On the Power of Gifts[*]

Rosario Macera[†] Vera L. te Velde [‡]

June 2016

(Preliminary Version)

Abstract

We study the efficacy and efficiency of gift-exchange as an incentive mechanism. We formalize two well-established psychological insights—the importance of expectations about upcoming outcomes to shape behavior, and the adaptation over time to constant stimuli—to analyze an important but up to now neglected aspect of monetary gifts: their surprising or anticipated nature. Our theory delivers four main messages, (1) the power of gifts is greater whenever they surprise workers, (2) the power of surprising gifts, however, wanes over time, (3) gifts are cursed as, once a firm grants one, it should grant it forever, and therefore (4) gifts are only profitable in short-term interactions. We conclude that to make the most of the power of monetary gifts to motivate workers, it is crucial to manage the expectations that gifts create.

JEL Classification: D030, J31, J33, M52

Keywords: Reference-dependent preferences, reciprocity, compensation, gift exchange.

[*]We thank Constanca Esteves-Sorenson for initial discussions and some of the quotes used in this paper.

[†]rosario.macera@uc.cl

[‡]v.tevelde@uq.edu.au.

Introduction

Whether paying above-market wages is an efficient mechanism to elicit employees' effort is an important economic question. If workers engage in gift exchange in response to above-market wages, i.e., if workers interpret higher wages as a gift and as a consequence they reciprocate above-market wages with above-minimum effort, as proposed by Akerlof's (1982), then firms can rely on gift exchange to motivate workers. Gift exchange, therefore, could save firms from the onerous costs of successfully designing and implementing pay-for-performance incentive schemes, especially in the common case where effort is difficult to measure or contract on.

Despite its potential importance as an incentive mechanism, several aspects of gift exchange remain unstudied. In particular, there are no significant attempts to study which aspects of the monetary gift are important to trigger reciprocal responses from workers. Rather, the literature has focused on studying the importance of market conditions such as the market size (Brandts et al. 2010, e.g.) and its structure (Schram, Brandts & Gërxhani 2010); the characteristics of who grants the gift (Charness 2004, e.g.); the agent's information about the firm's surplus (Hennig-Schmidt, Rockenbach & Sadrieh 2010, DellaVigna et al. 2016) or his ability to repay the gift (Englmaier & Leider 2012*b*); and even the demographics of the recipient agent such as gender (Chaudhuri & Sbai 2011, Hannan, Kagel & Moser 2002).[1] To the best of our knowledge, however, no study has focused on the properties of the monetary gift itself.

This paper advances the incentives literature by theoretically studying an until now largely neglected aspect of monetary gifts: their surprising or anticipated nature. To this end, we develop a model of reference-dependent reciprocity where the reference point corresponds to the worker's recent rational expectations. We use the model to study the implications of this aspect of monetary gifts to profitably elicit effort in one-shot and repeated principal-agent interactions. We discuss how this model organizes findings from the experimental literature, and the novel predictions it makes in other contexts.

Our model of reference-dependent reciprocity builds on two important psychological insights: expected outcomes as a key determinant of behavior and hedonic adaptation. First, the idea that expectations about upcoming outcomes determine the response to actual outcomes dates back to Vroom (1964), while in economics it was formalized by Kőszegi & Rabin (2006, 2007, 2009) who embedded the idea in a model of reference-dependent preferences with loss aversion as in Kahneman & Tversky (1979).[2] Building on expectations as reference points allow us to formalize the idea of a payment being surprising or anticipated, depending on

[1]These studies build on earlier demonstrations of gift exchange in the laboratory (Fehr, Kirchsteiger & Riedl 1993, Fehr, Kirchsteiger & Riedl 1998, Fehr et al. 1998).

[2]Vroom's Expectancy theory has been extensively studied and applied in psychology. For a description and a metanalysis of the evidence see Van Eerde & Thierry (1996) and Ambrose & Kulik (1999).

whether the agent foresaw the gift or not and thus whether he incorporated it into his payment expectations. Second, adaptation refers to the idea that the hedonic response to a constant stimuli is decreasing in time. First proposed in psychology by Helson (1964) and applied to economics by Frederick & Loewenstein (1999) and Frey & Stutzer (2002), among others, incorporating hedonic adaptation into our model allows us to predict the dynamic consequences of fixed wage increases.[3]

Our model delivers four novel messages on the efficacy and efficiency of monetary gifts,

(1) The power of gifts is greater whenever they are surprising. Because surprising gifts trigger a pleasing departure from payment expectations while anticipated payments do not, they elicit a larger effort responses. Similarly, surprising pay cuts lead to more costly retaliation from workers than anticipated cuts.

(2) The power of a surprising gift, however, wanes over time. Because a permanent fixed wage increase is a constant stimuli, we argue that a high fixed wage looses power as workers update their payment expectations: as expectation update, hedonic adaptation occurs and effort returns to its baseline level.

(3) Moreover, gifts are cursed: Firms should grant them forever or never grant them. The intuitive implication of the waning power of gift exchange is that firms should grant one-time monetary gifts randomly to repeatedly trigger reciprocal effort. We show, however, that this intuition is incorrect: if the transitory monetary gift creates the expectation of potential further gifts, agents will retaliate when these expectations are unfulfilled by exerting lower effort than if they had not received any gift. We show that this retaliation of expected but unfulfilled gifts always outweighs the benefits from the sporadic positive surprises. In fact, in shorter-term interactions we show that, if firms are unable to manage expectations, they always prefer to turn surprising bonuses into permanent wage raises, even though the raise is only able to raise effort immediately after it is granted.

(4) And even if a gift is surprising, it is only profitable in short interactions. Because monetary gifts lead to expectations of further gifts, and because both sporadic gifts and permanent raises are costly for the firm as soon as expectations adapt, then it must be the case that gifts are more likely to be profitable in shorter interactions.

Even though to our knowledge we are first to explore theoretically the implications of the surprising versus anticipated nature of monetary gifts on effort provision and gift profitability, our model is related to others studying the efficiency of gift exchange. Kranton (1996) showed that gift exchange can persist even if it is inefficient relative to a market interaction. Dur (2009) explores a different aspect of gift giving by focusing on how low wages can signal that an altruistic firm can pay attention to their workers, a resource which workers

[3]After Helson (1964), Brickman & Campbell (1971) coined the term *hedonic treadmill* to apply the concept of adaptation to happiness. See Diener, Lucas & Scollon (2006) for a good review and pending issues in the literature on adaptation.

value and thus reciprocate with higher effort. Netzer & Schmutzler (2014) showed that gift exchange does not arise in equilibrium if the agent has intentions-based reciprocal preferences as he cannot interpret any profitable effort increase as kind. Finally, Benjamin (2015) develops a model where workers have fairness concerns that depend on a reference transaction. He shows that it rationalizes several types of wage rigidities, such as wage insensitivity to market conditions.

This paper proceeds as follows. Section 2 presents the model. Sections 3 to 6 formalize the four messages and provide some interesting insights about the optimal size of gift and the fair-wage hypothesis (Akerlof & Yellen 1990). Section 7 relates our messages to the empirical evidence on gift exchange. Before concluding, Section 8 discusses some lessons extracted from our model predictions as well as avenues for future research.

The Model Set Up

A principal hires an agent to exert effort e for a fixed wage w. We specify the number of periods they interact and the level of the wage relative to the market wage $\underline{w}$ in each section. The firm is assumed to be a risk neutral profit maximizer with no "behavioral" components, with preferences defined by the profit function $\pi(e, w) = be - w$, $b > 0$.

(1) The employee's preferences. Inspired by Kőszegi & Rabin (2006), we assume the employee experiences utility from two sources: standard consumption utility from material outcomes (wages and effort costs) and reference-dependent utility or gain-loss utility from comparing actual outcomes to the expected ones. Whenever expected outcomes are stochastic, we also follow Kőszegi & Rabin (2006) by assuming the agent compares each point of the reference distribution to each possible outcome. In principal-agent models assuming reference-dependent preferences, it is standard to assume the agent compares wages and effort against a reference point (see Kőszegi (2014) for a review). In our model, however, since wages are fixed, we simplify by ignoring reference dependent utility in this dimension. We therefore assume the agent compares three outcomes against his recently held expectations: effort costs, the firm's kindness, and the firm's profits. Thus, the novel component in our model is that the employee's concern for the firm's kindness and profits will also be reference-dependent.[4] Assumption 1 describes the shape of consumption utility, while Assumption 2 describes gain-loss utility.

Assumption 1 *(Consumption Utility $m(\cdot)$)*

(i) *Consumption utility is linear in the wage w.*

(ii) *$\underline{w}$, the market wage, is the workers' reservation wage.*

[4] In a related model Khalmetski, Ockenfels & Werner (2015) model surprising gifts in the context of guilt aversion in the spirit of Battigalli & Dufwenberg (2007). They apply the model to a dictator game to show that dictator transfers can both, decrease and increase with the recipient's expectation, depending on the weight put on positive and negative surprises, respectively.

(iii) *Effort costs, given by $c(e)$, are minimized at a baseline level $\underline{e} > 0$, and cost is increasing and convex in either direction from there. For simplicity, specify $c(e) = \frac{\gamma}{2}(e - \underline{e})^2$, $\gamma > 0$.*

(iv) *Altogether, consumption utility is given by*

$$m(e, w) = w - \frac{\gamma}{2}(e - \underline{e})^2 \tag{1}$$

In a model with fixed wages, the linearity of consumption utility in wages is an innocuous simplification. The market wage $\underline{w}$ defines the outside option that is needed to determine incentive compatibility of the principal-agent interaction in Section 5. The assumption that the effort cost is quadratic is made for simplicity as it allow us to derive closed form solutions, which will help understanding the mechanisms. The assumption that the convex effort-cost function has a positive interior minimum is also mathematically immaterial but made to aid with intuition. Finally, the assumption that consumption utility is additive is standard in principal-agent models.

Notice that $m(\cdot)$ in Assumption 1 does not include a pure reciprocity component. Including non-reference-dependent social preferences in the utility function would introduce a baseline positive correlation between wages and effort even in the absence of surprising gifts, but it would not modify the dynamics of reciprocal behavior over time. To focus on the effect of surprises specifically, we develop the model without such term, though we do comment after the main propositions how the inclusion of baseline reciprocity would affect them.[5]

In addition to consumption utility, we assume the worker experiences gain-loss utility in effort and in reciprocity. Similarly to non-reference-dependent models of intentions-based social preferences such as Rabin (1993), Levine (1998), Dufwenberg & Kirchsteiger (2004) or Falk & Fischbacher (2006), we assume the worker puts positive weight on the firm's profits if the firm has been unexpectedly kind, and negative weight if the firm has been unexpectedly unkind.

Assumption 2 *(Gain-Loss Utility $n(\cdot|\cdot)$)*

(i) *The firm's kindness $K(w)$ is strictly increasing, strictly concave, and normalized to 0 at the market wage*

[5]Reciprocity has been shown to impact many different market interactions (see Malmendier, te Velde & Weber (2014) for a review), and substantial evidence exists to support social preferences in gift exchange as well. For instance, Charness (2000), Charness (2004), Charness & Haruvy (2002), Charness et al. (2012), and Huck, Seltzer & Wallace (2011) all run variants of the gift exchange game in the lab that indicate that social preferences, and intentions-based social preferences, are a key driver of gift exchange. Laboratory results are in fact so strong that the gift exchange game has become a workhorse for investigating related phenomena, such as endogenous formation of long-term trading partnerships (Brown, Falk & Fehr 2004, Brown, Falk & Fehr 2012), wage compression (Charness, Gross & Guo 2014, Charness & Kuhn 2007, Güth et al. 2001, Kocher, Luhan & Sutter 2012), wage rigidity (Fehr & Falk 1999), charitable giving (Koppel & Regner 2011), deferred compensation (Huck, Seltzer & Wallace 2011), group decision making (Kocher & Sutter 2003), and peer effects/social comparison (Abeler et al. 2010, Hennig-Schmidt, Rockenbach & Sadrieh 2010, Clark, Masclet & Villeval 2010, Cohn et al. 2014, Gächter & Thöni 2010, Siang, Requate & Waichman 2011).

$\underline{w}$.[6]

(ii) *The gain-loss utility function is $\eta\mu(x)$, where $\eta > 0$ is the relative weight put on gains and losses. $\mu(x)$ is linear with a slope of 1 for $x \geq 0$ and $\lambda > 1$ for $x < 0$, where λ is the loss aversion parameter.*

(iii) *Given a reference effort level $\tilde{e}$, gain-loss utility from effort is,*

$$n_e(e|\tilde{e}) = \eta\mu(-c(e) + c(\tilde{e})) \tag{2}$$

(iv) *Given a reference wage $\tilde{w}$ and a reciprocity parameter $\alpha > 0$, gain-loss utility from reciprocity is,*

$$n_k(w, e|\tilde{w}, \tilde{e}) = \alpha\eta\mu(K(w) - K(\tilde{w}))\mu(\pi(w, e) - \pi(\tilde{w}, \tilde{e})) \tag{3}$$

(v) *Overall gain-loss utility is given by $n(e, w|\tilde{e}, \tilde{w}) = n_e(e|\tilde{e}) + n_k(w, e|\tilde{w}, \tilde{e})$.*

A few points deserve note. First, the assumption that the gain-loss utility function is linear is standard in applications of the Kőszegi & Rabin preferences.[7] Second, recall that because in gift-exchange wages are fixed, we do not add gain-loss utility in payments as it does not impact the employee's effort decision. Lastly, the gain-loss utility from reciprocity in *(iv)* departs from Kőszegi & Rabin (2006), as their model assumes additive separability in all utility components.[8]

Finally, Assumption 3 puts together Assumption 1 and Assumption 2.

Assumption 3 *(Total Utility)*

Total utility is the sum of material and gain-loss utility, $U(e, w|\tilde{e}, \tilde{w}) = m(e, w) + n(e, w|\tilde{e}, \tilde{w})$.

(2) Timing. The exact timing of the principal-agent interaction will change in each section. The general timing, however, is as follows. At the start of any given period the employee forms rational expectations about the

[6] The assumption $K(\underline{w}) = 0$ states that the market wage is the fair wage. This assumption is supported by, for example, Kahneman, Knetsch & Thaler (1986). They asked phone survey respondents the following question: "A small photocopying shop has one employee who has worked in the shop for six months and earns \$9 per hour. Business continues to be satisfactory, but a factory in the area has closed and unemployment has increased. Other small shops have now hired reliable workers at \$7 an hour to perform jobs similar to those done by the photocopy shop employee. The current employee leaves, and the owner decides to pay a replacement \$7 an hour." Of the 125 respondents 73% found paying the going market wage either "Completely Fair" or "Acceptable" whereas only 27% found it "Unfair".

[7] See for instance applications to pricing (Heidhues & Kőszegi 2008), labor supply (Crawford & Meng 2011), effort provision (Abeler et al. 2011), sales (Heidhues & Kőszegi 2014), among others.

[8] Perhaps the most direct application of Kőszegi & Rabin's (2006) model would be to assume a reference-dependent reciprocity component of the form $\alpha\eta\mu(K(w)(be - w) - K(\tilde{w})(b\tilde{e} - \tilde{w}))$. This, however, makes the counterintuitive prediction that a firm who is kind, but not as kind as expected, will *increase* profits as workers try to make up for loss in overall reciprocity. We also could have used the specification $\alpha\eta\mu((K(w) - K(\tilde{w}))(b(e - \tilde{e}) - (w - \tilde{w})))$, which would generate very similar results. However, this specification would predict that workers reward and punish wage deviations to an equal degree, whereas the psychology of loss aversion, along with the empirical evidence (Engelmann & Ortmann 2009, Kube, Maréchal & Puppe 2013, e.g.), indicate that workers are more sensitive to wage cuts than wage gains.

wage that he is going to receive and the effort he is going to exert. These expectations constitute his reference point. Let $\tilde{w}$ denote the wage expectations, which are possibly stochastic, and let $\tilde{e}$ denote his contingent effort plan. Then the principal announces the wage, the worker exerts effort and the period ends.

