Journal of Economic Theory and Econometrics, Vol. 32, No. 1, Mar. 2021, 54–79

Optimal Wage Contracts under Reference-Dependent Preferences*

Gue Sung Choi† Jaeok Park‡

Abstract In this paper, we consider a labor market consisting of a firm and a worker and study employment contracts that the firm chooses when the worker has reference-dependent preferences with respect to wages. The firm offers an employment contract, which specifies the effort level and the wage in each period, and if the worker accepts the contract, she decides whether to continue working for the firm in each period. The worker forms a reference wage in each period based on the past wages, and we introduce gain–loss utility into the worker's payoffs from employment and unemployment. We show, among other results, that when the initial reference wage is low, the firm's optimal employment contract has initial high wages if the worker is loss averse and initial low wages if she is gain seeking. Our results provide explanations for signing bonuses and seniority-based pay systems based on reference-dependent preferences.

Keywords Reference Dependence, Wage Contracts, Gain–Loss Utility, Loss Aversion, Gain Seeking

JEL Classification D91, J31, C61, M50

*We are grateful to Jay Pil Choi, Chulyoung Kim, Sang-hyun Kim, and two anonymous referees for helpful comments and suggestions.

†Department of Economics, The University of Texas at Austin, Austin, TX 78712, United States. E-mail: gschoi@utexas.edu

‡Corresponding author. School of Economics, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea. E-mail: jaeok.park@yonsei.ac.kr

Received January 29, 2021, Revised February 22, 2021, Accepted February 23, 2021

1. INTRODUCTION

It is common for a person to assess the outcome of a choice by its deviation from a reference point as well as its intrinsic value. We can observe reference dependence in a variety of decision making scenarios, including a worker's assessment of wages. Provided that a worker's main source of income is wages, her consumption expenditure will be largely determined by her wages, and we can think of her reference wage as the wage level that is necessary to maintain her consumption habits. If she suddenly gets laid off or takes a big pay cut, she needs to find an alternative source of income to purchase necessity goods, which typically brings about monetary or psychological cost. In contrast, if she receives an unexpected huge bonus, she can spend part of it on luxury goods, traveling, or fine dining and obtain extra satisfaction from it.

The goal of this paper is to study employment contracts that a firm offers to a worker who has reference-dependent preferences with respect to wages. The worker's reference dependence creates two contrasting incentives for the firm. As the worker has a higher reference wage, she suffers a larger loss from unemployment. Hence, the worker with a higher reference wage is willing to accept a lower wage for a given level of effort, and the firm may want to offer high wages at the beginning in order to increase the worker's reference wage and then exploit her fear of unemployment. On the other hand, when the wage exceeds the reference wage, the worker obtains a larger gain from the wage as her reference wage is lower. In other words, given a wage higher than the reference wage, a lower reference wage increases the worker's satisfaction from employment and thus makes the worker willing to put more effort. Hence, the firm may want to set low wages at the beginning in order to reduce the worker's reference wage and then take advantage of her satisfaction from employment. In the real world, we can find wage schemes consistent with each of these two incentives. Offering high wages at the beginning can be implemented by a signing bonus, while paying low wages to early-career employees is widely observed in a pay scale of a firm.

To facilitate our analysis, we consider a simple labor market consisting of a firm and a worker and abstract away from the worker's consumption–savings problem and her moral hazard problem. The firm offers an employment contract, which specifies the effort level and the wage in each period, and if the worker accepts the contract, she decides whether to continue working for the firm in each period. The worker forms a reference wage in each period based on the past wages, and we introduce "gain–loss utility" as in Kőszegi and Rabin (2006) into the worker's payoffs from employment and unemployment. Regarding the worker's attitude towards gains and losses, we consider the two cases of loss aversion and gain seeking. Although loss aversion can be defined formally in different ways (see, for example, Abdellaoui *et al.*, 2007), it basically reflects the idea that losses loom larger than gains. In our context, a loss-averse worker is hurt more by a wage cut from her reference wage than she is satisfied with a wage increase by the same amount. Gain seeking is termed by Abdellaoui *et al.* (2007), and it expresses the opposite idea to loss aversion. While loss aversion is a widely accepted principle in economics, Schmidt and Traub (2002) and Brooks and Zank (2005) find that roughly a quarter of subjects are gain seeking in their experimental tests.

The main results of this paper can be summarized as follows. When the initial reference wage is low,[1] the firm's optimal employment contract has initial high wages if the worker is loss averse and initial low wages if she is gain seeking. On the other hand, when the initial reference wage is high, the firm's optimal employment contract has constant wages independent of the reference wage. The results about a low initial reference wage are intuitive in view of the aforementioned two incentives of the firm. When the worker is loss averse, her fear of unemployment is stronger than her satisfaction from employment, and the firm's first incentive dominates the second one. In contrast, when the worker is gain seeking, she reacts more sensitively to her satisfaction from employment, and the firm's second incentive shapes the wage structure. These results can be interpreted as providing explanations for signing bonuses and seniority-based pay systems based on reference-dependent preferences. When the initial reference wage is high, the worker experiences losses from both employment and unemployment, and the effects of a reference wage cancel out. Thus, the result about a high initial reference wage holds.

Reference dependence and loss aversion are central principles in prospect theory (Kahneman and Tversky, 1979), and they have been widely studied in economics from both theoretical and empirical perspectives. Among many studies on these topics, a notable one is Kőszegi and Rabin (2006), who provide a general framework to study decision making under reference-dependent preferences and loss aversion. We compare our model with the framework of Kőszegi and Rabin (2006) in Remark 1 at the end of Section 2. The implications of reference dependence and loss aversion have been explored in various contexts, including labor markets. There is a strand of literature that studies the role of a

[1] We can think of a worker with a low initial reference wage as an entry-level worker and one with a high initial reference wage as an experienced worker who enters the labor market after quitting a well-paid job.

reference (or "target") income level on the labor supply of New York City taxi drivers (Camerer *et al.*, 1997; Farber, 2005, 2008; Crawford and Meng, 2011). While this literature deals with a worker's decision on daily work hours or effort, in our model the firm dictates the worker's effort, and we focus on the structure of employment contracts the firm offers to the worker. There are also previous studies that adopt the idea that wages are habit-forming and take the past wage as a reference point. Clark (1999) and Grund and Sliwka (2007) show empirical evidence that wage increases improve job satisfaction and provide an explanation for rising wage profiles based on reference-dependent preferences. Grund and Sliwka (2007) also present a theoretical model to explain the empirical evidence. In their model, the worker myopically chooses an effort level, which determines the wage. In contrast, in our model, the firm chooses both effort and the wage, and both the firm and the worker are foresighted.

The remainder of this paper is organized as follows. We describe our model in Section 2 and characterize optimal employment contracts in Section 3. We conclude in Section 4. Proofs of all theorems and propositions are relegated to the Appendix.

2. THE MODEL

We consider a simple labor market consisting of a firm and a worker. Time is discrete and infinite, and it is denoted by $t = 0, 1, 2, \ldots$. Both the firm and the worker are infinitely lived and discount future payoffs by a discount factor $\delta \in (0, 1)$. At the beginning of period 0, the firm offers an employment contract $\{e_t, w_t\}_{t=0}^{\infty}$, where e_t and w_t are nonnegative real numbers for all $t = 0, 1, \ldots$. For all $t = 0, 1, \ldots$, e_t represents the effort level that the firm requires from the worker, and w_t denotes the wage that the firm pays to the worker. After observing the contract offered, the worker decides whether to accept the contract or not. If the worker accepts the contract, she is free to leave the firm at any point of time, but once she leaves, she cannot return and remains unemployed forever. If the worker is employed in period t, she exerts the effort level e_t and receives the wage w_t. That is, we abstract away from the worker's moral hazard problem, assuming that the firm can enforce any effort level it specifies. When the worker chooses the effort level e_t, she produces output $y(e_t)$ and incurs cost $c(e_t)$. We impose the following regularity assumptions on the functions y and c.

Assumption 1. *The functions $y : \mathbb{R}_+ \to \mathbb{R}_+$ and $c : \mathbb{R}_+ \to \mathbb{R}_+$ are twice differentiable on $\mathbb{R}_{++}$. They satisfy $y(0) = c(0) = 0$, $y'(e) > 0$, $y''(e) < 0$, $\lim_{e\to 0^+} y'(e) = \infty$, $\lim_{e\to\infty} y'(e) = 0$, $c'(e) > 0$, $c''(e) > 0$, $\lim_{e\to 0^+} c'(e) = 0$, and $\lim_{e\to\infty} c'(e) =$*

∞.

The functions y and c have value zero at $e = 0$ and are strictly increasing. The function y is strictly concave, while c is strictly convex.

