On the Power of Surprising Versus Anticipated Gifts*

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

August 5, 2017

Abstract

We study an important but largely neglected aspect of gift exchange: the surprising or anticipated nature of the wage raise. By varying the awareness agents have about possible upcoming wage increases in an expectations-based reference-dependent model of reciprocity, we deliver four main messages about the efficacy and efficiency of gift exchange: (1) gifts are more powerful when they fully surprise workers; (2) the power of fully surprising gifts, however, wanes over time; (3) gifts are cursed, as once a firm grants a fully surprising one, it should grant it forever; and therefore (4) gifts are most likely profitable in short-term interactions. We relate these predictions to the existing evidence and study the model's recommendations for the design of field experiments testing for gift exchange.

JEL Classification: D030, J31, J33, M52

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

*We thank Constanca Esteves-Sorenson for some of the quotes used in this paper. We also thank Botond Kőszegi and Stefano DellaVigna, and seminar participants at the Australian Economic Theory Workshop, Western Economic Association International annual meetings, University of Technology Sydney, University of New South Wales, Monash University, University of Sydney, and Technische Universität Berlin for helpful comments and suggestions. Rosario Macera specially thanks Fondecyt Grant Iniciación #11140101 for funding.

[†]rmacera@uandes.cl

[‡]v.tevelde@uq.edu.au.

1 Introduction

Whether paying above-market wages is an efficient mechanism to elicit employee effort is an important economic question. As proposed by Akerlof (1982), if workers interpret an above-market wage as a gift and thus reciprocate with above-minimum effort, then firms can rely on above-market fixed wages to motivate workers. Gift exchange, therefore, could save firms from the onerous costs of designing and implementing performance-contingent schemes, especially in the common case where output is difficult to measure or contract on.

This paper advances the incentives literature by theoretically studying an until recently overlooked aspect of monetary gifts: their surprising or anticipated nature. To this end, we develop a model of reference-dependent reciprocity where expected outcomes constitute the reference point. The idea that expectations about an upcoming outcome determine the response to the actual outcome dates back to Vroom (1964) but it was popularized by Kőszegi & Rabin (2006, 2007, 2009) who embedded expectations in a model of reference-dependent preferences with loss aversion as in Kahneman & Tversky (1979).[1] Our model builds on Kőszegi & Rabin (2006) by adding loss aversion into the agent's preferences, but crucially departs from it by varying the reciprocal agent's awareness about potential upcoming wage raises. This allows us to explore the implications of monetary gifts, ranging from fully surprising to partially and totally anticipated, on the agent's effort response and the profitability of the wage raise.

Our model delivers four novel messages on the efficacy and efficiency of monetary gifts in one-shot and repeated principal-agent interactions,

(1) The power of gifts is greater when they fully surprise workers. Because fully surprising gifts trigger a pleasing departure from payment expectations while anticipated gifts do not, they elicit a larger, and possibly profitable, effort response. Similarly, surprising wage cuts lead to greater effort decreases than anticipated cuts.

(2) The power of a fully surprising gift, however, wanes over time. Because workers update their payment expectations, effort returns to its baseline level and a permanent fixed wage increase loses power over time.

(3) Gifts are cursed: firms should never grant them or grant them forever. An 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 further gifts, agents will retaliate by exerting lower effort—relative to not receiving or expecting any gift—when these expectations are unfulfilled. We show that this retaliation of expected but unfulfilled gifts always outweighs the benefits from the sporadic positive surprises, rendering stochastic gifts unprofitable relative to committing to the market wage. Furthermore, in shorter-term interactions we show

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

that, if a firm grants a surprising gift, she always prefers to turn it into permanent wage raise, even though the raise is only able to increase effort one time, immediately after it is granted.

(4) Even if a gift is fully surprising, it is only profitable in short interactions. When deciding whether to grant a fully surprising gift, the firm must trade off its immediate benefits with the delayed costs of revealing the firm's inability to commit to the market wage, which forces her to grant unprofitable stochastic gifts thereafter. Gifts, therefore, are more likely to be profitable in shorter interactions where the benefits of the surprise are more likely outweighed by the following costs.

This paper makes four main contributions. First, to the best of our knowledge, we are first to show that the profitability of gift exchange can depend on the surprising versus anticipated nature of the monetary gift. Our model thus relates to others theoretically 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) showed that gift exchange can arise even with low wages if firms can couple low wages with attention, a valued resource for workers. Benjamin (2015) studies a gift exchange game where the worker has fairness concerns to show that the model rationalizes several types of wage rigidities. In a different vein, Netzer & Schmutzler (2014) showed that gift exchange does not arise in equilibrium if agents have intentions-based reciprocal preferences as they can not interpret profitable wage increases as kind. Unlike these papers, ours studies the efficiency of gift exchange by focusing on the properties of the gift itself rather than on the agents' preferences or the economic environment.[2]

Our paper also relates to a few others studying the importance of surprises in decision making. In a psychological game-theoretic model, Ruffle (1999) shows that whenever players use mixed strategies, emotions such as surprise, pride, embarrassment, and disappointment can arise in the equilibrium of a gift giving game where agents compare actual actions with their first-order beliefs. Khalmetski, Ockenfels & Werner (2015) model surprising gifts in the context of a dictator game where agents experience guilt aversion in the spirit of Battigalli & Dufwenberg (2007). They show that dictator transfers can both decrease and increase with recipients' expectations, depending on the weight put on positive and negative surprises, respectively. Unlike these papers, however, we specifically analyze the profitability of gift exchange by focusing on whether the gift was fully surprising or anticipated in one-shot and repeated principal-agent interactions.

Besides improving our knowledge of the profitability of gift exchange, this paper also sheds light on the

[2]To the best of our knowledge, there are only a handful of papers focusing on the properties of the gift. For instance, Kube, Maréchal & Puppe (2012) show that gifts in kind trigger a reciprocal effort response, while monetary gifts do not. Others are Gilchrist, Luca & Malhotra (2016) and Sliwka & Werner (2017), which we review in detail in Section 7. The literature focusing on agent preferences or the economic environment, however, is more developed. For instance Brandts et al. (2010) and Schram, Brandts & Gërxhani (2010) focus on the impact of the structure and size of the market; Charness (2004) studies the characteristics of who grants the gift; Hennig-Schmidt, Rockenbach & Sadrieh (2010), Englmaier & Leider (2012*a*) and DellaVigna et al. (2016) focus on the agent's information about the firm's surplus or others in his ability to repay the gift; while Chaudhuri & Sbai (2011), and Hannan, Kagel & Moser (2002) focus on the demographics of the recipient.

design of the standard field experiment testing for gift exchange. To avoid selection of abler workers, treated workers are hired at the market wage and then fully surprised with a wage increase right before they work for the first time. Our theory suggests, however, that the fully surprising nature of the gift can trigger a transitorily higher effort response than that of workers in real-world firms where above-market wages are commonly used. Moreover, our theory further highlights that such deviations can lead to future profit loses if the deviation from the initial agreement leads workers to expect additional wage raises that will not occur. These results highlight that future tests of gift exchange in the field should be careful about unintended effects of the treatment (a surprising wage raise) on beliefs, which can can create short as well as long-term effects muddling the estimation of gift exchange profitability.

Third, in addition to contributing to our understanding of the profitability of gift exchange and the design of the standard gift exchange experiment, this paper also contributes to the literature on expectations-based reference-dependent preferences by proposing a novel model of reference-dependent social preferences. Reference-dependent preferences have been shown to be relevant to economic behavior in a large array of domains, from financial decisions, insurance, saving, pricing, labor supply, etc. (see DellaVigna (2009), Barberis (2013) and Kőszegi (2014) 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) present laboratory evidence on the role of expectations as reference points, while Card & Giuliano (2013), Crawford & Meng (2011), Pope, Price & Wolfers (2011), and Lien, Peng & Zheng (2016) provide empirical evidence.[3] We argue that reference-dependent preferences can naturally be combined with social preferences and that this combination can improve our understanding of gift exchange, both in field experiments and real firms.

Our fourth contribution is to emphasize that a crucial aspect of the architecture of an incentive scheme is the effect of the scheme on expectations. Because gifts are most powerful when fully surprising and most damaging when expected but unfulfilled, firms should place as much attention on how an incentive is communicated as to its actual payments, since communication is crucial for shaping workers' expectations about future payments.

This paper proceeds as follows. Section 2 presents the model. Sections 3 to 6 formalize the four messages and provide some further insights about the optimal size of gifts. Section 7 relates our messages to the empirical evidence on gift exchange. Finally, Section 8 concludes by discussing some lessons extracted from our model predictions as well as avenues for future research.

[3] 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.

2 Model Set Up

(1) *Preferences.* A principal (the employer) hires an agent (the worker) to exert effort $e \in \mathbb{R}^+$ for a fixed wage $w \in \mathbb{R}^+$.

Assumption 1 *(Principal's preferences). The principal is assumed to be a risk neutral profit maximizer with no "behavioral" components. Her profit function is $\pi(e, w) = be - w$ where $b > 0$.*

The worker experiences utility from two sources: standard consumption utility from material outcomes (Assumption 2) and reference-dependent utility or gain-loss utility from comparing actual outcomes with a reference point (Assumption 3). Define $\underline{w}$ and w_h as the market wage and an above market-clearing wage, respectively, where $w_h > \underline{w} > 0$.

Assumption 2 *(Consumption Utility $m(\cdot)$). For a given wage w,*

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

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

(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, as stated in part (i), is an immaterial simplification. Part (ii) defines the market wage $\underline{w}$ as the outside option needed to determine the participation constraints of the principal-agent interaction in Section 5. Part (iii) states that the cost of effort is increasing and convex, a standard assumption in principal-agent models. The further assumption of a quadratic effort cost is made for simplicity as it allows us to derive closed form solutions, which helps with understanding the mechanisms. Moreover, the assumption that the convex effort-cost function has a strictly positive interior minimum is also mathematically immaterial but made to aid with intuition. Finally, part (iv) assumes that consumption utility is additive, again as standard in principal-agent models.

Notice that $m(\cdot)$ in Assumption 2 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 qualitatively modify the effect of surprising gifts on reciprocal behavior. To focus on the effect of surprises, we develop the model without such a term and comment after the main propositions how the inclusion of baseline reciprocity would affect them.[4]

Assumption 3 describes the second component of the workers' preferences: gain-loss utility from comparing actual outcomes to a reference point. In our model the two relevant domains are effort and reciprocity, the latter being the novel component. Similarly to models of intentions-based social preferences such as Rabin (1993), Levine (1998), Dufwenberg & Kirchsteiger (2004), or Falk & Fischbacher (2006), gain-loss utility from reciprocity assumes the worker puts positive weight on the employer's profits if she has been unexpectedly kind and negative weight if she has been unexpectedly unkind.