(3) Equilibrium. Following Kőszegi & Rabin (2006) we assume the worker must form and execute *credible* wage-contingent effort plans, i.e. effort plans he knows he will not want to deviate from once he has planned them. These credible plans, thus, will constitute the employee's reference point. If there are several credible effort plans, then the agent chooses the plan that provides him with the highest overall utility. This is the worker's preferred personal equilibrium (PPE).

Definition 1 *(The Agent's Preferred Personal Equilibrium (PPE))*
In any given period t actual effort e_t and a plan for future effort $\boldsymbol{\tilde{e}_{t,\tau}}$, $t < \tau \leqslant T$ correspond to a PPE if,

(i) $e_t \in \underset{e'_t \in E}{\operatorname{argmax}}\ EU_t(e'_t, w_t | \tilde{e}_{t-1,t}, \tilde{w}_{t-1,t})$ *and*

(ii) $\boldsymbol{\tilde{e}_{t,\tau}} \in \underset{\boldsymbol{e'_t} \in E^{PE}}{\operatorname{argmax}}\ EU_\tau(\boldsymbol{e'_t}, \boldsymbol{\tilde{w}_{t,\tau}} | \boldsymbol{e'_t}, \boldsymbol{\tilde{w}_{t,\tau}})\ \tau > t$

where (1) *E corresponds to the set of all possible efforts and E^{PE} corresponds to the set of all efforts that constitute an equilibrium for periods $\tau > t$ given $\boldsymbol{\tilde{w}_{t,\tau}}$,* (2) *$\tilde{e}_{t-1,t}$ constituted an equilibrium in period t-1 and* (3) *$\boldsymbol{\tilde{w}_{t,\tau}}$ and $\tilde{w}_{t-1,t}$ are rationally formed given the economic environment.*

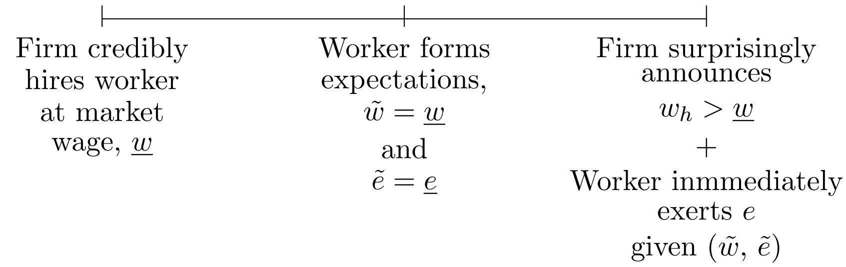
Part (i) in Definition 1 says that in each period the agent exerts effort so to maximize his utility given the period-t expectations about effort and wages he formed in period $t-1$. Part (ii) ensures that when forming his effort plans for the future periods—effort plans that will determine his expectations about the principal's payoff and kindness—he will only choose a plan that he knows he will be willing to follow given the outcome expectations these plans generate.

Before presenting our model's predictions, Lemma 1 presents one useful implication of this solution concept.

Lemma 1 *Suppose the agent expects his future wage to be a fixed wage $\tilde{w}$ with certainty. Then, he forms his effort plans as a consumption utility maximizer, yielding $\tilde{e} = e$.*

The intuition behind Lemma 1 is straightforward: since agents form plans rationally, absent uncertainty or information arrival it must be the case that their expectations will be realized. Because expectations are met and plans must be credible, they experience zero gain-loss utility and thus total utility reduces to consumption utility. Finally, notice that Lemma 1 holds in one-shot and in repeated interactions.

Figure 1: Timing of the one-shot interaction with a fully-surprising gift



The Power of Gifts is Greater Whenever They Are Surprising

"If your boss walked over your desk and handed you $10,000, would it make you work harder for the rest of the day, or the rest of the year? I think it would!"

Member in Style.Gather.com commenting on

Oprah's surprising bonus to her magazine employees in 2010.

In this section we explore the extent to which surprising versus anticipated gifts can elicit a reciprocal response and the firm's profit from granting such gifts. Aside from our main message of gifts being more effective whenever they are surprising, we explore other conditions such as the size of the gift and the impact of the kindness perception of the market wage in triggering reciprocal effort.

For this section, assume the principal and the agent interact during one period, and that at the beginning of the only period the agent is recruited by the principal at the market wage $\underline{w}$. Because the agent was recruited at the market wage, he forms his wage expectation as $\tilde{w} = \underline{w}$. Moreover, because the agent does not expect any deviation from the market wage, by Lemma 1 it must be the case that $\tilde{e} = \underline{e}$. Assume now that right before exerting work, the firm surprisingly raises the wage to $w_h > \underline{w}$. Figure 1 displays the resulting timeline.

Immediately after being surprised with a high wage w_h, the worker's problem is to find e^* to solve,

$$e^* \in \operatorname*{argmax}_e w - \frac{\gamma}{2}(e-\underline{e})^2 + \eta\mu\big(-\frac{\gamma}{2}(e-\underline{e})^2\big) + \alpha\eta\mu(K(w_h) - K(\tilde{w}))\mu(\pi(w_h, e) - \pi(\underline{w}, \underline{e})) \tag{4}$$

To understand how this re-optimization necessarily leads to an increase in effort relative to an anticipated gift, consider the possibilities that the worker faces in this optimization problem. First notice that he can't simply solve the first order condition, because his utility function has a kink in it at the effort level satisfying $\pi(e, w) = \pi(\underline{e}, \underline{w})$, where the firm's profits shifts from the loss domain to the gain domain.[9] Suppose his optimum

[9] Notice that no such kink exists where gain-loss effort costs shift from the loss to gain domain, however, since *any* deviation from the plan of minimal effort implies a loss due to the quadratic nature of the cost function.

does occur at a differentiable point in his utility function. Then the first order condition applies, and we can rearrange it to

$$e^* = \underline{e} + \frac{\alpha \eta K(w_h) \mu'_\pi b}{\gamma(1+\eta\lambda)} > \underline{e} \tag{5}$$

where the inequality holds because $\mu'_\pi = \mu'(\pi(w_h, e^*) - \pi(\underline{w}, \underline{e})) > 0$. The worker therefore exerts more effort than he would have absent the surprising gift in order to reciprocate the principal's unexpected kindness.

In case this optimization process runs into the kink in the utility function, however, the worker must compare any interior solutions to the utility he would get from exactly repaying the firm the costs of the gift ($be - w_h = b\underline{e} - \underline{w} \Rightarrow e = \underline{e} + (w_h - \underline{w})/b$). Details are relegated to the appendix, but the worker chooses this increased effort level for a nonzero range of gifts, but he *never* prefers to stick with his plan of choosing $\underline{e}$, and thus always exerts more effort than he planned to.

Recalling that Lemma 1 says that an unsurprising gift will lead to effort $\underline{e}$, we have thus proved the following proposition.

Proposition 1 *A surprising gift $w_h > \underline{w}$ leads to higher effort than an anticipated gift, that is $e^* > \underline{e}$.*

Before examining the profitability of this reciprocal response, we notice that this effort response not only conforms to Akerlof (1982) gift-exchange hypothesis, but also dovetails with the fair-wage hypothesis in Akerlof & Yellen (1990). According to this theory, workers only engage in reciprocal behavior in response to a wage increase if they perceive the market wage as unfair.[10] To see how is this related to our model, we relax the assumption that $K(\underline{w}) = 0$ and consider some further structure on the kindness function. In particular, take two strictly increasing, strictly concave functions K_1 and K_2, where $K_1(\underline{w}) > 0$, $K_2(\underline{w}) < 0$ and $K_1(w_h) = K_2(w_h)$. Then it must be the case that $K_1(w_h) - K_1(\underline{w}) < K_2(w_h) - K_2(\underline{w})$ and thus the excess effort in response to a gift, $\frac{\alpha\eta(K(w_h) - K(\underline{w}))\mu'_\pi b}{\gamma(1+\eta\lambda)}$, must be stronger under K_2 than under K_1. Intuitively, in our model the effort response is proportional to the size of the surprise in the kindness domain. For a given perceived kindness of the gift then, an unkind market wage implies a bigger surprise relative to a kind one, thus triggering also greater reciprocal effort.

Finally, notice that Proposition 1 is robust to incorporating baseline reciprocity in an additive form. Whenever additive, reciprocity in levels—triggered by the gift irrespective of its surprising or anticipated nature—occurs independently of the reference-dependent reciprocity, only the magnitude of the reciprocal response would be affected. In particular the effort response would be larger.

[10] Cohn, Fehr & Goette (n.d.) find supporting evidence for this prediction in a field experiment in which newspaper distributors received surprising wage increases, but the reciprocal response was attributable to those who perceived the baseline wage to be unfair.

Gifts must be small to be profitable

Delving more deeply into the worker's optimization problem will allow us to determine when his response will be in the loss domain of profits, when it will be in the gain domain, and when it will be at the kink at $be - w = b\underline{e} - \underline{w}$. This is of course of primary interest to the firm, which would like to choose a gift that elicits a profitable response, not just an increase in effort levels.

Consider first the case where the gift and the worker's optimal response leaves the firm's profits in the gain domain, i.e., $w_h < b(e - \underline{e}) + \underline{w}$. Denote e_g (for **g**ain) as the optimal effort *given* this assumption. Using equation (5) and the assumption about the firm's profits, we have that e_g and the range of profitable gifts are,

$$e_g = \underline{e} + \frac{\alpha\eta K(w_h)b}{\gamma(1+\eta\lambda)} \qquad \Rightarrow \qquad w_h < \frac{\alpha\eta K(w_h)b^2}{\gamma(1+\eta\lambda)} + \underline{w} \tag{6}$$

On the other hand, if profits are in the loss domain, $w_h > b(e - \underline{e}) + \underline{w}$, an analogous calculation shows that the optimal response e_l (for **l**oss) and the range of unprofitable gifts are,

$$e_l = \underline{e} + \frac{\alpha\eta\lambda K(w_h)b}{\gamma(1+\eta\lambda)} \qquad \Rightarrow \qquad w_h > \frac{\alpha\eta\lambda K(w_h)b^2}{\gamma(1+\eta\lambda)} + \underline{w} \tag{7}$$

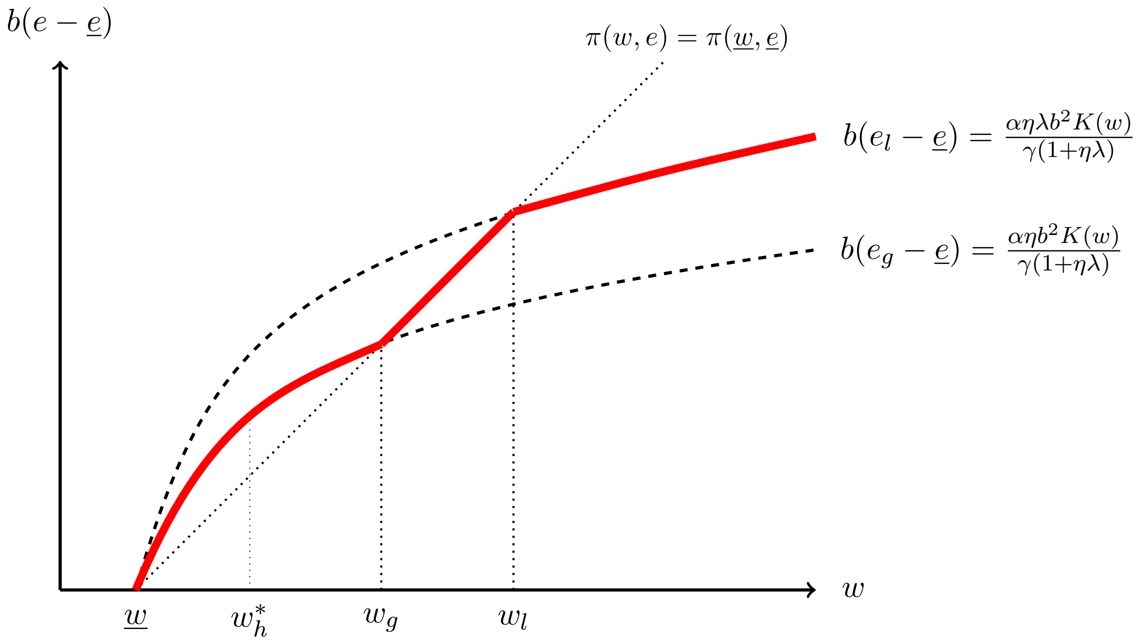
We relegate the details to the appendix, but it turns out that these responses are in fact optimal so long as the profit conditions they imply hold, and when neither hold, the worker will optimally respond with $be - w = b\underline{e} - \underline{w} \Leftrightarrow e = \underline{e} + (w_h - \underline{w})/b$.

Figure 2 shows with a solid thick red line the worker's response, in terms of increased revenue $b(e - \underline{e})$, as a function of the size of the gift. To understand the shape of this revenue-response function, start by noticing that the diagonal line is where increased revenue exactly compensates the firm for the cost of the wage increase. The lower dotted concave curve shows the revenue increase that occurs when the worker responds with e_g, and since he only does this when profits are in the gain domain, he follows this curve as long as it is above the diagonal, or equivalently, if gifts are smaller than a threshold $w_g = \underline{w} + \frac{\alpha\eta K(w_g)b}{\gamma(1+\eta\lambda)}$. Similarly, the upper dotted concave curve shows the revenue increase that occurs when the worker chooses e_l, which he does when this revenue does not compensate for the cost of the gift, placing the firm in the loss domain. Similarly to the gain-domain case, he follows this curve as long as long as he is in the loss domain, i.e., below the diagonal line or equivalently, if gifts are greater than a threshold $w_l = \underline{w} + \frac{\alpha\eta\lambda K(w_l)b}{\gamma(1+\eta\lambda)}$. In between these regions, which is a nonzero region due to the assumption that $\lambda > 1$ and the concavity of K, the worker chooses e at the kink in the profit domain where profits are unchanged relative to baseline.

The revenue-response function in Figure 2 thus shows that gifts are strictly profitable only if they are sufficiently small. Intuitively, if a gift is worth less than a cutoff value w_g, then it is small enough that the worker can easily reciprocate fully to keep the firm's profits in the gain domain. Gifts above the cutoff value w_l, however, are too big and thus too expensive in terms of effort cost, to reciprocate fully. Proposition 2 summarizes.

Proposition 2 *A surprising gift* $w_h > \underline{w}$ *is profitable only if it is small enough, that is, if* $w_h < w_g = \underline{w} + \frac{\alpha\eta K(w_g)b}{\gamma(1+\eta\lambda)}$.

Figure 2: Revenue-response function to a fully surprising gift



Wage cuts are more harmful when they are surprising

We use the same setup to analyze the extent of negative reciprocity. In particular, we show that the timing of information about a wage cut is important: managing expectations prior to a wage cut can prevent negative gain-loss reciprocity. This is important as, unsurprisingly, when firms cut wages, workers frequently respond by reducing effort (Lee & Rupp 2007, Krueger & Mas 2004). Moreover, some evidence exists that this negative effect can be mitigated by providing additional information that preserves the firm's image or controls expectations (Greenberg 1990, Kahneman, Knetsch & Thaler 1986).