A departure from a standard model is that the worker has reference-dependent preferences with respect to wages. In each period $t = 0, 1, \ldots$, the worker has a reference wage, denoted by $r_t \in \mathbb{R}_+$. A reference wage can be interpreted as the wage level that the worker is accustomed to. The initial reference wage r_0 is given, and the reference wages in later periods evolve according to the equation

$$r_{t+1} = \rho r_t + (1-\rho) w_t \tag{1}$$

for all $t = 0, 1, \ldots$, where $\rho \in (0, 1)$. Using (1), we can obtain the relationship

$$r_t = \rho^t r_0 + (1-\rho) \sum_{\tau=0}^{t-1} \rho^{t-1-\tau} w_\tau \tag{2}$$

for all $t = 1, 2, \ldots$. That is, the reference wage in a period is determined by the initial reference wage and the wages in the past periods. The worker receives utility from the absolute levels of wages while she is employed. In addition, the worker obtains a gain if the actual wage exceeds the reference wage, while she suffers a loss if the actual wage falls short of the reference wage. That is, the worker's utility from wages consists of two components: one that depends on absolute levels and the other that depends on differences from reference points. Let $R : \mathbb{R} \to \mathbb{R}$ be a function defined by

$$R(x) = \begin{cases} \lambda_g x & \text{if } x \geq 0, \\ \lambda_l x & \text{if } x < 0, \end{cases}$$

where $\lambda_g, \lambda_l \geq 0$. We represent the gain–loss component in period t by $R(w_t - r_t)$. We say that the worker is *loss averse* if $\lambda_l > \lambda_g$ and *gain seeking* if $\lambda_g > \lambda_l$.

The worker has no other source of income than receiving wages from the firm, and thus once she leaves the firm, her income becomes zero afterwards. If the worker becomes unemployed at the beginning of period t, it is as if she remains employed with the terms $(e_\tau, w_\tau) = (0, 0)$ for all $\tau \geq t$. Hence, it is without loss of generality to assume that the firm offers a contract that induces the worker to be employed in every period. If the worker is employed in every period given the contract $\{e_t, w_t\}_{t=0}^{\infty}$, the firm's payoff is given by $\sum_{t=0}^{\infty} \delta^t [y(e_t) - w_t]$ and the worker's payoff by $\sum_{t=0}^{\infty} \delta^t [w_t + R(w_t - r_t) - c(e_t)]$, where r_t is determined by (2) for all $t = 1, 2, \ldots$.

If the worker leaves the firm in period t with the reference wage r_t, the reference wage in period $\tau > t$ becomes $r_\tau = \rho^{\tau-t} r_t$, and thus her continuation payoff is given by

$$L(r_t) := -\frac{\lambda_l}{1-\delta\rho} r_t. \tag{3}$$

As the worker is accustomed to a higher wage level, she suffers more from unemployment. The worker's participation constraint in period τ can be written as

$$\sum_{t=\tau}^{\infty} \delta^{t-\tau}[w_t + R(w_t - r_t) - c(e_t)] \geq L(r_\tau).$$

The firm chooses an employment contract to maximize its payoff while satisfying the worker's participation constraint in every period. Hence, the firm's problem can be expressed as

$$\max_{\{e_t, w_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \delta^t [y(e_t) - w_t] \tag{4}$$

$$\text{subject to } \sum_{t=\tau}^{\infty} \delta^{t-\tau}[w_t + R(w_t - r_t) - c(e_t)] \geq L(r_\tau) \quad \forall \tau = 0, 1, \ldots, \tag{5}$$

$$r_{t+1} = \rho r_t + (1-\rho) w_t \quad \forall t = 0, 1, \ldots, \tag{6}$$

$$r_0 \geq 0 \text{ given}, \tag{7}$$

$$e_t, w_t \geq 0 \quad \forall t = 0, 1, \ldots. \tag{8}$$

We refer to a solution to the firm's problem as an *optimal (employment) contract*. If an optimal contract satisfies the participation constraint in every period with equality, we refer to it as an *optimal contract with binding participation constraints*. In the formulation of the firm's problem, it is implicitly assumed that the firm can commit to its contract, and thus it cannot change the terms of the contract after period 0. However, when an optimal contract satisfies all the participation constraints with equality, the firm does not gain by revising the contract after period 0. Hence, we can think of an optimal contract with binding participation constraints as a more robust solution than an optimal contract, and we can use it to select a desirable optimal contract.

Remark 1. A person's utility in the framework of Kőszegi and Rabin (2006) has two components, "consumption utility" and "gain–loss utility," and they are additively separable across consumption dimensions. The worker's payoff in our model has consumption utility from wages and effort, but gain–loss utility applies only to wages. That is, in the worker's payoff $\sum_{t=0}^{\infty}\delta^t[w_t+R(w_t-r_t)-c(e_t)]$, w_t and $-c(e_t)$ represent consumption utility from wages and effort, respectively, while $R(w_t-r_t)$ is gain–loss utility from wages. The particular form of the gain–loss function $R(x)$ with loss aversion in our model satisfies all the properties of a "universal gain–loss function" in Kőszegi and Rabin (2006) and is used in their two applications. A reference point in the framework of Kőszegi and Rabin (2006) is determined by rational expectations, whereas it is formed by adaptive expectations in our model. In our model, the wage profile in an employment contract is deterministic, and thus if the reference wage is formed by rational expectations, the worker sets the reference wage equal to the actual wage and gain–loss utility does not play a role.[2] By using adaptive expectations, we capture the worker's adaptation to past wages (or habit formation[3]). There is also an advantage of adaptive expectations in terms of mathematical tractability, as can be seen from our analysis in the next section. We can think of other kinds of reference wage formation processes such as $r_{t+1}=\max\{r_t,w_t\}$ and $r_{t+1}=\min\{r_t,w_t\}$, and it is also possible that a worker's reference wage is determined as the average wage of others in her peer group. We leave it for future research to study the effects of using different reference wage formation processes on optimal employment contracts and to identify the determinants of reference wages.

3. OPTIMAL CONTRACTS

In this section, we study optimal contracts considering the two cases of loss aversion and gain seeking. Below we introduce some notations that will be used to describe optimal contracts.

$$\alpha_l := 1+\lambda_l-\lambda_l\frac{\delta(1-\rho)}{1-\delta\rho}=1+\lambda_l\frac{1-\delta}{1-\delta\rho},$$
$$\alpha_g := 1+\lambda_g-\lambda_g\frac{\delta(1-\rho)}{1-\delta\rho}=1+\lambda_g\frac{1-\delta}{1-\delta\rho},$$

[2]Another situation in which gain–loss utility does not play a role occurs when we focus on a steady state under adaptive expectations, in which the reference wage coincides with the actual wage. Thus, in order to study the effects of gain–loss utility and the reference wage, we consider the discounted payoffs of the firm and the worker instead of their steady-state, or long-run, payoffs.

[3]There is a large literature on habit formation in consumption. For example, Won (2019) studies general equilibrium models where consumption habit is given exogenously or determined by current or past endowments.

$$
\begin{aligned}
\alpha_{gl} &:= 1+\lambda_g-\lambda_l\frac{\delta(1-\rho)}{1-\delta\rho},\\
\hat{e}(w) &:= c^{-1}(\alpha_l w) \quad \text{for all } w \geq 0,\\
e^l &\in \arg\max_{e\geq 0} y(e)-\frac{1}{\alpha_l}c(e),\\
e^g &\in \arg\max_{e\geq 0} y(e)-\frac{1}{\alpha_g}c(e),\\
w^l &:= \frac{1}{\alpha_l}c(e^l),\\
w^g &:= \frac{1}{\alpha_l}c(e^g).
\end{aligned}
$$

The numbers α_l and α_g represent the marginal effect of the wage in any period τ (i.e., w_τ) on the worker's total payoff in period τ (i.e., $\sum_{t=\tau}^{\infty}\delta^{t-\tau}[w_t+R(w_t-r_t)-c(e_t)]$) when there are a loss and a gain, respectively, in every period from period τ (i.e., $w_t<r_t$ for every $t\geq\tau$ and $w_t>r_t$ for every $t\geq\tau$, respectively). Note that, when λ_l and λ_g are positive, α_l and α_g are larger than 1, which means that gain–loss utility motivates the worker. When the participation constraints in periods τ and $\tau+1$ are binding, we can think of the period-τ component of the worker's payoff as $w_\tau+R(w_\tau-r_\tau)-c(e_\tau)+\delta L(r_{\tau+1})$. The number α_{gl} represents the marginal effect of the wage in any period τ on the period-τ component of the worker's total payoff when there is a gain in period τ. When the marginal payoff of the wage is α_l, the firm can induce the worker to put the effort level e satisfying $c(e)=\alpha_l w$ by paying the wage w. This effort level is denoted by $\hat{e}(w)$. Since the function c is strictly increasing with range $\mathbb{R}_+$, $\hat{e}(w)$ is well-defined and strictly increasing in w. The effort levels e^l and e^g can be interpreted as the optimal effort levels for the firm when the marginal payoffs of the wage are α_l and α_g, respectively. The wages w^l and w^g are the ones to compensate for the effort levels e^l and e^g when the marginal payoff of the wage is α_l. By Assumption 1, e^l and e^g are uniquely defined and positive, and thus w^l and w^g are positive. If $\lambda_l>\lambda_g$, we have $\alpha_l>\alpha_g>\alpha_{gl}$, $e^l>e^g$, and $w^l>w^g$. If $\lambda_g>\lambda_l$, we have $\alpha_{gl}>\alpha_g>\alpha_l$, $e^g>e^l$, and $w^g>w^l$.

