STRATHCLYDE

DISCUSSION PAPERS IN ECONOMICS

A THEORY OF WAGE SETTING BEHAVIOR

BY

MARCO FONGONI AND ALEX DICKSON

NO 15-05

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STRATHCLYDE
GLASGOW

A Theory of Wage Setting Behavior*†

Alex Dickson and Marco Fongoni‡

Department of Economics, University of Strathclyde, Glasgow, UK

February 17, 2016

Abstract

In this paper we provide a micro-foundation for wage rigidity in a simple and tractable model of wage setting behavior, inspired by a synthesis of recent convergent insights from anthropological and experimental research, and drawing on concepts advanced in the behavioral economics literature. The core principles underlying our theory are contractual incompleteness, fairness, reciprocity, and reference dependence and loss aversion in the evaluation of wage contracts by workers. The model establishes a wage-effort relationship that captures a worker's asymmetric reference-dependent reciprocity, in which loss aversion implies effort responds more strongly to wage changes below the reference wage than above it. This basic relationship gives rise to wage rigidity around a worker's reference wage. We explore these implications further in a simple dynamic stochastic environment in which a worker adapts their feelings of entitlement once they become employed. The model allows us to shed new light on the importance of anticipated negative reciprocity and the cost of wage rigidity for a farsighted firm's hiring and wage setting behavior.

JEL Codes: C78, J30, J41.

*We would like to thank Julia Darby for very helpful discussion on this work which has helped to shape the direction of our thinking. Seminar participants at the University of Edinburgh provided valuable discussion, and we particularly acknowledge the comments of Mike Elsby and Tim Worrall. Thanks are also due to Rod Cross and Rodolphe Desbordes for constructive feedback on the manuscript.

†This version of the paper is an update on the original version made available in July 2015 (which is available from the authors upon request).

‡Corresponding author: e-mail: marco.fongoni@strath.ac.uk.

1 Introduction

Virtually every macroeconomic model of the business cycle—be it based on efficiency wages, search frictions, New Keynesian imperfections, or a combination of these—requires wage rigidity to be built-in to sufficiently explain unemployment fluctuations. Whilst compelling reasons might be advanced to *assume* wage rigidity, a convincing microeconomic account of wage setting behavior that captures what we know of employment relationships is much more desirable. When discussing workers' basic predispositions against wage reductions Okun [1981, p.10] noted "This is a plausible story, even though it cannot be deduced from a basic utility function!" and some three decades later Elsby [2009, p.155, footnote 2] reminds us that this gap in the literature remains: "Given the empirical evidence for worker resistance to wage cuts, it is surprising that there has not yet been an explicit model of such wage rigidity in the literature". The aim of this paper is to provide a theoretical foundation for wage rigidity in a simple and tractable microeconomic model of wage setting, inspired by ideas advanced in the behavioral economics literature and a synthesis of recent convergent insights from anthropological and experimental research.

Our model accounts for workers assessing the fairness of the wage they are paid in relation to their reference 'fair' wage, and responding to this with their choice of effort. Consideration of the worker's perceptions of fairness and the particular impact of unfair wages on morale and productivity for loss averse workers influence the optimal wage setting policy of a firm, tempering the incentive to adjust the wage to exogenous changes in external market conditions and hence giving rise to wage rigidity. In addition, the paper makes two further key contributions: i) it offers a psychological foundation for asymmetric reciprocity, identifying loss aversion as the driver of negative reciprocity being stronger than positive reciprocity; and ii) it analyzes the implications of 'asymmetric reference-dependent reciprocity' and anticipated wage rigidity for a firm's hiring and wage setting behavior in a dynamic decision making environment.

The basic premise of our theory is that there is contractual incompleteness over effort in an employment relationship, which is at least in part discretionary. A worker evaluates the fairness of the wage they are paid relative to a reference 'fair' wage. A wage that exceeds the reference wage is perceived as a gift, whilst if it falls below the reference wage it is perceived as unfair. Central to our model is the inclusion of a 'morale function' in the worker's payoff, which measures their evaluation of the productive effort they undertake in light of the fairness of the wage they are paid: if the worker is paid a wage they perceive as a gift they will receive an increase in utility from increasing their effort (a gift to the firm); whilst if they perceive their wage to be unfair the worker's utility will increase by reducing their effort (reducing the firm's payoff). As such, a worker's payoff exhibits both positive and negative reciprocity, which stems from their reference-dependent preferences. In addition, by allowing for the worker to be loss averse, which implies that unfair wages generate a disproportionate decrease in payoff, our theory captures an asymmetry in reciprocity—negative reciprocity is stronger than positive reciprocity—that seems a prevalent behavioral feature of employment relationships.

The inclusion of reciprocity establishes a wage-effort relationship in which the optimal effort of the worker is increasing in the wage paid. If a worker is loss averse the asymmetry between the effect of negative and positive reciprocity implies there is a kink in the optimal effort function at the reference wage. We define this wage-effort relationship as the worker's 'asymmetric reference-dependent reciprocity'. If a firm is considering paying a loss averse worker below their reference wage there will be a relatively large negative impact on their effort and consequent reduction in the output produced, consideration of which gives rise to wage rigidity.

We explore the implications of our theory in a simple two-period dynamic stochastic environment in which the evolution of the match productivity is uncertain, and the worker adapts their feelings of entitlement once they become employed: whilst the worker's reference wage at the start of the employment relationship is exogenously given, in the subsequent period of employment it is endogenously determined by the wage paid in the initial employment contract. As the employment relationship passes from one period to the next the firm may seek to renegotiate the wage with the worker due to changes in the state of the economy: for instance, absent other considerations, the firm may want to cut the worker's wage following a decline in the match productivity. However, the consequences for morale, effort and productivity of doing so may outweigh the benefit. Such downward rigidity of the initial wage contract is therefore an inherent feature of our dynamic model, driven by the worker's adaptation of the reference wage and the relatively large cost to the firm of negative reciprocity that stems from loss aversion. A farsighted firm will anticipate the future cost of negative reciprocity, understanding the persistence and irreversibility of the initial wage within an employment relationship. Consequently, they will adjust their behavior in the initial period. A key insight that arises from our model is that the anticipation of the worker's negative reciprocity during the employment relationship puts downward pressure on the initial wage offered to a loss averse worker and, since loss aversion generally reduces the value of an employment contract, upward pressure on the firm's hiring reservation productivity.

Our model can support a narrative that accounts for several aspects of wage setting behavior in the employment relationship, explain stylized facts of labour markets, and also provide novel insights and interpretations of worker and firm behavior. For instance, our framework accounts for: asymmetries in intensity and persistence of reciprocity; the rationale for increasing wage profiles; the coexistence of upward and downward wage rigidity which depends on the initial employment conditions; and the importance of wage rigidity for hiring behavior by distinguishing newly hired workers' wages from those of incumbents.

Being based on established principles in behavioral economics, our approach has the advantage of being clear about the nature of the behavioral forces at play, their driving factors and their implications, which contributes to our understanding of wage rigidity in employment relationships. By considering that workers' evaluation of wages is reference dependent, our model captures both positive and negative reciprocity and identifies loss aversion as the source of wage rigidity. Akerlof [1982] captured what we now know as positive reciprocity in a model of gift exchange, and Akerlof and Yellen [1990] in their fair wage-effort hypothesis considered what we now know as negative reciprocity; by incorporating both and deriving, rather than assuming, the effort response of workers to wages offered, we can consider the relative strength of these two effects. The nature of the employment relationship that emerges from our analytical framework lends support to reduced-form relationships that have been assumed in important recent contributions to the literature, for example, the reduced-form effort function used by Elsby [2009] and the reduced-form reference-dependent production function implemented by Eliaz and Spiegler [2013]. Perhaps closest in spirit to our approach is the contribution of Danthine and Kurmann [2007] who present a macroeconomic model based on micro-foundations, where workers' preferences exhibit reciprocity à la Rabin [1993] from whence gift exchange can be derived, but they don't capture asymmetric reference-dependent reciprocity, which is crucial for our results.

In the next section we elucidate our synthesis of the employment relationship. We set out our model based on contractual incompleteness, fairness, reciprocity, and reference dependence and loss aversion in section 3, where we also derive the implications for the worker's optimal effort decision and the wage-setting rule of the firm. Section 4 explores the implications of our model in a simple dynamic framework. In section 5 we discuss the logical implications of our model in relation to the role of the worker's reference wage and the literature on labor markets fluctuations, and section 6 offers some concluding remarks. All proofs are contained in the appendix.

2 Morale, fairness and reciprocity in employment relationships

There is an emerging consensus in the literature that behavioral concerns such as fairness, workers' morale and reciprocity influence firms' wage setting behavior. These intrinsic aspects of the employment relationship are also considered to be key behavioral forces that underly the observation of downward wage rigidity. In this section we argue that these ideas have been considered in the literature at least since the turn of the twentieth century, and that it is thanks to recent convergent findings in anthropological and experimental research combined with theories from behavioral economics that a unified consensus has emerged. Hence we propose a synthesis of the literature that will provide the underlying conceptual framework for the theory and the analysis developed in this paper.