To formalize how expectations affect negative reciprocity, we need to extend K to negative values. In accordance with prospect theory, which holds that people have diminishing sensitivity to both gains and losses, we assume that K is convex over $\mathbb{R}^-$. In particular, for simplicity, assume that K is rotationally symmetric around $\underline{w}$: $K(\underline{w}+w) = -K(\underline{w}-w)$.

The analysis proceeds analogously to Section 3.1. The worker anticipates a wage of $\underline{w}$ and plans to exert effort $\underline{e}$, but is fully surprised by a wage cut $w_c < \underline{w}$. He thus reformulates his effort plan to maximize

$$w_c - \frac{\gamma}{2}(1+\eta\lambda)(e-\underline{e})^2 + \alpha\eta\lambda K(w_l)(b(e-\underline{e}) - (w_c - \underline{w})\mu'_\pi$$

where $\mu'_\pi = \mu'(b(e-\underline{e}) - (w_c - \underline{w}))$. Depending on whether the worker reduces effort enough to hurt the firm's profits on net (e_l) or not (e_g), the two potential interior solutions to this optimization problem (not at either kink) are derived from the FOC of this utility function. We have:

$$e_l - \underline{e} = \frac{\alpha\eta\lambda^2 bK(w_l)}{\gamma(1+\eta\lambda)} \qquad \text{requiring} \qquad \frac{\alpha\eta\lambda^2 K(w_l)b^2}{\gamma(1+\eta\lambda)} < w_l - \underline{w} \tag{8}$$

and

$$e_g - \underline{e} = \frac{\alpha\eta\lambda K(w_l)b}{\gamma(1+\eta\lambda)} \qquad \text{requiring} \qquad \frac{\alpha\eta\lambda K(w_l)b^2}{\gamma(1+\eta\lambda)} > w_l - \underline{w}. \tag{9}$$

As in Section 3.1, we can define cutoff values w_l and w_g such that e_l is a valid local optimum when $w_c > w_l$ and e_g is a valid local optimum when $w_c < w_g$. These cutoff values cause the profit constraints above to hold with equality. Figure 3 illustrates. Since $w_l < w_g$, at least one of these two responses is always an option that the worker must compare to the utility he receives at the two kinks at his reference points.

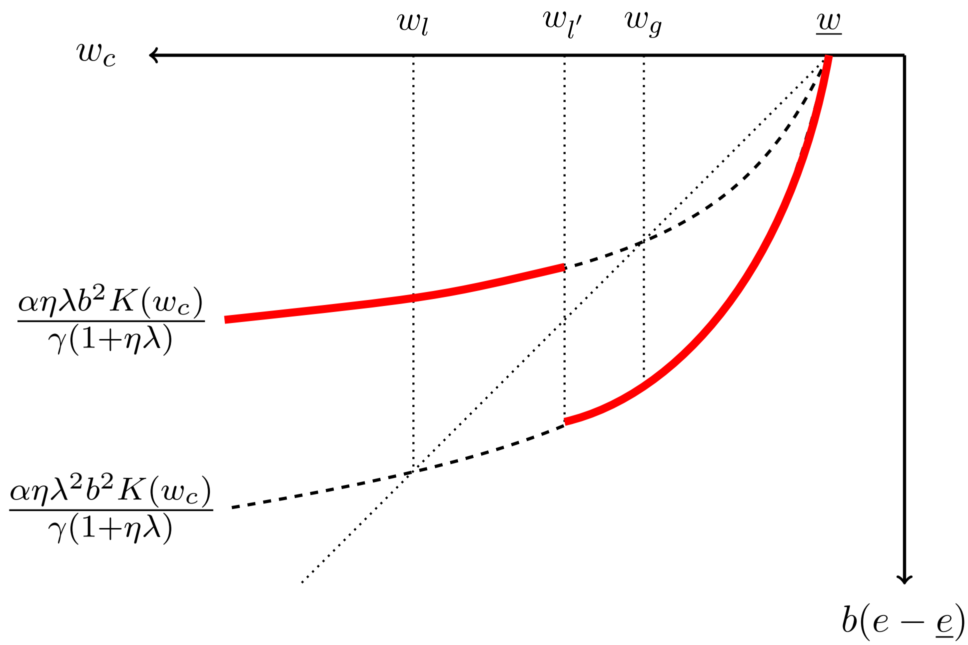
Details are left to the appendix, but e_l and e_g are always preferred to the effort at the kink $e = \underline{e} - (\underline{w} - w_l)/b$, so this is never chosen. We therefore have the revenue response function shown in red in Figure 3.

Whatever the size of the wage cut, the worker responds with a drop in effort. But had the wage cut been anticipated, Lemma 1 would have applied and effort would not have decreased at all. We can summarize this in the following result:

Proposition 3 *Retaliation for wage cuts is worse when cuts are surprising.*

Once again we must emphasize that in order to focus on the effect of surprises specifically, we have abstracted away from other forms of social preferences, including regular non-reference dependent reciprocity. Such preferences, if included, could lead to a correlation between wage cuts and retaliation even when wage cuts are fully anticipated. Nonetheless, Proposition 3 indicates that the situation is made worse, at least temporarily, by the

Figure 3: Revenue-response function to fully surprising wage cut



surprise factor.

We can also compare the size of the effort response between raises and equivalently sized cuts. This is most easily done by comparing Figure 2 to Figure 3 (rotated 180°). Note that the point labeled w_l in Figure 2 exactly corresponds to the point labeled w_g in Figure 3. It's then easy to see that small wage raises provoke a much smaller effort response than equivalently sized cuts, but large wage cuts and raises are responded to in equal magnitude. We have the following:

Proposition 4 *Negative wage shocks are reciprocated at least as strongly as equivalently sized positive wage shocks.*

In particular, surprising wage cuts, even small ones, can never be profitable for the firm. Looking to Figure 3, the revenue response function exceeds the diagonal isoprofit line on when wage cuts are very large. But intuitively, and as reflected in our assumption that $\underline{w}$ is the reservation wage, such large cuts should immediately cause workers to quit and find alternative employment. Even if job switching takes time and a worker cannot immediately quit, while this worker may therefore stick around during a temporary *small* wage cut, larger cuts are necessary to make retaliation too costly for complete reciprocation, and larger cuts are less likely to be tolerated at all.

Optimal gifts

Given the worker's effort responses, we now turn to the firm's choice problem. Not surprisingly, an optimal gift will be one that is able to elicit reciprocal effort while placing the firm in the gain domain. Proposition 5 presents the conditions for such a profitable gift to exist.

Proposition 5 *The optimal profitable full-surprise wage w_h^* is in the range $\underline{w} < w_h < w_g$, and satisfies $K'(w_h^*) = \frac{\gamma(1+\eta\lambda)}{\alpha\eta b^2}$. This optimal surprise wage exists so long as $K'(\underline{w} >) \frac{\gamma(1+\eta\lambda)}{\alpha\eta b^2}$.*

The optimal gift can be easily seen in Figure 2. The firm wants to maximize the additional revenue it will earn beyond the cost of the gift, thus it wants to maximize the distance between the part of the revenue-response function that is above the diagonal line and the diagonal line. The derivative of the kindness function at the market wage ensures that the marginal benefit of reciprocating the gift is sufficiently large, so that the revenue curve does actually rise above this diagonal. Intuitively, if the worker's reciprocal response to the gift is not sufficiently strong, then the gift (even though relatively small) will not be profitable for the firm, as the effort response will not compensate for the extra cost.

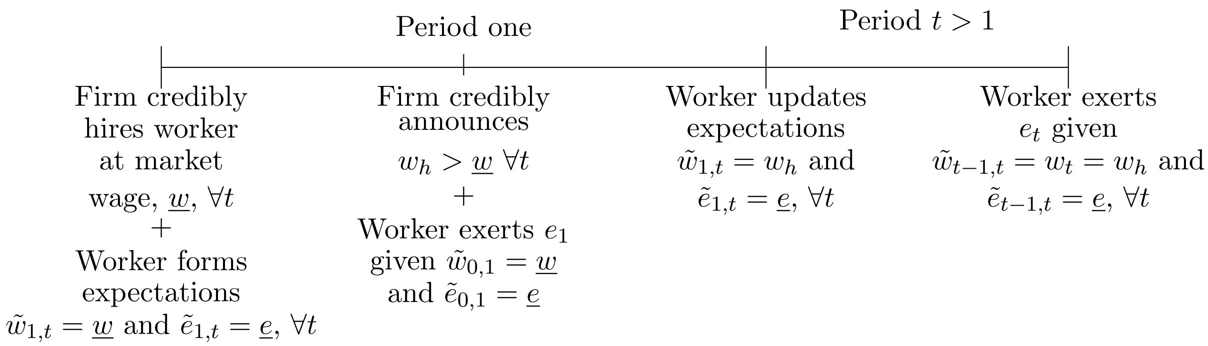
Corollary 1 describes how does the exact magnitude of the optimal gift depends on the model parameters.

Corollary 1 *The likelihood of existence of a profitable gift, the size of the optimal fully surprising gift, the size of the effort response to any gift and the optimal gift in particular, and the firm's profits after the surprise, are decreasing in γ and λ, and increasing in α, b, and η.*

The mechanism behind these relationships can be seen in Figure 2. As mentioned before, the optimal gift is the location where the gap between the effort response function for the firm's gain domain and the 45 degree "break even" line is the largest. In the case of the cost of effort parameter γ, a high γ increases the cost of reciprocation, both in terms of consumption and gain-loss utilities. This lowers the revenue-response curve in the gain domain, which in turn reduces the probability that a profitable gift exists, while if it exists, it decreases the size of it.

The optimal gift responds similarly to an increase in λ, which increases the size of losses relative to that of gains. Because the worker does not experience losses in the region of profitable gift exchange, $\underline{w} < w_h < w_g$, λ only comes into play in the effort domain. A higher λ thus, increases the cost of exerting higher effort to reciprocate the gift and thus it reduces the optimal gift in the same way that a higher γ does. To the contrary, α and b increase the value of reciprocation and thus they have a positive impact on the likelihood that an optimal gift exists, and if it exists, on its size. In particular, α has a direct impact by increasing the worker's sensitivity

Figure 4: Timing of the repeated interaction with permanent gift, fully surprising only in the first period



to a surprising gift, while b has an indirect impact through improving the returns to effort for the firm. Finally, η behaves differently since it is a relative weight on any kind of gain-loss utility, effort or reciprocity, compared to material utility. The increased value of reciprocity wins out though, and a higher η leads to a higher optimal gift. Intuitively, because K' is big enough, even though the gift is small, the response to it is large, and thus the gain in the reciprocity domain will be relatively more important than the loss in the effort domain.

The Power of Gifts Wanes Over Time

"A raise is only a raise for thirty days. After that, it's just somebody's salary."

Jim Goodnight, CEO of the SAS Institute.

In this section we show that reciprocal effort is not constant but rather it decreases over time. To this end, we now extend our analysis and consider the case where the principal and the agent interact repeatedly and, following Akerlof (1982), we assume that the gift takes the form of a fixed permanent wage increase. We preserve, however, the assumption that a gift in the first period can fully surprise the worker.

Assume there are infinitely many periods and the principal hires the agent in $t = 0$ to work at the market wage $\underline{w}$ in every period $t \geqslant 1$. Right before working for the first time, however, the principal fully surprises the worker with a wage raise, i.e., $w_t = w_h > \underline{w}$ for all t, and credibly commits to not raise the wage again.[11]

To analyze the temporal response of effort to a surprising permanent gift, start by noticing that the worker's effort decision in a given period is only linked to his previous decisions through the reference point. Let $\tilde{e}_{t-1,t}$ be the rational effort plan made in period $t-1$ for period t, and likewise, let $\tilde{w}_{t-1,t}$ be the wage that the worker expects in period $t-1$ to receive in period t.

[11] In Section 5 we relax the assumption that the principal can credibly commit to not raise the wage again and thus workers can create the expectation of a further gift.

Given the actual gift w_h, the period-t worker's problem is,

$$e_t^* \in \underset{e}{\operatorname{argmax}}\, w - \frac{\gamma}{2}(e-\underline{e})^2 + \eta\mu\big(-\frac{\gamma}{2}(e-\underline{e})^2 + \frac{\gamma}{2}(\tilde{e}_{t-1,t}-\underline{e})^2\big) + \alpha\eta\mu(K(w_h)-K(\tilde{w}_{t-1,t}))\mu(\pi(w_h,e)-\pi(\tilde{w}_{t-1,t},\tilde{e}_{t-1,t})) \tag{10}$$

The key and only difference between the worker's problem in the one shot case (see equation (4)) and that in equation (10) is that for each period t the reference point is updated in the previous period $t-1$ in case information arrives.[12] Aside from the reference-point, therefore, there is no link between periods.

(1) Period-one effort decision. Because the agent was credibly hired to work at the market wage in all periods, and thus the gift comes as a surprise, it must be the case that $\tilde{w}_{0,1} = \underline{w}$. Moreover, from Lemma 1 we know that because he does not expect any departures from expectations he forms his effort plans as a consumption utility maximizer and thus $\tilde{e}_{0,1} = \underline{e}$. The period-one decision thus reduces to that in the one-shot case and the worker increases his effort above his plan of $\underline{e}$ in response to the gift, as characterized in Proposition 1.

(2) Period-$t > 1$ effort decision. At the end of period one the agent is already aware that $w_t = w_h$ for all t. Therefore we assume that hedonic adaptation (Frederick & Loewenstein 1999, Frey & Stutzer 2002) occurs and thus $\tilde{w}_{t-1,t} = w_h$ for all $t > 1$. Following Lemma 1, the period-t, $t > 1$ decision thus reduces to maximize consumption utility and so $e_t^* = \underline{e}$.

Figure 4 displays the resulting timeline, while the following proposition summarizes,

Proposition 6 *After an increase in effort due to a surprising wage raise, effort wanes back to baseline in every period thereafter.*

An important comment on Proposition 6 is due as a final word. First, the prediction that effort spikes and then wanes would still hold had we added baseline non-reference dependent reciprocity to the worker's preference structure. The only difference would be that the level of effort after the raise would be higher than at the market wage.

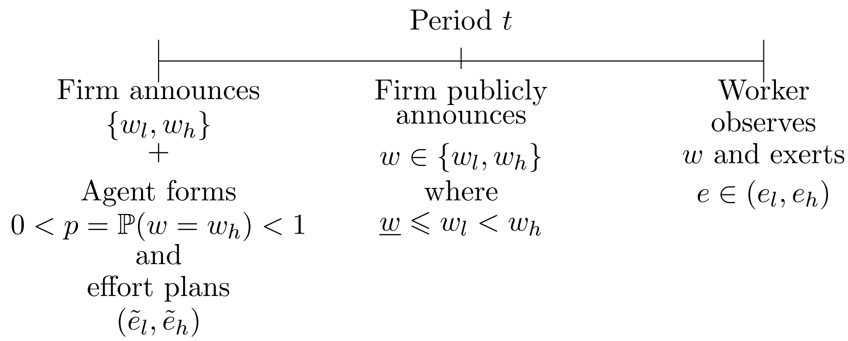
Gifts are Cursed: Grant Them Forever or Never Grant Them

> *"I've learned my lesson. I'm going on vacation every 3-4 months and I'm going to do the minimum."*
>
> Unknown worker in Yelp conversation "I didn't get a RAISE or BONUS this year".