Assumption 3 *(Gain-Loss Utility $n(\cdot|\cdot)$). For a given wage w,*

(i) *The firm's kindness $K(w)$ is strictly increasing, strictly concave, and it is normalized to zero at the market wage, $K(\underline{w}) = 0$.*

(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 piece-wise 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*

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

(iv) *Given a reference wage $\tilde{w}$, a reference effort $\tilde{e}$, 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 $n(e, w|\tilde{e}, \tilde{w}) = n_c(e|\tilde{e}) + n_k(w, e|\tilde{w}, \tilde{e})$.*

Part (i) assumes that the kindness function $K(\cdot)$ is strictly increasing and concave. This captures the intuition that the marginal utility of kindness is strictly decreasing but always positive, as usually assumed when modeling positive hedonic feelings. Normalizing the kindness function to zero at the market wage reflects the evidence

[4] Reciprocity has been shown to impact many different market interactions (see Malmendier, te Velde & Weber (2014) for a review), and substantial evidence exists that gift exchange is caused by social preferences 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 (Gross, Guo & Charness (2015), Charness & Kuhn (2007), Güth et al. (2001), Kocher, Luhan & Sutter (2012)), wage rigidity (Fehr & Falk (1999)), charitable giving (Koppel & Regner (2014)), deferred compensation (Huck, Seltzer & Wallace (2011)), group decision making (Kocher & Sutter (2007)), 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)).

that workers consider the market wage as the fair wage and thus neither kind nor unkind.[5] The assumption that the gain-loss utility function is piece-wise linear, as described in part (ii), is standard in applications of reference-dependent preferences and is made to highlight the role of loss aversion in the predictions.[6] Part (iii) and (iv) describe gain-loss utility in effort and reciprocity, the two relevant dimensions in a model with fixed wages.[7] In particular part (iii) describes gain-loss utility in effort where the worker uses the gain-loss utility function $\mu(\cdot)$ to compare the actual effort $e \in E$ against the reference effort $\tilde{e}$. Part (iv) describes gain-loss utility in reciprocity by assuming a multiplicative form for gain-loss utility in kindness and in the firm's profits.[8]

Finally, Assumption 4 defines the worker's total utility,

Assumption 4 *(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) *Reference point formation.* To complete the agent's gain-loss utility, we specify how the agent builds his effort and wage expectations.

2.1) *Effort expectations.* Given his wage expectations, we assume effort expectations are rational as in Kőszegi & Rabin (2006). Given an expected wage $\tilde{w}$, the rational agent forms an effort plan $\tilde{e}$ that he must be willing to implement once it constitutes his reference point. Definition 1 formally presents this effort plan in the simplest case where there is only one period and $\tilde{w}$ is deterministic,

Definition 1 *(The Agent's Preferred Personal Equilibrium (PPE) with deterministic wage expectations) Given a wage expectation $\tilde{w}$, an effort plan $\tilde{e} \in E$ corresponds to a preferred personal equilibrium (PPE) iff*

(i) $\tilde{e} \in \underset{e}{\operatorname{argmax}}\, U(e, \tilde{w}|\tilde{e}, \tilde{w})$ *and*

(ii) $\tilde{e} \in \underset{e \in E^*}{\operatorname{argmax}}\, U(e, \tilde{w}|e, \tilde{w})$

where $E^* = \{e \in E \mid e \text{ solves (i)}\}$.

Part i) states that, given $\tilde{w}$, the effort plan $\tilde{e}$ must be credible; that is, having planed it, the workers will want to follow through it. Part ii) states that, if there are multiple credible plans, then the worker chooses the

[5] Kahneman, Knetsch & Thaler (1986) present experimental evidence of this assumption.

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

[7] Whenever wages are fixed, gain-loss utility in the monetary dimension does not affect effort provision and thus, without loss of generality, we omit it from the analysis.

[8] This term departs from the natural extension of the Kőszegi & Rabin (2006) framework, $\alpha\eta\mu(K(w)(be-w)-K(\tilde{w})(b\tilde{e}-\tilde{w}))$. This functional form would make the counterintuitive prediction that a firm who is kind, but not as kind as expected, will *increase* worker effort as workers try to make up for loss in overall reciprocity. Another alternative would be $\alpha\eta\mu((K(w)-K(\tilde{w}))(b(e-\tilde{e})-(w-\tilde{w})))$, but this would predict that workers reward and punish wage deviations to an equal degree, whereas the psychology of loss aversion, along with the empirical evidence in Engelmann & Ortmann (2009) and Kube, Maréchal & Puppe (2013), indicate that workers are more sensitive to wage cuts than wage gains.

one providing the highest utility among all credible plans. This definition follows Kőszegi & Rabin (2006)'s preferred personal equilibrium concept.[9]

2.2) *Wage expectations.* We explore the worker's effort response and the gift profitability under two different assumptions about the formation of the wage expectation, $\tilde{w}$. To replicate beliefs in a standard gift exchange field experiment, Sections 3 and 4 assume the worker is fully surprised by the wage increase. Since, arguably, no real-world employer raises the wage the first day of work, workers in gift exchange field experiments—who are hired at the market wage to avoid selection of abler workers and whose wage is raised without warning before the first labor shift—expect to receive the market wage with certainty, as agreed upon recruitment.[10]

In Sections 5 and 6 we relax the assumption that gifts fully surprise the agent and build on the standard assumption that the agent foresees the possibility of a wage increase. In this case, agents incorporate the wage raise into his rational expectations about—the now possibly stochastic—wage expectation. We will show that this assumption produces starkly different predictions about effort responses and gift exchange profitability relative to the fully-surprising gift assumption, which thus has strong implications for how gift exchange experiments should deal with belief formation.

(3) *Timing.* The number of periods is specified in each section, but the general timing for a given period is as follows. The principal announces a (possibly stochastic) wage. Then the worker forms expectations about the wage he is going to receive, $\tilde{w}$, and forms a credible effort plan, $\tilde{e}$. Then the actual wage is revealed, and the worker immediately exerts his effort for the period.

3 The Power of Gifts is Greater Whenever They Are Fully Surprising

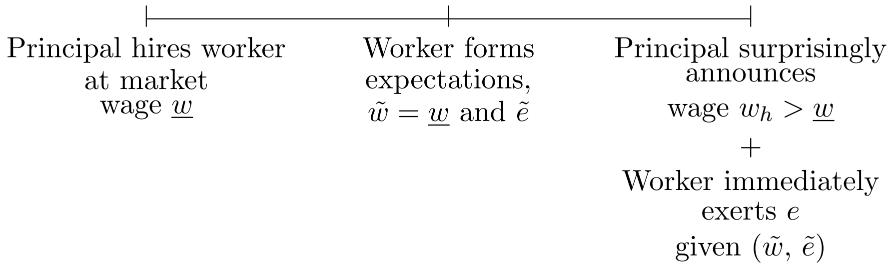
"If your boss walked over your desk and handed you $10,000, would it make you work harder for the rest of the day, or the rest of the year? I think it would!"
— Member in style.gather.com commenting on Oprah's surprising bonus to her magazine employees in 2010.

This section explores the extent to which fully surprising versus fully anticipated gifts can elicit a reciprocal response and increase the employer's profits in a one-shot principal-agent interaction. To create a fully surprising

[9] We generalize Definition 1 to the stochastic reference point case in Section 5.

[10] Assuming the agent does not foresee a (potentially) profitable gift is the equivalent to assuming that a rational agent has limited knowledge about other players' action sets, the core idea in the unawareness literature (e.g., Modica & Rustichini (1994), Dekel, Lipman & Rustichini (1998*a*), Dekel, Lipman & Rustichini (1998*b*), Aumann (2005)). After Li (2009) showed that the standard state space model could represent this unawareness, the idea that agents can take rational decisions with only partial knowledge of the relevant action spaces has flourished. See Kawamura (2005) for an application to competitive markets, Filiz-Ozbay (2012) for an application to insurance contracts, Auster (2013) for the implications of unawareness for moral hazard, Von Thadden & Zhao (2012) for the market's reaction to heterogenous levels of awareness in the agents' population, etc. Masatlioglu, Nakajima & Ozbay (2012) study how unawareness affects the revealed preference approach. There is also a related literature on incomplete contracts, which builds on the same idea. See, e.g., Maskin (2002) and Tirole (2009).

Figure 1: Timing and assumptions for the one-shot interaction with a fully surprising gift.



gift, we assume that the worker is unaware of the possibility of a wage raise and thus believes that he will surely be paid the market wage, as promised during recruitment. As mentioned before, this replicates the expectations of workers in gift exchange field experiments who, being unaware that they are part of a study, believe that the wage they were hired at will be the actual wage. Aside from the main message of gifts being more effective when they are fully surprising, we explore the impact of the perceived kindness of the market wage, the gift size, and the positive or negative nature of the surprise in triggering reciprocal effort.

Figure 1 summarizes this section's timing and assumptions. The principal hires the agent at the market wage, $\underline{w}$, to work for one period. Given this wage, the agent forms the belief that $\tilde{w} = \underline{w}$. Using this belief, he forms his effort expectation, $\tilde{e}$. Then the principal grants a fully surprising gift $w_h > \underline{w}$, the agent immediately exerts effort e, and the interaction ends.

Following the assumption that the agent forms his effort plans rationally given his wage expectations, we assume that the effort plan is credible and maximizes his ex ante utility as in Definition 1. Lemma 1 presents one simple but useful implication of the equilibrium effort whenever the reference point is deterministic,

Lemma 1 *For any fixed wage $\bar{w}$, suppose the agent expects $\tilde{w} = \bar{w}$ with certainty. Then, he forms his effort plans as a consumption utility maximizer, $\tilde{e} \in \text{argmax}\ m(e, \bar{w}) = \underline{e}$. Moreover, if the principal actually grants $\bar{w}$, the agent executes his plan, yielding $e = \underline{e}$.*

The intuition behind Lemma 1 is straightforward: since agents form effort plans rationally, if the wage is certain, expectations are met and agents experience zero gain-loss utility. Total utility therefore reduces to consumption utility and the worker plans to exert minimum effort for any fixed wage level. If no information arrives, the agent will implement this plan, as it is the unique credible plan.[11]

[11]Lemma 1 is the application of Proposition 3 in Kőszegi & Rabin (2006) to our setting.

3.1 The effort response depends on the surprising nature of the gift

Before turning to gift profitability, we start by comparing the effort response to anticipated versus surprising gifts. Before considering a fully surprising gift as described in Figure 1, consider the fully anticipated case: effort is not exerted immediately after the gift's announcement but rather there is a time gap between when the actual wage w_h is announced and effort is exerted. If this time gap is sufficiently long, then hedonic adaptation (Helson (1964), Frederick & Loewenstein (1999), Frey & Stutzer (2002)) occurs and the agent begins to work with a reference point that fully incorporates his expectations of the higher wage, $\tilde{w} = w_h$.[12] In this case, even if paid an above-market clearing wage, by Lemma 1 the agent behaves as a consumption utility maximizer and exerts minimum effort.