2.1 Early insights

That workers' morale is linked with their perceptions of fairness and their productivity, and that employers are concerned about these issues when deciding upon wage policies, has long been acknowledged by economists. Marshall [1890] often expressed the reasons why employers would pay workers high wages and discussed the negative impacts on 'efficiency' and work 'intensity' of otherwise lower wages. Slichter [1920] placed workers' feelings of being treated unfairly as one of the most important causes of low morale and the resulting non-cooperative behavior of workers towards the employer. Hicks [1963], Solow [1979] and Okun [1981] advanced similar arguments when discussing the possible sources of the Keynesian wage floor [Keynes, 1936]: they argued that resistance to cut nominal wages comes from employers, concerned about the effects of wage cuts on workers' morale, 'ability' and 'willingness to work' [Hicks, 1963, p. 94-95]. The gift-exchange model of Akerlof [1982] and the fair wage-effort hypothesis of Akerlof and Yellen [1990] provide the first contributions that formalize some of these insights, appealing to what has become known in the behavioral economics literature as positive and negative reciprocity.

2.2 Anthropological evidence

Within the last three decades, thanks to the ground-breaking work of several economists including Blinder and Choi [1990], Campbell and Kamlani [1997], Bewley [1999] and more recently Galuscak et al. [2012], Druant et al. [2012] and Du Caju et al. [2014], our understanding of the employment relationship has been greatly enhanced.[1] By interviewing firms' managers and labor leaders in several countries these studies provide insight into the validity of the behavioral assumptions advanced in the theoretical literature.

A central finding is that firms' managers are concerned about treating workers fairly. Wage reductions that are perceived as unfair damage morale, inducing grievance among workers who negatively reciprocate the employer with lower effort and productivity [Bewley, 2007]. On the other hand wage rises could generate improvements in effort and cooperation among workers. As such, these findings suggest the existence of a relationship between wage *changes* and workers' effort.

Such a relationship is not universally straightforward, however: for instance Campbell and Kamlani [1997] find that effort responds more intensely to wage cuts than to increases in wages, and that any positive effect of wage increases on effort is believed to be temporary by managers, since workers rapidly get used to the wage received. On the other hand, wage reductions without impacts on morale are also achievable by employers, though only when workers understand their necessity in avoiding the firm shutting down or to prevent mass layoffs [Bewley, 2007].

2.3 Experimental evidence

There is an additional stream of evidence that comes from laboratory and field experiments. Overall the most important finding is confirmation of the existence of reciprocal behavior in the employment relationship: when people receive extra pay in excess of their standards of fairness they reciprocate with higher effort (positive reciprocity); when people perceive they have been treated unfairly they reciprocate by exerting minimum or lower effort (negative reciprocity). However, field experiments document evidence that positive reciprocity is weaker than negative reciprocity (see, for instance, Malmendier et al. [2014]).

[1] Other anthropological studies include those by Kaufman [1984], Baker et al. [1994], Agell and Lundborg [1995, 2003] and Agell and Bennmarker [2007]. Reviews of this literature can be found in Howitt [2002] and Bewley [2007].

In a combined laboratory and field experiment Cohn et al. [2014] try to address this inconsistency. They infer that positive reciprocity exists but may quickly disappear, which is consistent with the previously discussed anthropological findings.

One interpretation attributes this result to the asymmetric nature of workers' reciprocity behavior: negative reciprocity is stronger than positive reciprocity [Fehr et al., 2009]. Another interpretation suggests that the weak, or temporary, response of effort to wage rises is the outcome of a shift of the workers' standards of fairness to the higher wage received [Gneezy and List, 2006].

Taken together, evidence from laboratory and field experiments offer complementary insights into understanding the impacts of wage changes on effort, by reinforcing the existence of an asymmetric wage-effort relationship.

2.4 The proposed synthesis

We propose a synthesis of a theory of wage setting behavior that captures the essential features of wage setting and the employment relationship that emerge from several ideas and perspectives. We build the theory around four core concepts: workers' morale; their perceptions of fairness; reciprocity; and contractual incompleteness.

Workers' Morale. Morale represents the workers' state of mind when performing a productive activity. As concluded by Bewley [2007], good morale is not related to happiness or job satisfaction, but with the willingness of workers to cooperate and work to achieve the firm's goals; and when morale is low workers tend to hold back cooperation and cease to identify themselves with the firm. We capture the idea that workers' willingness to exert effort is directly related to their morale: when morale is good, cooperation is enhanced and performing the productive activity generates a psychological benefit; when morale is low, workers are less motivated and the psychological cost of exerting effort increases.

Perceptions of Fairness. Changes in workers' morale depend on whether workers feel they are treated fairly by their employer. Following the standard approach in the literature we capture these perceptions within a reference 'fair' wage relative to which the fairness of a wage contract is evaluated: a wage below the reference wage is perceived as *unfair*, while a wage above is perceived as a *gift*.[2] Moreover, by incorporating the intrinsic psychological

[2]The reference 'fair' wage is an artifact that simplifies the broader concept of workers' perceptions of fairness. The literature has captured the same concept with other names such as 'fair wage' in Marshall [1890] and Hicks [1963], 'wage norm' in Lindbeck and Snower [1986], 'perceptions of entitlement' in Kahneman et al. [1986], 'feelings of entitlement' in Hart and Moore [2008] and 'the reference frame of fairness judgements' in Fehr et al. [2009].

aspect of human decision making of loss aversion [Kahneman and Tversky, 1979], we capture that morale is most affected when workers feel they are being treated unfairly. This implies that a wage cut below the reference wage (perceived as a loss) has a greater impact on morale than a wage rise of the same amount (perceived as a gain).[3]

Reciprocity. The idea that is perhaps most prominent from the literature discussed is that the employment relationship is based on a mutual understanding of reciprocal behavior. If a firm sets a wage contract that is considered unfair, workers start to feel a grievance against the firm and morale will decrease. As a consequence effort will become more psychologically costly and workers will negatively reciprocate the treatment they perceive as unfair by exerting less effort. A similar response, but in the opposite direction, would arise if the firm sets a wage that is above the workers' reference wage. Moreover, due to the assumption that unfair behavior has a stronger impact on workers' morale (loss aversion), effort will be more responsive to wage changes that are considered unfair as opposed to wage changes considered as gifts. Thus, workers are characterized by intentions-based reciprocity [e.g. Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006], while firms, although self-interested, are concerned about fairness because of the effect of workers' responses to wage changes on profitability.[4]

Contractual Incompleteness. When thinking about firms' wage setting behavior, and more generally about the employment contract, we consider a negotiation in which an employer (the buyer) offers a wage in exchange for productive activity by a worker (the seller). However, unlike in goods markets, the employer is not able to contract upon the 'quality' of workers' productive activity: effort is discretionary and therefore not contractible. This peculiarity of labor markets brought Okun [1981] to the conclusion that the employment relationship is governed by an 'invisible handshake' and Williamson [1985] to define the employment contract as an 'incomplete agreement'. According to Williamson [1985, p. 262-63], only the minimum job performance can be enforced by the contract (the 'perfunctory cooperation'), while workers 'enjoy discretion' about the quality of their service, in terms of cooperation, effort and efficiency (the 'consummate cooperation'). This latter aspect is influenced by the worker's evaluation of the fairness of the wage they are offered.

[3]This assumption is consistent with the evidence reported by surveys and experiments as discussed above. While wage rises have a weak impact on morale, unfair wage cuts damage workers' morale due to an 'insult effect' and a 'standard of living effect' [Bewley, 2007, p. 161]. Fehr et al. [2009, p. 377] argue that evidence of such behavior suggests the existence of 'reference-dependent fairness concerns'.

[4]This type of reciprocity is conceptually different from the idea of inequity aversion [Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000] used in a model of wage setting by Benjamin [2015]. Malmendier et al. [2014] compare the various models of reciprocity advanced in the literature.

The key intuition of our theory is that there is a cost associated with reducing wages. If workers perceive wage reductions to be unfair their effort and willingness to work will drop, reducing productivity and influencing the firm's profitability. Thus the firm's managers may refrain from cutting wages in light of adverse economic conditions if, at the margin, the related cost in the form of negative reciprocity is greater than the benefit of paying lower wages. This insight confirms the predictions of early prominent hypotheses, and offers a psychological foundation for it, drawing from recent anthropological and experimental research.

3 The model

We begin by considering the wage-setting behavior of a firm for a single period of employment with an established worker to illustrate ideas. The firm learns the match productivity q and the worker's exogenously-given reference wage r at the start of the employment period. It then decides whether to continue the employment relationship and, if so, the wage w to offer to the worker. The match productivity is a realization of the random variable Q which is distributed on $[0, \infty]$ with cumulative distribution function F and density function f, and captures the interaction between the firm's technology, per-worker capital and the intrinsic productivity of the worker. For simplicity, we assume the worker will accept any contract offered.[5] After considering the wage in relation to their reference wage, the worker decides on a (non-negative) level of effort e which generates output for the firm. Payoffs are then realized, the form of which we describe next.