Because workers adapt to fixed wages and thus higher wages only elicit extra effort whenever unexpected, a seemingly obvious corollary to Proposition 6 is that the optimal strategy for firms is not to grant a permanent

[12] The general setting allows for expectations to be updated at the end of every period. Since in this section we are assuming information arrives only in the first period, we refrain from emphasizing this possibility.

Figure 5: Timing of the interaction with stochastic gift



wage raise, but rather to randomize and grant gifts stochastically. Random gifts seem to be optimal as they could take advantage of the benefits of surprising agents to elicit higher effort.

In this section, however, we show that this intuition is incorrect: gifts are cursed because if granting them creates the expectation of further gifts, the decrease in expected revenue from retaliation when workers do not receive the gift outweighs the expected revenue in case the gift is granted.

The setting is as follows. We assume that after being surprised with a wage raise, the worker updates his belief and puts a positive probability $0 < p < 1$ on receiving a further gift in the next period. Because information arrives only in the first period—before the worker updates beliefs—without loss of generality we assume that this probability is constant across periods. Moreover, because in the long-run equilibrium of an infinitely repeated game these beliefs must coincide with the true probability, the firm thus must choose an optimal value of p in addition to the baseline wage level $w_l > \underline{w}$ and the gift wage level $w_h > w_l$. The worker's problem is therefore stationary and the worker's rational reference point is stochastic and equal to the firm's optimal strategy. We further assume that the worker exerts effort *after* observing whether the gift was granted, and so he must form credible contingent plans (e_l, e_h), which he will carry out after observing the wage realization. Figure 5 displays the timeline and assumptions.

To investigate whether firms might repeatedly use wage surprises to elicit higher effort, we will attempt to find a stochastic wage arrangement parameterized by (w_h, w_l, p) in which both the firm's and worker's participation constraints are met; that is, the firm must earn higher profits than if using a non-stochastic market wage $\underline{w}$, and the worker must earn higher expected utility than if taking an alternative non-stochastic market-wage job. As we mentioned before, however, we will show that such arrangement does not exist, and thus profitable random gifts can not occur in equilibrium.

On the way to characterizing such an equilibrium, we establish some conditions on what it must look like.

First, we formally state the straightforward result that workers do in fact positively reciprocate in the high wage state and negatively reciprocate in the low wage state.

Lemma 2 *Given a stochastic wage contract* (w_l, w_h, p)*, the worker's PPE* (e_l, e_h) *satisfies* $e_l < \underline{e}$ *and* $e_h > \underline{e}$*.*

Next, we find that any stochastic contract of this form will be responded to by workers in a way that ensures unequal profits in the two wage states.[13]

Lemma 3 *Given a stochastic wage contract* (w_l, w_h, p)*, the worker never responds with effort levels that ensure equal profits in either wage state.*

Next we show that if a firm in fact wants to offer a stochastic contract, workers will never respond with equal costs of effort (i.e. perfectly symmetric effort levels around $\underline{e}$).

Lemma 4 *Given a stochastic wage contract* (w_l, w_h, p) *that the firm is willing to offer instead of a certain market wage, the worker never responds with equal costs of effort in the two wage states.*

These two lemmas serve to eliminate the kinks in the worker's utility function as potential PPE plans. Finally, we can establish the main result:

Proposition 7 *A stochastic gift exchange arrangement* (w_h,w_l,p) *is never profitable for the firm compared to paying* $\underline{w}$ *with certainty.*

An important implication of Proposition 7 is that once workers learn to expect occasional gifts, it is no longer profitable to grant them. Gifts, therefore, can increase effort above the baseline if they are surprising (see Proposition 1), but inducing this excess effort is no free lunch because a surprise reveals the firm's inability to commit to a payment scheme, thereby creating the expectation of further payments, which are then no longer worthwhile to grant.

This result raises the question: If the firm never wants to grant stochastic gifts, shouldn't workers update their beliefs to discount this possibility? No: If workers did infer that $p = 0$, this would restore the opportunity for a fully surprising gift, which the firm of course would want to grant a second time (given that it granted the first). As soon as it grants a first gift, therefore, the firm has no option but to commit itself to an indefinite period of minimizing losses by setting $0 < p \leq 1$.

[13] Note that the usual principle of agents only being willing to randomize their strategies if the utilities of all strategies chosen with nonzero probability are equal does not hold here because the act of randomizing itself, and the relevant probabilities, change the payoffs involved through the worker's stochastic gain-loss utility.

Gifts are Only Profitable in Sufficiently Short Interactions

The curse of gifts, that any profitably surprising gift is inevitably followed by an infinite stream of expected losses for the firm, immediately implies the following:

Corollary 2 *In an infinite game, gift exchange is only profitable if the firm has a sufficiently high discount factor (or equivalently, if there is a sufficiently high probability of the game ending in each period.)*

But what about shorter term interactions? In the infinite game, the firm has to carefully control expectations about the probability of future gifts, but in finite interactions this incentive is limited. The analysis of the infinite period game in the previous section therefore finds a stationary solution that no longer applies when the game is finite. So can a firm profitably grant one or multiple gifts in a finite horizon version of the game? Unfortunately, this is also difficult for the firm to do over more than a single period.

In the finite-period case, the logic unravels through backwards induction. In the final interaction between a worker and the firm, the worker must expect some positive probability of a gift, because an expectation of $p = 0$ immediately opens the door for a profitable fully-surprising gift. But without the incentive to maintain low expectations of further gifts, the firm therefore unambiguously wants to choose the wage that leads to highest profits. In any stochastic contract, Lemma 3 tells us that the firm will always strictly prefer either the high or the low wage, and will not be willing to randomize between them. The firm must therefore choose a wage that they can credibly promise with certainty, which the worker will anticipate and respond to with minimal effort by Lemma 1. The firm's optimal behavior is to grant the lowest wage from which it has no profitable fully surprising alternative wage, which turns out to be the optimal fully surprising gift wage w_h, found in Section 3.3. By backwards induction this same process happens in each earlier round as well, and we have the following result:

Proposition 8 *In a finite game, the firm chooses the optimal fully surprising gift in the first round, and this turns into a permanent raise in later rounds.*

Similarly to Corollary 2, this also makes it difficult for a firm to profitably use surprises in long interactions. In an N-period game, the firm's initial profit from the surprise must outweigh the losses of having to pay a higher wage for the remaining periods. Clearly, gift exchange is more likely to be profitable the shorter the interaction.

As before, a caveat is warranted: Our model abstracts from other forms of social preferences in order to focus specifically on the effect of surprises. Clearly if workers are reciprocal at baseline, a positive correlation between wages and effort may exist in the long run, which could improve the profitability of gift exchange.

Related Evidence

In this section we relate our model predictions to the existing evidence on gift-exchange and derive some implications of our model for experimental designs whose identification strategy rely on surprises.

(1) The power of gifts is greater whenever they surprise workers. Proposition 1 showed that surprising gifts elicit more effort than anticipated ones. Direct evidence of this prediction can be found in Gilchrist, Luca & Malhotra (forthcoming). They hired three groups of oDesk workers for a one-time data-entry task, all of whom requested wages of less than \$3 per hour in their profiles. The first group is hired and paid \$3, the second is hired and paid \$4, while the third is hired at \$3 but is surprised, right before work, with a \$1 increase per hour. The find that paying \$3+\$1 yields a 20% increase in productivity compared to paying \$4, while there are no productivity differences between the \$3 and \$4 groups. Relatedly, Ockenfels, Sliwka & Werner (2015) implement a field experiment where subjects are hired for a one-time job of inserting adhesive labels into books of a University library. They hire all subjects for the market wage of 8 Euros. Aside from the baseline condition, right before work one group received one surprising increase to 12 Euros per hour, meanwhile the third group received two surprising increases to 10 and 14 Euros respectively. They find that that workers' performance is about 11% higher for the same total wage when their wage is increased in two surprising steps relative to just one of them.

Moreover, there is also empirical support for our model's prediction that the power of a surprising gift is increasing in the marginal return to effort, b, and in the marginal cost of effort, γ (see equation (5) and Corollary 1). For instance, Hennig-Schmidt, Rockenbach & Sadrieh (2010) find that clear information about the employer surplus is needed for gift exchange to arise in the lab, and Charness, Frechette & Kagel (2004) show that including a payoff table *reduces* gift exchange, but only among high-effort workers for whom marginal returns to effort are low. Englmaier & Leider (2010) also show that higher returns to effort induce more reciprocation, and Englmaier & Leider (2012*a*) find a similar result in the field simply by saying that the manager would be getting a bonus if the job was finished within the week, hinting thus higher returns to effort. Interestingly, DellaVigna et al. (2016), in an envelope-stuffing field experiment where the returns to the employer are systematically varied, find that workers do exert more effort when the return to work for the employer is positive, even though their response is not sensitive to the precise size of the return.

With respect to the effort cost, even though evidence is thin as laboratory experiments typically do not use real-effort tasks, and those using real-effort tasks usually don't measure effort costs, there are exceptions. In particular, Gneezy (2002), using a real-effort task (solve mazes), find that when difficult "level 5" mazes are used instead of "level-2" mazes, gift exchange is significantly reduced.[14]

(2) The power of gifts wanes over time. The reaction to surprises as a driver of gift exchange is also consistent with findings that observe a short-lived or waning effect after workers have become accustomed to the new higher wage. Evidence of this can be found in the two gift-exchange studies in Gneezy & List (2006). In a data-entry study, they recruited workers for the market wage of \$8 and then surprised them with a large permanent increase to \$12. They find that the initially significant effort response of 27% falls to an insignificant 11%. In a second study on door-to-door fundraising, where subjects were hired at \$10 dollars and then fully surprised with a 100% increase, effort initially spiked by 72% to then fall to an insignificant 6%. Similarly, Bellemare & Shearer (2009) found that effects on productivity were concentrated on the day a surprise bonus was given to workers in a tree-planting firm.

One important concern with these studies, however, is that in the field waning of gift exchange can be confounded with fatigue as work after the raise takes place in the same or consecutive days. In a large laboratory study, however, Sliwka & Werner (forthcoming) replicates the same finding while ruling out fatigue. In particular, Sliwka & Werner (forthcoming) document that an unanticipated wage increase in the form of a permanent raise only increases effort relative to the constant-wage baseline profile in the period the raise is granted. To the contrary, they find that an unanticipated wage profile with the same overall cost, which sets smaller regular raises instead of one big permanent raise, does increase overall effort. Most importantly, the effort response to the permanent raise is of a similar magnitude to that of a smaller raise in the same period, which is subsequently followed by other effort increases relative to the baseline. Thus, it can not be the case that subjects were unable due to fatigue to reciprocate the permanent rise in the period subsequent to the permanent rise.

The fact that surprising incentives are powerful but only in the short run raises a word of caution for designs that rely on surprises in their identification strategy. Even though the element of surprise is included for a good reason—in the case of gift exchange to avoid the selection of more productive workers with higher reservation wages (Lazear 2000)—our model predicts that it can overestimate the extent of the hypothesis under study. Thus our model recommends that, whenever possible, the effects of the intervention should be assessed in longer spans.[15]

[14]Interestingly, Gneezy (2002) also finds that the increase in effort when returns to effort are high is *smaller* than when returns are low, but that the former case is profitable, unlike the latter. This is also compatible with a model of reference-dependent preferences, because the worker may be trying to exactly "repay" the employer, and no more. That is, he may be choosing effort to land on the kink in his utility function where he experiences no gains or losses in reciprocity. If he is at that kink (that is $e = \underline{e} + (w_h - \underline{w})/b$), then an increase in b is in fact expected to marginally reduce effort.

[15]The surprising element of an intervention is not confined to the gift-exchange literature. For instance Karlan & Zinman (2009) implement a randomized experiment on a South-African micro-lender to distinguish the effects of moral hazard and adverse selection on default rates. By surprisingly reducing the interest rate to a group of clients the day they show up to take the loan, they

(3) Gifts are cursed: grant them forever or never grant them. Propositions 7 through 8 show that if firms want to use gifts as an instrument to motivate workers in the long run, they will have to either suffer the losses from retaliation whenever failing to give further gifts, or make the initial gift a permanent raise. This result built on two key aspects of our model, (1) the assumption that one-time gifts create the expectation of further gifts, and (2) whenever these expectation are unmet, it triggers negative reciprocity that is stronger than positive reciprocity, rendering gifts unprofitable in expectation.

Even though to the best of our knowledge there is no evidence in economics of monetary gifts leading to the expectation of further monetary gifts, this assumption is intuitive and relates to the literature of learning under ambiguity. For instance, Epstein & Schneider (2007) develop a model where players do not know the true probability of an event, and thus their subjective probabilities change as the game progresses and players learn. The idea that under ambiguity players learn and update their state space is thus a common one in the literature.[16]

The second aspect, the asymmetry between positive and negative reciprocity when responding to deviations from an expected payment, also reflected in the one-shot interaction analyzed in Proposition 4, is well grounded in the experimental evidence. Hannan (2005) modify a standard laboratory gift-exchange experiment to add an exogenous shock to the firms' profit after which firms and workers could adjust their previous choices. She finds that adjusting wages downwards has a negative impact on effort choice, which is twice as large as the effect of a wage increase of the same magnitude. Engelmann & Ortmann (2009) similarly find little evidence of positive reciprocity, but do find negative reciprocity, in an experimental design that modifies the traditional setup by moving equilibrium away from a corner solution. In the field, there are similar results. Kube, Maréchal & Puppe (2013) hired workers for a data-entry task for a projected wage of 15 Euros. The day of the experiment a group of workers were actually paid 10 Euros and the other 20 Euros. They found that cutting the expected payment reduced average output by 20% relative to the control that received the expected 15 Euros, while the wage increase did not increase output even though effort was not bounded above, which was verified by using a piece rate.[17] These results are in line with the well stablished stylized fact that firms are reluctant to cut wages to avoid hurting worker's "morale" (e.g., Bewley (2009)).

separately identify the extent of moral hazard and adverse selection on default rates.Our model suggests that Karlan & Zinman's (2009) methodology can underestimate the role of both moral hazard and adverse selection on default rates as the client's departure from expectations about his own costs and benefits and those of the firm, could have triggered reference-dependent reciprocal behavior that led to improvements of their ex-post behavior. Our model further predicts that, even if the clients' behavior is biased upwards by their reference-dependent social preferences, once they acclimatize the firm's gift, their default rates should increase.

[16] Relatedly, there is a growing literature in contract theory studying the shape of the optimal incentive scheme whenever workers do not know the whole action space and thus they can be fully surprised by actions that they were not initially aware off. See Von Thadden & Zhao (2012) and Carroll (2015).

[17] Workers, however, were recruited under the piece rate, opening the possibility of selection.

As mentioned in the text, an important caveat on our results pertaining to the profitability of gift exchange is warranted: If workers have non-reference dependent reciprocity in their preferences, they may reciprocate gifts even in the long run, which could in turn improve the profitability of repeated gift exchange. We do not claim to rule out this possibility. If, however, evidence on the waning size of gift exchange (Gneezy & List 2006, Bellemare & Shearer 2009) or the possibly weak nature of gift exchange in the field (Kube, Maréchal & Puppe 2012, Hennig-Schmidt, Rockenbach & Sadrieh 2010, List 2006, Esteves-Sorenson & Macera 2014, DellaVigna et al. 2016) is interpreted as evidence that this baseline, non-reference dependent reciprocity is relatively weak, our results should apply directly without readjusting for omitted forms of social preferences.

Discussion

In this section we discuss some implications of our model predictions and venues for future research.