As in Figure 1, suppose now that the employer announces the wage gift immediately prior to the worker exerting effort, so that hedonic adaptation does not occur. As a consequence $\tilde{w} = \underline{w}$, and from Lemma 1 his effort plan is to exert minimum effort, $\tilde{e} = \underline{e}$. The wage news $w_h > \underline{w}$ force the agent to maximizes his total consumption and gain-loss utility *given* his wage and effort expectations. The worker's immediately chooses an effort level e^* solving

$$e^* \in \underset{e}{\operatorname{argmax}}\, w_h - \frac{\gamma}{2}(e-\underline{e})^2 + \eta\mu\big(-\frac{\gamma}{2}(e-\underline{e})^2\big) + \alpha\eta\mu(K(w_h) - K(\underline{w}))\mu((be - w_h) - (b\underline{e} - \underline{w})). \quad (4)$$

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 cannot simply solve the first order condition because his utility function has a kink at the effort level satisfying $\pi(e,w) = \pi(\underline{e},\underline{w})$, where the firm's profits shift from the loss domain to the gain domain.[13] But if his optimal effort does happen to occur at a differentiable point in his utility function, then the first-order condition applies, and we can rearrange it to show that effort corresponds to

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

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 addition to these two possible interior solutions, the worker must also consider the utility he would get at the kink in his utility function, where he exactly repays the firm the costs of the gift ($be - w_h = b\underline{e} - \underline{w} \Rightarrow$

[12] We comment on the speed of adaptation of reference points in Section 4.

[13] Notice that no such kink exists in the effort domain since *any* deviation from the plan of minimal effort implies a loss.

$e = \underline{e} + (w_h - \underline{w})/b)$. Consider first the case where the gift and the worker's optimal response to it increases profits, i.e., $w_h < b(e - \underline{e}) + \underline{w}$. Denote e_g (where the subindex stands for gain) as the optimal effort *given* this assumption. Using equation (5) and the requirement that the firm's profits increase with the gift, we have that e_g and the range of profitable gifts are

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

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 (where the subindex stands for loss) and the range of unprofitable gifts are

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

Details are relegated to the appendix in the proof of Proposition 1, 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 $be - w_h = b\underline{e} - \underline{w} \Leftrightarrow e = \underline{e} + (w_h - \underline{w})/b$. The worker thus always exerts more effort than he planned to.

Recalling that Lemma 1 says that any anticipated gift will lead to effort $\underline{e}$, we have thus sketched the proof for the following proposition.

Proposition 1 *A fully surprising gift $w_h > \underline{w}$ leads to higher effort than a fully anticipated gift of the same magnitude, that is $e^* > \underline{e}$.*

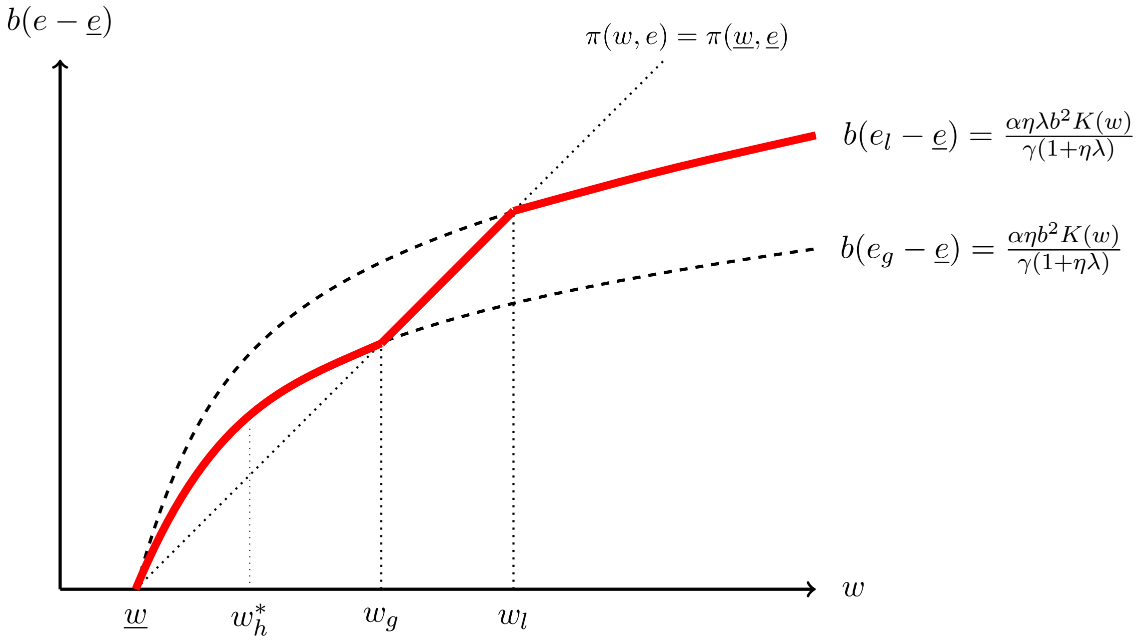
Finally, note that Proposition 1 is robust to incorporating baseline reciprocity in an additive form. Adding such a component to the agent's preferences would trigger extra reciprocal effort depending only on the gift's magnitude, regardless of whether the gift is anticipated or surprising. The surprise, however, would still trigger an additional increase in effort above the baseline reciprocal effort.

3.2 Fully surprising gifts must be small to be profitable

We now turn to the profitability of gift giving when gifts are fully surprising. The analysis in Section 3.1 determines when the worker's optimal effort response will increase or decrease profits, or neither in the case the worker chooses effort at the kink in utility where $be - w = b\underline{e} - \underline{w}$. 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 the effort level.

Building on Proposition 1, 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 gift wage. To understand the shape of this revenue-response

Figure 2: Revenue response function to a fully surprising gift in a one-time interaction.



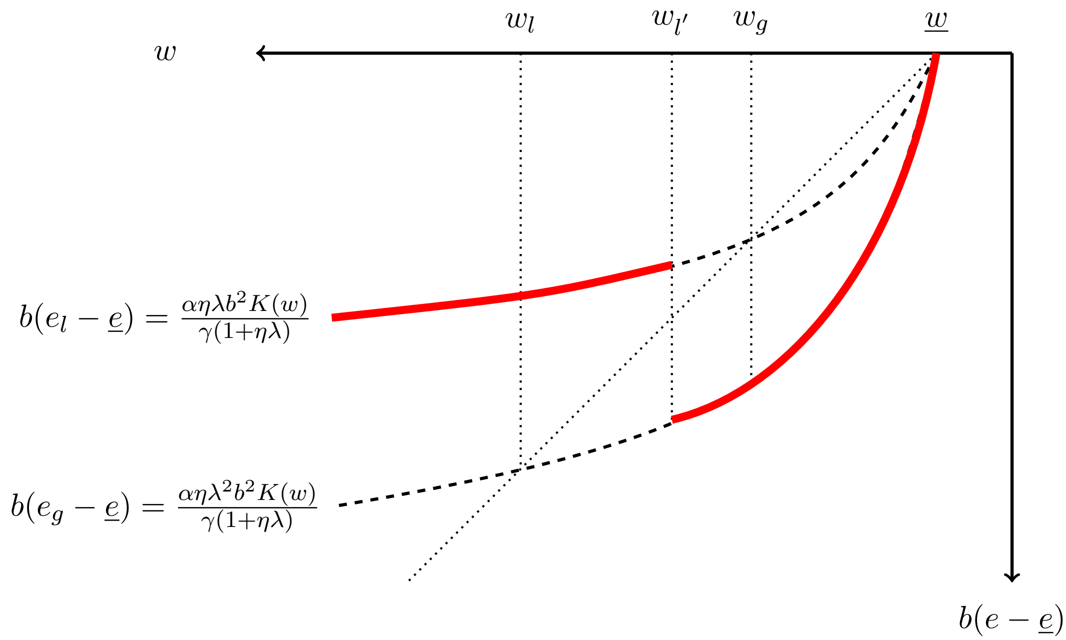
function, start by noticing that the diagonal line is where increased revenue exactly compensates the firm for the cost of the wage increase; above the line the firm experiences profits and below it it experiences losses. The lower dotted concave curve shows the revenue increase that occurs when the worker responds with e_g. Since he only does this when profits increase as a result of the gift, 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 increase does not compensate for the cost of the gift, decreasing profits. Similarly to the gain-domain case, he follows this curve 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 loss aversion ($\lambda > 1$) and the concavity of K, the worker chooses e at the kink in his utility function where profits are unchanged relative to baseline.

The position of the thick red revenue-response function in Figure 2 relative to the 45 degree line shows that gifts are strictly profitable only if they are sufficiently small. When gifts are smaller than a cutoff value w_g, the revenue response function lies above the diagonal line and thus the firm experiences profits. Intuitively, only gifts that are smaller than w_g are small enough that the worker can easily reciprocate enough to keep the firm's profits in the gain domain. To the contrary, gifts above the cutoff value w_l are too expensive in terms of effort cost to reciprocate fully. Proposition 2 summarizes.

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

3.3 Wage cuts are more harmful when they are surprising

Figure 3: Revenue response function to a fully surprising wage cut in a one-time interaction.



The same approach of Section 3.2 allows us to analyze negative reciprocity to fully surprising wage cuts. In particular, we show that the timing of information about a wage cut matters: managing expectations prior to the cut can ameliorate the effort decrease induced by negative reference-dependent reciprocity. This is important as, unsurprisingly, when firms cut wages, workers frequently respond by reducing effort (Lee & Rupp 2007, Krueger & Mas 2004, Kube, Maréchal & Puppe 2013).[14]

To formalize how the surprising nature of a wage cut can induce negative reciprocity, we 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 $\underline{w}$: $K(\underline{w} + w) = -K(\underline{w} - w)$ for all w.

The analysis proceeds analogously to Section 3.2. The worker anticipates $\underline{w}$ and plans to exert effort $\underline{e}$, but he is fully surprised by a wage cut $w_c < \underline{w}$. With this information, the worker's problem is to immediately

[14]The effort reduction in response to a wage cut does not seem to depend on the nature of the employer-employee relationship, however. Chen & Horton (2016) show that wage cuts harm effort even in online labor markets where the employment contract resembles a spot contract more than a labor contract.

choose an effort level e^* to maximize his utility given this (unrealized) reference point. This e^* solves

$$e^* \in \underset{e}{\operatorname{argmax}}\, w_c - \frac{\gamma}{2}(e-\underline{e})^2 + \eta\mu(-\frac{\gamma}{2}(e-\underline{e})) + \alpha\eta\mu(K(w_c)-K(\underline{w}))\mu(b(e-\underline{e})-(w_c-\underline{w})), \tag{8}$$

where $K(w_c) - K(\underline{w}) < 0$. 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 the kink) are derived from the first-order condition of this utility function. These effort levels, and the wage cut sizes for which they indeed correspond to local maxima, are

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

and

$$e_l - \underline{e} = \frac{\alpha\eta\lambda^2 bK(w_c)}{\gamma(1+\eta\lambda)} \qquad \Leftrightarrow \qquad \frac{\alpha\eta\lambda^2 K(w_c)b^2}{\gamma(1+\eta\lambda)} < w_c - \underline{w} \tag{10}$$

.