3.1 Payoffs

The per-worker output in an employment relationship (the price of which is normalized to one) is a function of both the match productivity and the effort chosen by the worker, and is denoted $y(q, e)$. The per-worker cost of production is $s(w)$ where w is the wage paid to the worker. The firm is materially motivated by profit:

$$\pi(w; q, e) = y(q, e) - s(w). \tag{1}$$

We make the following assumptions:

F1. $s'(w) > 0$ and $s''(w) \geq 0$ for all w.

[5] We choose not to include a reservation utility for the worker but it is straightforward to do so, the effect of which is to add an additional threshold to the model.

F2. $y^e(q,e), y^q(q,e) > 0$, $y^{ee}(q,e), y^{qq}(q,e) \leq 0$ and $y^{qe}(q,e) > 0$ for all e and q.

Notice that assumption F2 implies the marginal product of effort is increasing in the match productivity.

To capture behavior consistent with our synthesis—namely: reference dependence and loss aversion of the wage in relation to the reference wage; and reciprocity in relation to the worker's perception of the fairness of the wage—we specify the worker's preferences by an additively separable utility function composed of a *wage utility*, u, and an *effort utility*, v:

$$U(e; w, r) = u(w, r) + v(e; w, r).$$

The *wage utility* $u(w, r)$ represents the worker's perceived utility from wage evaluations which we suppose is comprised of a standard utility function and a gain-loss function $n(w|r)$:

$$u(w, r) = m(w) + \xi n(w|r).$$

The function $m(w)$ captures the effect of absolute wage levels on the worker's utility, and $n(w|r)$ captures the worker's evaluation of the wage relative to the reference wage, the functional form of which we define shortly. The parameter $\xi \geq 0$ measures the worker's subjective weight of relative wage comparisons in the wage utility. We assume:

W1. $m'(w) > 0$ and $m''(w) < 0$ for all w.

The *effort utility* $v(e; w, r)$ that the worker derives from engaging in productive activity takes the form

$$v(e; w, r) = b(e) - c(e) + M(e; w, r),$$

where $b(e)$ represents the worker's intrinsic psychological benefit of being productive, and $c(e)$ their intrinsic psychological and physical cost of productive activity.[6] We assume:

W2. $b'(e) > 0$ and $b''(e) \leq 0$ for all e; $c'(e) > 0$ and $c''(e) > 0$ for all e; and $b'(0) > c'(0)$.

[6] In our model we want to capture the idea that 'normal effort' (i.e. effort that a worker would exert absent morale considerations) is not zero. This necessitates the inclusion of intrinsic benefits and costs of productive activity. Whilst this approach contrasts with, for example, shirking models in the efficiency wage literature (e.g. Shapiro and Stiglitz [1984]) it is consistent with the idea that workers perceive positive satisfaction from engaging with productive activity (see, for example, the discussion in Altmann et al. [2014, Appendix]). Inspired by the findings reported in Bewley [2007], that it is not wage levels but changes in wages that influence effort, normal effort should be a non-pecuniary concept and is therefore modeled as being independent of the wage.

The function $M(e; w, r)$ is the 'morale function', a key component of our model that lies at the heart of many of our results. We specify

$$M(e; w, r) \equiv g(e)n(w|r) \tag{2}$$

and make the following assumption:

W3. $g'(e) > 0$ and $g''(e) = 0$ for all e.

This allows us to define $g'(e) \equiv \zeta > 0$, and so we can write $M(e; w, r) = \zeta e n(w|r)$.[7]

Morale depends on the worker's evaluation of the wage in relation to the reference wage, the functional form of which we now define. We assume that $n(w|r) \equiv \mu(m(w) - m(r))$ where $\mu(\cdot)$ is a gain-loss value function that exhibits loss aversion in the spirit of Kahneman and Tversky [1979]. As such, for a loss averse worker the evaluation of utility differences between the wage and the reference wage will be steeper for wages below the reference wage than for those above it. We make the following assumption about the functional form of gain-loss utility.[8]

W4. $\mu(0) = 0$ and, whilst $\mu(x)$ is non-differentiable at $x = 0$, it is continuous with $\mu'(x) > 0$ and $\mu''(x) = 0$ for all $x \neq 0$. Moreover, for any $x > 0$, $\mu'(-x)/\mu'(x) \equiv \lambda \geq 1$.

Under this assumption, it follows that the gain-loss utility is piecewise-linear:

$$n(w|r) \equiv \mu(m(w) - m(r)) = \begin{cases} \eta(m(w) - m(r)) & \text{if } w \geq r \text{ and} \\ \lambda\eta(m(w) - m(r)) & \text{if } w < r \end{cases} \tag{3}$$

where $\eta > 0$ is a scaling parameter that represents the importance of gain-loss utility for the worker, and $\lambda \geq 1$ represents the worker's degree of loss aversion.[9]

The morale function, being dependent on $n(w|r)$, captures an additional psychological cost/benefit of productive effort associated with the worker's perception of fairness. If the wage exceeds the reference wage (it is perceived as a 'gift') the worker gains some additional benefit of productive effort and an increase in effort (a 'gift' to the firm) will increase utility.

[7]From a modelling perspective, the functional form of $M(\cdot)$ is similar to the function used by Danthine and Kurmann [2007] to derive the workers' reciprocal gift in terms of effort, in the spirit of Rabin [1993]. Our specification is fundamentally different, however, as we incorporate a gain-loss function in morale that allows for negative, as well as positive, reciprocity, and asymmetries between the two.

[8]Assumption W4 closely resembles the assumptions of Kőszegi and Rabin [2006] over the properties of their 'universal gain-loss function', except we do not have diminishing sensitivity.

[9]The gain-loss function enters the worker's utility twice, crucially in the morale function but also in the wage utility. Whilst the latter is not important for our theory (indeed, we allow $\xi = 0$), we believe a realistic model of wage evaluation should allow loss aversion to influence the utility received from a wage offer. This would become important if we modeled the worker's reservation utility.

If the wage falls short of the reference wage (it is perceived as 'unfair') there is an additional psychological cost of productive effort and a reduction in effort increases utility. As such, the morale function implies the worker's payoff exhibits reciprocity, and since morale is linked to loss aversion, negative reciprocity is stronger than positive reciprocity.

Since we specify that morale depends on effort, these same considerations apply to the margins of the worker's payoff function: for a loss averse worker the reduction in the marginal utility of effort for an unfair wage will be larger than the increase in marginal utility for a wage an equivalent amount above the reference wage. Consequently, the effect on optimal effort, which is determined by these margins, will be asymmetric. We now turn to derive the relationship between effort and the wage.

3.2 The worker's choice of effort

A worker's choice of effort after the wage has been set will depend on their evaluation of the wage in relation to their reference wage. Given a reference wage r and a wage offer w the worker will seek to

$$\max_{e \geq 0} U(e; w, r).$$

Recalling that

$$U(e; w, r) = m(w) + \xi\mu(m(w) - m(r)) + b(e) - c(e) + \zeta e\mu(m(w) - m(r)),$$

and from (3) that $\mu(m(w) - m(r))$ is piecewise-linear, we denote by $\tilde{e}(w, r, \lambda)$ the utility-maximizing effort, which satisfies the first-order condition

$$\Omega(e, w, r, \lambda) \equiv b'(e) - c'(e) + \zeta\mu(m(w) - m(r)) \leq 0, \tag{4}$$

in which the inequality is replaced with an equality if $e > 0$.[10]

The following theorem defines the properties of the worker's optimal effort function, which exhibits 'asymmetric reference-dependent reciprocity'.

Theorem 1. *The worker's optimal productive effort function takes the form*

$$\tilde{e}(w, r, \lambda) = \begin{cases} \tilde{e}(w, r)^{+} & \textit{if } w > r \\ \tilde{e}_n & \textit{if } w = r \\ \tilde{e}(w, r, \lambda)^{-} & \textit{if } w < r \end{cases} \tag{5}$$

[10]The second-order sufficient condition is $b''(e) - c''(e) < 0$, which is satisfied under assumption W2.

where $\tilde{e}_n$ is 'normal' effort given by the solution to $b'(e) = c'(e)$; $\tilde{e}(w,r)^+(> \tilde{e}_n)$ is implicitly defined by $b'(e) + \zeta\eta(m(w) - m(r)) = c'(e)$; and $\tilde{e}(w,r,\lambda)^-(< \tilde{e}_n)$ is implicitly defined by $b'(e) \leq c'(e) + |\zeta\lambda\eta(m(w) - m(r))|$ (with equality if $e > 0$, which is true for all $w > \underline{w}(r,\lambda)$, defined in the proof).

For a given r, the optimal effort $\tilde{e}(w,r,\lambda)$ is a continuous function of w with $\tilde{e}^w(w,r,\lambda) > 0$ and $\tilde{e}^{ww}(w,r,\lambda) < 0$ for all $w \neq r$. Moreover,[11]

$$\lim_{\epsilon \to 0} \tilde{e}^w(r-\epsilon, r, \lambda)^- = \lambda \cdot \lim_{\epsilon \to 0} \tilde{e}^w(r+\epsilon, r)^+,$$

so the effort function has a kink at $w = r$ if $\lambda > 1$. For a given w, the optimal effort $\tilde{e}(w,r,\lambda)$ is a continuous function of r with $\tilde{e}^r(w,r,\lambda) < 0$ for all $w \neq r$. Finally, whilst optimal effort above the reference wage is independent of λ, below the reference wage $\tilde{e}^\lambda(w,r,\lambda)^- < 0$ and $\tilde{e}^{w\lambda}(w,r,\lambda)^- > 0$.