3.1. LOSS-AVERSE WORKER

In the following theorem, we characterize the firm's optimal contract when the worker is loss averse.

Theorem 1. *Suppose that*

$$\lambda_g < \lambda_l < \frac{1-\delta\rho}{1-\rho}(\lambda_g+\rho).$$

(i) If $r_0 > w^l$, then there exists a unique optimal contract with binding participation constraints $\{e_t^*, w_t^*\}_{t=0}^{\infty}$, and it is given by $e_t^* = e^l$ and $w_t^* = w^l$ for all $t = 0, 1, \ldots$.

(ii) If $r_0 < w^g$, then there exists a unique optimal contract with binding participation constraints $\{e_t^*, w_t^*\}_{t=0}^{\infty}$, and it is given by $e_t^* = e^g$ and

$$w_t^* = \frac{1}{\alpha_{gl}}\left[c(e^g) - (\lambda_l - \lambda_g)r_t\right] \tag{9}$$

for all $t = 0, 1, \ldots$.

(iii) If $w^g \le r_0 \le w^l$, then there exists a unique optimal contract $\{e_t^*, w_t^*\}_{t=0}^{\infty}$, and it is given by $e_t^* = \hat{e}(r_0)$ and $w_t^* = r_0$ for all $t = 0, 1, \ldots$.

The firm's payoff at an optimal contract is given by

$$\Pi(r_0) = \begin{cases} \frac{1}{1-\delta}\left[y(e^l) - \frac{1}{\alpha_l}c(e^l)\right] & \text{if } r_0 > w^l, \\ \frac{1}{1-\delta}\left[y(\hat{e}(r_0)) - r_0\right] & \text{if } w^g \le r_0 \le w^l, \\ \frac{1}{1-\delta}\left[y(e^g) - \frac{1}{\alpha_g}c(e^g)\right] + \frac{\lambda_l - \lambda_g}{\alpha_g(1-\delta\rho)}r_0 & \text{if } r_0 < w^g, \end{cases} \tag{10}$$

and the worker's payoff by $L(r_0)$.

In Theorem 1, we impose the assumption that $\lambda_g < \lambda_l < (1-\delta\rho)(\lambda_g+\rho)/(1-\rho)$. The first part of the assumption (i.e., $\lambda_g < \lambda_l$) means that the worker is loss averse. The second part (i.e., $\lambda_l < (1-\delta\rho)(\lambda_g+\rho)/(1-\rho)$) is imposed in order to guarantee the existence of an optimal contract. When λ_l is too large, the worker with a high reference wage suffers a huge loss from unemployment, and the firm may want to keep increasing the wage and thus the reference wage without bound in order to exploit the worker's fear of unemployment. Note that $\lambda_g < (1-\delta\rho)(\lambda_g+\rho)/(1-\rho)$ holds and thus the interval $(\lambda_g, (1-\delta\rho)(\lambda_g+\rho)/(1-\rho))$ is nontrivial for any $\lambda_g \ge 0$. Also, the assumption that $\lambda_l < (1-\delta\rho)(\lambda_g+\rho)/(1-\rho)$ implies $\alpha_{gl} > 0$.

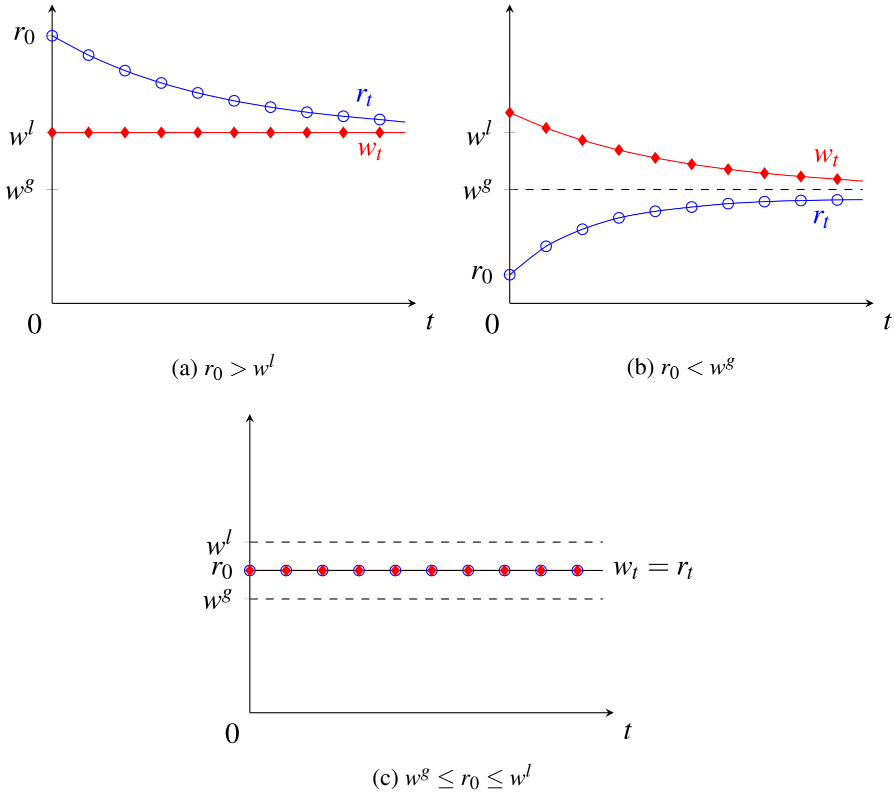
First, consider the case where the initial reference wage is sufficiently high (i.e., $r_0 > w^l$). In this case, an optimal contract involves a loss in every period (i.e., $w_t < r_t$ for all t). Then increasing the wage in a period by a unit reduces the loss by λ_l in that period. At the same time, it induces higher reference wages and thus larger losses in the future periods, whose aggregate effect in terms of the current payoff is given by $-\lambda_l\delta(1-\rho)/(1-\delta\rho)$. Hence, an extra unit of the wage increases the worker's payoff by $\alpha_l = 1+\lambda_l-\lambda_l\delta(1-\rho)/(1-\delta\rho)$. As a result, it is optimal for the firm to set the effort level at e^l, which maximizes $y(e)-c(e)/\alpha_l$, in every period.[4] There are different ways of paying wages over time to compensate for this effort level while satisfying the participation constraint in every period. By setting all the participation constraints with equality, we can pin down the wage in every period as w^l. The initial reference wage affects the losses from being employed and unemployed, and these two effects cancel out exactly. Hence, the initial reference wage has no impact on the optimal contract with binding participation constraints.

Next, consider the case where the initial reference wage is sufficiently low (i.e., $r_0 < w^g$). In this case, an optimal contract involves a gain in every period (i.e., $w_t > r_t$ for all t). Analogously to the previous case, an extra unit of the wage increases the worker's payoff by α_g, and it is optimal for the firm to select the effort level as e^g, which maximizes $y(e)-c(e)/\alpha_g$, in every period. Again, there are many ways of choosing wages over time to satisfy all the participation constraints, but setting all the participation constraints with equality yields $w_t^* = [c(e^g)-(\lambda_l-\lambda_g)r_t]/\alpha_{gl}$ for all t. Because the worker is loss averse, her fear of unemployment dominates the gains she obtains while being employed. As the worker has a higher reference wage, she becomes more afraid of being unemployed, and thus the firm can induce the same level of effort with a lower wage. In other words, a higher reference wage results in a larger wage discount, as can be seen from w_t^* in (9) decreasing in r_t. Thus, unlike in the previous case, the initial reference wage affects the optimal contract with binding participation constraints. Since $w_t^* > r_t$ for all t, r_t is strictly increasing, and thus w_t^* is strictly decreasing over time. Both $\{w_t^*\}$ and $\{r_t\}$ converge to w^g. In the optimal contract $\{e_t^*, w_t^*\}_{t=0}^{\infty}$, the firm pays high wages in early periods in order to induce high reference wages, even though it requires the worker to put a constant level of effort. In the real world, this kind of wage paths can be implemented by having a signing bonus paid over time.[5] In the optimal contract $\{e_t^*, w_t^*\}_{t=0}^{\infty}$, we can think of w^g as the base wage and $w_t^* - w^g$ as the fraction of the signing bonus paid in period t.