Figure 3 shows with a solid thick red line the worker's response, in terms of decreased revenue $b(e-\underline{e})$, as a function of the wage. The diagonal line is where decreased revenue exactly compensates the savings from the wage cut; above the line it is profitable to cut wages and below the line it is not. The first dotted concave curve (closer to the x-axis) shows the revenue decrease that occurs when the worker responds with e_g. Since he only does this when profits increase as a result of the wage cut, he follows this curve as long as it is below the diagonal, or equivalently, if the wage is smaller than the threshold $w_g = \underline{w} + \frac{\alpha\eta K(w_g)b}{\gamma(1+\eta\lambda)}$. Similarly, the second dotted concave curve (the one further below from the x-axis) shows the revenue decrease that occurs when the worker chooses e_l, which he does when lost revenue more than compensates for the savings from the cut, decreasing profits. The worker follows this curve as long as profits are in the loss domain, i.e., below the diagonal line, which is true as long as $w_c > w_l = \underline{w} + \frac{\alpha\eta\lambda K(w_l)b}{\gamma(1+\eta\lambda)}$. Since $w_l < w_g$, notice that—to the contrary of Figure 2—for wages between w_l and w_g, both e_l and e_g, are local best responses. The proof of Proposition 3 shows that for this range there exists a $w_{l'}$, with $w_l < w_{l'} < w_g$, such that e_g is optimal when $w_c < w_{l'}$ and e_l is optimal when $w_c > w_{l'}$. Thus, the agent never prefers effort at the kink $e = \underline{e} - (\underline{w} - w_l)/b$.

The resulting revenue response function in Figure 3 shows that, 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. Proposition 3 summarizes,

Proposition 3 *A fully surprising wage cut* $w_c < \underline{w}$ *leads to lower effort than an anticipated wage cut of the same magnitude, that is* $e^* < \underline{e}$.

Once again we must emphasize that in order to focus on the effect of surprises, we have abstracted away from other forms of social preferences, including 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 is robust to the inclusion of such preferences and indicates that the situation is made worse by the surprise factor. Indeed, if a surprising wage cut occurs when workers are being paid above-market wages and exerting above-minimum effort, retaliation is even cheaper so the surprise is more damaging for the firm.

We can also compare the size of the effort response to 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. Then it is 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.[15] We have thus the following proposition,

Proposition 4 *Fully surprising wage cuts are reciprocated at least as strongly as equivalently sized fully surprising wage increases.*

3.4 The firm's fully-surprising gift choice

Given the worker's effort response and the set of profitable gifts, we now turn to the firm's choice problem. The optimal fully-surprising gift is of course the one that maximizes the firm's profits among all profitable gifts. Proposition 5 describes this gift.

Proposition 5 *The optimal fully-surprising gift* w_h^* *exists if* $K'(\underline{w}) > \frac{\gamma(1+\eta\lambda)}{\alpha\eta b^2}$. *Moreover,* w_h^* *is in the range* $\underline{w} < w_h < w_g = \frac{\alpha\eta b^2 K(w)}{\gamma(1+\eta\lambda)}$, *and satisfies* $K'(w_h^*) = \frac{\gamma(1+\eta\lambda)}{\alpha\eta b^2}$.

The optimal gift can be easily seen in Figure 2. Since the firm wants to maximize the additional revenue it will earn beyond the cost of the gift, the optimal gift maximizes the distance above the diagonal line of the revenue-response function. If the derivative of the kindness function at the market wage is large enough, i.e., if the marginal benefit of reciprocating the gift is sufficiently large, the revenue curve will in fact rise above this

[15] The asymmetric response of effort to wage increases vis-a-vis wage decreases is well established in the empirical literature (Offerman (2002), Al-Ubaydli & Lee (2009), Kube, Maréchal & Puppe (2013)). Moreover, this asymmetry has also been established in prices in general (Ahrens, Pirschel & Snower (2017)).

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 the comparative statics.

Corollary 1 *The existence of a profitable gift, the size of the optimal fully surprising gift, the size of the effort response to any gift and to the optimal gift in particular, and the firm's profits after the surprise, are all 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 maximizes the gap between the effort response function in the firm's gain domain and the 45 degree line—where the cost and benefit of the gift are the same. First consider 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, which in turn reduces the probability that a profitable gift exists, and if it does exist, it decreases the size of it.

The optimal gift responds similarly to an increase in loss aversion (λ), which measures the impact of losses relative to gains. Because the worker does not experience losses in reciprocity 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 increases the cost of exerting higher effort to reciprocate the gift and thus 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 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 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 consumption utility. Increasing η therefore asymmetrically increases the importance of reciprocity compared to effort. As shown in the proof of Corollary 1, the gain in the reciprocity domain wins out and causes w_h^* to rise.

4 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 decreases over time as agents come to expect a wage raise with certainty. To this end, we assume that the principal and agent interact repeatedly and that the gift takes the

Figure 4: Timing and assumptions for the repeated interaction with a permanent gift that is only fully surprising in the first period.

Period $t = 0$	Period $t = 1$	Periods $t > 1$
Principal hires worker at market wage $\underline{w}$ for all t + Worker forms expectations $\tilde{w}_t = \underline{w}$ and $\tilde{e}_t$ for all t	Principal surprisingly announces wage $w_h > \underline{w}$ for all t + Worker immediately exerts e_1 given $\tilde{w}_1 = \underline{w}$ and $\tilde{e}_1$ + Worker updates $\tilde{w}_t = w_h$ and $\tilde{e}_t$ for $t > 1$	Worker exerts e_t given $\tilde{w}_t = w_h$ and $\tilde{e}_t$

form of a fixed permanent wage increase. We maintain the assumption that in the first period the worker is fully surprised by the wage increase. As in Section 3, this assumption replicates the expectation of workers in a gift exchange experiment who, being unaware that they are part of a study, expect to be paid the market wage in the first period. The assumption that after a permanent wage increase the agent comes to expect the gift with certainty follows the psychological insight on the adaptation literature, positing that the hedonic response to a constant stimuli is decreasing in time. Hedonic adaptation to constant simuli was first proposed in psychology by Helson (1964) and applied to economics, among others, by Frederick & Loewenstein (1999) and Frey & Stutzer (2002). Incorporating hedonic adaptation into our model (via updating of expectations) allows us to predict the dynamic consequences of wage increases.[16] To isolate the role of adaptation, we further assume that the agent correctly believes that there will be no further raises and thus from period two onwards he correctly expects to receive the higher wage. In Section 6 we extend this analysis by assuming that, after a surprising wage increase, the agent foresees the possibility of a further raise because he realizes that the principal cannot commit to a wage.

Figure 4 summarizes this section's timing and assumptions. There are infinite periods. In period zero the principal hires the agent at the market wage $\underline{w}$ for all upcoming periods. Given this wage, the agent forms the belief that he will be paid the market wage for all periods, $\tilde{w}_t = \underline{w}$. Using this belief, he forms his effort expectation for all upcoming periods, $\tilde{e}_t$. In period one, the first working period, the principal permanently raises the wage to $w_h > \underline{w}$ and the agent immediately exerts effort e_1. At the end of the first period the agent

[16]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 review.

updates his wage and effort expectations for periods two onwards. Without further announcements from the principal, in all upcoming periods the agent exerts effort given his wage expectations $\tilde{w}_t = w_h$ for all $t > 1$, which have already acclimatized to the wage increase. [17]

To analyze the temporal response of effort to a surprising permanent gift, start by noticing that since in every period the wage expectation is a fixed wage, by Lemma 1 the agent rationally plans to exert minimum effort in every period. Thus, given the actual gift w_h, the period $t \geqslant 1$ worker's problem is

$$e_t^* \in \underset{e}{\operatorname{argmax}}\, w_h - \frac{\gamma}{2}(e-\underline{e})^2 + \eta\mu\big(-\frac{\gamma}{2}(e-\underline{e})^2 + \frac{\gamma}{2}(\tilde{e}_t-\underline{e})^2\big) + \alpha\eta\mu(K(w_h)-K(\tilde{w}_t))\mu(\pi(w_h,e)-\pi(\tilde{w}_t,\tilde{e}_t)). \quad (11)$$

Notice that the key and only difference between the worker's problem in the one shot case (see equation (4)) and that in the repeated-interaction case (equation (11)), is that for each period t the reference point is potentially updated. Because the principal only delivers wage news in period one and utility across periods is only linked through the reference point, the effort decision reduces to a choice in period one and one choice that applies in every period from two onwards.

(1) Period-one effort decision. Since $\tilde{w}_1 = \underline{w}$, from Lemma 1 the agent forms his plans for future effort as a consumption utility maximizer and thus $\tilde{e}_1 = \underline{e}$. The period-one decision is thus identical to the choice characterized by Proposition 1. At the end of this period the agent updates expectations and forms $\tilde{w}_t = w_h$ $\tilde{e}_t = \underline{e}$ for all $t > 1$ (see Lemma 1).

(2) Period $t > 1$ effort decision. As mentioned above, in the absence of further news the worker's expectations are exactly met and he is left to maximize consumption utility. Lemma 1 applies and $e_t = \underline{e}$.

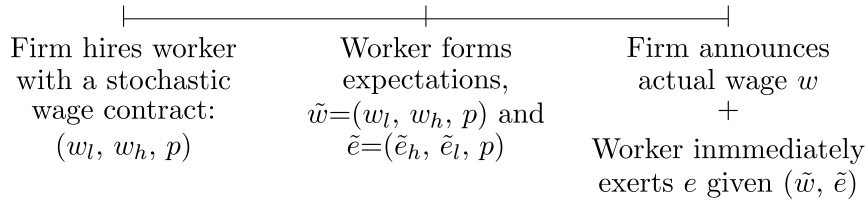
Proposition 6 summarizes:

Proposition 6 *After a first-period effort increase due to a fully surprising permanent wage raise $w_h > \underline{w}$ for all t, effort wanes back to baseline in every period thereafter, that is $e_1 > \underline{e}$ but $e_t = \underline{e}$ for $t > 1$.*

Two important comments on Proposition 6 are due. First, the prediction that effort spikes and then wanes would still hold had we added baseline non-reference-dependent reciprocity to the worker's preferences. The only difference would be that the level of effort after the raise would be higher than that at the market wage. Second, how fast effort goes back to its baseline level depends on the speed of adaptation of the reference point. Even though more research is needed to explore how fast expectations adapt in the field, evidence suggests that expectations do adapt fairly quickly but not immediately. In a laboratory experiment varying the time gap that subjects are given between payment information and the execution of a task measuring loss aversion, Song

[17] Notice that to simplify the exposition we do not explicitly assume that the agent updates expectations at the end of periods $t > 1$. We do so at no cost because there no more news besides those occurring in period one.

Figure 5: Timing and assumptions for the stage game with a stochastic gift.



(2016) finds that loss aversion is decreasing in the time gap. Thakral & Tô (2017)) provide field evidence of adaptive reference points using data on New York city cab drivers by showing that labor-supply reductions are stronger in response to earnings that accumulate more recently.

5 Gifts are Cursed: Never Grant Them or Grant Them Forever

"I would have loved to give out bonuses, but the thought of people being mad that they didn't get a bonus—a bonus!—when we didn't make enough money to justify it, doesn't make me feel great. If I wanted to be cynical, I could argue that I should have never started giving bonuses in the first place."