Theorem 1 establishes that if a worker is paid their reference wage then they will exert normal effort, $\tilde{e}_n$. We identify a positive relationship between the wage and effort where the effect on effort from a change in the wage is asymmetric for a loss averse worker due to a kink in the effort function at the reference wage. This is illustrated in Figure I.

FIGURE I.

The positive relationship between effort and the wage is driven in our model by the morale function: an increase in the wage gives a higher marginal utility of effort which consequently results in higher optimal effort. The asymmetric nature of effort responses has the particular implication that for changes in the wage from an initial wage equal to the reference wage, the effect of negative reciprocity that results from a reduction in the wage will be greater than the effect of positive reciprocity resulting from an increase in the wage. The extent of this 'asymmetric reference-dependent reciprocity' depends on the worker's degree of loss aversion: whilst when the wage exceeds the reference wage optimal effort is independent of λ, below the reference wage more loss averse workers exert less effort which decreases faster as the wage gets further from the reference wage. Indeed, if a worker is not loss averse ($\lambda = 1$), reciprocity is symmetric.

[11] Throughout the paper, where sequences of ϵ are considered over which limits are taken we specify that $\{\epsilon_n\}_{n=1}^{\infty} \subset \mathbb{R}_+$, meaning that where the wage is specified to be $r - \epsilon$ and we take the limit as $\epsilon \to 0$ we consider the wage increasing to the reference wage, and likewise when the wage is specified to be $r + \epsilon$ and we take $\epsilon \to 0$ we consider the wage decreasing to the reference wage.

The derived effort relationship establishes the model as providing a micro-foundation for effort functions that exhibit asymmetric reciprocity that are commonly assumed in the literature, but that are not explicitly modeled.[12] This micro-foundation is based on perceptions of fairness coupled with loss aversion in the employment relationship, and is consistent with the evidence that suggests that workers are subject to such forces that was summarized in Section 2.[13]

3.3 The firm's wage setting rule

Next we consider the firm's problem in setting the wage given that it deduces the behavior of the worker in response to the wage offer. Suppose a firm is facing a worker who has an exogenously-given reference wage r and for whom the match productivity is q. The firm will seek to maximize its payoff given in (1) where the worker's effort is determined as in (5). As such, the firm's problem is to

$$\max_{w \geq 0} \pi(w; q, \tilde{e}(w, r, \lambda))$$

where $\pi(w; q, e) = y(q, e) - s(w)$. The firm will continue the employment relationship with the worker at the optimal wage if it is profitable to do so (which, as noted, we assume the worker will accept), otherwise the employment relationship will be terminated.

The firm's output depends on the match productivity and on the effort of the worker, and the firm seeks to balance the marginal product of labor with the per-worker marginal cost. Consider two workers with differing match productivity that are otherwise identical (and in particular have the same reference wage). Conventional thinking implies that the worker with the lower productivity should be paid a lower wage. The model outlined here captures the idea that if the firm paid the worker its preferred wage for a given match productivity and this falls below the worker's reference wage, the worker's morale may be affected that will impact their effort and therefore the output produced from the employment relationship. This implies a cost of reducing the wage below the reference wage borne from the effect of negative reciprocity which may be large if the worker is

[12] For example, our model is consistent with the reduced-form effort function assumed by Elsby [2009] (except that normal effort depends on the wage which, as discussed, we have chosen not to model), and with the reduced-form reference-dependent output function assumed by Eliaz and Spiegler [2013]. It is also consistent with the assumed effort function of Akerlof [1982] in his gift exchange model (when $w > r$) and the fair wage-effort hypothesis of Akerlof and Yellen [1990] (when $w < r$).

[13] The idea that loss aversion gives rise to asymmetric reciprocity in the effort relationship is consistent with the intuition presented in, for example, Campbell and Kamlani [1997], Mas [2006], Bewley [2007] and Fehr et al. [2009].

loss averse, and must be considered in relation to the benefit of paying the lower wage: asymmetric reference-dependent reciprocity on the part of workers influences the marginal considerations of the firm.

We showed in Theorem 1 that the worker's optimal effort function $\tilde{e}(w, r, \lambda)$ is continuous in the wage, but that there is a kink at $w = r$ for a loss averse worker. The firm's payoff function, being otherwise smooth, will inherit this property. The fact that the profit function is continuous but has a kink at $w = r$ allows us to derive the optimal wage setting rule, that we explain intuitively below before formally presenting it in Theorem 2. For $w \neq r$ define the marginal profit as

$$\Psi(w; q, r, \lambda) = \frac{\mathrm{d}\pi(w; q; \tilde{e}(w, r, \lambda))}{\mathrm{d}w} = y^e(q, \tilde{e}(w, r, \lambda))\tilde{e}^w(w, r, \lambda) - s'(w),$$

and note that the profit function is concave[14]. Subject to the optimal wage contract being profitable, the optimal wage setting rule is characterized by two productivity thresholds, q_l and q_u. The upper threshold q_u is derived such that if $q > q_u$ then profit is increasing in the wage at wages just above the reference wage so concavity implies the optimal wage must exceed r and will be characterized by $\Psi(w; q, r, \lambda) = 0$ (in which $\tilde{e}(w, r, \lambda) = \tilde{e}(w, r)^+$ since $w > r$). The lower threshold q_l is derived such that if $q < q_l$ then profit is decreasing in the wage for wages just smaller than the reference wage, which by concavity implies the optimal wage will be below r and will be characterized by $\Psi(w; q, r, \lambda) = 0$ (in which $\tilde{e}(w, r, \lambda) = \tilde{e}(w, r, \lambda)^-$ since $w < r$). For $q_l \leq q \leq q_u$ it will be the case that for all wages below the reference wage $\Psi(w; q, r, \lambda) > 0$ and for all wages above the reference wage $\Psi(w; q, r, \lambda) < 0$, and therefore profit will be maximized with a wage equal to the reference wage: any worker whose reference wage is r and whose productivity lies in this range will be paid the same wage, giving rise to what we call a 'range of rigidity'. If the optimal wage contract is unprofitable, which will be the case if the match productivity falls below a threshold that is influenced by the reference wage, the firm will end the employment relationship.[15]

Theorem 2. *The firm's optimal wage $\tilde{w}(r, q, \lambda)$ is a continuous function of q and r and*

[14] This is established by noting that profit is continuous, and for $w \neq r$, $\Psi^w(w; q, r, \lambda) = y^{ee}(q, \tilde{e}(w, r))\tilde{e}^w(w, r)^2 + y^e(q, \tilde{e}(w, r))\tilde{e}^{ww}(w, r) - s''(w) < 0$ under Assumptions F1 and F2 and the results of Theorem 1.

[15] In the Appendix following the proof of Theorem 2 we explore the implications of the support of the distribution of match productivity being $[\underline{q}, \bar{q}]$, where conceivably a worker may have a reference wage such that the firm may wish to pay below the reference wage no matter how large the match productivity, or always pay above the reference wage no matter how small the match productivity.

is characterized by

$$\tilde{w}(r,q,\lambda) = \begin{cases} \tilde{w}(r,q)^+ & \text{if } q > q_u(r) \\ r & \text{if } q_l(r,\lambda) \leq q \leq q_u(r) \\ \tilde{w}(r,q,\lambda)^- & \text{if } q < q_l(r,\lambda), \end{cases}$$

so long as $q \geq \underset{\sim}{q}(r,\lambda)$*, the firm's reservation productivity, where*

$$\begin{aligned} q_l(r,\lambda) &= \{q : \lim_{\epsilon\to 0} \Psi(r-\epsilon; q, r, \lambda) = 0\}, \\ q_u(r) &= \{q : \lim_{\epsilon\to 0} \Psi(r+\epsilon; q, r, \lambda) = 0\}, \ \text{and} \\ \underset{\sim}{q}(r,\lambda) &= \max\{0, q : \pi(\tilde{w}(r,q,\lambda); q, \tilde{e}(\tilde{w}(r,q,\lambda), r, \lambda)) = 0\} \end{aligned}$$

(all singletons).

