_L^{ST} \leq \bar{t} < 0$

From (13) and (11), it follows that the principal's expected utility from inducing low effort is given by

$$U_P(t_L^0, t_H^0, 0) = (p_1 + p_2)V_L + p_3 V_H - \bar{t}\left(1 - \frac{p_1 + p_2}{\pi_1 + \pi_2}\right), \tag{A6}$$

and her expected utility from inducing high effort is given by

$$U_P(t_L^1, t_H^1, 1) = p_1 V_L + (p_2 + p_3)V_H - \bar{t} - \frac{p_2 + p_3}{\pi_2}\psi. \tag{A7}$$

Comparison of (A6) and (A7) reveals that $U_P(t_L^1, t_H^1, 1) \geq U_P(t_L^0, t_H^0, 0)$ if and only if

$$V_H - V_L \geq \frac{p_2 + p_3}{\pi_2}\Delta^{FB} + \frac{p_1 + p_2}{(\pi_1 + \pi_2)p_2}\bar{t}. \tag{A8}$$

Case 3: $\bar{t} < t_L^{ST}$

From (13) and (11), it follows that the principal's expected utility from inducing low effort is given by

$$U_P(t_L^0, t_H^0, 0) = (p_1 + p_2)V_L + p_3 V_H - \bar{t}\left(1 - \frac{p_1 + p_2}{\pi_1 + \pi_2}\right), \tag{A9}$$

and her expected utility from inducing high effort is given by

$$U_P(t_L^1, t_H^1, 1) = p_1 V_L + (p_2 + p_3)V_H - \frac{p_2 + p_3}{\pi_2 + \pi_3}\psi - \frac{p_1 - \pi_1}{1 - \pi_1}\bar{t}. \tag{A10}$$

Comparison of (A9) and (A10) reveals that $U_P(t_L^1, t_H^1, 1) \geq U_P(t_L^0, t_H^0, 0)$ if and only if

$$V_H - V_L \geq \frac{p_2 + p_3}{\pi_2 + \pi_3}\Delta^{FB} - \frac{\bar{t}}{p_2}\left(\frac{1 - p_1}{1 - \pi_1} - \frac{p_1 + p_2}{\pi_1 + \pi_2}\right). \tag{A11}$$

$\square$

Proof of Lemma 1. With $p_2 = p$ and $p_1 = p_3 = \frac{1-p}{2}$, we have $\Delta^{FB} = \frac{\psi}{p}$. Given the specification of $\Delta(\bar{t})$ identified in (14), we have to distinguish whether $0 \leq \bar{t}$, $t_L^{ST} \leq \bar{t} < 0$, or $\bar{t} < t_L^{ST}$. We consider each of these cases in turn.

Case 1: $0 \leq \bar{t}$

As $p_2 = p$ and $p_1 = p_3 = \frac{1-p}{2}$ imply that $\pi_2 = \frac{p}{p+\delta(1-p)}$, from (14) it follows that

$$\Delta(\bar{t}) = \frac{(1+p)[p+\delta(1-p)]}{2p}\Delta^{FB}. \tag{A12}$$

As a consequence, $\Delta(\bar{t}) \lesseqgtr \Delta^{FB}$ if and only if $\delta \lesseqgtr \frac{p}{1+p}$, where $0 < \frac{p}{1+p} < 1$ as $p \in (0,1)$.

Case 2: $t_L^{ST} \leq \bar{t} < 0$

As $p_2 = p$ and $p_1 = p_3 = \frac{1-p}{2}$ imply that $\pi_2 = \frac{p}{p+\delta(1-p)}$ and $\pi_1 + \pi_2 = \frac{\delta(1-p)+2p}{2[p+\delta(1-p)]}$, from (14) it follows that

$$\Delta(\bar{t}) = \frac{(1+p)[p+\delta(1-p)]}{2p}\left[\Delta^{FB} + \frac{2}{\delta(1-p)+2p}\bar{t}\right]. \tag{A13}$$

As a consequence, $\Delta(\bar{t}) \lesseqgtr \Delta^{FB}$ if and only if $Q(\delta) \lesseqgtr 0$, where

$$Q(\delta) := \delta^2(1-p^2)\psi + \delta p[\psi(1+3p)+2\bar{t}(1+p)] + \frac{2p^2[\bar{t}(1+p)-\psi(1-p)]}{1-p}. \tag{A14}$$