Managing expectations is crucial to make the most of gifts

> *"[I] Wonder if we will see a Christmas bonus again. I called the Partner Contact Center and they have [stated that] it was a one time deal because the company made so much money last year."*
>
> Starbucks employee on the surprising bonus of $250
> granted in 2004 (http://starbucksgossip.typepad.com.)

The messages delivered in Section 3 to Section 6 point out the limitations of using gift exchange as an incentive mechanism. Our model predicts that the most powerful gifts are those that are surprising, but surprises only trigger temporary excess effort. Moreover, surprises most probably create the expectation of further gifts, which whenever unfulfilled are harmful for the firm's profits. Firms, therefore, cannot repeatedly exploit the profitable effort response to surprising gifts.

We do not believe, however, that gifts have no scope to motivate workers. Rather, our analysis leads us to the conclusion that leveraging expectations is crucial to rely on gifts to motivate workers. Firms that are able to leverage employees expectations by credibly convincing workers that gifts are only one time—or if the gift is in the form of a higher permanent wage, by credibly convincing workers that there will be no further raises—will be able to exploit the benefits of unexpected gift giving. Intuitively, if firms can manage workers' expectations, then monetary gifts can be a useful instrument to motivate employees while avoiding future retaliation against a failure to repeat those gifts.

The idea that leveraging expectations is crucial for the efficacy and efficiency of gift exchange opens new and interesting questions on what are the mechanisms that firms use to manage expectations of further raises or of new transitory gifts. In a related paper, we take this question to the realm of gifts in kind and explore whether one nice property of this type of gifts (relative to monetary gifts) is that they do not create the expectation of a further gift. We speculate that this property might be at the heart of the widespread use of this type of incentives in the workplace.

The idea that managing expectations is crucial for gift exchange to occur speaks to an often neglected aspect of incentives in the workplace: their implementation. Our model suggests that any firm considering using gifts to motivate its workforce should not only focus on the amount, frequency, and type of the gift, but also how it will be delivered to workers. Clearly from the quote above, this is one of the issues that Starbucks forgot to design when they decided to grant an unusual Christmas bonus to their employees.[18]

The temporal structure of optimal gifts

Our model also speaks to another issue that has not received due attention in the gift-exchange literature: the temporal structure of the monetary gift. Indeed, most of the experimental literature (specially the laboratory evidence) has abstracted from this issue by assuming one-shot principal-agent interactions. Even though necessary in order to avoid confounding reciprocity with reputation confounds, this assumption has drawn attention away from studying the properties of gifts in repeated interactions.

Building on Akerlof (1982), most of the literature on gift exchange (both in the field and the few papers in the laboratory in which repeated interaction takes place) has assumed that gift exchange takes place through a permanent wage raise, which aims to mimic an above-market wage payment.[19] Our analysis enriches this view by studying a possible alternative to a permanent wage raise, namely, random one-period gifts. We show that if firms want to motivate workers through gifts, a permanent wage raise is actually optimal relative to random gifts as the former manages expectations and avoids having workers draw harmful inferences about the likelihood of future gifts.

Anecdotal evidence, however, suggests that random monetary gifts are sometimes observed in real-world firms (see for instance the well-know case when Oprah surprisingly gave $10,000 to her O-Magazine employees along with some gifts in kind).[20] None of this anecdotal evidence, however, seems to point out to a repeated use of this type of incentives. This suggests that firms did use some mechanism in order to avoid the creating of harmful expectations. We therefore interpret this evidence as one time events that are exploited by firms for workers to "develop sentiment" for the firm so that a standard gift-exchange mechanism as proposed by Akerlof (1982) can take place. Whether one-time gifts can serve this purpose is a venue for future research.

[18] An other important part of implementation has to do with taxes: surprising bonuses are considered discretionary payments and thus excluded from overtime pay calculation.

[19] To the best of our knowledge, Dufwenberg & Kirchsteiger (2004) is the only theoretical attempt to also study reciprocity in a repeated setting.

[20] See the news in `http://marquee.blogs.cnn.com/2010/06/16/oprah-doles-out-thousands-to-magazine-staff/`.

The laboratory versus field tests debate on the efficacy of gift exchange

It is now well known that field test of the gift-exchange hypothesis have not been able to replicate the large wage-effort elasticities found in laboratory tests using the gift-exchange game (Fehr, Kirchsteiger & Riedl (1993)). Esteves-Sorenson & Macera (2014) summarizes to show that in the lab wage-effort elasticities are always above one. In the field, however, the most common elasticity is zero. For instance, Hennig-Schmidt, Rockenbach & Sadrieh (2010) find no gift exchange at all among students hired to type research abstracts, and Englmaier & Leider (2012*a*) find, if anything, negative gift exchange in their baseline treatment using temp workers in a data entry job. Similarly, Kube, Maréchal & Puppe (2013) find no evidence of positive reciprocity to monetary gifts, Al-Ubaydli et al. (2015) observe minor and insignificant gift exchange among temp workers hired to stuff envelopes, and Ockenfels, Sliwka & Werner (2015) obtain a similar outcome using temp workers inserting tags into library books. More recently, Esteves-Sorenson & Macera (2014) find no significant gift exchange at all with a data-entry task despite larger than usual gifts, while DellaVigna et al. (2016) also fails to find reciprocal effort in response to monetary gifts in a stuffing-envelope field experiment with an impressive sample size.[21]

The disparity in findings have lead to a heated debate. Levitt & List (2007) catalogues various reasons why laboratory experiments on social preferences may not extend to the field in general, with specific discussion on the gift exchange literature. Camerer (2011) responds, and on the specific issue of gift exchange shows that one experiment directly comparing field to lab outcomes, List (2006), finds comparable outcomes in the two settings on aggregate, with heterogeneity based on whether field actors are usually active in the same local market. Esteves-Sorenson & Macera (2014) also identifies an array of differences that can rationalize the disparity in findings.

Our model points out a new aspect that can shed light on the modest elasticities observed in the field: the lack of credibility of gifts due to the experimenter's out-of-equilibrium play. To see this, recall that our model shows that reciprocal effort is short lived and that as a result firms will incur in losses after the initial effort spike. This renders gifts profitable only if they are sufficiently small and if the interaction with the principal is short enough. Because workers with rational expectations are aware of this equilibrium behavior, surprising gifts given in field tests—which are often of a large magnitude—might be viewed by subjects as an irrational move by the hiring firm. The key question is then, can reciprocal behavior arise if workers judge firms as irrational? May be not. A gift that is perceived as unprofitable, and thus not likely observed in the real-world, might damage the basic nature of a labor relationship, which is based on the mutual benefits of the interaction.

The fact that workers in field tests might perceive the surprising gift as non profitable could damage the very

[21] DellaVigna et al. (2016) and Esteves-Sorenson & Macera (2014) both summarize the existing field evidence.

purpose of a field experiment: its external validity. Surprising gifts do not seem to be common in real-world labor markets (for good reasons according to our model predictions), even less in newly stablished, short-term interactions between employers and employees. Thus, with the purpose of avoiding selection of abler workers field-experiment designs might be damaging their very purpose of examining reciprocal behavior in a naturally occurring environment.

We conclude that the future challenge of field tests in gift exchange is to employ designs that consider the equilibrium play of the firm to achieve identification. This implies that future experimental designs should not only be guided by models that focus on the worker's response to gift giving, but which also focus on the equilibrium behavior of firms, in order to ensure the credibility of gift-giving in the field.

Conclusion

This paper studies the power of gift exchange to elicit excess effort. We extend the approach in Kőszegi & Rabin (2006) to model workers with reference-dependent reciprocal preferences, incorporating the well established psychological principle that people adapt to any constant stimuli. Under this assumption, we show that gifts are the most powerful whenever they are surprising, and the excess effort they elicit wanes overtime. Moreover, we show that in equilibrium firms cannot repeatedly surprise workers with gifts, and thus the only profitable gift is cursed as it must be granted forever. We also study the profitability of gifts to find that only small gifts are profitable and only in short-term interactions.

In addition to contributing to the incentives literature by studying the scope of above-market wages as in incentive mechanism, and to the gift-exchange literature by theoretically studying the unstudied surprising or expected aspect of monetary gifts, this paper also contributes to the theoretical literature on expectations-based reference-dependent preferences by proposing a novel model of reference-dependent preferences. Reference-dependent preferences have been shown to be relevant to economic behavior in a large array of domains, from financial decisions, insurance, consumption and saving, pricing, labor supply, etc. (see DellaVigna (2009) and Barberis (2013) for reviews). Moreover, recently, the laboratory and empirical evidence on expectations as reference points has also flourished. For instance Abeler et al. (2011), Gill & Prowse (2012) Marzilli Ericson & Fuster (2011), and Karle, Kirchsteiger & Peitz (2015) all present laboratory evidence of the role of expectations as reference points, while Card & Giuliano (2011), Crawford & Meng (2011), Pope, Price & Wolfers (2011) and Lien, Peng & Zheng (2015) provide empirical evidence.[22] We argue that these principles can naturally combine with social preferences, and demonstrate how this combination concisely organizes evidence on gift exchange.

[22] For opposing evidence of the role of expectation as the reference point see Heffetz & List (2014) in the context of the endowment effect and Zimmermann (2015) in the context of the timing of information arrival.

References

Abeler, Johannes, Armin Falk, Lorenz Goette, and David Huffman. 2011. "Reference points and effort provision." *American Economic Review*, 101(2): 470–492.

Abeler, Johannes, Steffen Altmann, Sebastian Kube, and Matthias Wibral. 2010. "Gift exchange and workers' fairness concerns: When equality is unfair." *Journal of the European Economic Association*, 8(6): 1299–1324.

Akerlof, George A. 1982. "Labor contracts as partial gift exchange." *Quarterly Journal of Economics*, 97(4): 543–569.

Akerlof, George A., and Janet L. Yellen. 1990. "The fair wage-effort hypothesis and unemployment." *Quarterly Journal of Economics*, 105(2): 255–283.

Al-Ubaydli, Omar, Steffen Andersen, Uri Gneezy, and John A. List. 2015. "Carrots that look like sticks: Toward an understanding of multitasking incentive schemes." *Southern Economic Journal*, 81(3): 538–561.

Ambrose, Maureen L., and Carol T. Kulik. 1999. "Old friends, new faces: Motivation research in the 1990s." *Journal of Management*, 25(3): 231–292.

Barberis, Nicholas C. 2013. "The Psychology of Tail Events: Progress and Challenges." *American Economic Review*, 103(3): 611–616.

Battigalli, Pierpaolo, and Martin Dufwenberg. 2007. "Guilt in games." *American Economic Review*, 97(2): 170–176.

Bellemare, Charles, and Bruce Shearer. 2009. "Gift giving and worker productivity: Evidence from a firm-level experiment." *Games and Economic Behavior*, 67(1): 233–244.

Benjamin, Daniel J. 2015. "A theory of fairness in labor markets." *Japanese Economic Review*, 66(2): 182–225.

Bewley, Truman F. 2009. *Why Wages Don't Fall During a Recession.* Harvard University Press.

Brandts, Jordi, Klarita Gërxhani, Arthur Schram, and Jolanda Ygosse-Battisti. 2010. "Size doesn't matter! Gift exchange in experimental labor markets." *Journal of Economic Behavior & Organization*, 76(3): 544–548.

Brickman, Philip, and Donald T. Campbell. 1971. "Hedonic relativism and planning the good society." In *Adaptation-level theory: A symposium.* , ed. Mortimer Herbert Appley, 287–305. Academic Press.

Brown, Martin, Armin Falk, and Ernst Fehr. 2004. "Relational contracts and the nature of market interactions." *Econometrica*, 72(3): 747–780.

Brown, Martin, Armin Falk, and Ernst Fehr. 2012. "Competition and relational contracts: The role of unemployment as a disciplinary device." *Journal of the European Economic Association*, 10(4): 887–907.

Camerer, Colin F. 2011. "The promise and success of lab-field generalizability in experimental economics: A critical reply to Levitt and List." *SSRN Electronic Journal*.

Card, David, and Laura Giuliano. 2011. "Peer Effects and Multiple Equilibria in the Risky Behavior of Friends." NBER Working Paper Series No. 17088.

Carroll, Gabriel. 2015. "Robustness and linear contracts." *American Economic Review*, 105(2): 536–563.

Charness, Gary B. 2000. "Responsibility and effort in an experimental labor market." *Journal of Economic Behavior & Organization*, 42(3): 375–384.

Charness, Gary B. 2004. "Attribution and reciprocity in an experimental labor market." *Journal of Labor Economics*, 22(3): 665–688.

Charness, Gary B., and Ernan Haruvy. 2002. "Altruism, equity, and reciprocity in a gift-exchange experiment: an encompassing approach." *Games and Economic Behavior*, 40(2): 203–231.

Charness, Gary B., and Peter J. Kuhn. 2007. "Does pay inequality affect worker effort? Experimental evidence." *Journal of Labor Economics*, 25(4): 693–723.

Charness, Gary B., Guillaume R. Frechette, and John H. Kagel. 2004. "How robust is laboratory gift exchange?" *Experimental Economics*, 7(2): 189–205.

Charness, Gary B., Ramón Cobo-Reyes, Natalia Jiménez, Juan A. Lacomba, and Francisco Lagos. 2012. "The hidden advantage of delegation: Pareto improvements in a gift exchange game." *American Economic Review*, 102(5): 2358–2379.

Charness, Gary B., Till Gross, and Christopher Guo. 2014. "Merit pay and wage compression with productivity differences and uncertainty." Working Paper.

Chaudhuri, Ananish, and Erwann Sbai. 2011. "Gender differences in trust and reciprocity in repeated gift exchange games." *New Zealand Economic Papers*, 45(1-2): 81–95.

Clark, Andrew E., David Masclet, and Marie-Claire Villeval. 2010. "Effort and comparison income: Experimental and survey evidence." *Industrial and Labor Relations Review*, 63(3).

Cohn, Alain, Ernst Fehr, and Lorenz F. Goette. "Fair wages and effort provision: Combining evidence from the lab and the field." Forthcoming in *Management Science*.

Cohn, Alain, Ernst Fehr, Benedikt Herrmann, and Frédéric Schneider. 2014. "Social comparison and effort provision: Evidence from a field experiment." *Journal of the European Economic Association*, 12(4): 877–898.

Crawford, Vincent P., and Juanjuan Meng. 2011. "New York City cabdrivers' labor supply revisited: reference-dependence preferences with rational-expectations targets for hours and income." *American Economic Review*, 101(5): 1912 – 1932.

DellaVigna, Stefano. 2009. "Psychology and Economics: Evidence from the Field." *Journal of Economic Literature*, 47(2): 315–372.

DellaVigna, Stefano, John A. List, Ulrike Malmendier, and Gautam Rao. 2016. "Estimating social preferences and gift exchange at work."

Diener, Ed, Richard E. Lucas, and Christie Napa Scollon. 2006. "Beyond the hedonic treadmill: Revising the adaptation theory of well-being." *American Psychologist*, 61(4): 305–314.

Dufwenberg, Martin, and Georg Kirchsteiger. 2004. "A theory of sequential reciprocity." *Games and Economic Behavior*, 47(2): 268–298.

Dur, Robert. 2009. "Gift exchange in the workplace: Money or attention?" *Journal of the European Economic Association*, 7(2-3): 550–560.

Engelmann, Dirk, and Andreas Ortmann. 2009. "The robustness of laboratory gift exchange: A reconsideration." Working Paper.

Englmaier, Florian, and Stephen G. Leider. 2010. "Gift exchange in the lab: It is not (only) how much you give..." CESifo Working Paper No. 2944.