$\tilde{w}(r,q)^+$ *(*$> r$*) is the solution to* $\Psi(w;q,r,\lambda) = 0$ *in which* $\tilde{e}(w,r,\lambda) = \tilde{e}(w,r)^+$*, and* $\tilde{w}(r,q,\lambda)^-$ *(*$< r$*) is the solution to* $\Psi(w;q,r,\lambda) \leq 0$*, with equality if* $w > \underline{w}(r,\lambda)$*, in which* $\tilde{e}(w,r,\lambda) = \tilde{e}(w,r,\lambda)^-$*.* $\tilde{w}^q(r,q,\lambda) > 0$ *for all* $q \in [\underset{\sim}{q}(r,\lambda), \infty] \setminus [q_l(r,\lambda), q_u(r)]$*;* $\tilde{w}^r(r,q,\lambda) > 0$ *for all* $\underset{\sim}{q}(r,\lambda) \leq q \leq \infty$*; and* $\tilde{w}^\lambda(r,q,\lambda)^- > 0$ *for all* $\underset{\sim}{q}(r,\lambda) \leq q < q_l(r,\lambda)$ *(if this range is non-empty).*

The productivity thresholds defining the wage setting rule satisfy

a) $q_l(r,1) = q_u(r)$ *and* $q_l^\lambda(r,\lambda) < 0$ *implying* $q_l(r,\lambda) < q_u(r)$ *for all* $\lambda > 1$*; and*

b) $q_u'(r) > 0$ *and* $q_l^r(r,\lambda) > 0$*,*

and the reservation productivity has the properties that $\underset{\sim}{q}^r(r,\lambda) > 0$ *and* $\underset{\sim}{q}^\lambda(r,\lambda) \geq 0$*, where the final inequality is strict if* $\tilde{w}(r, \underset{\sim}{q}(r,\lambda), \lambda) < r$*.*

FIGURE II.

Theorem 2 elucidates the features of the firm's optimal wage setting rule when facing a loss averse worker, which is illustrated in Figure II. The optimal wage is non-decreasing in the match productivity and if the worker is loss averse there is a range of match productivity within which the wage is not adjusted. The lower and upper thresholds of this range depend on the worker's reference wage. If the match productivity is $q_u(r)$ the worker will be paid their reference wage which is perceived to be fair and so will exert normal effort. If the match productivity exceeds $q_u(r)$ then the firm will find it profitable to pay above the reference wage since such a wage, seen as a gift, will be positively reciprocated with higher than normal effort which is valuable from these relatively productive workers. If the match productivity is slightly less than $q_u(r)$ the firm would look to reduce the wage, but understands that paying a wage below the reference wage will be seen as unfair and generate a reduction in effort, the value of which exceeds the wage cut despite the relatively low match productivity. This is true for all $q \in [q_l(r, \lambda), q_u(r)]$, identifying the *range of rigidity* which is non-empty if the worker is loss averse. If the match productivity is below $q_l(r, \lambda)$ then the very low match productivity warrants a wage below the reference wage despite the negative impact on effort this will have through the worker perceiving this wage as unfair.

The nature of the wage setting rule hints at wage rigidity not only between workers that have different match productivity but that are otherwise identical, but also for a worker over time. Consider a simple dynamic framework in which a worker's reference wage remains the same between periods. Then the worker may be paid the same wage in consecutive periods despite a reduction in the match productivity as the firm seeks to avoid inciting negative reciprocity that would come from cutting the wage in response to a lower match productivity. We further explore a dynamic version of the model in the sequel.

If a worker has a greater degree of loss aversion then the effect of negative reciprocity borne from paying a wage below the reference wage is stronger, resulting in a greater reduction in effort which is more costly to the firm. As such, for a more loss averse worker, the lower threshold of the range of rigidity is lower since the firm is unwilling to suffer the relatively high cost of negative reciprocity even if the match productivity is relatively low: $q_l^\lambda(r, \lambda) < 0$. If the match productivity is low enough that the firm wishes to pay a wage below the reference wage, that wage will be higher the more loss averse a worker is as the firm has an incentive to attenuate some of the effect on effort from stronger reciprocity: $\tilde{w}^\lambda(r, q, \lambda)^- > 0$.[16] Negative reciprocity not only tempers the firm's incentive to cut the wage, it also reduces the extent to which the wage is cut. Finally, the more loss averse a worker is, the higher the match productivity the firm requires from the employment relationship for it to be profitable: $\underset{\sim}{q}^\lambda(r, \lambda) > 0$. This follows since the firm pays a higher wage in an attempt to mitigate the effect of negative reciprocity on

[16] This implication has been theoretically derived and empirically corroborated by Holden and Wulfsberg [2014], who show that "even if the wage is cut, the resulting wage will be higher than if the wage-setting process had been completely flexible". In contrast with their theoretical model, our theory attributes this result to the worker's extent of negative reciprocity.

effort but optimally doesn't fully mitigate it leading to lower effort from the worker and reduced profit; hence the reservation productivity above which the firm becomes profitable increases.

The effect of a worker feeling entitled to a higher wage (i.e. a higher reference wage) increases the optimal wage offered by the firm and increases both the threshold above which the wage is higher than the reference wage, and the threshold below which the wage is less than the reference wage. In addition, since for any match productivity the firm receives lower profit, the range of match productivity over which the employment relationship continues shrinks.

The features of the wage setting rule highlighted in Theorem 2 imply that when a worker is loss averse a range of rigidity exists and it is larger for individuals that are more loss averse. For a worker that is not loss averse ($\lambda = 1$) there is no range of rigidity and the wage responds smoothly to productivity.

4 Dynamic wage setting

We now turn to explore the implications of capturing reciprocity and reference-dependence in the employment relationship in a dynamic environment that is subject to uncertainty and, inspired by the literature that suggests reference points are influenced by previous contractual arrangements, workers adapt their reference wage to the wage they have been paid in the past.[17]

The timing of the model is as follows (we augment our notation in an obvious way with time subscripts). At time 0 the productivity characterizing the match, q_0, is observed, as is the exogenously given reference wage of the worker, r_0. Knowing these, the firm then decides whether to offer a wage contract to the worker and start the employment relationship. For simplicity we assume that any offer is accepted by the worker. If an employment relationship is established, then at the end of the first employment period the match productivity for the subsequent period, q_1 (which is independent of q_0), is randomly

[17] According to Kahneman et al. [1986] when workers enter a firm there is a shift in their feelings of entitlement and the most recent negotiated wage is adopted as the standard of fairness. This sort of adaptation is believed to be an active behavioral feature of workers' perceptions of fairness, supported by the anthropological evidence surveyed by Bewley [2007] and by laboratory and field experiments of the employment relationship (see, for instance, Chemin and Kurmann [2012] and Koch [2014]). In contract theory the idea of "contracts as reference points" has been analyzed by Hart and Moore [2008] and further explored by Herweg and Schmidt [2015]. The laboratory experiments of Fehr et al. [2011, 2014] and Bartling and Schmidt [2015] provide strong support for this assumption, which also reflects the idea that past experience and adaptation play a significant role in the process of individuals' reference point formation (see Herz and Taubinsky [2014] and Smith [2015] for evidence of this hypothesis, and Stommel [2013] for a review of the literature).

re-drawn from the same distribution F. In addition, the worker adapts their reference wage to the wage paid in the initial period of employment. After observing q_1, and inferring the worker's new reference wage $r_1 = \tilde{w}_0$, the firm considers whether it wants to continue the employment relationship and, if so, whether to adjust the wage of the worker in light of the change in the match productivity. The timing of the employment relationship is illustrated in Figure III.

FIGURE III.

The firm therefore faces a two-period dynamic stochastic optimization problem in which it seeks to maximize the sum of its expected discounted profits from the employment relationship. Letting δ represent the firm's discount factor, this is characterized by

$$\begin{aligned} J_0(r_0, q_0) = \quad & \max_{w_0} \pi(w_0; q_0, e_0) + \delta \mathbb{E}_0[J_1(r_1, q_1)] \\ & \text{s.t. } r_1 = w_0 \text{ and} \\ & e_t = \tilde{e}(w_t, r_t, \lambda), \\ & \text{where } J_1(r_1, q_1) = \max_{w_1} \pi(w_1; q_1, e_1) \end{aligned} \tag{6}$$

which is solved recursively. Whilst we consider that the firm is forward looking, absent a link between the initial and the subsequent employment period, the worker will choose productive effort to maximise their per-period utility, in accordance with the derivation of optimal productive effort in Theorem 1.[18]

The analysis that follows is divided in two parts: first we consider the case of a myopic firm ($\delta = 0$) that ignores the link between the two employment periods, where we highlight that wage rigidity, both downward and upward, may occur as a result of workers being loss averse. Then we consider a farsighted firm ($\delta > 0$) where we characterize the optimal employment contract that solves the dynamic problem in (6), and explore its properties.

4.1 Myopic firm

If the firm is myopic then, ignoring the fact that the wage becomes the worker's subsequent reference wage, it will set the wage in each employment period to maximize the per-period profit given the state of the employment relationship (the match productivity and the

[18] Forward-looking behavior could be incorporated with expectations-based reference points, as in Kőszegi and Rabin [2006]. However, the evidence previously cited strongly suggests that workers' reference wages are backward looking, as in our model.

worker's reference wage). As such, the firm will adopt the standard wage setting rule as derived in Theorem 2, so in each period $t = 0, 1$ the optimal wage takes the form

$$\tilde{w}_t = \tilde{w}(r_t, q_t, \lambda) = \begin{cases} \tilde{w}(r_t, q_t)^+ & \text{if } q_t > q_u(r_t) \\ r_t & \text{if } q_l(r_t, \lambda) \leq q_t \leq q_u(r_t) \\ \tilde{w}(r_t, q_t, \lambda)^- & \text{if } q_t < q_l(r_t, \lambda) \end{cases}$$

so long as $q_t \geq \underset{\sim}{q}(r_t, \lambda)$, otherwise the firm will not offer, or renegotiate, the employment contract and the employment relationship will be over.