If the quadratic equation $Q(\delta) = 0$ has one (or two) real solution(s), its (their) discriminant is given by

$$D(\bar{t}) := p^2\left(\psi^2(p+3)^2 - 4\bar{t}\psi(1-p^2) + 4\bar{t}^2(1+p)^2\right).$$

Differentiation reveals that $D(\cdot)$ is a strictly convex function that attains its minimum at $t^D = \frac{\psi(1-p)}{2(1+p)}$. As $D(t^D) = p^2(p+3)\psi > 0$ implies that $D(\bar{t}) > 0$ for all $\bar{t} \in \mathbb{R}$, the quadratic equation $Q(\delta) = 0$ has two real solutions given by

$$\delta^-(\bar{t}) = \frac{-p[\psi(1+3p)+2\bar{t}(1+p)] - \sqrt{D(\bar{t})}}{2\psi(1-p^2)} \tag{A15}$$

and

$$\delta^+(\bar{t}) = \frac{-p[\psi(1+3p)+2\bar{t}(1+p)] + \sqrt{D(\bar{t})}}{2\psi(1-p^2)}. \tag{A16}$$

Calculations reveal that $\delta^-(\bar{t}) < 0 < \delta^+(\bar{t}) < 1$ if and only if $-\frac{\psi(1-p)}{2p} < \bar{t} < \frac{\psi(1-p)}{1+p}$, where the latter condition is satisfied because $t_L^{ST} \leq \bar{t} < 0$ by hypothesis, and $\delta \in (0,1)$ implies $-\frac{\psi(1-p)}{2p} < t_L^{ST}$. Hence, with $Q''(\delta) > 0$, it follows that $\Delta(\bar{t}) \lesseqgtr \Delta^{FB}$ if and only if $\delta \lesseqgtr \delta^+$.

Case 3: $\bar{t} < t_L^{ST}$

As $p_2 = p$ and $p_1 = p_3 = \frac{1-p}{2}$ imply that $\pi_2 + \pi_3 = \frac{2p+\delta(1-p)}{2[p+\delta(1-p)]}$, from (14) it follows that

$$\Delta(\bar{t}) = \frac{(1+p)[p+\delta(1-p)]}{2p+\delta(1-p)}\Delta^{FB}. \tag{A17}$$

As a consequence, $\Delta(\bar{t}) \lesseqgtr \Delta^{FB}$ if and only if $\delta \lesseqgtr 1$.

The above observations imply that $\Delta(\bar{t}) \lesseqqgtr \Delta^{FB}$ if and only if $\delta \lesseqqgtr \hat{\delta}(\bar{t})$ with

$$
\hat{\delta}(\bar{t}) := \begin{cases} \frac{p}{1+p} & \text{if} \quad 0 \le \bar{t}, \\ \delta^+(\bar{t}) & \text{if} \quad t_L^{ST} \le \bar{t} < 0, \\ 1 & \text{if} \quad \bar{t} < t_L^{ST}. \end{cases} \tag{A18}
$$

From (A18) we obtain that

$$
\lim_{\bar{t} \nearrow 0} \hat{\delta}(\bar{t}) = \frac{p\left[-3\psi p - \psi + \sqrt{\psi^2(p+3)^2}\right]}{2\psi(1-p^2)} = \frac{2p\psi(1-p)}{2\psi(1-p)(1+p)} = \frac{p}{1+p} = \hat{\delta}(0)
$$

and

$$
\hat{\delta}(t_L^{ST}) = \frac{\psi - \psi p - 4\psi p^2 + \sqrt{\psi^2(1+p+2p^2)^2}}{2\psi(1-p^2)} = 1 = \lim_{\bar{t} \nearrow t_L^{ST}} \hat{\delta}(\bar{t})
$$

and

$$
\frac{d\hat{\delta}(\bar{t})}{d\bar{t}} = \begin{cases} 0 & \text{if} \quad 0 < \bar{t}, \\ -\frac{p}{\psi(1-p^2)}\left[1 + p + \frac{\psi(1-p^2)-2(1+p)^2\bar{t}}{\sqrt{\psi^2(p+3)^2 - 4\bar{t}\psi(1-p^2)+4\bar{t}^2(1+p)^2}}\right] < 0 & \text{if} \quad t_L^{ST} < \bar{t} < 0, \\ 0 & \text{if} \quad \bar{t} < t_L^{ST}, \end{cases}
$$

which establishes that $\hat{\delta}(\cdot)$ is a continuous and weakly decreasing function. $\square$