— Jay Goltz, owner of five small businesses in Chicago, writing for the New York Times blog on running a small business.

In this section we relax the fully-surprising assumption of Sections 3 and 4 to explore how the effort response and the gift profitability change if the agent anticipates deviations from the market wage. Assuming that the principal and agent interact indefinitely and the agent foresees wage increases conforms to standard game-theoretical analysis and is a reasonable assumption when firms have previously used surprising gifts and workers have come to expect further raises.

Since higher wages only elicit extra effort when unexpected, a seemingly obvious implication of Proposition 6 is that firms should grant gifts stochastically. At first glance, random gifts might be optimal as they could take advantage of the benefits of surprising agents—at least probabilistically—to elicit higher effort without inducing hedonic adaptation. We show, however, that this intuition is incorrect: gifts are cursed when they are expected but uncertain because the decrease in expected revenue due to retaliation when workers do not receive the gift outweighs the revenue boost when the gift is granted.

The analysis in this section focuses on stationary equilibria of the infinite principal-agent interaction. In a stationary equilibrium the principal offers the same compensation scheme in every period, either the same fixed wage or the same probability of a gift across periods. Considering non-stationary contracts—e.g., by allowing each period's fixed wage or the probability of getting a gift to be contingent on the game history—could confound the motivation coming from reciprocal preferences with those arising from consumption utility, i.e. career concerns. We thus focus on equilibria where the payment scheme is constant across stages, which is closer to the gift exchange principle that no dynamic considerations are needed to motivate workers.

Figure 5 summarizes the stage game timing and assumptions. The stage contract offered by the firm is now possibly stochastic: she will pay a gift wage w_h with probability $p > 0$, and a baseline wage w_l with probability $1 - p$, where $w_h > w_l \geq \underline{w}$.[18] Since the agent is rational, the contract (w_l, w_h, p) defines the agent's wage expectations $\tilde{w}$=(w_l, w_h, p). Using this stochastic belief, he forms contingent effort plans $\tilde{e_l}$ and $\tilde{e_h}$ in case w_l or w_h is realized, respectively. We summarize this contingency plan as the triple $(\tilde{e_l}, \tilde{e_h}, p)$ where p is, as before, the probability that w_h is realized. Then the principal announces the actual wage $w \in \{w_h, w_l\}$, the agent immediately exerts effort e, and the interaction ends.[19]

Since the gift is now possibly stochastic, Definition 2 formally extends the equilibrium plan and effort decision in Definition 1 to stochastic reference points and contingent effort plans. In Definition 2, exerted effort contingent on the wage and effort expectations must maximize utility given the wage realization, and the full contingency plan $(\tilde{e_l}, \tilde{e_h}, p)$ must maximize expected utility among all such credible plans. Furthermore, as in Kőszegi & Rabin (2006), we assume that realized outcomes are compared to all possible unrealized outcomes and the overall sense of gain or loss is the probability-weighted average of these comparisons.

Definition 2 *(The Agent's Preferred Personal Equilibrium (PPE) with stochastic wage expectations)* *Given a stochastic wage expectation* $\tilde{w}$=(w_l, w_h, p)*, a contingent effort plan* $(\tilde{e_l}, \tilde{e_h}, p)$ *is a preferred personal equilibrium (PPE) iff*

(i) $\tilde{e}_l \in \underset{e_l}{\operatorname{argmax}}\ EU(e_l, w_l|\tilde{e}, \tilde{w}) = m(w_l, e_l) + EU(n(e_l, w_l|\tilde{e}, \tilde{w}))$

(ii) $\tilde{e}_h \in \underset{e_h}{\operatorname{argmax}}\ EU(e_h, w_h|\tilde{e}, \tilde{w}) = m(w_h, e_h) + EU(n(e_h, w_h|\tilde{e}, \tilde{w}))$ *and*

(iii) $(\tilde{e}_l, \tilde{e}_h) \in \underset{(e_l \in E_l^*, e_h \in E_h^*)}{\operatorname{argmax}}\ EU(e_l, e_h, \tilde{w}|e_l, e_h, \tilde{w})$

where $E_l^* = \{e \in E | e$ *solves (i)*$\}$ *and* $E_h^* = \{e \in E | e$ *solves (ii)*$\}$*. Actual effort is* $e = \tilde{e}_l$ *if* w_l *is observed and* $e = \tilde{e}_h$ *if* w_h *is observed.*

[18] A different framing for this assumption is that the principal grants w_l with certainty and offers the possibility of a (stochastic) gift, or that the firm hires the worker at w_l but that the worker knows the firm is unable to commit to paying this wage and forms rational expectations about the probability of a gift. Allowing for the base wage to be higher than the market wage, $w_l > \underline{w}$, gives the firm the freedom to compensate workers for uncertainty overall if the ex ante utility of the market wage plus probabilistic gift does not satisfy the worker's participation constraint.

[19] For tractability, we focus on the worker's pure strategy contingency plans and actions.

With stochastic gifts, the firm's problem is to optimally choose w_l, w_h and p in order to maximize profits with the knowledge that the worker will anticipate these choices when setting his wage expectations and choosing his effort rationally, i.e. following Definition 2. The worker then only accepts the job if it meets his participation constraint; that is, if this stochastic contract provides higher expected utility than his reservation wage $\underline{w}$. The firm, likewise, only offers this contract if it meets her participation constraint; that is, if it leads to higher expected profits than credibly commiting to pay $\underline{w}$ with certainty.

Given the worker's optimal reaction to the stochastic contract (w_l, w_h, p) and the firm's problem, we now explore whether the firm can improve profits above those she would get when credibly commiting to pay the market wage with certainty. By narrowing down in Lemmas 2 through 4 what the stochastic contract must look like, we will show that this is in fact never possible.

First, in Lemma 2 we formally state the result that workers do in fact positively reciprocate in the high-wage state and negatively reciprocate in the low-wage state.

Lemma 2 *Given any stochastic wage contract (w_l, w_h, p) with $p \in (0,1)$, 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.[20]

Lemma 3 *Given any stochastic wage contract (w_l, w_h, p) with $p \in (0,1)$, the worker's PPE is never such that $\pi(w_l, e_l) = \pi(w_h, e_h)$.*

Next, we show that under any such contract that satisfies the firm's participation constraint (relative to credibly offering $\underline{w}$ with certainty), workers will respond with unequal costs of effort in the two wage states (that is, with asymmetric effort levels around $\underline{e}$). Lemma 3 and Lemma 4 serve to eliminate the kinks in the worker's utility function as potential PPE plans, which enables the derivation of Proposition 7 below.

Lemma 4 *Given any stochastic wage contract (w_l, w_h, p) with $p \in (0,1)$ that meets the firm's participation constraint, the worker's PPE is never such that $c(e_l) = c(e_h)$.*

We can finally establish the main result:

Proposition 7 *Any stochastic gift exchange arrangement (w_h,w_l,p) with $p \in (0,1)$ is never preferred by the firm to credibly commiting to pay $\underline{w}$ with certainty.*

[20]Therefore, 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.

Figure 6: Timing and assumptions of the repeated interaction with a gift that is fully surprising in the first period, and that leads to the expectation of further gifts.

Period $t = 0$	Period $t = 1$	Periods $t > 1$
Principal hires worker at market wage $\underline{w}$ for all t + Worker forms expectations $\tilde{w}_t = \underline{w}$ and $\tilde{e}_t$ for all t	Principal surprisingly announces wage $w_h > \underline{w}$ for $t = 1$ + Worker immediately exerts e_1 given $\tilde{w}_1 = \underline{w}$ and $\tilde{e}_1$	Worker forms expectations, $\tilde{w}_t$=(w_l, w_h, p) and $\tilde{e}_t$=($\tilde{e}_h$, $\tilde{e}_l$, p) for $t > 1$ + Firm announces actual wage $w_t \in \{w_l, w_h\}$ + Worker inmmediately exerts e_t given ($\tilde{w}_t$, $\tilde{e}_t$)

Proposition 7 implies that whenever agents anticipate the possibility of a gift, the best the firm can do is to offer stochastic gifts even though this leads to lower profits than offering $\underline{w}$ with certainty. Once workers become aware of the possibility of gifts, the firm is left with no option but to commit to an indefinite period of minimizing losses by setting $0 < p \leq 1$. This highlights the importance and difficulty of managing expectations in real-world gift exchange; we discuss this further in Section 8.

6 Gifts Can Be Profitable in Sufficiently Short Interactions

In this section we show that when the principal-agent interaction is sufficiently short, if the principal is able to fully surprise the agent, a gift might be profitable even if it generates the expectation of further gifts. To this end, we merge the belief-formation assumptions in Sections 3 and 5 and assume that at $t = 0$ the agent is fully surprised by a wage increase and that as a consequence, from $t = 1$ onwards, he assigns a strictly positive probability to receiving a further gift. These assumptions replicate anecdotal evidence of real-world firms actually granting surprising discretionary bonuses and allow us to explore the dynamic consequences of such gifts, i.e., workers expecting further gifts.[21,22]

Figure 6 summarizes this section's timing and assumptions. There are $T > 1$ periods. In period zero the

[21] Several cases of surprising discretionary bonuses are found in the popular press. For instance, Oprah Winfrey surprised staff members of *O*, the Oprah Magazine, with a $10,000 dollar check (see Oprah's story here) and the Grenda family, owners of a Melbourne-based bus company, unexpectedly gave an average bonus of $8,500 to its employees (see the Grenda family's story here).

[22] As explained in footnote 10, our preferred interpretation of full surprises is that there is an asymmetry of information about the principal's action set. Intuitively, workers are fully surprised by the wage increase, not because they failed to see that such raise is profitable for the firm, but simply because they were not aware that such a gift was feasible. This is intuitive given that in the labor market these gifts are not common (so uncommon that that they make it to the news! See anecdotal evidence in footnote 21.)

principal hires the agent at the market wage $\underline{w}$ for all upcoming periods. The agent forms the belief that he will be paid the market wage for all periods, $\tilde{w}_t = \underline{w}$, as in Section 3. Using this belief, he forms his effort expectation for all upcoming periods, $\tilde{e}_t$. In period one, the first working period, the principal raises the period-one wage to $w_h > \underline{w}$ and the agent immediately exerts effort e_1, following Definition 1. In period two onwards, the agent updates his wages and effort expectations for the upcoming period by assigning probability p to receiving the same gift again, correctly anticipating the firm's optimal gift-giving strategy as in Section 5. In all of the following periods the principal then announces the actual wage and the worker exerts effort, following Definition 2.[23]

6.1 The infinite-period case

Whenever the principal-agent relationship is infinitely repeated, the principal's decision of whether to use a fully-surprising gift compares the benefits triggered by the full surprise (Proposition 5) to the costs of revealing the inability to commit to paying $\underline{w}$. Revealing the possibility of a wage raise leads to the curse of gifts: non-profitable stochastic gift exchange ad infinitum since in every period the benefits of granting a gift are outweighed by the costs of unmet gift expectations (Proposition 7).