To illustrate the dynamics of wage setting behavior for a myopic firm, we now consider different scenarios the firm may face during the employment relationship assuming that in both the initial and subsequent employment periods the match productivity exceeds the firm's reservation productivity $\underset{\sim}{q}(r_0, \lambda)$ and $\underset{\sim}{q}(r_1, \lambda)$ respectively, reserving commentary on when this is not the case as a postscript.

FIGURE IV.

Suppose that, in the initial employment period, $q_0 > q_u(r_0)$ so the firm offers a wage that exceeds the reference wage to appeal to the worker's positive reciprocity: $\tilde{w}_0 = \tilde{w}(r_0, q_0)^+$ and the worker will exert supra-normal effort. As the employment relationship passes from the initial period into the subsequent period, the worker adjusts their feelings of entitlement, adapting their reference wage to their initial wage: $r_1 = \tilde{w}_0$. This 'shifts' the wage-setting rule, as illustrated in Figure IV: the reference wage increases, the lower threshold increases, and the upper threshold increases to somewhere between its previous value and the initial match productivity. If the production function exhibits constant returns to effort, $q_u(\tilde{w}_0)$ increases to exactly q_0.[19]

If $q_1 > q_0 (\geq q_u(\tilde{w}_0))$ then the match productivity in the subsequent employment period exceeds the upper threshold of the wage setting rule and the firm will increase the worker's wage to benefit from the gift being reciprocated with supra-normal effort. If $q_1 = q_0$ then $\tilde{w}_1 \geq \tilde{w}_0$: whilst a wage of $\tilde{w}_0$ was positively reciprocated by the worker when compared

[19] In Theorem 2 we demonstrated that $q_l^r > 0$ and $q_u' > 0$. Recall that $q_u(r)$ is the value of q where $\lim_{\epsilon \to 0} \Psi(r + \epsilon; q, r, \lambda) = 0$. If $\tilde{w}_0 = \tilde{w}^+$ then the first-order condition is satisfied with equality at the optimal wage in the initial contract: $\Psi(\tilde{w}_0; q_0, r_0, \lambda) = 0$. Recall (from the preliminaries in the proof of Theorem 2) that $\Psi^r = y^{ee}\tilde{e}^r\tilde{e}^w \geq 0$. Then, in the subsequent period where the worker's reference wage increases to $\tilde{w}_0$, we find that $\lim_{\epsilon \to 0} \Psi(\tilde{w}_0 + \epsilon; q_0, \tilde{w}_0, \lambda) \geq 0$, and therefore, since $\Psi^q > 0$, the value of q that regains equality of this expression with 0, which is precisely $q_u(\tilde{w}_0)$, will not exceed q_0. If $y^{ee} = 0$ then $\Psi^r = 0$ and $\lim_{\epsilon \to 0} \Psi(\tilde{w}_0 + \epsilon; q_0, \tilde{w}_0, \lambda) = 0$ so $q_u(\tilde{w}_0) = q_0$.

to a reference wage of r_0, with the worker's updated sense of entitlement effort with this wage will be merely normal—positive reciprocity is a temporary phenomenon when the worker adapts their reference wage[20]— and so the firm optimally pays at least this wage in the subsequent contract.

If $q_1 < q_0$ then the employment renegotiation depends on by how much the match productivity reduces: only if $q_1 < q_l(\tilde{w}_0, \lambda)$ will the firm implement a wage cut, by reducing the wage below wage paid in the initial employment period. Indeed, reductions in the match productivity may be associated with an increase in the wage (if they are small and in the range $[q_u(\tilde{w}_0), q_0]$, which is non-empty only with decreasing returns to effort). As such, a fall in match productivity over time ($q_1 < q_0$) is not necessarily followed by a wage cut: the worker's adaptation to a wage consistent with a match productivity of q_0 implies that, if the match productivity only moderately decreases, the firm will keep the wage equal to the worker's reference wage. The negative effect of what is now perceived as an unfair wage, borne through negative reciprocity, will be larger than the benefit of paying the lower wage and hence the firm will avoid inciting such negative reciprocity and will freeze the wage. Downward wage rigidity is an inherent feature of the employment contract in a dynamic environment, the key features that drive which are the worker's adaptation of the reference wage ($r_1 = \tilde{w}_0$), and the relatively large cost to the firm of negative reciprocity that derives from loss aversion ($\lambda > 1$).

In the case where the initial match productivity is such that $q_0 < q_l(r_0, \lambda)$ the firm will pay an initial wage below the worker's reference wage, despite the implied negative reciprocity. In the following period, the worker's reference wage reduces to the initial wage level,[21] and the firm's wage setting rule adjusts accordingly. In this case, the lower threshold of the range of rigidity reduces to somewhere between q_0 and $q_l(\tilde{w}_0, \lambda)$ with the implication that the employment relationship will exhibit *upward* wage rigidity if the subsequent match productivity increases but is less than $q_u(\tilde{w}_0, \lambda)$, since within this range the firm will pay a wage no higher than in the initial employment period. The intuition is that in the initial employment period the firm essentially 'over-paid' the worker in an optimal trade-off of wage vs negative reciprocity; in the subsequent period the firm will

[20] This implication of the model is consistent with the evidence reported by field research (experiments and surveys) that the positive effects of a wage gift on morale and effort are believed to be weak and only temporary by firms' managers. It also supports the interpretation of this evidence according to which positive reciprocity quickly disappears as workers get used to the wage they receive (see, for instance, Campbell and Kamlani [1997], Bewley [1999], Gneezy and List [2006] and Cohn et al. [2014]).

[21] Hence, adaptation is symmetric. In the discussion section we consider the implications of asymmetric adaptation of the reference wage.

pay the same wage even if the match is more productive, and will only increase the wage if the match becomes substantially more productive and the benefits of positive reciprocity warrant a wage rise.

FIGURE V.

Finally, if, in the initial employment period the match productivity is such that $q_l(r_0, \lambda) \leq q_0 \leq q_u(r_0)$ then the worker will be paid their reference wage, which consequently does not change between periods. As such, the wage setting rule in the subsequent employment period is exactly the same as in the initial period, implying that there may be both downward and upward wage rigidity if the subsequent productivity draw is such that $q_l(r_0, \lambda) \leq q_1 \leq q_u(r_0)$.

Consider the evolution of the firm's reservation productivity between periods; the initial determining hiring behavior, and the subsequent capturing the firm's layoff decision. If $q_0 < \underset{\sim}{q}(r_0, \lambda)$ then no contract is offered. Otherwise a contract is offered and, since the initial wage becomes the subsequent reference wage, when the firm comes to renegotiate the contract the reservation productivity will depend on the initial wage. We deduced in Theorem 2 that $\underset{\sim}{q}^r(r, \lambda) > 0$. This implies, in particular, that if the firm has initially paid a wage that exceeds the worker's reference wage, then in the subsequent period the firm's reservation productivity increases, implying the firm will re-contract only over a reduced subset of the support of the match productivity distribution. As such, it is not inconceivable for an employment relationship to be characterized by the same match productivity in both periods but, whilst the worker is hired and paid a wage above their reference wage in the initial period, in the subsequent period the firm doesn't renegotiate the contract.

Due to workers' loss aversion and adaptation of the reference wage, the value of the subsequent period employment contract is lower if a worker was initially paid a wage in excess of their reference wage. This implies that if a worker is paid a higher initial wage contract, there is a higher probability that the subsequent period match productivity will not be high enough to make the relationship profitable, increasing the likelihood of the worker being laid off. We further consider the implications of our model for layoff decisions and unemployment in Section 5.

4.2 Farsighted firm

If a firm is farsighted it will consider the link between the initial and the subsequent wage negotiation that comes from the worker's adaptation of their reference wage during the employment period to the wage in the initial contract. This influences the worker's future effort response as it is relative to the reference wage $r_1 = \tilde{w}_0$ that subsequent wage offers will be evaluated, which in turn influences the value of the continuing employment relationship for the firm.