Notes

1   In the taxonomy of Hart and Holmstrom [1], *moral hazard* refers to post-contractual private information, (in contrast to *adverse selection*, which refers to pre-contractual private information), and *hidden action* refers to a party's private information regarding actions taken by this party (as opposed to *hidden information*, which refers to a party's private information regarding its own type or some other payoff-relevant parameter).

2   The moral hazard model with hidden action, where both the principal and the wealth-constrained agent have linear utility functions for money, was applied to address questions of organization design [5,6], sales force compensation [7–9], job design [10–14], team compensation [15], delegation [16,17], lawyer compensation [18], human capital accumulation [8], and privacy protection at the workplace [19]. Experimental evidence for the predictions of the moral hazard model are provided by Hoppe and Kusterer [11], Nieken and Schmitz [20], and Hoppe and Schmitz [21].

3   For example, Harnois and Gabriel [23] predicted that, by 2020, clinical depression would outrank cancer as the second greatest cause of death and disability worldwide. Similarly, conducting a representative study of how children felt about their employed parents, Galinsky [24] reports that most children wished for their parents to be less stressed and less tired from their respective work.

4   Notably, the conceptual concerns regarding salience theory voiced by Kontek [26] and Bako and Neszveda [27] do not apply to our analysis.

5   In a dynamic setting where the same agent and the same principal interact over several periods, Ohlendorf and Schmitz [31] show that the usual finding of underwork prevails.

6   Complementary to our findings for post-contractual hidden action, Goldlücke and Schmitz [33] show that overwork can prevail in a model with post-contractual hidden information.

7   For an extensive overview of these contributions, see Herweg et al. [35].

8   This binary-effort specification, which is widely used in contract-theoretic contributions [6,7,15,40], is primarily made to ease exposition. First, a continuous effort specification would have required a rather technical discussion of the applicability of the first-order approach. Second, with more than two effort levels and, thus, more than two choice options, salience-theoretic preferences may exhibit intransitive choice behavior, which might complicate the analysis regarding incentive compatibility. See Laffont and Martimort [41] or Salanié [42] for a textbook treatment of the binary-effort model.

9   As usual, and also the case in our model set-up, the probability of a high-value V_H being generated is strictly higher if A exerts high effort $e = 1$ than if A exerts low effort $e = 0$. With the application of salience theory requiring a state-based description of the uncertainty that underlies the production process, we capture this aspect by assuming that the probability distribution over value realizations induced by $e = 0$ is dominated state-wise by the probability distribution over value realizations induced by $e = 1$. While the assumption of state-wise dominance is not necessary to derive our results, it simplifies the exposition, as it allows the restriction of attention to only three potential states of the world.

10 By now several applications of salience theory for choice under risk exist [25] that rely on the specification of the decision weight in (4). For example, regarding investor choice and market equilibrium, Bordalo et al. [43] show that salience theory can account for several empirically well-documented puzzles in the finance literature, and Cosemans and Frehen [44] find strong empirical support for these predictions using cross-section data of US stocks. Königsheim et al. [45] empirically estimate the "local-thinking" parameter of salience theory and find substantial heterogeneity with regard to whether a lottery's upside or downside is salient. By implementing manipulations of salience in both a choice study and an eye-tracking study, Alós-Ferrer and Ritschel [46] investigate to what extend the preference reversal phenomenon can be explained by salience theory for choice under risk.

11 Making the assumption that A's outside option does not affect the misperception of probabilities seems particularly plausible if the associated reservation utility has a non-monetary origin such as the stigma or shame of being unemployed. If, on the other hand, A's reservation utility has a monetary origin, such as a fixed wage at the next-best occupation or unemployment benefits, then A's outside option might well affect salience. We comment on this below in Section 4.