Finally, consider the case where the initial reference wage is intermediate (i.e., $w^g \le r_0 \le w^l$). In this case, the firm sets the wage equal to the initial reference wage in every period in the optimal contract. The kink in the graph

[4]Although it is not stated in Theorem 1, the effort level is uniquely determined by an optimal contract.

[5]A signing bonus is typically a one-time payment, but in sports contracts, it is sometimes paid over several years.



(a) $r_0 > w^l$ (b) $r_0 < w^g$

(c) $w^g \le r_0 \le w^l$

Figure 1: Paths of wages and reference wages in optimal contracts in the case of a loss-averse worker

of the gain–loss function $R(x)$ at the origin creates the status quo bias. The gain coefficient λ_g is not high enough for the firm to induce gains from the initial reference wage, while the loss coefficient λ_l is too high for the firm to induce losses. As a result, the firm creates neither a gain nor a loss and matches the wage to the initial reference wage in every period, and the reference wage remains constant over time. Because the fear of unemployment motivates the worker, the firm can sustain the effort level $\hat{e}(r_0)$ with wage r_0. In Figure 1, we illustrate the paths of wages and reference wages in the optimal contract with binding participation constraints in each of the three cases.

We compare the effort level and the agents' payoffs with those in the benchmark case of no gain–loss utility (i.e., $\lambda_g = \lambda_l = 0$). When the worker has no gain–loss utility, the optimal contract with binding participation constraints is given by $e_t = e^o$ and $w_t = c(e^o)$ for all $t = 0, 1, \ldots$, where e^o denotes the unique maximizer of $y(e) - c(e)$. The firm's payoff at this optimal contract is $[y(e^o) - c(e^o)]/(1-\delta)$, and that of the worker is 0. As we have seen in the above discussions, gain–loss utility further motivates the worker, which makes the firm select an effort level higher than e^o in an optimal contract. An optimal contract makes the worker indifferent between accepting it and rejecting it. As the worker expects a larger loss from unemployment, she is willing to accept a less attractive contract. Hence, the worker's payoff at an optimal contract gets lower as she has a higher initial reference wage and as she becomes more loss averse (i.e., the loss coefficient λ_l gets larger). When the initial reference wage is low, the firm can exploit the worker's fear of unemployment. Hence, the firm's payoff at an optimal contract increases with r_0 for $r_0 \leq w^l$ and is constant for $r_0 \geq w^l$. When $r_0 = 0$, the worker suffers no loss from unemployment, but even in this case, as long as $\lambda_g > 0$, gain–loss utility improves the firm's payoff by motivating the worker.

We can evaluate an employment contract from society's point of view by different criteria such as the total payoff (i.e., $\sum_{t=0}^{\infty} \delta^t [y(e_t) - c(e_t) + R(w_t - r_t)]$), the total payoff without gain–loss utility (i.e., $\sum_{t=0}^{\infty} \delta^t [y(e_t) - c(e_t)]$), and the total output (i.e., $\sum_{t=0}^{\infty} \delta^t y(e_t)$). In the following proposition, we study these measures at an optimal contract when the worker is loss averse.

Proposition 1. *Suppose that*

$$0 < \lambda_g < \lambda_l < \frac{1-\delta\rho}{1-\rho}(\lambda_g + \rho).$$

(i) The total payoff at an optimal contract is continuous and strictly decreasing in r_0, and it is larger than $[y(e^o) - c(e^o)]/(1-\delta)$ at $r_0 = 0$ and decreases without bound as r_0 goes to infinity.

(ii) The total payoff without gain–loss utility at an optimal contract is continuous in r_0, strictly decreasing in r_0 on $[w^g, w^l]$, and constant in r_0 elsewhere.

(iii) The total output at an optimal contract is continuous in r_0, strictly increasing in r_0 on $[w^g, w^l]$, and constant in r_0 elsewhere.

When the worker is loss averse, she is mainly motivated by her fear of unemployment, which increases with her reference wage. Hence, as the worker has a higher initial reference wage, the firm obtains a higher payoff, and the worker is exploited more severely. The total payoff at an optimal contract decreases with the initial reference wage, which means that the loss suffered by the worker outweighs the benefit in the firm's payoff. As mentioned above, when $r_0 = 0$, the worker has no fear of unemployment, while she is still motivated by gain–loss utility. As a result, the total payoff at $r_0 = 0$ exceeds the optimal total payoff without gain–loss utility, $[y(e^o) - c(e^o)]/(1-\delta)$. The total payoff without gain–loss utility is maximized by setting $e_t = e^o$ for all t, and it decreases as e_t moves away from e^o. As the worker has a higher initial reference wage, she is induced to put more effort at a level higher than e^o, which reduces the total payoff without gain–loss utility and increases the total output.

3.2. GAIN-SEEKING WORKER

Now we consider a gain-seeking worker. When the worker is gain seeking, the gain–loss function $R(x)$ is convex, and the firm may benefit from alternating between gains and losses. In order to make the firm's problem more tractable, we restrict attention to employment contracts that induce monotone sequences of reference wages. That is, we add the constraint that the sequence $\{r_t\}$ is monotone (i.e., $r_{t+1} \geq r_t$ for all $t = 0, 1, \ldots$ or $r_{t+1} \leq r_t$ for all $t = 0, 1, \ldots$) to the firm's problem (4)–(8) and refer to a solution to this restricted problem as an *optimal monotone contract*. If an optimal monotone contract satisfies all the participation constraints with equality, we refer to it as an *optimal monotone contract with binding participation constraints*. As before, we can regard an optimal monotone contract with binding participation constraints as a robust solution in the sense that the firm has no incentive to revise the contract after period 0.

In the following theorem, we characterize the firm's optimal monotone contract when the worker is gain seeking.

Theorem 2. *Suppose that $\lambda_g > \lambda_l$. There exists a unique value of $\tilde{r} \in (w^l, w^g)$ such that the following statements hold.*

(i) If $r_0 > \tilde{r}$, then there exists a unique optimal monotone contract with binding participation constraints $\{e_t^*, w_t^*\}_{t=0}^{\infty}$, and it is given by $e_t^* = e^l$ and $w_t^* = w^l$ for all $t = 0, 1, \ldots$.

(ii) If $r_0 < \tilde{r}$, then there exists a unique optimal monotone contract with binding participation constraints $\{e_t^*, w_t^*\}_{t=0}^{\infty}$, and it is given by $e_t^* = e^g$ and

$$w_t^* = \frac{1}{\alpha_{gl}} [c(e^g) + (\lambda_g - \lambda_l) r_t] \tag{11}$$

for all $t = 0, 1, \ldots$.

(iii) If $r_0 = \tilde{r}$, then there are exactly two optimal monotone contracts with binding participation constraints, and they are the contracts in (i) and (ii).

The firm's payoff at an optimal monotone contract is given by

$$\Pi(r_0) = \begin{cases} \frac{1}{1-\delta}\left[y(e^l) - \frac{1}{\alpha_l}c(e^l)\right] & \text{if } r_0 \geq \tilde{r}, \\ \frac{1}{1-\delta}\left[y(e^g) - \frac{1}{\alpha_g}c(e^g)\right] - \frac{\lambda_g - \lambda_l}{\alpha_g(1-\delta\rho)} r_0 & \text{if } r_0 < \tilde{r}, \end{cases} \tag{12}$$

and the worker's payoff by $L(r_0)$.

Unlike in Theorem 1, we do not need to assume that $\lambda_l < (1-\delta\rho)(\lambda_g + \rho)/(1-\rho)$ in Theorem 2, because it is implied by $\lambda_l < \lambda_g$. Since a monotone sequence is either decreasing or increasing, we search for optimal contracts among contracts that yield a decreasing sequence of reference wages and those that yield an increasing one, which we call optimal decreasing contracts and optimal increasing contracts, respectively. If $r_0 > w^l$, the optimal decreasing contract with binding participation constraints is obtained as in Theorem 1(i). If $r_0 \leq w^l$, the constraint of decreasing $\{r_t\}$ is binding, and the optimal decreasing contract generates constant reference wages and is characterized as in Theorem 1(iii). Similarly, the optimal increasing contract with binding participation constraints is obtained as in Theorem 1(ii) if $r_0 < w^g$ and as in Theorem 1(iii) if $r_0 \geq w^g$.[6]

As explained following Theorem 1, when the contract involves a loss in every period, the effects of the initial reference wage on the worker's payoffs from employment and unemployment cancel out exactly. Thus, the initial reference wage does not affect the value of the optimal decreasing contract if $r_0 > w^l$. On the other hand, when the contract involves a gain in every period, the worker is more sensitive about the gain she obtains when she is employed than the loss she suffers when unemployed. As the worker has a higher reference wage, she obtains a smaller gain, and the firm needs to offer a higher wage in order to motivate her. That is, a higher reference wage leads to a higher wage, as can be seen from w_t^* in (11) increasing in r_t. We can interpret this as the firm paying a wage premium to the worker with a high reference wage. As a result, the value of the optimal increasing contract decreases with the initial reference wage if $r_0 < w^g$. In the optimal increasing contract for $r_0 < w^g$, the worker has a low reference wage initially and thus receives a low wage, but as she is employed longer, her reference wage gets higher and she earns a higher wage, even though the effort level remains constant. In the real world, this kind of wage paths can be generated by offering gradual pay raises according to the worker's seniority in a pay scale.