The tradeoff between the immediate benefits and the delayed but infinite costs of granting a fully surprising gifts, immediately implies the following corollary to Proposition 7.

Corollary 2 *In an infinite interaction, 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.)*

6.2 The finite-period case

In the finite-period case, the logic unravels through backward induction. In the final period, the equilibrium belief p cannot be zero, as this creates the possibility of profitably fully surprising the agent with a wage raise. And if gifts have probability $p \in (0, 1)$, we know from Lemma 3 that workers will respond in such a way that the firm will have unequal profits in the two wage states. In the absense of the necessity of maintaining expectations for the future, the firm will therefore not be willing to randomize between wages, but will choose the wage leading to higher profits for sure. This contradicts the assumption of $p \in (0, 1)$, however, so we are left with $p = 1$ as the sole possibility. Through backwards induction, the same occurs in every earlier period

[23]Notice we do not explicitly assume that the agent updates expectations at the end of periods $t > 1$. As explained in Section 5, stationary contracts are the most reasonable assumption in a gift exchange environment, where the size of the gift and the likelihood of getting it are constant across periods. This is because making the gift size or its probability depend on the history of the game could create other forms of incentives that would be confounded with incentives from reciprocity.

as well. Proposition 8 shows that the optimal wage level to which the firm can credibly commit is in fact the same as the optimal fully surprising gift found in Proposition 5.

Proposition 8 *In a finite interaction, if the firm chooses to grant a fully surprising gift, she grants the optimal gift in the first round, and then turns the gift into a permanent raise in later periods.*

Similarly to Corollary 2, Proposition 8 also highlights how difficult it is for a firm to profitably use surprises in long interactions. In an finite-period game, the firm's initial profit from the surprise must outweigh the losses of having to pay a higher wage that does not trigger reciprocal effort due to its anticipated nature for the remaining periods. Clearly, gift exchange is more likely to be profitable the shorter the interaction.

As before, a caveat is warranted: In order to focus on the effect of surprises, our model abstracts from non-reference-dependent reciprocity. 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 whenever gifts are anticipated. When gifts are stochastic, however, baseline reciprocity could render gifts even less profitable relative to credibly committing to the market wage. If the worker is reciprocating with above-minimal effort to begin with at some wage $w > \underline{w}$, a surprising gift w_h is now more difficult to reciprocate because the worker is starting off at a higher marginal cost of effort. Conversely, unfulfilled expected gifts are extremely cheap to negatively reciprocate, as the worker can simply withdraw his previous high effort. The risk to the firm from granting surprising gifts is thus amplified when baseline reciprocity is assumed.

7 Related Evidence

In this section we relate our model predictions to the existing evidence on gift exchange.[24]

(1) The power of gifts is greater when 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 (2016). 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 was hired and paid \$3, the second was hired and paid \$4, while the third was hired at \$3 but was surprised, right before work, with a \$1 increase per hour. They 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, in a laboratory experiment Sliwka & Werner (2017) find that an increasing wage profile that is not communicated in advance to subjects elicits higher effort relative to a constant wage profile with equal aggregated cost.[25]

The empirical evidence also supports the prediction that the power of a surprising gift is increasing in the marginal return to effort, b (see equation (5) and Corollary 1). First of all, Hennig-Schmidt, Rockenbach & Sadrieh (2010) find that clear information about the employer surplus is needed for gift exchange to arise in the lab.[26] Englmaier & Leider (2012*a*) explicitly varied the returns to effort in a gift exchange laboratory game to show that higher returns to effort induce more reciprocation, while Englmaier & Leider (2012*b*) find a similar result in the field where they hinted high returns to effort by saying that the manager would be getting a bonus if the job was finished within the week.[27]

There is also evidence that reciprocal evidence is decreasing in the marginal cost of effort, γ. Even though evidence is thin as laboratory experiments typically do not use real-effort tasks, there are exceptions. In particular, Gneezy (2002), using a real-effort task (solving mazes), find that when difficult “level 5” mazes are used instead of “level-2” mazes, gift exchange is significantly reduced.[28] Engelmann & Ortmann (2009), while

[24]We focus on the papers that are most relevant for our model predictions. For reviews focusing on the existence of gift exchange, see DellaVigna et al. (2016) and Esteves-Sorenson (2017).

[25]It could be the case, however, that subjects in Sliwka & Werner (2017) did come to expect future wage raises after the wage increase in the first period. Their result, therefore, is compatible with our predictions only if subjects did not update their gift expectations during the eight experimental periods, each lasting 250 seconds.

[26]Somewhat contrarily, Charness, Frechette & Kagel (2004) show that including a clear payoff table reduces gift exchange, but only among high-effort workers for whom marginal returns to effort are low.

[27]DellaVigna et al. (2016), in the largest gift exchange field experiment to date, also vary systematically the returns to the employer. They find that workers do exert more effort when the return for the employer in the stuffing-envelope task is higher, even though their response is not sensitive to the precise size of the return. However, since they cannot document a significant effort response to a monetary gift, they are unable to test whether the return for the employer affects gift exchange in the field.

[28]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.

not using a real effort task, also find that higher efficiency gains to effort increase both wage offers and effort exerted.

(2) The power of gifts wanes over time. Our emphasis on surprises is also consistent with findings of short-lived or waning effects of gift exchange. Evidence of this can be found in the two gift exchange studies in Gneezy & List (2006). In a data-entry study, they recruit workers for the market wage of $8 and then surprise 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 six hour period. In a second study on door-to-door fundraising, subjects were hired at $10 dollars and then fully surprised with a 100% increase. Effort initially spiked by 72% but then fell back to an insignificant 6%. Similarly, Bellemare & Shearer (2009) find that effects on productivity were concentrated on the day a surprise bonus was given to workers in a tree-planting firm. The waning of gift exchange has also been replicated in the laboratory with real-effort tasks. For instance, in a multiple-period experiment, Sliwka & Werner (2017) examine how reciprocal effort is affected by the timing of wage increases. They document that reciprocal effort does wane back to its baseline level in response to a surprising permanent wage increase, rendering the permanent raise ineffective for increasing overall productivity.[29]

In the existing literature, the primary alternative explanation proposed for such a waning pattern is fatigue: workers are simply too tired to continue reciprocating after having increased effort in response to the gift. Suggestive evidence ruling out this hypothesis, however, can be found in Sliwka & Werner (2017), as previously described. They find that workers who are given a sequence of small wage increases instead of a fully surprising wage increase, do not display a decreasing pattern of effort, despite the former induces higher aggregated effort at the same overall cost. Other suggestive evidence ruling out fatigue comes from comparing the effect of gifts relative to that of piece rates, where piece rates yield higher effort even late in the work period (Esteves-Sorenson 2017, DellaVigna et al. 2016).

(3) Gifts are cursed: grant them forever or never grant them. Proposition 7 through Corollary 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. These results build 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

[29]Most of the studies that do not find a waning effect are studies with short-term interactions in which subjects likely don't have time to grow accustomed to the gift. Kube, Maréchal & Puppe (2012), for example, does not observe significant waning of effort in a three-hour work period.

expectation of further monetary gifts, this assumption is intuitive and relates to the literature where subjects learn even under ambiguity. For instance, Epstein & Schneider (2007) develop a model in which players do not know the relevant probability distributions but nonetheless update the relevant probability space.[30]

The second aspect, the asymmetry between positive and negative reciprocity when responding to deviations from an expected payment, which is also reflected in the one-shot interaction analyzed in Proposition 4, is well grounded in the experimental evidence. Hannan (2005) modifies a standard laboratory gift exchange experiment to add an exogenous shock to firms' profits after which firms and workers can adjust their wage and effort 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. This asymmetry has also been documented in the field. Kube, Maréchal & Puppe (2013) hired workers for a data-entry task for a "projected" wage of 15 Euros. Right before work, a group of workers received a wage cut to 10 Euros and the other group a wage raise to 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 effort even though effort did respond positively to monetary incentives.[31] These results are in line with the well established stylized fact that firms are reluctant to cut wages to avoid hurting workers' "morale" (e.g., Bewley (2009)).

As mentioned in the text, an important caveat 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, improving in turn the profitability of 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), Sliwka & Werner (2017)) or the possibly weak nature of gift exchange in the field (Kube, Maréchal & Puppe (2012), Hennig-Schmidt, Rockenbach & Sadrieh (2010), List (2006), Esteves-Sorenson (2017), DellaVigna et al. (2016)) is interpreted as evidence that this baseline, non-reference-dependent reciprocity is relatively weak—at least in the workplace—our results should apply directly without readjusting for omitted forms of social preferences.

8 Conclusion and Discussion

We conclude by discussing some lessons and implications of our predictions. Whenever possible, we present venues for future research.

[30]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 can thus be fully surprised by actions that they were not initially aware off. See Von Thadden & Zhao (2012) and Carroll (2015).

[31]To test whether the lack of positive reciprocity was due to workers being unable to reciprocate, they hired workers for a piece rate and verified that there was room for a productivity increase above the baseline. Workers, however, were recruited under the piece rate, opening the possibility that the subjects working for the piece rate where more productive workers relative to those hired for a fixed wage.

8.1 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.

Extending the approach in Kőszegi & Rabin (2006) to model workers with reference-dependent reciprocal preferences, our model predicts that the most powerful gifts are those that are surprising (Proposition 1), but surprises only trigger temporary excess effort (Proposition 6). Moreover, surprises probably create the expectation of further gifts, which are harmful for the firm's profits when unfulfilled (Proposition 7). Firms, therefore, cannot repeatedly surprise workers with gifts, and thus the only profitable gift is cursed as it must be granted forever.

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.

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 transitory gifts. For instance, one pending question in the literature of gift exchange is why gifts in kind trigger reciprocal effort (Maréchal & Thöni (2016)) to a greater extent than monetary gifts even when they are less valued (Kube, Maréchal & Puppe (2012)).[32] In the context of our model, one possible answer is that, in contrast to monetary gifts, gifts in kind do not create the expectation of a further gift as their usage is usually more sporadic. Our theory thus predicts that this property of gifts in kind 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 communicated and delivered to workers. Clearly from the quote above, this is one of the issues that Starbucks forgot to think about when they decided to grant an unusual Christmas bonus to their employees.[33]

[32] DellaVigna et al. (2016) find that workers reciprocate neither monetary gifts nor gifts in kind.

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

8.2 Using surprise incentives in experimental designs

The fact that surprising incentives are powerful but only in the short run raises a word of caution for experimental designs that rely on surprises in their identification strategy. Even though the element of surprise is usually 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 this feature 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 timespans.[34]

8.3 Optimal gifts in repeated interactions

Our model also speaks to another issue that has not received due attention in the gift exchange literature: gift exchange 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.[35] Even though necessary in order to avoid confounding reciprocity with reputation, this assumption has drawn attention away from studying the properties of gifts in repeated interactions.

Our analysis enriches our knowledge of gift exchange 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 tension exists between the temporary profit increases from improved effort and the difficulty of managing workers' expectations about the likelihood of further gifts.