12 In addition to the ordering property, Bordalo et al. [25] require the function $\sigma(\cdot, \cdot)$ to be continuous and bounded and to satisfy two further properties called diminishing sensitivity and reflection. Neither of these assumptions is of relevance for our analysis.

13 Here, the problem faced by P is reminiscent of contracting problems with non-common priors [50,51]. In our model, however, the different probability assessments are an endogenous result induced by the optimal contract.

14 Notably, without salience bias there is a continuum of optimal contracts in the case where the participation constraint is binding. More specifically, all transfer combinations that satisfy the participation constraint with equality, which also satisfy limited liability and the incentive compatibility constraint, are part of an optimal contract. Graphically, this follows from the fact that, without salience bias, the slope of the isocost curves and the slope of the $\Phi_1^P(\cdot)$-curve are identical. For $\delta \to 1$, the optimal contract for a salience-biased agent converges to one particular element of the continuum of contracts that are optimal in the absence of salience bias—the contract where t_L is set as low as possible and t_H is specified such that the participation constraint binds.

15 Notable alternative theories are regret theory (e.g., Loomes and Sugden [54]) and theories of additive differences (e.g., Kőszegi and Szeidl [55]). For a comprehensive overview and qualitative comparison of such theories see Bhatia et al. [56]. According to their classification of key properties, we applied a theory with (i) *between-option interactions* and (ii) *weight transformations* that are based on (iii) *similarity and dissimilarity*. While models of *value transformations* may lead to similar findings, their welfare implications might be rather different ([56], p. 1352).

16 Note that the necessary minimum wage might well be strictly positive if A's reservation utility is strictly positive, or exerting low effort is associated with a strictly positive effort cost.

References

1. Hart, O.; Holmstrom, B. The Theory of Contracts. In *Advances in Economics and Econometrics, Econometric Society Monographs, Fifth World Congress*; Bewely, T., Ed.; Camebridge University Press: Camebridge, UK, 1987; pp. 71–155.
2. Innes, R.D. Limited Liability and Incentive Contracting with Ex-Ante Action Choices. *J. Econ. Theory* **1990**, *52*, 45–67. [CrossRef]
3. Baliga, S.; Sjöström, T. Decentralization and Collusion. *J. Econ. Theory* **1998**, *83*, 196–232. [CrossRef]
4. Pitchford, R. Moral Hazard and Limited Liability: The Real Effects of Contract Bargaining. *Econ. Lett.* **1998**, *61*, 251–259. [CrossRef]
5. Berkovitch, E.; Israel, R.; Spiegel, Y. A Double Moral Hazard Model of Organization Design. *J. Econ. Manag. Strategy* **2010**, *19*, 55–85. [CrossRef]
6. Kaya, A.; Vereshchagina, G. Partnerships versus Corporations: Moral Hazard, Sorting, and Ownership Structure. *Am. Econ. Rev.* **2001**, *104*, 291–307. [CrossRef]
7. Dai, T.; Jerath, K. Salesforce Compensation with Inventory Considerations. *Manag. Sci.* **2013**, *59*, 2490–2501. [CrossRef]
8. Kräkel, M.; Schöttner, A. Optimal Sales Force Compensation. *J. Econ. Behav. Organ.* **2016**, *126*, 179–195. [CrossRef]
9. Schöttner, A. Optimal Sales Force Compensation in Dynamic Settings: Commissions vs. Bonuses. *Manag. Sci.* **2017**, *63*, 1529–1544. [CrossRef]
10. Schmitz, P.W. Allocating Control in Agency Problems with Limited Liability and Sequential Hidden Actions. *RAND J. Econ.* **2005**, *36*, 318–336.
11. Hoppe, E.I.; Kusterer, D.J. Conflicting Tasks and Moral Hazard: Theory and Experimental Evidence. *Eur. Econ. Rev.* **2011**, *55*, 1094–1108. [CrossRef]
12. Kragl, J.; Schöttner, A. Wage Floors, Imperfect Performance Measures, and Optimal Job Design. *Int. Econ. Rev.* **2014**, *55*, 525–550. [CrossRef]
13. Pi, J. Job Design with Sequential Tasks and Outcome Externalities Revisited. *Econ. Lett.* **2014**, *125*, 274–277. [CrossRef]
14. Pi, J. Another Look at Job Design with Conflicting Tasks. *Aust. Econ. Pap.* **2018**, *57*, 427–434. [CrossRef]
15. Che, Y.K.; Yoo, S.W. Optimal Incentives for Teams. *Am. Econ. Rev.* **2001**, *93*, 525–541. [CrossRef]
16. Tamada, Y.; Tsai, T.S. Delegating the Decision-Making Authority to Terminate a Sequential Project. *J. Econ. Behav. Organ.* **2014**, *99*, 178–194. [CrossRef]
17. Kräkel, M.; Schöttner, A. Delegating Pricing Authority to Sales Agents: The Impact of Kickbacks. *Manag. Sci.* **2019**, *66*, 2686–2705. [CrossRef]
18. At, C.; Friehe, T.; Gabuthy, Y. On Lawyer Compensation when Appeals are Possible. *BE J. Econ. Anal. Policy* **2019**, *19*, 1–11. [CrossRef]
19. Schmitz, P.W. Workplace Surveillance, Privacy Protection, and Efficiency Wages. *Labour Econ.* **2005**, *12*, 727–738. [CrossRef]