[6]As in Theorem 1, the effort level is uniquely determined by an optimal increasing or decreasing contract.

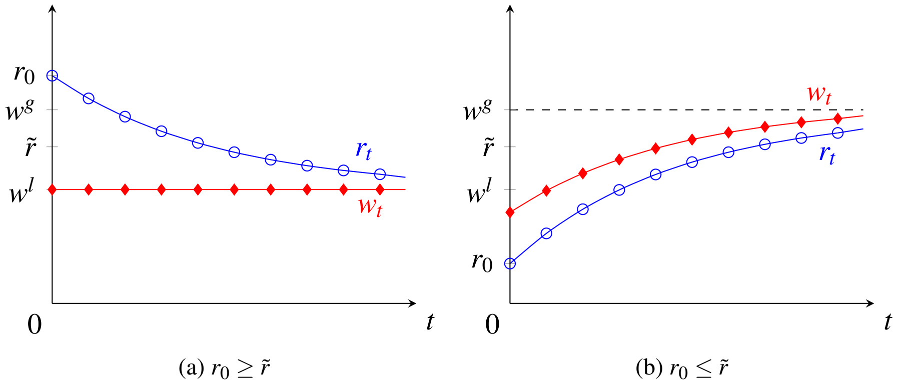


Figure 2: Paths of wages and reference wages in optimal monotone contracts in the case of a gain-seeking worker

Since $\lambda_l < \lambda_g$, we have $w^l < w^g$. Thus, for any $r_0 \in (w^l, w^g)$, the optimal decreasing contract yields a strictly decreasing sequence of reference wages, while the optimal increasing contract yields a strictly increasing one. As noted above, the value of the optimal decreasing contract is independent of r_0 if $r_0 > w^l$, and that of the optimal increasing contract decreases with r_0 if $r_0 < w^g$. We can show that there exists unique $\tilde{r} \in (w^l, w^g)$ at which the value of the optimal decreasing contract is equal to that of the optimal increasing contract. Then the optimal decreasing contract yields a higher value than the optimal increasing contract for any $r_0 > \tilde{r}$, the optimal increasing contract yields a higher value for any $r_0 < \tilde{r}$, and the firm is indifferent between the two contracts for $r_0 = \tilde{r}$. Hence, if $r_0 > \tilde{r}$, the optimal monotone contract is given by the optimal decreasing contract, which involves constant wages and decreasing reference wages. Also, if $r_0 < \tilde{r}$, the optimal monotone contract is given by the optimal increasing contract, which involves increasing wages and increasing reference wages. In Figure 2, we illustrate the paths of wages and reference wages in the optimal monotone contract with binding participation constraints.

As in the case of loss aversion, the presence of gain–loss utility results in a higher effort level, a lower payoff of the worker, and a higher payoff of the firm than those in the benchmark case of no gain–loss utility. In the following proposition, we study the social evaluation of an optimal monotone contract when the worker is gain seeking.

Proposition 2. *Suppose that $\lambda_g > \lambda_l > 0$. Let $\tilde{r}$ be the value obtained in Theorem 2.*

(i) The total payoff at an optimal monotone contract is continuous and strictly decreasing in r_0, and it is larger than $[y(e^o)-c(e^o)]/(1-\delta)$ at $r_0=0$ and decreases without bound as r_0 goes to infinity.

(ii) The total payoff without gain–loss utility at an optimal monotone contract is constant in r_0 on $[0,\tilde{r})$ and $(\tilde{r},\infty)$, and it takes a smaller value at $r_0<\tilde{r}$ than at $r_0>\tilde{r}$.

(iii) The total output at an optimal monotone contract is constant in r_0 on $[0,\tilde{r})$ and $(\tilde{r},\infty)$, and it takes a larger value at $r_0<\tilde{r}$ than at $r_0>\tilde{r}$.

When the worker is gain seeking, she is mainly motivated by her satisfaction from employment, which decreases with her reference wage. Hence, the firm can take advantage of a low initial reference wage. Specifically, when $r_0<\tilde{r}$, the wage increases with the initial reference wage while the effort level is fixed. As a result, the firm's payoff at an optimal monotone contract is strictly decreasing in r_0 on $[0,\tilde{r})$ and is constant on $(\tilde{r},\infty)$. Meanwhile, the worker's payoff at an optimal monotone contract is determined by $L(r_0)$, which is strictly decreasing in r_0. Hence, the total payoff at an optimal monotone contract is strictly decreasing in r_0. As in the case of loss aversion, the total payoff at $r_0=0$ exceeds $[y(e^o)-c(e^o)]/(1-\delta)$. As the worker with a lower initial reference wage is more motivated, the effort level for $r_0<\tilde{r}$ is higher than that for $r_0>\tilde{r}$, while both exceed e^o. Hence, the total payoff without gain–loss utility is lower and the total output is higher for $r_0<\tilde{r}$ than for $r_0>\tilde{r}$.

4. CONCLUSION

In this paper, we have studied the effects of a worker's reference-dependent preferences on a firm's design of wage contracts. By using a simple model, we have provided explanations for decreasing and increasing wage profiles based on reference dependence. We have made many simplifying assumptions in order to keep our model tractable, and extending our model in different directions will help us better understand the effects of reference dependence on a worker's behavior and a firm's optimal contract. For example, we can allow the worker's moral hazard problem by having stochastic output, introduce the consumption–savings decisions of the worker, have multiple, possibly heterogeneous firms and workers, and consider different forms of the gain–loss function and the reference wage formation process.

While a person's attitude towards gains and losses can be regarded as her innate trait, it can also be considered as a trait that can be shaped by training or education. Our results show that the worker is hurt by her loss aversion as it reduces her reservation payoff. At the same time, the firm benefits from the worker's consideration of gains and losses because it motivates her to work harder. Hence, if the worker can be trained to be less sensitive to gains and losses, it will make the worker better off and the firm worse off. Whether such training is desirable from society's perspective depends on how to measure social welfare. If social welfare is measured by $\sum_{t=0}^{\infty} \delta^t [y(e_t) - c(e_t)]$, such training is socially desirable. In contrast, if social welfare is measured by output, cultivating the worker's sensation of gains and losses benefits society. Given that gains can contribute to social welfare by increasing the worker's job satisfaction, one may search for a socially optimal level of the worker's sensation of gains and losses. In a similar vein, if the worker's formation of reference wages can be affected by training or education, one may search for a reference wage formation process that is desirable for the worker, the firm, or society.

APPENDIX: PROOFS

Proof of Theorem 1: (i) Let us modify the firm's problem (4)–(8) by replacing the function $R(x)$ in the participation constraints in (5) with $\lambda_l x$ for all $x \in \mathbb{R}$. The participation constraint in period 0 must be binding at any optimal contract for this modified problem, because otherwise we can increase e_0 to improve the objective value while satisfying all the constraints in the problem. Using (2), the binding participation constraint in period 0 can be rewritten as

$$\sum_{t=0}^{\infty} \delta^t w_t = \frac{1}{\alpha_l} \sum_{t=0}^{\infty} \delta^t c(e_t). \tag{13}$$

Using (13), the objective function in (4) can be rewritten as

$$\sum_{t=0}^{\infty} \delta^t \left[y(e_t) - \frac{1}{\alpha_l} c(e_t) \right].$$

The objective function is maximized by setting $e_t = e^l$ for all $t = 0, 1, \ldots$. Hence, any wage sequence $\{w_t\}_{t=0}^{\infty}$ that satisfies all the constraints of the modified problem with $e_t = e^l$ for all $t = 0, 1, \ldots$ constitutes an optimal contract for the modified problem. There exists such a wage sequence. For example, setting all the participation constraints with equality yields $w_t = w^l$ for all $t = 0, 1, \ldots$.