Anecdotal evidence, however, suggests that random monetary gifts are sometimes observed in real-world firms (see footnote 21). None of this anecdotal evidence, however, seems to point out a repeated use of this type of incentives. This suggests that firms who granted surprising gifts do 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.

[34]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 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.

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

8.4 Gift exchange and classical mechanisms behind efficiency wages

Our model also agrees with the idea that gift exchange is a complementary explanation to efficiency wages rather than a substitute to classical mechanisms behind the optimality of above-market wages. Shapiro & Stiglitz (1984) propose that it can optimal for firms to pay above market-clearing wages to increase workers' opportunity cost of losing the job to join the unemployment pool; Stiglitz (1974) and Salop (1979) argue that paying above market wages is profitable as it saves firms the onerous costs of labor turnover, while Weiss (1980) proposed that paying higher wages leads to selection of better workers as ability and reservation wages are correlated. As gift exchange builds on the agent's preference structure, it can, along with classical explanations, provide a more complete picture of why workers' marginal productivity might depend on the wage (Yellen (1984)).

8.5 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 (2017) summarizes, showing that wage-effort elasticities in the lab are large, but are small or non-significant in the field. For instance, Hennig-Schmidt, Rockenbach & Sadrieh (2010) find no gift exchange among students hired to type research abstracts, and Englmaier & Leider (2012*b*) find, if anything, negative gift exchange in their baseline treatment using workers in a data-entry job. Similarly, Kube, Maréchal & Puppe (2013) find no evidence of positive reciprocity to monetary gifts, while Al-Ubaydli et al. (2015) observe minor and insignificant gift exchange among temp workers hired to stuff envelopes. More recently, Esteves-Sorenson (2017) finds no significant gift exchange with a data-entry task despite larger than usual gifts, while DellaVigna et al. (2016) also fail to find reciprocal effort in response to monetary gifts in an envelope-stuffing field experiment with an impressive sample size.

The disparity in findings in laboratory versus field studies has led to a heated debate. Levitt & List (2007) catalogue 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 (2015) responds, and on the specific issue of gift exchange, shows that one experiment directly comparing field to lab outcomes, that of List (2006), finds comparable outcomes in the two settings. Esteves-Sorenson (2017) 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 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 gifts as irrational? Maybe 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 unprofitable could damage the very 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), and even less so in newly established, short-term interactions between employers and employees. Thus, to avoid selecting abler workers, field-experiment designs might be damaging their very purpose of examining gift exchange 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 should also focus on the equilibrium behavior of firms, in order to ensure the credibility of gift giving in the field.

References

Abeler, Johannes, Armin Falk, Lorenz F. 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.

Ahrens, Steffen, Inske Pirschel, and Dennis J. Snower. 2017. "A theory of price adjustment under loss aversion." *Journal of Economic Behavior and Organization*, 134: 78–95.

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

Al-Ubaydli, Omar, and Min Sok Lee. 2009. "An experimental study of asymmetric reciprocity." *Journal of Economic Behavior and Organization*, 72(2): 738–749.

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.

Aumann, Robert J. 2005. "Musings on information and knowledge." *Econ Journal Watch*, 2(1): 88–96.

Auster, Sarah. 2013. "Asymmetric awareness and moral hazard." *Games and Economic Behavior*, 82: 503–521.

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. 2015. "The promise and success of lab-field generalizability in experimental economics: A critical reply to Levitt and List." In *Handbook of Experimental Economic Methodology.* , ed. Guillaume R. Frechette and Andrew Schotter, Chapter 14. Oxford University Press.

Card, David, and Laura Giuliano. 2013. "Peer effects and multiple equilibria in the risky behavior of friends." *Review of Economics and Statistics*, 95(4).

Carroll, Gabriel D. 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.

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.

Chen, Daniel L., and John J. Horton. 2016. "Are online labor markets spot markets for tasks? A field experiment on the behavioral response to wage cuts." *Information Systems Research*, 27(2): 403–423.

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, 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.

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini. 1998*a*. "Recent developments in modeling unforeseen contingencies." *European Economic Review*, 42(3-5): 523–542.

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini. 1998*b*. "Standard state-space models preclude unawareness." *Econometrica*, 66(1): 159–173.

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." Working Paper.

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. 2012*a*. "Gift exchange in the lab: It is not (only) how much you give..." Working Paper.

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. 2017. "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, Georg Kirchsteiger, and Arno Riedl. 1993. "Does fairness prevent market clearing? An experimental investigation." *Quarterly Journal of Economics*, 108(2): 437–459.

Filiz-Ozbay, Emel. 2012. "Incorporating unawareness into contract theory." *Games and Economic Behavior*, 76(1): 181–194.

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. 2016. "When 3+1 > 4: Gift Structure and Reciprocity in the Field." *Management Science*, 62(9): 2639 – 2650.

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.

Gross, Till, Christopher Guo, and Gary B. Charness. 2015. "Merit pay and wage compression with productivity differences and uncertainty." *Journal of Economic Behavior & Organization*, 117: 233–247.

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 an expectations 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.

Kawamura, Enrique. 2005. "Competitive equilibrium with unawareness in economies with production." *Journal of Economic Theory*, 121(2): 167–191.

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. 2007. "Individual versus group behavior and the role of the decision making procerure in gift-exchange experiments." *Empirica*, 34(1): 63– 88.

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. 2014. "Corporate social responsibility in the work place: Experimental evidence from a gift-exchange game." *Experimental Economics*, 17: 347–370.

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.

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.

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. 2016. "Major earthquake experience and present-focused expenditures." Working Paper.

Li, Jing. 2009. "Information structures with unawareness." *Journal of Economic Theory*, 144(3): 977–993.

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.

Maréchal, Michel André, and Christian Thöni. 2016. "Hidden persuaders: Do small gifts lubricate business?" Working Paper.

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.

Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y. Ozbay. 2012. "Revealed attention." *American Economic Review*, 102(5): 2183–2205.

Maskin, Eric. 2002. "On indescribable contingencies and incomplete contracts." *European Economic Review*, 46(4-5): 725–733.

Modica, Salvatore, and Aldo Rustichini. 1994. "Awareness and partitional information structures." *Theory and Decision*, 37(1): 107–124.

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

Offerman, Theo. 2002. "Hurting hurts more than helping helps." *European Economic Review*, 46(8): 1423–1437.

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.

Ruffle, Bradley J. 1999. "Gift giving with emotions." *Journal of Economic Behavior & Organization*, 39(4): 399–420.

Salop, Steven C. 1979. "A model of the natural rate of unemployment." *American Economic Review*, 69(1): 117–125.

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

Shapiro, Carl, and Joseph E. Stiglitz. 1984. "Equilibrium unemployment as a worker discipline device." *American Economic Review*, 74(3): 433–44.

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. 2017. "Wage increases and the dynamics of reciprocity." *Journal of Labor Economics*, 35(2): 299–344.

Song, Changcheng. 2016. "An experiment on reference points and expectations." Working Paper.

Stiglitz, Joseph E. 1974. “Alternative theories of wage determination and unemployment in LDC’s: The labor turnover model.” *Quarterly Journal of Economics*, 88(2): 194–227.

Thakral, Neil, and Linh T. Tô. 2017. “Daily labor supply and adaptive reference points.” Working Paper.

Tirole, Jean. 2009. “Cognition and incomplete contracts.” *American Economic Review*, 99(1): 265–294.

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.

Weiss, Andrew. 1980. “Job queues and layoffs in labor markets with flexible wages.” *Journal of Political Economy*, 88(3): 526.

Yellen, Janet L. 1984. “Efficiency wage models of unemployment.” *American Economic Review*, 74(2): 200–205.

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

A Appendix: Proofs

Proof of Lemma 1

Following Definition 1, the agent forms his effort plans rationally to maximize $EU(e, \tilde{w}|\tilde{e}, \tilde{w})$, where $\tilde{e} \in \underset{e \in E}{\text{argmax}}\ EU(e, \tilde{w}|e, \tilde{w})$. Because $\mu(K(\tilde{w}) - K(\tilde{w}))$ is independent of $\tilde{e}$, so gain-loss reciprocity $n_k(\tilde{w}, e|\tilde{w}, \tilde{e}) = 0$ for any $\tilde{e}$. Therefore the agent's problem reduces to $\tilde{e} \in \underset{e \in E}{\text{argmax}}\ m(e, \tilde{w}) + n_c(e|\tilde{e}) \Rightarrow \tilde{e} = \underline{e}$. If no further information arrives, the agent implements this effort plan as it is the only credible plan and thus the unique PPE.

Proof of Proposition 1

By Lemma 1, since $\tilde{w} = \underline{w}$, 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:

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

and

$$e_l = \underline{e} + \frac{\alpha\eta\lambda K(w_h)b}{\gamma(1+\eta\lambda)} \qquad \text{requiring} \qquad w_h > \frac{\alpha\eta\lambda K(w_h)b^2}{\gamma(1+\eta\lambda)} + \underline{w}. \tag{13}$$

Figure 2 in the text 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 local 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 happens only where profits are exactly equal to expectations; note that there is no additional kink in the utility function where effort costs are exactly equal to expectations because effort costs are always in the loss domain when $\tilde{e} = \underline{e}$. Additionally, $\underline{e}$ is never preferred because $U(\underline{e}+\epsilon|\underline{e},\underline{w}) > U(\underline{e}|\underline{e},\underline{w})$ for small ϵ.

If the worker were to choose effort at the kink-point in his utility function, this would yield

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

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

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

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_h and/or w_g are equal to $\underline{w}$ are straightforward to account for.

Proof of Proposition 2

As described 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

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

and

$$e_l = \underline{e} + \frac{\alpha\eta\lambda^2 K(w_c)b}{\gamma(1+\eta\lambda)} \qquad \text{requiring} \qquad w_c - \underline{w} > \frac{\alpha\eta\lambda^2 K(w_c)b^2}{\gamma(1+\eta\lambda)} + \underline{w} \tag{15}$$

Figure 7 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. That is, $w_l < w_g$, to the contrary of Proposition 1.

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.