20. Nieken, P.; Schmitz, P.W. Repeated Moral Hazard and Contracts with Memory: A Laboratory Experiment. *Games Econ. Behav.* **2012**, *75*, 1000–1008. [CrossRef]
21. Hoppe, E.I.; Schmitz, P.W. Hidden Action and Outcome Contractibility: An Experimental Test of Moral Hazard Theory. *Games Econ. Behav.* **2018**, *109*, 544–564. [CrossRef]
22. Galinsky, E.; Bond, J.T.; Kim, S.; Backon, L.; Brownfield, E.; Sakai, K. *Overwork in America: When the Way We Work Becomes Too Much*; Families and Work Institute: New York, NY, USA, 2005.
23. Harnois, G.; Gabriel, P. *Mental Health and Work: Impact, Issues and Good Practices*; World Health Organization and International Labour Organisation: Geneva, Switzerland, 2000; WHO/MSD/MPS/00.2.
24. Galinsky, E. *Ask the Children: What America's Children Really Think about Working Parents*; William Morrow and Company, Inc.: New York, NY, USA, 1999.
25. Bordalo, P.; Gennaioli, N.; Shleifer, A. Salience Theory for Choice under Risk. *Q. J. Econ.* **2012**, *127*, 1243–1285. [CrossRef]
26. Kontek, K. A Critical Note on Salience Theory of Choice under Risk. *Econ. Lett.* **2016**, *143*, 168–171. [CrossRef]
27. Bako, B.; Neszveda, G. The Achilles' Heel of Salience Theory and How to Fix It. *Econ. Lett.* **2020**, *193*, 109265. [CrossRef]
28. Holmström, B. Moral Hazard and Observability. *Bell J. Econ.* **1979**, *10*, 74–91. [CrossRef]
29. Grossman, S.J.; Hart, O.D. An Analysis of the principal–agent Problem. *Econometrica* **1983**, *51*, 7–46. [CrossRef]
30. Kräkel, M.; Schöttner, A. Minimum Wages and Excessive Labor Supply. *Econ. Lett.* **2010**, *108*, 341–344. [CrossRef]
31. Ohlendorf, S.; Schmitz, P.W. Repeated Moral Hazard and Contracts with Memory: The Case of Risk-Neutrality. *Int. Econ. Rev.* **2012**, *53*, 433–452. [CrossRef]
32. Englmaier, F.; Fahn, M.; Schwarz, M.A. *Long-Term Employment Relations When Agents Are Present Biased*; CEPR Discussion Paper No. DP13227; 2018. Available online: https://econpapers.repec.org/paper/cprceprdp/13227.htm (accessed on 6 December 2021).
33. Goldlücke, S.; Schmitz, P.W. Pollution Claim Settlements Reconsidered: Hidden Information and Bounded Payments. *Eur. Econ. Rev.* **2018**, *110*, 211–222. [CrossRef]
34. Bordalo, P.; Gennaioli, N.; Shleifer, A. Salience and Consumer Choice. *J. Political Econ.* **2013**, *127*, 803–843. [CrossRef]
35. Herweg, F.; Müller, D.; Weinschenk, P. Salience in Markets. In *Handbook of Behavioral Industrial Organization*; Edward Elgar Publishing: Cheltenham, UK, 2018.
36. Adrian, N. *Price Discrimination and Salience-Driven Consumer Preferences*; Technical Report, Discussion Papers; University of Bern: Bern, Switzerland, 2019.