Englmaier, Florian, and Stephen G. Leider. 2012*a*. "Contractual and organizational structure with reciprocal agents." *American Economic Journal: Microeconomics*, 4(2): 146–183.

Englmaier, Florian, and Stephen G. Leider. 2012*b*. "Managerial payoff and gift exchange in the field." CESifo working paper: Behavioral Economics No. 3707.

Epstein, Larry G., and Martin Schneider. 2007. "Learning under ambiguity." *Review of Economic Studies*, 74(4): 1275–1303.

Esteves-Sorenson, Constança, and Rosario Macera. 2014. "Gift exchange in the workplace: Addressing the conflicting evidence with a careful test." Working Paper.

Falk, Armin, and Urs Fischbacher. 2006. "A theory of reciprocity." *Games and Economic Behavior*, 54(2): 293–315.

Fehr, Ernst, and Armin Falk. 1999. "Wage rigidity in a competitive incomplete contract market." *Journal of Political Economy*, 107(1): 106–134.

Fehr, Ernst, Erich Kirchler, Andreas Weichbold, and Simon Gächter. 1998. "When social norms overpower competition: Gift exchange in experimental labor markets." *Journal of Labor Economics*, 16(2): 324–351.

Fehr, Ernst, Georg Kirchsteiger, and Arno Riedl. 1993. "Does fairness prevent market clearing? An experimental investigation." *Quarterly Journal of Economics*, 108(2): 437–459.

Fehr, Ernst, Georg Kirchsteiger, and Arno Riedl. 1998. "Gift exchange and reciprocity in competitive experimental markets." *European Economic Review*, 42(1): 1–34.

Frederick, Shane, and George F. Loewenstein. 1999. "Hedonic adaptation." In *Well-being: The foundations of hedonic psychology.* , ed. Daniel Kahneman, Ed Diener and Norbert Schwarz, 302–329. New York:Russell Sage Foundation.

Frey, Bruno S., and Alois Stutzer. 2002. "What can economists learn from happiness research?" *Journal of Economic Literature*, 40(2): 402–435.

Gächter, Simon, and Christian Thöni. 2010. "Social comparison and performance: Experimental evidence on the fair wage–effort hypothesis." *Journal of Economic Behavior & Organization*, 76(3): 531–543.

Gilchrist, Duncan S., Michael Luca, and Deepak Malhotra. forthcoming. "When $3 + 1 > 4$: Gift structure and reciprocity in the field." *Management Science*. Forthcoming in *Management Science*.

Gill, David, and Victoria Prowse. 2012. "A structural analysis of disappointment aversion in a real effort competition." *American Economic Review*, 102(1): 469–503.

Gneezy, Uri. 2002. "Does high wage lead to high profits? An experimental study of reciprocity using real effort." Working Paper.

Gneezy, Uri, and John A. List. 2006. "Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments." *Econometrica*, 74(5): 1365–1384.

Greenberg, Jerald. 1990. "Employee theft as a reaction to underpayment inequity: The hidden cost of pay cuts." *Journal of Applied Psychology*, 75(5): 561–568.

Güth, Werner, Manfred Königstein, Judit Kovács, and Enikõ Zala-Mezõ. 2001. "Fairness within firms: The case of one principal and multiple agents." *Schmalenbach Business Review*, 53(April): 82–101.

Hannan, R. Lynn. 2005. "The combined effect of wages and firm profit on employee effort." *Accounting Review*, 80(1): 167–188.

Hannan, R. Lynn, John H. Kagel, and Donald V. Moser. 2002. "Partial gift exchange in an experimental labor market: Impact of subject population differences, productivity differences , and effort requests on behavior." *Journal of Labor Economics*, 20(4): 923–951.

Heffetz, Ori, and John A. List. 2014. "Is the endowment effect a reference effect?" *Journal of the European Economic Association*, 12(5): 1396–1422.

Heidhues, Paul, and Botond Kőszegi. 2008. "Competition and price variation when consumers are loss averse." *American Economic Review*, 98(4): 1245–1268.

Heidhues, Paul, and Botond Kőszegi. 2014. "Regular prices and sales." *Theoretical Economics*, 9(1): 217–251.

Helson, Harry. 1964. *Adaptation-level Theory.* Oxford:Harper and Rowe.

Hennig-Schmidt, Heike, Bettina Rockenbach, and Abdolkarim Sadrieh. 2010. "In search of workers' real effort reciprocity — A field and a laboratory experiment." *Journal of the European Economic Association*, 8(4): 817–837.

Huck, Steffen, Andrew J. Seltzer, and Brian Wallace. 2011. "Deferred compensation in multiperiod labor contracts: An experimental test of Lazear's model." *American Economic Review*, 101(2): 819–843.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect theory: An analysis of decision under risk." *Econometrica*, 47(2): 263–292.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1986. "Fairness as a constraint on profit seeking: Entitlements in the market." *American Economic Review*, 76(4): 728–741.

Karlan, Dean, and Jonathan Zinman. 2009. "Observing unobservables: Identifying information asymmetries with a consumer credit field experiment." *Econometrica*, 77(6): 1993–2008.

Karle, Heiko, Georg Kirchsteiger, and Martin Peitz. 2015. "Loss aversion and consumption choice: Theory and experimental evidence." *American Economic Journal: Microeconomics*, 7(2): 101–120.

Khalmetski, Kiryl, Axel Ockenfels, and Peter Werner. 2015. "Surprising gifts: Theory and laboratory evidence." *Journal of Economic Theory*, 159: 163–208.

Kocher, Martin G., and Matthias Sutter. 2003. "Individual versus group behavior and the role of the decision making procedure in gift-exchange experiments." Working Paper.

Kocher, Martin G., Wolfgang J. Luhan, and Matthias Sutter. 2012. "Testing a forgotten aspect of Akerlof 's gift exchange hypothesis: Relational contracts with individual and uniform wages." IZA Discussion Paper No. 6415.

Koppel, Hannes, and Tobias Regner. 2011. "Corporate social responsibility in the work place: Experimental evidence on CSR from a gift-exchange game." Jena Economic Research Papers No. 2011,030.

Kőszegi, Botond. 2014. "Behavioral contract theory." *Journal of Economic Literature*, 52(4): 1075–1118.

Kőszegi, Botond, and Matthew Rabin. 2006. "A model of reference-dependent preferences." *Quarterly Journal of Economics*, 121(4): 1133–1165.

Kőszegi, Botond, and Matthew Rabin. 2007. "Reference-dependent risk attitudes." *The American Economic Review*, 97(4): 1047–1073.

Kőszegi, Botond, and Matthew Rabin. 2009. "Reference-dependent consumption plans." *American Economic Review*, 99(3): 909–936.

Kranton, Rachel E. 1996. "Reciprocal exchange: a self-sustaining system." *American Economic Review*, 86(4): 830–851.

Krueger, Alan B., and Alexandre Mas. 2004. "Strikes, Scabs, and Tread Separations: Labor Strife and the Production of Defective Bridgestone/Firestone Tires." *Journal of Political Economy*, 112(2): 253–289.

Kube, Sebastian, Michel André Maréchal, and Clemens Puppe. 2012. "The currency of reciprocity: Gift exchange in the workplace." *American Economic Review*, 102(4): 1644–1662.

Kube, Sebastian, Michel André Maréchal, and Clemens Puppe. 2013. "Do wage cuts damage work morale? Evidence from a natural field experiment." *Journal of the European Economic Association*, 11(4): 853–870.

Lazear, Edward P. 2000. "Performance pay and productivity." *American Economic Review*, 90(5): 1346–1361.

Lee, Darin, and Nicholas G. Rupp. 2007. "Retracting a gift: How does employee effort respond to wage reductions?" *Journal of Labor Economics*, 25(4): 725–761.

Levine, David K. 1998. "Modeling altruism and spitefulness in experiments." *Review of Economic Dynamics*, 1(3): 593–622.

Levitt, Steven D., and John A. List. 2007. "What do laboratory experiments measuring social preferences reveal about the real world?" *Journal of Economic Perspectives*, 21(2): 153–174.

Lien, Jaimie W., Qingqing Peng, and Jie Zheng. 2015. "Major Earthquakes and Time Preference Indicators." Working Paper.

List, John A. 2006. "The behavioralist meets the market: Measuring social preferences and reputation effects in actual transactions." *Journal of Political Economy*, 114(1): 1–37.

Malmendier, Ulrike, Vera L. te Velde, and Roberto A. Weber. 2014. "Rethinking reciprocity." *Annual Review of Economics*, 6: 849–874.

Marzilli Ericson, Keith M., and Andreas Fuster. 2011. "Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments." *Quarterly Journal of Economics*, 126(4): 1879–1907.

Netzer, Nick, and Armin Schmutzler. 2014. "Explaining gift-exchange - The limits of good intentions." *Journal of the European Economic Association.*

Ockenfels, Axel, Dirk Sliwka, and Peter Werner. 2015. "Timing of kindness – Evidence from a field experiment." *Journal of Economic Behavior & Organization*, 111: 79–87.

Pope, Devin G., Joseph P. Price, and Justin Wolfers. 2011. "Awareness reduces racial bias." NBER Working Paper Series No. 19765.

Rabin, Matthew. 1993. "Incorporating fairness into game theory and economics." *American Economic Review*, 83(5): 1281–1302.

Schram, Arthur, Jordi Brandts, and Klarita Gërxhani. 2010. "Information, bilateral negotiations, and worker recruitment." *European Economic Review*, 54(8): 1035–1058.

Siang, Ch'ng Kean, Till Requate, and Israel Waichman. 2011. "On the role of social wage comparisons in gift-exchange experiments." *Economics Letters*, 112(1): 75–78.

Sliwka, Dirk, and Peter Werner. forthcoming. "How do agents react to dynamic wage increases? An experimental study." *Journal of Labor Economics.*

Van Eerde, Wendelien, and Henk Thierry. 1996. "Vroom's expectancy models and work-related criteria: A meta-analysis." *Journal of Applied Psychology*, 81(5): 575–586.

Von Thadden, Ernst Ludwig, and Xiaojian Zhao. 2012. "Incentives for unaware agents." *Review of Economic Studies*, 79(3): 1151–1174.

Vroom, Victor Harold. 1964. *Work and Motivation.* New York:Wiley.

Zimmermann, Florian. 2015. "Clumped or piecewise? Evidence on preferences for information." *Management Science*, 61(4): 740–753.

Appendix: Proofs

Proof of Lemma 1

The worker forms a plan $\tilde{e}$ and then chooses actual effort to maximize $U(e, \tilde{w}|\tilde{e}, \tilde{w})$. With no surprises there is no gain-loss utility, and utility is clearly maximized by the minimum-cost effort level $\underline{e}$. Therefore, this is the only credible plan and the unique PPE.

Proof of Proposition 1

By lemma 1, the employee's effort plan is $\tilde{e} = \underline{e}$. The possible actions after hearing about the wage surprise are to increase effort to e_g so that profits are in the gain domain, to choose another effort e_l so that profits are in the loss domain, to stick with the plan $\underline{e}$, or to exactly compensate the firm for their profit losses due to the wage increase: $be - w_h = b\underline{e} - \underline{w} \Rightarrow e = (w_h - \underline{w})/b + \underline{e}$.

As shown in the text, the first two possibilities are given by

$$e_g = \underline{e} + \frac{\alpha\eta K(w_h)b}{\gamma(1+\eta\lambda)} \qquad \text{requiring} \qquad w_h < \frac{\alpha\eta K(w_h)b^2}{\gamma(1+\eta\lambda)} + \underline{w}. \tag{11}$$

and

$$e_l = \underline{e} + \frac{\alpha\eta\lambda K(w_h)b}{\gamma(1+\eta\lambda)} \qquad \text{requiring} \qquad w_h > \frac{\alpha\eta\lambda K(w_h)b^2}{\gamma(1+\eta\lambda)} + \underline{w} \tag{12}$$

Figure 2 shows the curves defined by the RHS of these profit constraints, so that the inequalities hold with equality when $w_h = w_g$ and $w_h = w_l$ respectively, so that e_g is a valid local optimum when $w_h < w_g$ and e_l is a valid local optimum when $w_h > w_l$.

At a given wage gift w_h the worker must check whether these optima exist and compare the one(s) that does to the utility he would get from the kinked point in his utility function, where the first order condition doesn't exist. This yields

$$U\left(\frac{w_h - \underline{w}}{b} + \underline{e}, w_h|\underline{e}, \underline{w}\right) = w_h - \frac{\gamma}{2}(1+\eta\lambda)\left(\frac{w_h - \underline{w}}{b}\right)^2.$$

Likewise, the utilities resulting from e_h and e_l when they are true local optima are respectively

$$U(e_g, w_h|\underline{e}, \underline{w}) = w_h + \frac{(\alpha\eta b K(w_h))^2}{2\gamma(1+\eta\lambda)} - \alpha\eta K(w_h)(w_h - \underline{w})$$

and

$$U(e_l, w_h|\underline{e}, \underline{w}) = w_h + \frac{(\alpha\eta\lambda bK(w_h))^2}{2\gamma(1+\eta\lambda)} - \alpha\eta\lambda K(w_h)(w_h - \underline{w}).$$

Comparing e_g to $e = \underline{e} + (w_h - \underline{w})/b$, we find that e_g is preferred when

$$\left(w_h - \underline{w} - \frac{b^2\alpha\eta K(w_h)}{\gamma(1+\eta\lambda)}\right)^2 > 0$$

which is of course always true, and e_l is similarly always preferred when it exists. Therefore, as shown in Figure 2, e_g is chosen when $w_h < w_g$, e_l is chosen when $w_h > w_l$, and in the region where neither is true, the kink point is chosen.

Extreme or corner cases in which w_l and/or w_g are zero are straightforward to account for.

Proof of Proposition 2

In the text.

Proof of Proposition 3

Proceeding as in the proof of Proposition 1, as shown in the text, the possible optima that are not at the kink in the utility function, and the profit constraints that they require/imply, are given by

$$e_g = \underline{e} + \frac{\alpha\eta\lambda K(w_c)b}{\gamma(1+\eta\lambda)} \qquad \text{requiring} \qquad w_c - \underline{w} < \frac{\alpha\eta\lambda K(w_c)b^2}{\gamma(1+\eta\lambda)}. \tag{13}$$

and

$$e_l = \underline{e} + \frac{\alpha\eta\lambda^2 K(w_c)b}{\gamma(1+\eta\lambda)} \qquad \text{requiring} \qquad w_c - \underline{w} > \frac{\alpha\eta\lambda^2 K(w_h)b^2}{\gamma(1+\eta\lambda)} + \underline{w} \tag{14}$$

Figure 6 is a more detailed version of Figure 3 that shows the curves defined by the RHS of these profit constraints, so that the inequalities hold with equality when $w_c = w_g$ and $w_c = w_l$ respectively, so that e_g is a valid local optimum when $w_c < w_g$ and e_l is a valid local optimum when $w_c > w_l$.

The worker must then compare these options, when they exist, to the kink in his utility function. The utilities of all three options are derived similarly to the positive reciprocity case. Comparing e_g or e_l to $\underline{e} + (w_c - \underline{w})/b$, we find that the utility at the kink is never optimal, similarly to the demonstration in the proof of Proposition 1;

the difference between these propositions is that either e_g or e_l is always an available option in the negative surprise case, so that $\underline{e} + (w_c - \underline{w})/b$ is in fact never chosen.