Since $\lambda_g < \lambda_l$, the feasible set of the original problem is a subset of that of the modified problem. Hence, any optimal contract for the modified problem that is feasible for the original problem is optimal for the original problem. Let $e_t^* = e^l$ and $w_t^* = w^l$ for all $t = 0, 1, \ldots$. As noted above, the contract $\{e_t^*, w_t^*\}_{t=0}^{\infty}$

is an optimal contract for the modified problem. Suppose that $r_0 > w^l$. Then $w_t^* < r_t$ is satisfied for all $t = 0, 1, \ldots$. The contract $\{e_t^*, w_t^*\}_{t=0}^{\infty}$ is feasible for the original problem and thus optimal for the original problem. It also satisfies all the participation constraints of the original problem with equality. If a contract has $w_t > r_t$ for some t, we have $\sum_{t=0}^{\infty} \delta^t w_t > \sum_{t=0}^{\infty} \delta^t c(e_t)/\alpha_l$, and thus the objective function value of the contract is strictly lower than that of a contract with $e_t = e^l$ for all $t = 0, 1, \ldots$. Hence, it must be that $w_t \leq r_t$ for all t at any optimal contract for the original problem, and $\{e_t^*, w_t^*\}_{t=0}^{\infty}$ is the unique optimal contract with binding participation constraints.

(ii) Now let us modify the firm's problem (4)–(8) by replacing the function $R(x)$ in the participation constraints in (5) with $\lambda_g x$ for all $x \in \mathbb{R}$. As in (i), the participation constraint in period 0 must be binding at any optimal contract for this modified problem, and the binding participation constraint in period 0 can be rewritten as

$$\sum_{t=0}^{\infty} \delta^t w_t = \frac{1}{\alpha_g} \left[\sum_{t=0}^{\infty} \delta^t c(e_t) - (\lambda_l - \lambda_g) \frac{1}{1-\delta\rho} r_0 \right]. \tag{14}$$

Using (14), the objective function in (4) can be rewritten as

$$\sum_{t=0}^{\infty} \delta^t \left[y(e_t) - \frac{1}{\alpha_g} c(e_t) \right] + \frac{\lambda_l - \lambda_g}{\alpha_g} \frac{1}{1-\delta\rho} r_0.$$

The objective function is maximized by setting $e_t = e^g$ for all $t = 0, 1, \ldots$. Setting all the participation constraints with equality yields

$$w_t = \frac{1}{\alpha_{gl}} \left[c(e^g) - (\lambda_l - \lambda_g) r_t \right] \tag{15}$$

and

$$r_{t+1} = \left[\rho - \frac{(1-\rho)(\lambda_l - \lambda_g)}{\alpha_{gl}} \right] r_t + \frac{1-\rho}{\alpha_{gl}} c(e^g)$$

for all $t = 0, 1, \ldots$. Since $\lambda_l < (1-\delta\rho)(\lambda_g + \rho)/(1-\rho)$, we have $\alpha_{gl} > 0$ and $0 < \rho - (1-\rho)(\lambda_l - \lambda_g)/\alpha_{gl} < 1$. Hence, starting from any r_0, we obtain the sequences $\{r_t\}$ and $\{w_t\}$ that are bounded and converge to w^g. Let $e_t^* = e^g$ and $w_t^* = w_t$ as in (15) for all $t = 0, 1, \ldots$. Then the contract $\{e_t^*, w_t^*\}_{t=0}^{\infty}$ is an optimal contract for the modified problem. Suppose that $r_0 < w^g$. Then $w_t^* > r_t$ is satisfied for all $t = 0, 1, \ldots$. Since $\lambda_g < \lambda_l$, the feasible set of the original problem is a subset of that of the modified problem. Following an analogous argument to the one in (i), we can conclude that $\{e_t^*, w_t^*\}_{t=0}^{\infty}$ is the unique optimal contract with binding participation constraints for the original problem.

(iii) Let us modify the firm's problem (4)–(8) by replacing the right-hand side of the participation constraint in period 0 with $s + L(r_0)$, where $s \geq 0$. The participation constraints in all the later periods remain unchanged. We denote the optimal value of the modified problem by $V(r_0, s)$. Using similar arguments to those in (i) and (ii), we can obtain

$$V(r,s) = \begin{cases} \frac{1}{1-\delta}\left[y(e^l) - \frac{1}{\alpha_l}c(e^l)\right] - \frac{1}{\alpha_l}s & \text{if } r > w^l + \frac{1-\delta}{\alpha_l}s, \\ \frac{1}{1-\delta}\left[y(e^g) - \frac{1}{\alpha_g}c(e^g)\right] + \frac{\lambda_l - \lambda_g}{\alpha_g(1-\delta\rho)}r - \frac{1}{\alpha_g}s & \text{if } r < w^g + \frac{1-\delta}{\alpha_l}s. \end{cases} \tag{16}$$

Suppose that $r_0 \in [w^g + (1-\delta)s/\alpha_l, w^l + (1-\delta)s/\alpha_l]$. We claim that we have $w_t = r_0$ for all $t = 0, 1, \ldots$ in any optimal contract for the modified problem. Plugging $w_t = r_0$ for all $t = 0, 1, \ldots$ into the objective function yields $\sum_{t=0}^{\infty} \delta^t y(e_t) - r_0/(1-\delta)$, while the binding participation constraint in period 0 can be rewritten as

$$\sum_{t=0}^{\infty} \delta^t c(e_t) = \frac{\alpha_l}{1-\delta} r_0 - s. \tag{17}$$

Since $y(e)$ is concave and $c(e)$ is convex, it is optimal to set e_t as a constant. Let $e(r,s)$ denote the unique value of e satisfying $c(e) = \alpha_l r - (1-\delta)s$ for any (r,s) such that $w^g + (1-\delta)s/\alpha_l \leq r \leq w^l + (1-\delta)s/\alpha_l$. For any such (r,s), we have $e(r,s) \in [e^g, e^l]$. Setting $e_t = e(r_0, s)$ for all $t = 0, 1, \ldots$ satisfies (17). Moreover, when $w_t = r_0$ and $e_t = e(r_0, s)$ for all $t = 0, 1, \ldots$, all the participation constraints in periods $t = 1, 2, \ldots$ are satisfied as well. Thus, for any $r \in [w^g + (1-\delta)s/\alpha_l, w^l + (1-\delta)s/\alpha_l]$, if setting $w_t = r$ for all $t = 0, 1, \ldots$ is optimal, the optimal value is given by

$$V(r,s) = \frac{1}{1-\delta}\left[y(e(r,s)) - \frac{1}{\alpha_l}c(e(r,s))\right] - \frac{1}{\alpha_l}s. \tag{18}$$

Using a dynamic programming approach, we consider the following problem:

$$\max_{e,w,s} y(e) - w + \delta V(r', s) \tag{19}$$

$$\text{subject to } w + R(w-r) - c(e) = -\frac{\lambda_l}{1-\delta\rho}(r - \delta r') - \delta s, \tag{20}$$

$$r' = \rho r + (1-\rho)w, \tag{21}$$

$$r \in [w^g, w^l] \text{ given}, \tag{22}$$

$$e, w, s \geq 0, \tag{23}$$

where the function $V(r,s)$ in (19) is given by (16) and (18). At any solution to the firm's problem (4)–(8), the participation constraint in period 0 must be binding, but the participation constraints in later periods may be slack. The variable s in the value function $V(r,s)$ represents the slack of the participation constraint in the curent period. In the above dynamic programming problem, the firm chooses the slack s in the next period while having no slack in the current period, and thus the value of the problem can be expressed as $V(r,0)$.

We first show that, for any given $s \geq 0$, $w > r$ cannot be optimal for the dynamic programming problem. Using (20) and (21), we can obtain the relationship

$$w = \begin{cases} \frac{1}{\alpha_l}[c(e) - \delta s] & \text{if } w \leq r, \\ \frac{1}{\alpha_{gl}}[c(e) - (\lambda_l - \lambda_g)r - \delta s] & \text{if } w > r. \end{cases} \tag{24}$$

Let $\tilde{e}(w,r,s)$ be the unique value of e satisfying (24) given (w,r,s). By plugging $\tilde{e}(w,r,s)$ into the objective function in (19), we can regard the objective function as a function of (w,r,s), and its partial derivative with respect to $w > r$ is given by

$$\begin{aligned} & y'(\tilde{e}(w,r,s))\frac{\partial \tilde{e}(w,r,s)}{\partial w} - 1 + \delta(1-\rho)V_1(r',s) \\ =& \alpha_{gl}\frac{y'(\tilde{e}(w,r,s))}{c'(\tilde{e}(w,r,s))} - 1 + \delta(1-\rho)V_1(r',s), \end{aligned}$$

where

$$V_1(r,s) = \frac{\partial V(r,s)}{\partial r} = \begin{cases} 0 & \text{if } r > w^l + \frac{1-\delta}{\alpha_l}s, \\ \frac{1}{1-\delta}\left[\alpha_l \frac{y'(e(r,s))}{c'(e(r,s))} - 1\right] & \text{if } r \in [w^g + \frac{1-\delta}{\alpha_l}s, w^l + \frac{1-\delta}{\alpha_l}s], \\ \frac{\lambda_l - \lambda_g}{\alpha_g(1-\delta\rho)} & \text{if } r < w^g + \frac{1-\delta}{\alpha_l}s. \end{cases}$$