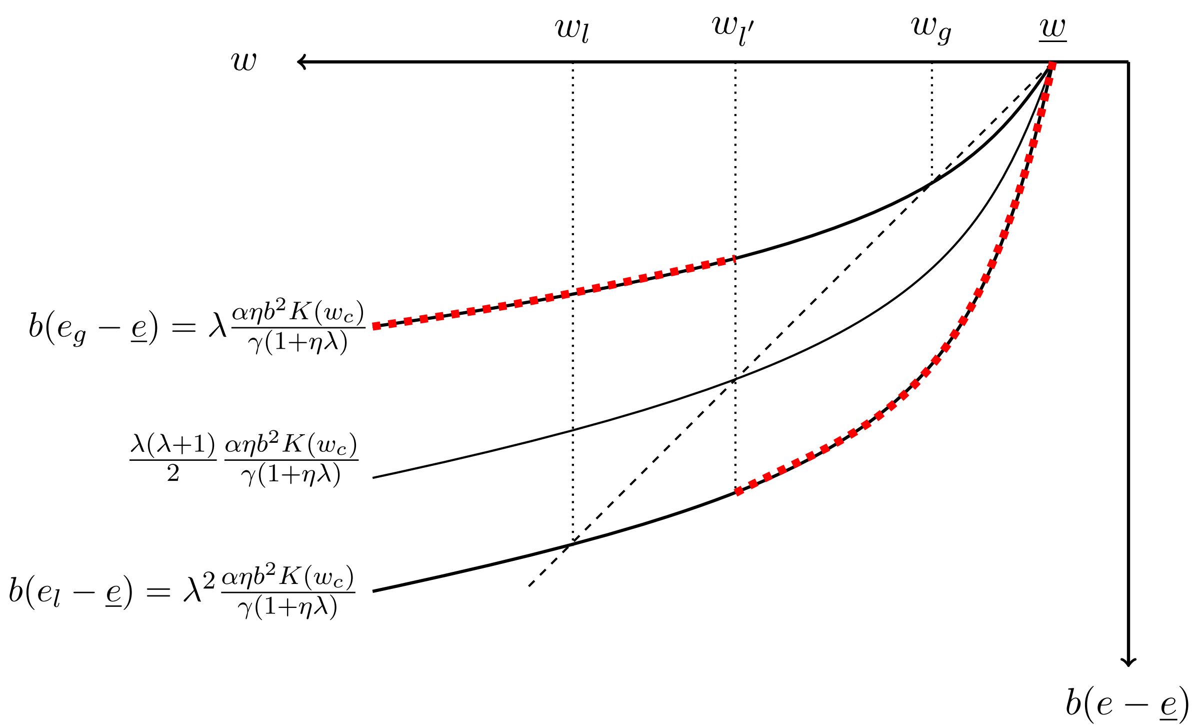
This relationship is a multiple of the revenue response curves that also determine the validity of the profit constraints above, so they are shown on Figure 7 together. By noticing that because $\lambda > 1$ and $1 < \frac{\lambda+1}{2} < \lambda$, the aggregate set of conditions imply that e_l is chosen for $w_c > w_{l'}$, and e_g is chosen otherwise. Regardless, a surprising wage cut is negatively reciprocated.

Extreme or corner cases in which w_c and/or w_g/w_l are $\underline{w}$ are straightforward to account for.

Proof of Proposition 4

As described in text.

Figure 7: Revenue response function to fully surprising wage cut



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 $K'(\underline{w}) > \frac{\gamma(1+\eta\lambda)}{\alpha\eta b^2}$ since K is concave, 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

As described in text.

Proof of Lemma 2

The worker formulates a contingency plan $\tilde{e} = (\tilde{e}_h, \tilde{e}_l)$ and then 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

$$\begin{aligned} U(e_h, w_h|\tilde{e}, \tilde{w}) &= m(w_h, e_h) + EU(n(e_h, w_h|\tilde{e}, \tilde{w}) \\ &= w_h - c(e_h) \\ &\quad + p\eta\mu(-c(e_h) + c(\tilde{e}_h)) + (1-p)\eta\mu(-c(e_h) + c(\tilde{e}_l)) \\ &\quad + (1-p)\alpha\eta\mu(K(w_h) - K(w_l))\mu(be_h - w_h - b\tilde{e}_l + w_l) \\ &= -\tfrac{\gamma}{2}(1 + (1-p)\eta\mu'_{chl} + p\eta\mu'_{chh})(e_h - \underline{e})^2 + (1-p)\eta\mu'_{chl}c(\tilde{e}_l) + p\eta\mu'_{chh}c(\tilde{e_h}) \\ &\quad + (1-p)\alpha\eta\Delta K\mu'_{\pi hl}(be_h - w_h - b\tilde{e}_l + w_l) \end{aligned} \tag{16}$$

where $\Delta K = K(w_h) - K(w_l)$ and

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

and

$$\mu'_{\imath hl} = \mu'(\pi(e_h, w_h) - \pi(\tilde{e}_l), w_l) = \begin{cases} 1 & \text{if } \pi(e_h, w_h) > \pi(\tilde{e}_l, w_l) \\ \lambda & \text{otherwise} \end{cases}$$

and μ'_{chh} and $\mu'_{\pi hh}$ are defined similarly. 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 the low wage state, the worker maximizes

$$\begin{aligned} U(e_l, w_l|\tilde{e}, \tilde{w}) &= m(w_l, e_l) + EU(n(e_l, w_l|\tilde{e}, \tilde{w}) \\ &= w_l - c(e_l) \\ &\quad + p\eta\mu(-c(e_l) + c(\tilde{e}_h)) + (1-p)\eta\mu(-c(e_l) + c(\tilde{e}_l)) \\ &\quad + p\alpha\eta\mu(K(w_l) - K(w_h))\mu(be_l - w_l - b\tilde{e}_h + w_h) \\ &= -\tfrac{\gamma}{2}(1 + (1-p)\eta\mu'_{cll} + p\eta\mu'_{clh})(e_l - \underline{e})^2 + (1-p)\eta\mu'_{cll}c(\tilde{e}_l) + p\eta\mu'_{clh}c(\tilde{e_h}) \\ &\quad - p\alpha\eta\lambda\Delta K\mu'_{\pi lh}(be_l - w_l - b\tilde{e}_h + w_h) \end{aligned} \tag{17}$$

with $\mu'_{.}$ defined analogously to as above.

The result can be trivially seen by considering that since the marginal cost of effort at $e = \underline{e}$ is zero according to Assumption 2, it is negligibly inexpensive to negatively reciprocate with some $e_l < \underline{e}$ when wages are unkind relative to expectations, and negligibly inexpensive to positively reciprocate with some $e_h > \underline{e}$ when wages are kind relative to expectations. This holds regardless of what the effort plan was, as slightly reducing effort from $\underline{e}$ will always decrease the gains or increase the losses to profits in the negative reciprocity case, and likewise for slightly increasing effort from $\underline{e}$ in the positive reciprocity case.

Proof of Lemma 3

Referring to equations 16 and 17 above, the first order conditions with respect to e_h and e_l in the high and low wage states respectively are

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

and

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

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 (either with $\pi(e_h,w_h) = \pi(e_l,w_l)$ or with $c(e_h) = c(e_l)$), 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 when evaluated 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 $\tilde{e} = (\tilde{e}_l,\tilde{e}_h)$, and define $\tilde{\mu}'_{chl} = \mu'(-c(\tilde{e}_h) + c(\tilde{e}_l))$ and $\tilde{\mu}'_{clh} = \mu'(-c(\tilde{e}_l) + c(\tilde{e}_h))$. Note that neither of these are 0 because profits and effort costs cannot simultaneously be equal in the two wage states, due to the symmetry of c. In the low wage state, $U'(e_l)$ just left of the discontinuity is given by

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

relying on the fact that 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 and Assumption 2. Similarly, just to the right of the kink the derivative is given by

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

In order for $U(e_l,w_l|\tilde{e},\tilde{w})$ to have a local optimum at $\tilde{e}_l$ we thus need $p\alpha\eta\lambda^2 b\Delta K \leq -\gamma(1+p\eta\tilde{\mu}'_{clh} + \lambda\eta(1-p))(\tilde{e}_l - \underline{e})$ and $p\alpha\eta\lambda b\Delta K \geq -\gamma(1+p\eta\tilde{\mu}'_{clh} + \eta(1-p))(\tilde{e}_l - \underline{e})$. This requires that $\lambda(1+p\eta\tilde{\mu}'_{clh} + \eta(1-p)) \leq$

$(1+p\eta\tilde{\mu}'_{clh}+\lambda\eta(1-p))$, or $\lambda-1 \leq (1-\lambda)p\eta\tilde{\mu}'_{clh}$. 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 $\tilde{e} = (\tilde{e}_l, \tilde{e}_h)$, such that $c(\tilde{e}_l) = c(\tilde{e}_h)$. For this to be true, $U(e_l, w_l|\tilde{e}, \tilde{w})$ (referring to (17) above) must have a local optimum over e_l at $\tilde{e}_l$ and $U(e_h, w_h|\tilde{e}, \tilde{w})$ (from (16)) must have a local optimum over e_h at $\tilde{e}_h$. Define $\tilde{\mu}'_{\pi lh} = \mu'(\pi(\tilde{e}_l, w_l) - \pi(\tilde{e}_h, w_h))$ and similarly for $\tilde{\mu}'_{\pi hl}$. In low wage case, the value of U' just to the left of the kink at $\tilde{e}_l$ is then given by

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

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

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

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

In the latter high wage case we similarly have

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

and

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

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

We can combine these two sets of inequalities, written as

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

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 $\frac{p\lambda}{1-p}\frac{\tilde{\mu}'_{\pi lh}}{\tilde{\mu}'_{\pi hl}}$. 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 $\tilde{\mu}'_{\pi hl} = \lambda$ and $\tilde{\mu}'_{\pi lh} = 1$. 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, $\tilde{\mu}'_{\pi hl} = 1$, $\tilde{\mu}'_{\pi lh} = \lambda$, 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

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

and

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

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 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:

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

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

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

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

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

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

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

We can break these constraints into four cases, with $\pi(\tilde{e}_l, w_l) \lesseqgtr \pi(\tilde{e}_h, w_h)$ and $c(\tilde{e}_h) \lesseqgtr c(\tilde{e}_l)$. 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

$$\begin{aligned} & p(b\tilde{e}_h - w_h) + (1-p)(b\tilde{e}_l - w_l) & > \; & b\underline{e} - \underline{w} \\ \Rightarrow \; & \tfrac{\eta\alpha b^2\Delta K(1-p)p}{\gamma(1+\eta\lambda(1-p)+\eta p)} - \tfrac{\eta\lambda^2\alpha b^2\Delta K(1-p)p}{\gamma(1+\eta p+\eta\lambda(1-p))} - pw_h - (1-p)w_l & > \; & -\underline{w} \\ \Rightarrow \; & \tfrac{\eta\alpha b^2\Delta K(1-p)p}{\gamma}\left(\tfrac{1-\lambda^2}{1+\eta\lambda(1-p)+\eta p}\right) & > \; & pw_h + (1-p)w_l - \underline{w} \end{aligned} \tag{18}$$

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

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

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

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

$$\begin{aligned}
p(b\tilde{e}_h - w_h) + (1-p)(b\tilde{e}_l - w_l) &> b\underline{e} - \underline{w} \\
\Rightarrow \quad \tfrac{\eta\alpha b^2\Delta K(1-p)p}{\gamma(1+\eta)} - \tfrac{\eta\lambda^2\alpha b^2\Delta K(1-p)p}{\gamma(1+\eta\lambda)} - pw_h - (1-p)w_l &> -\underline{w} \\
\Rightarrow \quad \tfrac{\eta\alpha b^2\Delta K(1-p)p}{\gamma}\left(\tfrac{1}{1+\eta} - \tfrac{\lambda^2}{1+\eta\lambda}\right) &> pw_h + (1-p)w_l - \underline{w}
\end{aligned} \tag{19}$$

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

As described in text.

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. If workers accurately anticipate a certain gift wage, this wage must 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.