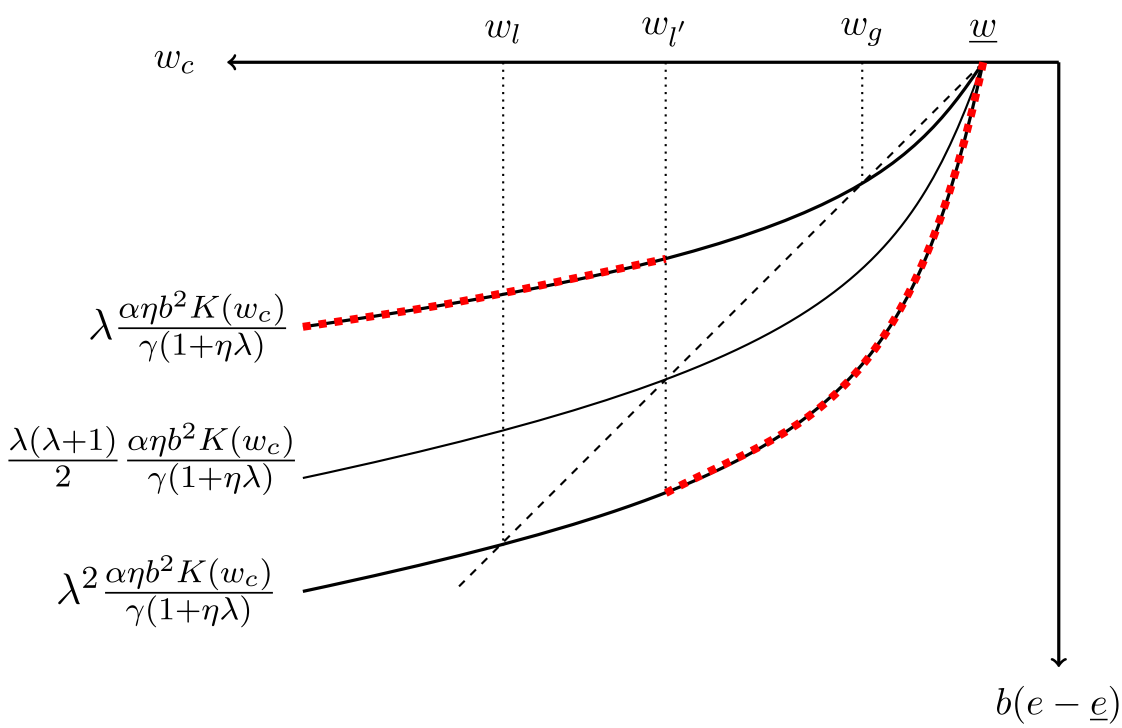
In the region between w_g and w_l where both e_l and e_g are valid optima, the worker prefers e_l to e_g only when

$$w_c - \underline{w} > \frac{\lambda(\lambda+1)}{2} \frac{\alpha\eta b^2 K(w_l)}{\gamma(1+\eta\lambda)}.$$

Define $w_{l'}$ as the value of w_c that makes this relationship hold with equality.

This relationship is a multiple of the revenue response curves that also determine the validity of the profit constraints above, so they are shown on Figure 6 together. By noticing that because $\lambda > 1$, $1 < \frac{\lambda+1}{2} < \lambda$, the aggregate set of conditions imply that e_l is chosen for $w_c > w_{l'}$, and e_g is chosen otherwise. Regardless, a surprising wage cut is negatively reciprocated.

Figure 6: Effort response function to fully surprising gift wage



Extreme or corner cases in which w_l and/or w_g are zero are straightforward to account for.

Proof of Proposition 4

As described in text.

Proof of Proposition 5

Continuing from the proof of Proposition 1, and referring to Figure 2, we can see that the range of profitable gifts, with $\underline{w} < w_h < w_g$, exists as long as the revenue response curve shown rises above the diagonal. That is, the slope of this curve at $\underline{w}$ must be greater than 1. This condition is equivalent to $K'(\underline{w}) > \frac{\gamma(1+\eta\lambda)}{\alpha\eta b^2}$, proving the first part of the proposition statement.

Given that a profitable gift exists, the optimal gift is the one that maximizes additional profits, which is the point where the revenue response curve is farthest above the wage cost diagonal. The firm's profits are

$$be_g - w_h = \frac{\alpha\eta b^2 K(w_h)}{\gamma(1+\eta\lambda)}$$

which has an FOC equivalent to the stated implicit definition of w_h^*.

Proof of Corollary 1

Existence of a profitable gift occurs when the condition stated in the proof of Proposition 5 holds, and this inequality trivially behaves as stated.

Implicitly differentiating the FOC from proposition 5 gives us

$$\frac{\partial w_h^*}{\partial \gamma} = \frac{1+\eta\lambda}{\alpha\eta b^2 K''(w_h)} < 0$$

$$\frac{\partial w_h^*}{\partial \lambda} = \frac{\gamma\eta}{\alpha\eta b^2 K''(w_h)} < 0$$

$$\frac{\partial w_h^*}{\partial \alpha} = \frac{-\gamma(1+\eta\lambda)}{\alpha^2\eta b^2 K''(w_h)} > 0$$

$$\frac{\partial w_h^*}{\partial b} = \frac{-2\gamma(1+\eta\lambda)}{\alpha\eta b^3 K''(w_h)} > 0$$

$$\frac{\partial w_h^*}{\partial \eta} = \frac{-\gamma}{\alpha\eta^2 b^2 K''(w_h)} > 0$$

The envelope theorem gives us:

$$\frac{\partial \Pi^*(w_h)}{\partial \gamma} = -\frac{\alpha\eta b^2}{\gamma(1+\eta\lambda)} K(w_h^*) < 0$$

$$\frac{\partial \Pi^*(w_h)}{\partial \lambda} = -\frac{\alpha\eta^2 b^2 K(w_h^*)}{\gamma(1+\eta\lambda)^2} < 0$$

$$\frac{\partial \Pi^*(w_h)}{\partial \alpha} = \frac{\eta b^2 K(w_h^*)}{\gamma(1+\eta\lambda)} > 0$$

$$\frac{\partial \Pi^*(w_h)}{\partial b} = \underline{e} + \frac{2\alpha\eta b K(w_h^*)}{\gamma(1+\eta\lambda)} > 0$$

$$\frac{\partial \Pi^*(w_h)}{\partial \eta} = \frac{\alpha b^2 K(w_h^*)}{\gamma(1+\eta\lambda)^2} > 0$$

The size of the effort response to the optimal gift is $e_h - \underline{e} = \frac{\alpha\eta b^2 K(w_h^*)}{\gamma(1+\eta\lambda)}$, which when partially differentiated using the partial derivatives of w_h^* calculated above, similarly yields the stated comparative statics. The response to a particular gift is similar.

Proof of Proposition 6

In the text.

Proof of Lemma 2

This can be trivially seen by comparing the utility in the high wage state, given any effort plan, of $e_h \leq \underline{e}$ versus $e_h = \underline{e} + \epsilon$ for some small $\epsilon > 0$. Similarly, $e_l < \underline{e}$.

Proof of Lemma 3

The worker formulates a plan $(\tilde{e}_h, \tilde{e}_l)$ and finds out the true wage w_h or w_l. He then chooses effort level e_h or e_l by maximizing his utility given his plan. In the high wage state he maximizes

$$\begin{aligned}
U(e_h|w_h, \tilde{e}_l, \tilde{e}_h) &= w_h - c(e_h) \\
&\quad + p\eta\mu(-c(e_h) + c(\tilde{e}_h)) + (1-p)\eta\mu(-c(e_h) + c(\tilde{e}_l)) \\
&\quad + (1-p)\alpha\eta\mu(K(w_h) - K(w_l))\mu(be_h - w_h - b\tilde{e}_l + w_l) \\
&= -\tfrac{\gamma}{2}(1 + (1-p)\eta\mu'_{chl} + p\eta\mu'_{chh})(e_h - \underline{e})^2 + (1-p)\eta\mu'_{chl}c(\tilde{e}_l) + p\eta\mu'_{chh}c(\tilde{e_h}) \\
&\quad + (1-p)\alpha\eta\Delta K\mu'_{\pi hl}(be_h - w_h - b\tilde{e}_l + w_l)
\end{aligned}$$

where

$$\mu'_{chl} = \mu'(-c(e_h) + c(\tilde{e}_l)) = \begin{cases} 1 & \text{if } c(e_h) > c(\tilde{e}_l) \\ \lambda & \text{otherwise} \end{cases}$$

and μ'_{chh} and $\mu'_{\pi hl}$ are defined similarly, and $\Delta K = K(w_h) - K(w_l)$; note that because $K(w_h) - K(w_h) = 0$, the term for reference dependent reciprocity relative to the expected high wage state is zero in this utility function.

In the low wage state, the worker maximizes

$$\begin{aligned} U(e_l, |w_l, \tilde{e}_l, \tilde{e}_h) \quad &= w_l - c(e_l) \\ &\quad + p\eta\mu(-c(e_l) + c(\tilde{e}_h)) + (1-p)\eta\mu(-c(e_l) + c(\tilde{e}_l)) \\ &\quad + p\alpha\eta\mu(K(w_l) - K(w_h))\mu(be_l - w_l - b\tilde{e}_h + w_h) \\ &= -\tfrac{\gamma}{2}(1 + (1-p)\eta\mu'_{cll} + p\eta\mu'_{clh})(e_l - \underline{e})^2 + (1-p)\eta\mu'_{cll}c(\tilde{e}_l) + p\eta\mu'_{clh}c(\tilde{e_h}) \\ &\quad - p\alpha\eta\lambda\Delta K\mu'_{\pi lh}(be_l - w_l - b\tilde{e}_h + w_h) \end{aligned}$$

with $\mu'_{\cdot}$ defined analogously to μ'_{chl} above.

The first order conditions in the high and low wage states respectively are

$$U'(e_h|w_h, \tilde{e}_l, \tilde{e}_h) = -\gamma(1 + (1-p)\eta\mu'_{chl} + p\eta\mu'_{chh})(e_h - \underline{e}) + (1-p)\alpha\eta b\Delta K\mu'_{\pi hl}$$

and

$$U'(e_l|w_l, \tilde{e}_l, \tilde{e}_h) = -\gamma(1 + (1-p)\eta\mu'_{cll} + p\eta\mu'_{clh})(e_l - \underline{e}) - p\alpha\eta\lambda b\Delta K\mu'_{\pi lh}$$

which are of course only defined where $\mu'_{\cdot}$ is defined.

To determine whether an effort level at a kink in the utility function(s) is part of a valid plan, we can consider how the utility function(s) conditional on planning to exert this effort looks and determine whether following through on that plan is at least a local optimum, with U' switching from positive to negative at that kink. We can, in this way, show that planning to exert effort such that $\pi(e_h, w_h) = \pi(e_l, w_l)$ is never a valid plan.

Assume towards contradiction that such a plan exists, denoted by $(\tilde{e}_l, \tilde{e}_h)$, and define $\mu'_{cl} = \mu'(-c(\tilde{e}_h) + c(\tilde{e}_l))$ and $\mu'_{ch} = \mu'(-c(\tilde{e}_l) + c(\tilde{e}_h))$. In the low wage state, $U'(e_l)$ just left of the discontinuity is given by

$$\lim_{e_l \to^- \tilde{e}_l} U'(e_l|w_l, \tilde{e}_l, \tilde{e}_h) = -\gamma(1 + p\eta\mu'_{ch} + \lambda\eta(1-p))(\tilde{e}_l - \underline{e}) - p\alpha\eta\lambda^2 b\Delta K,$$

relying on the fact, for any $\epsilon > 0$, $\pi(\tilde{e}_l - \epsilon, w_l) < \pi(\tilde{e}_l, w_l) = \pi(\tilde{e}_h, w_h)$ by hypothesis and $c(\tilde{e}_l - \epsilon) > c(\tilde{e}_l)$ by Lemma 2. Similarly, just to the right of the kink the derivative is given by

$$\lim_{e_l \to^+ \tilde{e}_l} U'(e_l|w_l, \tilde{e}_l, \tilde{e}_h) = -\gamma(1 + p\eta\mu'_{ch} + \eta(1-p))(\tilde{e}_l - \underline{e}) - p\alpha\eta\lambda b\Delta K.$$

In order for $U(e_l|w_l, \tilde{e}_l, \tilde{e}_h)$ to have a local optimum at $\tilde{e}_l$ we thus need $p\alpha\eta\lambda^2 b\Delta K \leq -\gamma(1+p\eta\mu'_{ch}+\lambda\eta(1-p))(\tilde{e}_l - \underline{e})$ and $p\alpha\eta\lambda b\Delta K \geq -\gamma(1+p\eta\mu'_{ch}+\eta(1-p))(\tilde{e}_l-\underline{e})$. This requires that $\lambda(1+p\eta\mu'_{ch}+\eta(1-p)) \leq (1+p\eta\mu'_{ch}+\lambda\eta(1-p))$, or $\lambda-1 \leq (1-\lambda)p\eta\mu'_{ch}$. But since $\lambda > 1$, the LHS is positive and the RHS is negative, so this statement is never true, and $(\tilde{e}_l, \tilde{e}_h)$ is never a PPE.

Proof of Lemma 4

As in the proof of Lemma 3, assume towards contradiction that a personal equilibrium plan exists, denoted by $(\tilde{e}_l, \tilde{e}_h)$, such that $c(\tilde{e}_l) = c(\tilde{e}_h)$. For this to be true, $U(e_l|w_l, \tilde{e}_l, \tilde{e}_h)$ must have a local optimum at $\tilde{e}_l$ and $U(e_h|w_h, \tilde{e}_l, \tilde{e}_h)$ must have a local optimum at $\tilde{e}_h$. In the former low wage case, the value of U' just to the left of the kink at $\tilde{e}_l$ is given by

$$\lim_{e_l \to^- \tilde{e}_l} U'(e_l|w_l, \tilde{e}_l, \tilde{e}_h) = -\gamma(1+\eta\lambda p+\eta\lambda(1-p))(\tilde{e}_l-\underline{e}) - p\alpha\eta\lambda b\Delta K\mu'_{\pi h},$$

and the value of U' just to the right of the kink is

$$\lim_{e_l \to^+ \tilde{e}_l} U'(e_l|w_l, \tilde{e}_l, \tilde{e}_h) = -\gamma(1+\eta p+\eta(1-p))(\tilde{e}_l-\underline{e}) - p\alpha\eta\lambda b\Delta K\mu'_{\pi h},$$

where both expressions rely on Lemma 2. Putting these together, for $\tilde{e}_l$ to be a local optimum of the utility function after the low wage is realized, it is necessary that $-(1+\eta)\gamma(\tilde{e}_l-\underline{e}) \leq p\alpha\eta\lambda b\Delta K\mu'_{\pi h} \leq -(1+\eta\lambda)(\tilde{e}_l-\underline{e})$. Additionally, because $c(\tilde{e}_h) = c(\tilde{e}_l) \Rightarrow \tilde{e}_l = 2\underline{e} - \tilde{e}_h$, we can rewrite this as $(1+\eta)\gamma(\tilde{e}_h-\underline{e}) \leq p\alpha\eta\lambda b\Delta K\mu'_{\pi h} \leq (1+\eta\lambda)\gamma(\tilde{e}_h-\underline{e})$.