Since $w > r \geq w^g$ and $s \geq 0$, we have $\tilde{e}(w,r,s) > e^g$. Hence, we have

$$\begin{aligned}
&\alpha_{gl}\frac{y'(\tilde{e}(w,r,s))}{c'(\tilde{e}(w,r,s))} - 1 + \delta(1-\rho)V_1(r',s) \\
&< \frac{\alpha_{gl}}{\alpha_g} - 1 + \delta(1-\rho)\frac{\lambda_l - \lambda_g}{\alpha_g(1-\delta\rho)} \\
&= \frac{1}{\alpha_g}\left[\alpha_{gl} - \alpha_g + (\lambda_l - \lambda_g)\frac{\delta(1-\rho)}{(1-\delta\rho)}\right] = 0.
\end{aligned}$$

This shows that $w > r$ cannot be optimal for any $s \geq 0$, and thus we focus on $w \leq r$ in the following. Now fix $w \leq r$. Then the partial derivative of the objective function with respect to s is given by

$$y'(\tilde{e}(w,r,s))\frac{\partial \tilde{e}(w,r,s)}{\partial s} + \delta V_2(r',s) = \delta\frac{y'(\tilde{e}(w,r,s))}{c'(\tilde{e}(w,r,s))} + \delta V_2(r',s), \tag{25}$$

where

$$V_2(r,s) = \frac{\partial V(r,s)}{\partial s} = \begin{cases} -\frac{1}{\alpha_l} & \text{if } s < \frac{\alpha_l}{1-\delta}(r - w^l), \\ -\frac{y'(e(r,s))}{c'(e(r,s))} & \text{if } s \in [\frac{\alpha_l}{1-\delta}(r-w^l), \frac{\alpha_l}{1-\delta}(r-w^g)], \\ -\frac{1}{\alpha_g} & \text{if } s > \frac{\alpha_l}{1-\delta}(r-w^g). \end{cases}$$

Since $w \leq r$, we have $r' \leq r \leq w^l$. Hence, the case of $s < \frac{\alpha_l}{1-\delta}(r' - w^l)$ cannot occur. Let us set $e(r,s) = e^g$ if $r < w^g + (1-\delta)s/\alpha_l$. Since $w \leq r'$, the expression in (25) is nonnegative when $s = 0$, and the optimal value of s given (w,r), which we denote by $s^*(w,r)$, can be obtained by equating $\tilde{e}(w,r,s)$ and $e(r',s)$, which yields

$$s^*(w,r) = \begin{cases} \alpha_l\rho(r-w) & \text{if } w \geq \frac{w^g - \delta\rho r}{1-\delta\rho}, \\ \frac{\alpha_l}{\delta}(w^g - w) & \text{if } w < \frac{w^g - \delta\rho r}{1-\delta\rho}. \end{cases}$$

Let $e^*(w,r) = \tilde{e}(w,r,s^*(w,r))$. We can express the objective function in (19) as

$$y(e^*(w,r)) - w + \delta V(\rho r + (1-\rho)w, s^*(w,r)), \tag{26}$$

which depends on (w,r). Its partial derivative with respect to $w \in ((w^g - \delta\rho r)/(1-$

$\delta\rho), r)$ is given by

$$\begin{aligned}
&y'(e^*(w,r))\frac{\partial e^*(w,r)}{\partial w} - 1 + \delta(1-\rho)V_1(r', s^*(w,r)) + \delta V_2(r', s^*(w,r))\frac{\partial s^*(w,r)}{\partial w} \\
=&(1-\delta\rho)\alpha_l \frac{y'(e^*(w,r))}{c'(e^*(w,r))} - 1 + \frac{\delta(1-\rho)}{1-\delta}\left[\alpha_l \frac{y'(e(r', s^*(w,r)))}{c'(e(r', s^*(w,r)))} - 1\right] \\
&+ \delta\rho\alpha_l \frac{y'(e(r', s^*(w,r)))}{c'(e(r', s^*(w,r)))} \\
=&\frac{1-\delta\rho}{1-\delta}\left[\alpha_l \frac{y'(e^*(w,r))}{c'(e^*(w,r))} - 1\right] > 0,
\end{aligned}$$

because $e^*(w,r) = e(r', s^*(w,r)) < e^l$. Also, its partial derivative with respect to $w < (w^g - \delta\rho r)/(1-\delta\rho)$ is given by

$$\begin{aligned}
&y'(e^*(w,r))\frac{\partial e^*(w,r)}{\partial w} - 1 + \delta(1-\rho)V_1(r', s^*(w,r)) + \delta V_2(r', s^*(w,r))\frac{\partial s^*(w,r)}{\partial w} \\
=&-1 + \frac{\delta(1-\rho)}{1-\delta\rho}\frac{\lambda_l - \lambda_g}{\alpha_g} + \frac{\alpha_l}{\alpha_g} \\
=&\frac{1}{\alpha_g}\left[\alpha_l - \alpha_g + (\lambda_l - \lambda_g)\frac{\delta(1-\rho)}{1-\delta\rho}\right] = \frac{\lambda_l - \lambda_g}{\alpha_g} > 0.
\end{aligned}$$

Hence, the objective function in (26) is strictly increasing for $w \leq r$, and thus setting $w = r$ is uniquely optimal. The optimal value of s at $w = r$ is $s^*(r,r) = 0$. This proves that $w = r$, $s = 0$, and $e = \tilde{e}(r,r,0) = \hat{e}(r)$ is the unique solution to the dynamic programming problem. This implies that in any optimal contract we have $w_t = r_t$, $e_t = \hat{e}(r_t)$, and the binding participation constraint in period $t+1$, for all $t = 0, 1, \ldots$. Hence, the contract with $w_t = r_0$ and $e_t = \hat{e}(r_0)$ for all $t = 0, 1, \ldots$ is the unique optimal contract. □

Proof of Proposition 1: (i) Let $T(r_0) = \Pi(r_0) + L(r_0)$ for all $r_0 \geq 0$, where the functions $\Pi(\cdot)$ and $L(\cdot)$ are defined by (10) and (3), respectively. Then $T(r_0)$ represents the total payoff at an optimal contract. It can be checked that both $\Pi(\cdot)$ and $L(\cdot)$ are continuous, which implies that $T(\cdot)$ is continuous. It can also be checked that the derivative of $\Pi(r_0)$ is weakly decreasing and that of $L(r_0)$ is constant, and thus we have

$$T'(r_0) \leq \frac{\lambda_l - \lambda_g}{\alpha_g(1-\delta\rho)} - \frac{\lambda_l}{1-\delta\rho} = -\frac{\alpha_l}{\alpha_g}\frac{\lambda_g}{1-\delta\rho} < 0$$

for all $r_0 \geq 0$. Hence, the total payoff at an optimal contract is strictly decreasing in r_0, and it decreases without bound as r_0 goes to infinity. Since e^g is the unique maximizer of $y(e) - c(e)/\alpha_g$ and $\alpha_g > 1$, we have

$$\begin{aligned} T(0) &= \frac{1}{1-\delta}\left[y(e^g) - \frac{1}{\alpha_g}c(e^g)\right] > \frac{1}{1-\delta}\left[y(e^o) - \frac{1}{\alpha_g}c(e^o)\right] \\ &> \frac{1}{1-\delta}\left[y(e^o) - c(e^o)\right]. \end{aligned}$$

(ii), (iii) As shown in the proof of Theorem 1, in any optimal contract we have $e_t = e^g$ for all t if $r_0 < w^g$, $e_t = \hat{e}(r_0)$ for all t if $w^g \leq r_0 \leq w^l$, and $e_t = e^l$ for all t if $r_0 > w^l$. Note that $\hat{e}(r_0)$ is continuous and strictly increasing in r_0 with $\hat{e}(w^g) = e^g$ and $\hat{e}(w^l) = e^l$. Let $e_t = e$ for all t. Then the total payoff without gain–loss utility, $[y(e) - c(e)]/(1-\delta)$, is maximized at $e = e^o$, and it decreases as e moves away from e^o. Also, the total output, $y(e)/(1-\delta)$ is strictly increasing in e. Since $e^l > e^g > e^o$, the results follow. □

Proof of Theorem 2: We define an optimal decreasing contract as a solution to the firm's problem (4)–(8) with the additional constraint that the sequence $\{r_t\}$ is decreasing (i.e., $r_{t+1} \leq r_t$ for all $t = 0, 1, \ldots$). Similarly, we define an optimal increasing contract as a solution to the firm's problem with the additional constraint that the sequence $\{r_t\}$ is increasing (i.e., $r_{t+1} \geq r_t$ for all $t = 0, 1, \ldots$). Since a monotone sequence is either decreasing or increasing, we can find an optimal monotone contract by comparing the objective function values of optimal decreasing contracts and optimal increasing contracts. If optimal decreasing contracts yield a higher value than optimal increasing contracts, then optimal decreasing contracts are optimal monotone contracts, and vice versa.