Since $r_1 = w_0$, the optimization problem for a farsighted firm when setting the initial wage contract is

$$\max_{w_0} \pi_0(w_0; q_0, \tilde{e}(w_0, r_0, \lambda)) + \delta \mathbb{E}_0[J_1(w_0, q_1)]. \tag{7}$$

The expected value of future profit $\mathbb{E}_0[J_1(w_0, q_1)] = \int_{\underset{\sim}{q}(w_0,\lambda)}^{\infty} J_1(w_0, q_1) \mathrm{d}F(q_1)$ now also depends on the initial wage. Recognizing that the reservation productivity for this contract, below which the firm would lay off the worker, may fall anywhere in the support of the distribution of match productivity, this can be expressed as[22]

$$\begin{aligned} \mathbb{E}_0[J_1(w_0, q_1)] = & \int_{\underset{\sim}{q}(w_0,\lambda)}^{\max\{\underset{\sim}{q}(w_0,\lambda), q_l(w_0,\lambda)\}} J_1(w_0, q_1)^{-} \mathrm{d}F(q_1) \\ + & \int_{\max\{\underset{\sim}{q}(w_0,\lambda), q_l(w_0,\lambda)\}}^{\max\{\underset{\sim}{q}(w_0,\lambda), q_u(w_0)\}} J_1(w_0, q_1)^{=} \mathrm{d}F(q_1) + \int_{\max\{\underset{\sim}{q}(w_0,\lambda), q_u(w_0)\}}^{\infty} J_1(w_0, q_1)^{+} \mathrm{d}F(q_1). \end{aligned} \tag{8}$$

This expression highlights that the firm faces different realizations of future profit when setting the initial wage contract w_0, depending on whether the subsequent match productivity q_1 is below, within or above the range of rigidity defined by $q_l(w_0, \lambda)$ and $q_u(w_0)$. Attentive observation of equation (8) allows us to infer two important insights: when setting the wage in the initial employment period the firm influences: i) the level of the expected value of future profit $J_1(w_0, q_1)$; and ii) the range of the distribution of the future match productivity within which the firm will subsequently cut or freeze the wage at time 1, and lay off the worker by ceasing an unprofitable match.

Let $\theta(w_0, \lambda)$ be the marginal effect of a wage increase in the initial wage contract on the expected profit in period 1, then we have the following result.

Proposition 1.

$$\theta(w_0, \lambda) \equiv \frac{\mathrm{d}}{\mathrm{d}w_0} \int_{\underset{\sim}{q}(w_0,\lambda)}^{\infty} J_1(w_0, q_1) \mathrm{d}F(q_1) < 0.$$

[22] $J_1(w_0, q_1)^{-;=;+}$ represents the optimized profit when $w_1 < w_0; w_1 = w_0; w_1 > w_0$, in which effort is given by $\tilde{e}(w_1, w_0, \lambda)^{-}; \tilde{e}_n; \tilde{e}(w_1, w_0)^{+}$.

When setting the initial wage contract in a dynamic environment a farsighted firm will account for an additional expected future cost: a higher initial wage influences the worker's feelings of entitlement in the subsequent renegotiation, which consequently influences the worker's effort and the value of the contract to the firm. A marginal increase in the initial wage lowers the expected value of profit because if, relative to the initial wage, the firm wants to lower the wage then the effect of negative reciprocity is greater; if it wishes to freeze the wage then the wage paid is simply higher; and if it wants to increase the wage then the effect of positive reciprocity is lower.

We now turn to deduce the optimal wage contract of a farsighted firm, which we denote $\tilde{w}_0$. So long as $\tilde{w}_0 \neq r_0$ the necessary first-order condition that must be satisfied is

$$\Psi(w_0; q_0, r_0, \lambda) + \delta\theta(w_0, \lambda) = 0.$$

As we note in the proof of the following theorem that presents the properties of the optimal contract, an additional assumption is required to proceed with the analysis which is that the magnitude of the 'current direct effect' of a change in the wage on marginal profit in the initial contract, $|\Psi^w|$, is always larger than the 'discounted expected future indirect effect' on marginal profit that results from the initial wage becoming the reference wage, captured by $\delta\theta^{w_0}$.[23] From this point on we also make two further innocuous assumptions to ease notational burden.[24]

Theorem 3. *Consider a farsighted firm for whom $\delta > 0$ and assume $\Psi^w + \delta\theta^{w_0} < 0$. Then the optimal wage setting rule for the initial employment contract is*

$$\tilde{w}_0 = \hat{w}(r_0, q_0, \lambda, \delta) = \begin{cases} \hat{w}(r_0, q_0, \lambda, \delta)^+ & \text{if } q_0 > \hat{q}_u(r_0, \lambda, \delta) \\ r_0 & \text{if } \hat{q}_l(r_0, \lambda, \delta) \leq q_0 \leq \hat{q}_u(r_0, \lambda, \delta) \\ \hat{w}(r_0, q_0, \lambda, \delta)^- & \text{if } q_0 < \hat{q}_l(r_0, \lambda, \delta), \end{cases}$$

[23] These effects could work in opposite directions as the sign of the second derivative of the expected future profit function with respect to the initial wage θ^{w_0} remains undetermined. We believe that in our model current direct effects of wage changes will dominate the expected future indirect effects through the influence of the initial wage on the reference wage, but nevertheless note that the concavity of Ψ established in Theorem 2 implies the assumed inequality will hold if the firm is sufficiently impatient.

[24] These are: 1) any contract offered by a firm is not constrained by the lower bound on effort, i.e. the wage always exceeds $\underline{w}(r, \lambda)$, which implies that unless the optimal wage is equal to the reference wage the firm's first-order condition is satisfied with equality; and 2) in the second employment period the firm's reservation productivity $\underset{\sim}{q}(w_0, \lambda)$, which determines layoffs, is always less than the lower threshold of the range of rigidity $q_l(w_0, \lambda)$.

so long as $q_0 \geq \underset{\sim}{\hat{q}}(r_0, \lambda, \delta)$, *where*

$$\hat{q}_l(r_0, \lambda, \delta) = \{q_0 : \lim_{\epsilon \to 0} \Psi(r_0 - \epsilon; q_0, r_0, \lambda) + \delta\theta(r_0, \lambda) = 0\},$$
$$\hat{q}_u(r_0, \lambda, \delta) = \{q_0 : \lim_{\epsilon \to 0} \Psi(r_0 + \epsilon; q_0, r_0, \lambda) + \delta\theta(r_0, \lambda) = 0\} \text{ and}$$
$$\underset{\sim}{\hat{q}}(r_0, \lambda, \delta) = \max\{0, q_0 : J_0(r_0, q_0) = 0\}.$$

$\underset{\sim}{\hat{q}}(r_0, \lambda, \delta)$ *is the reservation productivity that governs hiring: if* $q_0 < \underset{\sim}{\hat{q}}(r_0, \lambda, \delta)$ *no contract is offered to the worker.*

The optimal wage $\hat{w}(r_0, q_0, \lambda, \delta)^+$ *(*$> r_0$*) is the solution to* $\Psi(w_0; q_0, r_0, \lambda) + \delta\theta(w_0, \lambda) = 0$ *in which* Ψ *is evaluated using* $\tilde{e}(w_0, r_0)^+$*, and* $\hat{w}(r_0, q_0, \lambda, \delta)^-$ *(*$< r_0$*) is the solution to* $\Psi(w_0; q_0, r_0, \lambda) + \delta\theta(w_0, \lambda) = 0$*, in which* Ψ *is evaluated using* $\tilde{e}(w_0, r_0, \lambda)^-$.

For all $q_0 \in [\underset{\sim}{\hat{q}}(r_0, \lambda, \delta), \infty] \backslash [\hat{q}_l(r_0, \lambda, \delta), \hat{q}_u(r_0, \lambda, \delta)]$, $\hat{w}^{q_0}(r_0, q_0, \lambda, \delta) > 0$ *and* $\hat{w}^{r_0}(r_0, q_0, \lambda, \delta) > 0$.

What is the effect of the link between the initial wage and the subsequent reference wage on a firm's wage setting behavior? Intuitively, since a farsighted firm perceives a future cost of raising the current wage due to a higher future reference wage and a higher probability of being in a situation to enact a costly wage cut/freeze, they will set a lower initial wage compared to a myopic firm in an otherwise identical employment relationship. The following proposition clarifies.

Proposition 2. *Consider a firm for whom* $\delta > 0$ *and* $\Psi^w + \delta\theta^{w_0} < 0$ *that faces a loss averse worker. Then* $\hat{w}(r_0, q_0, \lambda, \delta) \leq \tilde{w}(r_0, q_0, \lambda)$*, with a strict inequality whenever the wage is not equal to the reference wage* r_0.

Proposition 2 concludes that if a firm is farsighted it will have an incentive to compress the initial wage contract for a newly hired worker if downward wage rigidity may be a feature of the following employment period.[25] Wage compression is thus a general result of our theory of wage setting behavior.[26]

We now turn to investigate the effect of loss aversion on the nature of the dynamic employment contract, by considering its effects on the initial wage contract $\hat{w}(r_0, q_0, \lambda, \delta)$

[25] Indeed, implicit differentiation of the wage setting rule reveals

$$\hat{w}^\delta = -\frac{\theta}{\Psi^w + \delta\theta^{w_0}} < 0,$$

so the more a firm cares about the future the lower will be the initial wage offered.

[26] Elsby [2009] obtains a similar result in an infinite-horizon dynamic model for an ongoing employment relationship when downward wage rigidity binds; he also provides empirical evidence of this result, which is further corroborated by Stüber and Beissinger [2012]. Our theory is consistent with these findings and provides a framework that elucidates the behavioral forces at play.

and on the firm's hiring reservation productivity $\hat{\underset{\sim}{q}}(r_0, \lambda, \delta)$, and its layoff reservation productivity $\underset{\sim}{q}(\tilde{w}_0, \lambda)$. Recall that the worker's degree of loss aversion influences the strength of negative reciprocity if the firm finds itself in a position where it optimally pays a wage below the worker's reference wage; therefore, the cost of enacting wage cuts is directly related to the degree of loss aversion, which has an effect in both employment periods.