In the latter high wage case we similarly have

$$\lim_{e_h \to^- \tilde{e}_h} U'(e_h|w_h, \tilde{e}_l, \tilde{e}_h) = -\gamma(1+\eta p+\eta\mu'_{cl}(1-p))(\tilde{e}_h-\underline{e}) + (1-p)\alpha\eta b\Delta K\mu'_{\pi l}$$

and

$$\lim_{e_h \to^+ \tilde{e}_h} U'(e_h|w_h, \tilde{e}_l, \tilde{e}_h) = -\gamma(1+\eta\lambda p+\eta\mu'_{cl}(1-p))(\tilde{e}_h-\underline{e}) + (1-p)\alpha\eta b\Delta K\mu'_{\pi l}$$

which together mean that for $\tilde{e}_h$ to be a potential component of a personal equilibrium, it's necessary that $(1+\eta)\gamma(\tilde{e}_h-\underline{e}) \leq (1-p)\alpha\eta b\Delta K\mu'_{\pi l} \leq (1+\eta\lambda)\gamma(\tilde{e}_h-\underline{e})$.

We can combine these two sets of inequalities, written as

$$(1+\eta)\gamma(\tilde{e}_h-\underline{e}) \leq \{(1-p)\alpha\eta b\Delta K\mu'_{\pi l},\ p\alpha\eta\lambda b\Delta K\mu'_{\pi h}\} \leq (1+\eta\lambda)\gamma(\tilde{e}_h-\underline{e}).$$

Note that the difference between the LHS and RHS is a factor of $(1+\eta\lambda)/(1+\eta)$, and the difference between the two intermediate values is a factor of $\frac{p\lambda}{1-p}\frac{\mu'_{\pi h}}{\mu'_{\pi l}}$. For this set of inequalities to be satisfiable, we therefore require that the latter quantity, along with its reciprocal, are both less than $(1+\eta\lambda)/(1+\eta)$.

We must break this into two cases: First, if $\pi(e_h, w_h) < \pi(e_l, w_l)$ then $\mu'_{\pi l} = \lambda$ and $\mu'_{\pi h} = 1$. But we can show that even if the worker would respond to a stochastic contract with a effort (e_l, e_h) such that $c(e_h) = c(e_l)$ and $\pi(e_h, w_h) < \pi(e_l, w_l)$, the firm would never want to offer such a contract. To see this, recall that the firm's participation constraint is that $p\pi(e_h, w_h) + (1-p)\pi(e_l, w_l) > b\underline{e} - \underline{w}$. This becomes $pbe_h + (1-p)be_l - b\underline{e} > pw_h + (1-p)w_l - \underline{w}$, which is equivalent to $(2p-1)b(e_h - \underline{e}) > pw_h + (1-p)w_l - \underline{w}$ when we restrict $e_l = 2\underline{e} - e_h$. Then, note that the condition for profits to be higher in the low wage state is equivalent to $w_h - w_l > 2b(e_h - \underline{e})$ (again using $e_l = 2\underline{e} - e_h$), which we can substitute into the participation constraint to get the necessary condition that $(2p-1)b(e_h - \underline{e}) > 2pb(e_h - \underline{e}) + w_l - \underline{w} \Leftrightarrow b\underline{e} - w_l > be_h - \underline{w}$. But by Lemma 2 and the assumption that $w_l \geq \underline{w}$, this can never be true.

So we are left with the possibility that the worker responds with $c(e_l) = c(e_h)$ and $\pi(e_h, w_h) > \pi(e_l, w_l)$. In this case, $\mu'_{\pi l} = 1$, $\mu'_{\pi h} = \lambda$, and the condition for this plan to in fact be a local optimum after the wage state is realized is that $\frac{1+\eta\lambda}{1+\eta} \geq \frac{p\lambda^2}{1-p}$ and $\frac{1+\eta\lambda}{1+\eta} \geq \frac{1-p}{p\lambda^2}$. The former is equivalent to $p \leq \frac{1+\eta\lambda}{1+\eta\lambda+(1+\eta)\lambda^2}$. Since the RHS is decreasing in λ and $\lambda > 1$, this requires that $p < \frac{1}{2}$.

But, returning to the firm's participation constraint, we have $(2p-1)b(e_h - \underline{e}) > pw_h + (1-p)w_l - \underline{w}$. Because the RHS is positive (since $w_h > w_l \geq \underline{w}$), and $e_h > \underline{e}$ by Lemma 2, this only holds if $p > 1/2$, contradicting the requirement that $p < 1/2$ for the worker to respond with these effort levels.

Proof of Proposition 7

Referring back to the proof of Lemma 3, the marginal utility functions for the worker upon realizing the wage state are

$$U'(e_h|w_h, \tilde{e}_l, \tilde{e}_h) = -\gamma(1 + (1-p)\eta\mu'_{chl} + p\eta\mu'_{chh})(e_h - \underline{e}) + (1-p)\alpha\eta b\Delta K\mu'_{\pi hl}$$

and

$$U'(e_l|w_l, \tilde{e}_l, \tilde{e}_h) = -\gamma(1 + (1-p)\eta\mu'_{cll} + p\eta\mu'_{clh})(e_l - \underline{e}) - p\alpha\eta\lambda b\Delta K\mu'_{\pi lh}$$

which are defined anywhere where μ' is defined. Lemmas 3 and 4 prove that the PPE can never occur at the kinks in these utility functions where $\pi(e_l, w_l) = \pi(e_h, w_h)$ or $c(e_h) = c(e_l)$, but by the definition of a PPE, it *must* occur at the kinks where $e_h = \tilde{e}_h$ and $e_l = \tilde{e}_l$. The workers utility functions given the wage state and effort contingency plan therefore must have local maxima at these points, which means the marginal utility functions must change signs accordingly. Any PPE must therefore have the following characteristics:

$$\lim_{e_h \to^- \tilde{e}_h} U'(e_h|w_h, \tilde{e}_l, \tilde{e}_h) = -\gamma(1 + \eta p + \eta\mu'_{cl}(1-p))(\tilde{e}_h - \underline{e}) + (1-p)\alpha\eta b\Delta K \mu'_{\pi l} \geq 0$$

$$\lim_{e_h \to^+ \tilde{e}_h} U'(e_h|w_h, \tilde{e}_l, \tilde{e}_h) = -\gamma(1 + \eta\lambda p + \eta\mu'_{cl}(1-p))(\tilde{e}_h - \underline{e}) + (1-p)\alpha\eta b\Delta K \mu'_{\pi l} \leq 0$$

$$\lim_{e_l \to^- \tilde{e}_l} U'(e_l|w_l, \tilde{e}_l, \tilde{e}_h) = -\gamma(1 + \eta\lambda p + \eta\lambda(1-p))(\tilde{e}_l - \underline{e}) - p\alpha\eta\lambda b\Delta K \mu'_{\pi h} \geq 0$$

$$\lim_{e_l \to^+ \tilde{e}_l} U'(e_l|w_l, \tilde{e}_l, \tilde{e}_h) = -\gamma(1 + \eta p + \eta(1-p))(\tilde{e}_l - \underline{e}) - p\alpha\eta\lambda b\Delta K \mu'_{\pi h} \leq 0,$$

which reduce to the following two constraints on the equilibrium effort and effort plan:

$$\frac{\eta\alpha b\Delta K \mu'_{\pi l}(1-p)}{\gamma(1 + \eta\mu'_l(1-p) + \eta\lambda p)} \leq \tilde{e}_h - \underline{e} \leq \frac{\eta\alpha b\Delta K \mu'_{\pi l}(1-p)}{\gamma(1 + \eta\mu'_l(1-p) + \eta p)}$$

$$\frac{\eta\lambda\alpha b\Delta K \mu'_{\pi h} p}{\gamma(1 + \eta\mu'_h p + \eta\lambda(1-p))} \leq -(\tilde{e}_l - \underline{e}) \leq \frac{\eta\lambda\alpha b\Delta K \mu'_{\pi h} p}{\gamma(1 + \eta\mu'_h p + \eta(1-p))}.$$

We can break these constraints into four cases, with $\pi(\tilde{e}_l, w_l) \lessgtr \pi(\tilde{e}_h, w_h)$ and $c(\tilde{e}_h) \lessgtr c(\tilde{e}_l)$. The two cases in which $\pi(\tilde{e}_l, w_l) > \pi(\tilde{e}_h, w_h)$ can be dispensed with easily, however: If this were the case, the firm would certainly not with to offer this contract, as they would earn greater profits by offering $\underline{w}$ for sure than in either wage state in the stochastic contract, by Lemma 2. The remaining two cases are considered in turn.

Case 1: $c(\tilde{e}_h) > c(\tilde{e}_l)$. In this case, the range of possible values for effort levels above become respectively

$$\frac{\eta\alpha b\Delta K(1-p)}{\gamma(1+\eta\lambda)} \leq \tilde{e}_h - \underline{e} \leq \frac{\eta\alpha b\Delta K(1-p)}{\gamma(1 + \eta\lambda(1-p) + \eta p)}$$

$$\frac{\eta\lambda^2\alpha b\Delta K p}{\gamma(1 + \eta p + \eta\lambda(1-p))} \leq -(\tilde{e}_l - \underline{e}) \leq \frac{\eta\lambda^2\alpha b\Delta K p}{\gamma(1+\eta)}.$$

Will a firm willingly offer a stochastic contract that elicits effort in these ranges? In the best case scenario, the condition that the firm's profits exceed their profits when offering the market wage with certainty becomes

$$\begin{array}{rrcl} & p(b\tilde{e}_h - w_h) + (1-p)(b\tilde{e}_l - w_l) & > & b\underline{e} - \underline{w} \\ \Rightarrow & \frac{\eta\alpha b^2\Delta K(1-p)p}{\gamma(1+\eta\lambda(1-p)+\eta p)} - \frac{\eta\lambda^2\alpha b^2\Delta K(1-p)p}{\gamma(1+\eta p+\eta\lambda(1-p))} - pw_h - (1-p)w_l & > & -\underline{w} \\ \Rightarrow & \frac{\eta\alpha b^2\Delta K(1-p)p}{\gamma}\left(\frac{1-\lambda^2}{1+\eta\lambda(1-p)+\eta p}\right) & > & pw_h + (1-p)w_l - \underline{w} \end{array} \tag{15}$$

Because the RHS and the first term on the LHS are both strictly positive, and because the denominator in the parentheses is positive and the numerator is negative, this inequality never holds, so the firm will never want to offer a contract that elicits these effort levels.

Case 2: $c(\tilde{e}_l) > c(\tilde{e}_h)$. In this case, the range of possible effort values are respectively

$$\frac{\eta\alpha b\Delta K(1-p)}{\gamma(1+\eta(1-p)+\eta\lambda p)} \leq \tilde{e}_h - \underline{e} \leq \frac{\eta\alpha b\Delta K(1-p)}{\gamma(1+\eta)}$$

$$\frac{\eta\lambda^2\alpha b\Delta K p}{\gamma(1+\eta\lambda)} \leq -(\tilde{e}_l - \underline{e}) \leq \frac{\eta\lambda^2\alpha b\Delta K p}{\gamma(1+\eta\lambda p+\eta(1-p))}.$$

The best chance the firm has to meet their participation constraint is then

$$\begin{aligned} p(b\tilde{e}_h - w_h) + (1-p)(b\tilde{e}_l - w_l) &> b\underline{e} - \underline{w} \\ \Rightarrow \quad \tfrac{\eta\alpha b^2\Delta K(1-p)p}{\gamma(1+\eta)} - \tfrac{\eta\lambda^2\alpha b^2\Delta K(1-p)p}{\gamma(1+\eta\lambda)} - pw_h - (1-p)w_l &> -\underline{w} \\ \Rightarrow \quad \tfrac{\eta\alpha b^2\Delta K(1-p)p}{\gamma}\left(\tfrac{1}{1+\eta} - \tfrac{\lambda^2}{1+\eta\lambda}\right) &> pw_h + (1-p)w_l - \underline{w} \end{aligned} \tag{16}$$

Because the RHS and first term on the LHS are positive, a necessary condition for this to hold is that

$$\begin{aligned} \tfrac{1}{1+\eta} &> \tfrac{\lambda^2}{1+\eta\lambda} \\ \Rightarrow \quad 0 &> \lambda^2(\eta+1) - \eta\lambda - 1 \end{aligned}$$

Because the RHS is increasing in λ and $lambda > 1$, the RHS is actually > 0, a contradiction. So the firm never wishes to offer a contract that elicits these effort levels either.

Proof of Corollary 2

An initial (profitable) gift reveals the inability of the firm to commit to paying $\underline{w}$, along with their desire to pay a higher wage if expectations of p are low enough. This prevents expectations from meeting $p = 0$, and any higher p leads to non-profitable stochastic gift exchange repeated ad infinitum. The initial boost to profits is thus clearly only worthwhile if the infinite stream of losses is sufficiently discounted.

Proof of Proposition 8

After a surprise gift, which reveals the inability of the employer to commit to a wage, the employee must infer a non-zero probability of a further gift: if they did infer a zero probability, the profitability of the first gift implies that a second fully surprising gift would also be profitable, contradicting the inference that the employer would not want to give a further gift.

As we saw in the proof of lemma 3, any stochastic contract (w_l, w_h, p) will result in unequal wages in the two wage states, causing the firm to strictly prefer the higher profit wage as long as the game is ending after this interaction so that there is no incentive to maintain the worker's expectations about future wages. But of course in a rational expectations equilibrium, the worker will anticipate this and form an extreme belief with $p = 0$ or $p = 1$.

Since $p = 0$ would immediately renew the possibility of profitable gift exchange, which the firm would certainly want to take, the only remaining possibility for a rational expectations equilibrium is that workers will expect a gift wage with probability 1. We must therefore find which gifts satisfy the firm's PPE conditions. That is, which wage can the firm plan to give that they will not wish to deviate from? This is equivalent to asking: what is the lowest wage from which there is no profitable fully surprising gift? As in the proof of Proposition 5, no profitable gift exists, starting from a base wage of w, if $K'(w) \leq \frac{\gamma(1+\eta\lambda)}{\alpha\eta b^2}$. Since K is concave, the lowest such wage is the one for which this relationship holds with equality. But the w satisfying this equation is exactly w_h^*, the optimal fully surprising gift.

Is there additionally no profitable deviation in the other direction? That is, if w_h^* is anticipated fully but w with $\underline{w} < w < w_h^*$ is actually paid, will this ever lead to higher profits than fulfilling the expectation of w_h^*? It turns out this is never true. To see this, similarly to the proof of Proposition 3, the worker might reciprocate by either keeping profits equal to expectations (which would not constitute a profitable deviation), or by exerting effort $\underline{e}$ as planned (which is easy to show is never preferred to slightly negatively reciprocating), or by negatively reciprocating in part or in full. The optimal effort levels for the latter two are, as before, given by

$$e_g - \underline{e} = \frac{\alpha\eta\lambda(K(w) - K(w_h^*)b}{\gamma(1+\eta\lambda)}$$

and

$$e_l - \underline{e} = \frac{\alpha\eta\lambda^2(K(w) - K(w_h^*)b}{\gamma(1+\eta\lambda)}.$$

Unlike in Proposition 3 though, since K is concave at the base wage w_h^*, the response of e_g is only a valid optimum, if at all, for small cuts relative to w_h^*. As in Proposition 5, it is never a valid optimum if $\frac{\alpha\eta\lambda K'(w_h^*)b}{\gamma(1+\eta\lambda)} > 1$. But by definition of w_h^*, this quantity is equal to $\lambda > 1$, so it is in fact never an optimal response. The firm can therefore never gain by choosing $\underline{w} < w < w_h^*$, and w_h^* is therefore the firm's PPE.

By backwards induction, the same occurs in every earlier period as well. And so, if the firm chooses the optimal fully surprising gift w_h^*, it must choose the same wage in every further period, turning it into a permanent raise.