First, let us consider the firm's problem with the constraint of decreasing $\{r_t\}$ added. Note that $r_{t+1} \leq r_t$ if and only if $w_t \leq r_t$, and thus any contract that yields decreasing $\{r_t\}$ induces a loss in every period. When the contract induces a loss in every period, it is optimal for the firm to set $e_t = e^l$ for all $t = 0, 1, \ldots$, as shown in the proof of Theorem 1(i). With $e_t = e^l$ for all $t = 0, 1, \ldots$, it is possible to satisfy decreasing $\{r_t\}$ as well as all the other constraints if and only if $r_0 \geq w^l$. We obtain $w_t = w^l$ for all $t = 0, 1, \ldots$ by setting all the participation constraints with equality. If $r_0 < w^l$, then the constraint of decreasing $\{r_t\}$ is binding, and we can show that it is optimal for the firm to set $w_t = r_0$ and $e_t = \hat{e}(r_0)$ for all $t = 0, 1, \ldots$, as in the proof of Theorem 1(iii). In summary, if $r_0 \geq w^l$, the unique optimal decreasing contract with binding participation constraints is given by $e_t = e^l$ and $w_t = w^l$ for all $t = 0, 1, \ldots$, and if $r_0 < w^l$, the unique optimal decreasing contract is given by $e_t = \hat{e}(r_0)$ and $w_t = r_0$ for all $t = 0, 1, \ldots$. Let $\Pi_l(r)$ be the objective function value of an optimal decreasing contract when the initial reference wage is given by r. Then we have

$$\Pi_l(r) = \begin{cases} \frac{1}{1-\delta}\left[y(e^l) - \frac{1}{\alpha_l}c(e^l)\right] & \text{if } r \geq w^l, \\ \frac{1}{1-\delta}\left[y(\hat{e}(r)) - r\right] & \text{if } r < w^l. \end{cases}$$

Next, let us consider the firm's problem with the constraint of increasing $\{r_t\}$ added. Note that $r_{t+1} \geq r_t$ if and only if $w_t \geq r_t$, and thus any contract that yields increasing $\{r_t\}$ induces a gain in every period. Note also that $\lambda_l < (1-\delta\rho)(\lambda_g+\rho)/(1-\rho)$ holds because $\lambda_l < \lambda_g$ and $\lambda_g < (1-\delta\rho)(\lambda_g+\rho)/(1-\rho)$. When the contract induces a gain in every period, it is optimal for the firm to set $e_t = e^g$ for all $t = 0,1,\ldots$, as shown in the proof of Theorem 1(ii). With $e_t = e^g$ for all $t = 0,1,\ldots$, it is possible to satisfy increasing $\{r_t\}$ as well as all the other constraints if and only if $r_0 \leq w^g$. We obtain $w_t = [c(e^g) + (\lambda_g - \lambda_l)r_t]/\alpha_{gl}$ for all $t = 0,1,\ldots$ by setting all the participation constraints with equality. If $r_0 > w^g$, then the constraint of increasing $\{r_t\}$ is binding, and we can show that it is optimal for the firm to set $w_t = r_0$ and $e_t = \hat{e}(r_0)$ for all $t = 0,1,\ldots$, as in the proof of Theorem 1(iii). In summary, if $r_0 \leq w^g$, the unique optimal increasing contract with binding participation constraints is given by $e_t = e^g$ and $w_t = [c(e^g) + (\lambda_g - \lambda_l)r_t]/\alpha_{gl}$ for all $t = 0,1,\ldots$, and if $r_0 > w^g$, the unique optimal increasing contract is given by $e_t = \hat{e}(r_0)$ and $w_t = r_0$ for all $t = 0,1,\ldots$. Let $\Pi_g(r)$ be the objective function value of an optimal increasing contract when the initial reference wage is given by r. Then we have

$$\Pi_g(r) = \begin{cases} \frac{1}{1-\delta}\left[y(e^g) - \frac{1}{\alpha_g}c(e^g)\right] - \frac{\lambda_g-\lambda_l}{\alpha_g(1-\delta\rho)}r & \text{if } r \leq w^g, \\ \frac{1}{1-\delta}\left[y(\hat{e}(r)) - r\right] & \text{if } r > w^g. \end{cases}$$

Since $\lambda_g > \lambda_l$, we have $w^g > w^l$. It can be checked that $\Pi_g(r) - \Pi_l(r)$ is continuous and strictly decreasing and that it is positive at $r = w^l$ and negative at $r = w^g$. Hence, there exists a unique value of r satisfying $\Pi_l(r) = \Pi_g(r)$, which we denote by $\tilde{r}$, and it belongs to the interval (w^l, w^g). If $r_0 < \tilde{r}$, then the optimal increasing contract with binding participation constraints is the unique optimal monotone contract with binding participation constraints. If $r_0 > \tilde{r}$, then the optimal decreasing contract with binding participation constraints is the unique optimal monotone contract with binding participation constraints. If $r_0 = \tilde{r}$, then both the optimal increasing contract with binding participation constraints and the optimal decreasing contract with binding participation constraints are optimal monotone contracts with binding participation constraints, and there is no other optimal monotone contract with binding participation constraints. $\square$

Proof of Proposition 2: (i) Let $T(r_0) = \Pi(r_0) + L(r_0)$ for all $r_0 \geq 0$, where the functions $\Pi(\cdot)$ and $L(\cdot)$ are defined by (12) and (3), respectively. Then $T(r_0)$ represents the total payoff at an optimal monotone contract. It can be checked that both $\Pi(\cdot)$ and $L(\cdot)$ are continuous, which implies that $T(\cdot)$ is continuous. It can also be checked that $\Pi(r_0)$ is strictly decreasing on $[0, \tilde{r}]$ and constant on $[\tilde{r}, \infty)$ and that $L(r_0)$ is strictly decreasing with $L'(r_0) = -\lambda_l/(1-\delta\rho) < 0$ for all $r_0 \geq 0$. Hence, the total payoff at an optimal monotone contract is strictly decreasing in r_0, and it decreases without bound as r_0 goes to infinity. We can show that $T(0) > [y(e^o) - c(e^o)]/(1-\delta)$ by the same argument as in the proof of Proposition 1.

(ii), (iii) As shown in the proof of Theorem 2, in any optimal monotone contract we have $e_t = e^g$ for all t if $r_0 < \tilde{r}$ and $e_t = e^l$ for all t if $r_0 > \tilde{r}$. Since $e^g > e^l > e^o$, the results follow. □

REFERENCES

Abdellaoui, M., Bleichrodt, H., and C. Paraschiv (2007). “Loss Aversion under Prospect Theory: A Parameter-Free Measurement,” *Management Science* 53, 1659-1674.

Brooks, P. and B. Zank (2005). “Loss Averse Behavior,” *Journal of Risk and Uncertainty* 31, 301–325.

Camerer, C., Babcock, L., Loewenstein, G., and R. Thaler (1997). “Labor Supply of New York City Cabdrivers: One Day at a Time,” *Quarterly Journal of Economics* 112, 407-441.

Clark, A.E. (1999). “Are Wages Habit-Forming? Evidence from Micro Data,” *Journal of Economic Behavior & Organization* 39, 179-200.

Crawford, V. and J. Meng (2011). “New York City Cab Drivers’ Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income,” *American Economic Review* 101, 1912-1932.

Farber, H.S. (2005). “Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers,” *Journal of Political Economy* 113, 46–82.

Farber, H.S. (2008). “Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers,” *American Economic Review* 98, 1069-1082.

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* 163, 313-335.

Kahneman, D. and A. Tversky (1979). “Prospect Theory: An Analysis of Decision under Risk,” *Econometrica* 47, 263-292.

Kőszegi, B. and M. Rabin (2006). “A Model of Reference-Dependent Preferences,” *Quarterly Journal of Economics* 121, 1133-1165.

Schmidt, U. and S. Traub (2002). “An Experimental Test of Loss Aversion,” *Journal of Risk and Uncertainty* 25, 233-249.

Won, D.C. (2019). “On the Micro-Level Irrelevance of Habit Formation in Complete Markets,” *Journal of Economic Theory and Econometrics* 30, 1-19.