37. Friehe, T.; Pham, C.L. Settling with Salience-biased Defendants. *Econ. Lett.* **2020**, *192*, 109235. [CrossRef]
38. Mungan, M.C. Salience and the Severity versus the Certainty of Punishment. *Int. Rev. Law Econ.* **2019**, *57*, 96–100. [CrossRef]
39. de la Rosa, L.E. Overconfidence and Moral Hazard. *Games Econ. Behav.* **2011**, *73*, 429–451. [CrossRef]
40. Müller, D.; Schmitz, P.W. The Right to Quit Work: An Efficiency Rationale for Restricting the Freedom of Contract. *J. Econ. Behav. Organ.* **2021**, *184*, 653–669. [CrossRef]
41. Laffont, J.J.; Martimort, D. *The Theory of Incentives: The Principal–Agent Model*; Princeton University Press: Princeton, NJ, USA, 2002.
42. Salanié, B. *The Economics of Contracts: A Primer*; MIT Press: Cambridge, MA, USA, 2005.
43. Bordalo, P.; Gennaioli, N.; Shleifer, A. Salience and Asset Prices. *Am. Econ. Rev. Pap. Proc.* **2013**, *103*, 623–628. [CrossRef]
44. Cosemans, M.; Frehen, R. Salience Theory and Stock Prices: Empirical Evidence. *J. Financ. Econ.* **2021**, *140*, 460–483. [CrossRef]
45. Königsheim, C.; Lukas, M.; Nöth, M. Salience Theory: Calibration and Heterogeneity in Probability Distortion. *J. Econ. Behav. Organ.* **2019**, *157*, 477–495. [CrossRef]
46. Alós-Ferrer, C.; Ritschel, A. Attention and Salience in Preference Reversals. *Exp. Econ.* **2022**. [CrossRef]
47. Sandroni, A.; Squintani, F. Overconfidence, Insurance, and Paternalism. *Am. Econ. Rev.* **2007**, *97*, 617–622. [CrossRef]
48. Grubb, M. Selling to Overconfident Consumers. *Am. Econ. Rev.* **2009**, *99*, 1770–1807. [CrossRef]
49. Herweg, F.; Müller, D. Overconfidence in the Markets for Lemons. *Scand. J. Econ.* **2016**, *118*, 354–371. [CrossRef]
50. Eliaz, K.; Spiegler, R. Contracting with Diversely Naive Agents. *Rev. Econ. Stud.* **2006**, *73*, 689–714. [CrossRef]
51. Eliaz, K.; Spiegler, R. Consumer Optimism and Price Discrimination. *Theor. Econ.* **2008**, *3*, 459–497.
52. Luce, R.D. *Individual Choice Behavior: A Theoretical Analysis*; Wiley: New York, NY, USA, 1959.
53. Tversky, A. Features of Similarity. *Psychol. Rev.* **1977**, *84*, 327. [CrossRef]
54. Loomes, G.; Sugden, R. Regret Theory: An Alternative Theory of Rational Choice under Uncertainty. *Econ. J.* **1982**, *92*, 805–824. [CrossRef]
55. Kőszegi, B.; Szeidl, A. A Model of Focusing in Economic Choice. *Q. J. Econ.* **2013**, *128*, 53–104. [CrossRef]
56. Bhatia, S.; Loomes, G.; Read, D. Establishing the Laws of Preferential Choice Behavior. *Judgm. Decis. Mak.* **2021**, *16*, 1324–1369.
57. Kahneman, D.; Tversky, A. On the Study of Statistical Intuitions. *Cognition* **1982**, *11*, 123–141. [CrossRef]