For a more loss averse worker the firm has a stronger incentive to reduce the gap between the wage paid and the reference wage, to attenuate the stronger effects of negative reciprocity whenever $\tilde{w}_0 < r_0$. In the initial employment period where the reference wage is fixed, this puts upward pressure on the wage if the firm faces a worker whom they would want to pay below their reference wage. However, the fact that this initial wage becomes the reference wage in the subsequent employment period puts downward pressure on the initial wage because the expected effect of negative reciprocity is larger for a more loss averse worker, and a lower reference wage reduces the magnitude of this effect.[27]

If the firm is considering an employment contract that is characterized by a sufficiently high initial match productivity such that the firm optimally pays at least the reference wage (i.e. $q_0 \geq \hat{q}_l$), then there is no current direct effect of loss aversion and the expected future cost of increased negative reciprocity means the firm will pay a worker with a higher degree of loss aversion a lower initial wage. If the firm is facing an employment relationship with an initial match productivity such that the optimal contract in the initial period calls for a wage below the reference wage (if $q_0 < \hat{q}_l$), the overall effect of a higher degree of loss aversion depends on the balance of the two effects, and how much the firm cares about the future cost of negative reciprocity. If the current effect dominates the expected future effect, then the initial wage will be higher for a more loss averse worker; if it doesn't and the firm is not too impatient then the initial wage will be lower for a more loss averse worker.[28]

We now discuss the effect of loss aversion on the determination of the firm's reservation productivity governing both hiring and layoff decisions. For hiring decisions, this depends on how loss aversion influences the firm's current and expected future profit, whilst for

[27] The conclusion that $\theta^{\lambda} < 0$ is subject to the qualification that the firm's lay-off reservation productivity doesn't increase too much with the degree of loss aversion: as we showed in Theorem 2 the firm's reservation productivity in the final period is increasing in the worker's degree of loss aversion so a more loss averse worker is more likely to be laid off which reduces the probability of the firm having to enact a wage cut, partially offsetting the increased expected cost of doing so.

[28] We deduced in Theorem 1 that $\tilde{e}^{w\lambda-} > 0$, so the negative effect of loss aversion on effort is stronger for wages just below the reference wage than when substantial reductions below the reference wage are considered. As such, the current effect will be larger the closer is the match productivity to the lower threshold $\hat{q}_l$.

layoff decisions it only depends on the profit in the subsequent employment contract. Loss aversion influences profit in two ways. First, there is a direct negative effect on effort if the wage is below the reference wage which acts to reduce profit and is present in both the initial and subsequent employment periods. Second, there is an indirect effect in the subsequent employment period that comes from the firm changing the initial wage that becomes the reference wage: if the initial wage increases this provides a compounding negative effect on effort in the subsequent contract if the wage is below the reference wage which lowers profit; whilst if the initial wage reduces there is an at least partially offsetting positive effect on effort which increases profit.

In the subsequent employment period, profit is influenced by both the direct and indirect effects described above. Consequently, the reservation productivity governing layoff decisions will increase either if the initial wage is higher (which lowers profit), or the negative direct effect dominates the positive indirect effect when the initial wage is lower.

In the initial employment period the firm accounts for the current direct effect of more loss aversion (which reduces profit), and the expected future direct and indirect effects discussed in the previous paragraph. We show in the following proposition that increased loss aversion unambiguously increases the hiring reservation productivity, where we also summarize the other effects of loss aversion discussed above.

Proposition 3. *Consider a firm for whom $\delta > 0$ and $\Psi^w + \delta\theta^{w_0} \leq 0$, and suppose that $\underset{\sim}{q}^{\lambda}$ is sufficiently small. Then $\theta^{\lambda}(w_0, \lambda) < 0$ and for $\lambda' > \lambda$,*

$$\hat{w}(r_0, q_0, \lambda', \delta) \gtreqless \hat{w}(r_0, q_0, \lambda, \delta) \Leftrightarrow \Psi^{\lambda} + \delta\theta^{\lambda} \gtreqless 0.$$

In addition,

a) $\underset{\sim}{\hat{q}}(r_0, \lambda', \delta) > \underset{\sim}{\hat{q}}(r_0, \lambda, \delta)$*; and*

b) $\underset{\sim}{q}(\hat{w}, \lambda') > \underset{\sim}{q}(\hat{w}, \lambda)$ *if either the wage in the initial contract increases, or the direct effect of loss aversion on effort exceeds the indirect effect.*

Proposition 3 highlights some ambiguity in how the firm will change its behavior when faced with a more loss averse worker, which is natural in this dynamic environment given the opposing effects when the wage becomes the subsequent reference wage. However, we conjecture that the current effect of loss aversion dominates the discounted expected future effect,[29] and that in the subsequent contract the direct effect of loss aversion on

[29]This will hold for sufficiently impatient firms, but we believe it has intuitive merit.

effort is larger than any indirect effect that comes through a change in the reference wage. If our conjecture is correct we can draw unambiguous conclusions: if $q_0 \geq \hat{q}_l$ then (even without our conjecture) the firm will set a lower initial wage; if $q_0 < \hat{q}_l$ the firm will set a higher initial wage; and both the hiring reservation productivity and the layoff reservation productivity increase: a greater degree of loss aversion leads to compression of wages above reference wages, as well as compression in hiring.

5 Discussion

The Role of the Reference Wage. One of the key features of our dynamic analysis is the assumption defining the worker's adaptation of the reference wage, $r_1 = \tilde{w}_0$. Although being consistent with a large body of evidence in the literature, this assumption abstracts from several other aspects of workers' reference wage formation: i) adaptation may not be symmetric; and ii) workers' perceptions of fairness at each negotiation date, namely r_0 and r_1, may be influenced by more than just past wages.

We have so far assumed that adaptation to the reference wage is symmetric, which implies both positive and negative reciprocity are only temporary phenomena. There is evidence to suggest that negative reciprocity is not only stronger as our theory predicts, but also more persistent than positive reciprocity. We could easily capture this in our model with an asymmetric adaptation rule where workers adapt more readily to wage rises than to wage cuts,[30] for example

$$r_1(\tilde{w}_0, r_0) = \begin{cases} \tilde{w}_0 & \text{if } \tilde{w}_0 > r_0 \\ r_0 & \text{if } \tilde{w}_0 \leq r_0. \end{cases}$$

With this specification the worker's positive reciprocity will only last one employment period if the wage is perceived as a gift at time 0, while if a wage contract is perceived as unfair there will be no adaptation of the reference wage and the same wage in the subsequent employment contract would be equally negatively reciprocated. Asymmetric adaptation will exacerbate the firm's cost of negative reciprocity when facing a loss averse worker, which reinforces the predictions established in Section 4.

We could also consider alternative formulations of the worker's reference wage that may differ depending on whether the worker is a new hire (at time 0) or an incumbent (at

[30] This hypothesis is supported by the experimental evidence on reference point adaptation provided by Arkes et al. [2008, 2010]: individuals (in our case workers) adapt more rapidly to gains (gifts) than to losses (unfair wages).

time 1). For instance a newly hired worker's reference wage r_0 could be influenced by the state of the labor market (as in Akerlof [1982] and Summers [1988]), the most recent wage contract paid in the *previous* employment relationship (as Koenig et al. [2014] suggest), or the wage of incumbent workers employed by the same firm (as explored by Snell and Thomas [2010]). On the other hand an employed worker's reference wage r_1 might be influenced by the wage of his peers outside the firm (as in Keynes [1936], Bhaskar [1990] or Driscoll and Holden [2004]), by expectations (as in Eliaz and Spiegler [2013]) or by the firm's ability to pay (as in Danthine and Kurmann [2007]).For example, if a worker considers the firm's ability to pay, they may revise their perceptions of fairness accordingly and accept a lower wage in periods of adverse economic conditions, without loss of morale and negative reciprocity. This insight could shed some new light on the importance of information disclosure by firms, and its influence on workers' perceptions of fairness as discussed by Kahneman et al. [1986] and Bewley [1999].

Our model, being derived from first principles and being transparent in the link between assumptions and conclusions, provides a tractable framework within which to consider these issues and understand their relevance in employment relationships.

History Dependence, Wage Dynamics, Hiring and Lay Offs. The dynamic analysis of Section 4 highlights two additional important insights of our theory that can be informative to ongoing theoretical and empirical research concerned with labour market fluctuations. Both insights are driven by the worker's adaptation of the reference wage — which carries the information contained in the wage contract $\hat{w}(r_0, q_0, \lambda, \delta)$ from the initial employment period into the subsequent one — and by the worker's asymmetric reference-dependent reciprocity.

First, the model reveals that the initial conditions characterizing the state of the economy when a new employment contract is offered (captured by the information contained in the two state variables r_0 and q_0) persist into the employment relationship and influence the subsequent wage setting and layoff decisions of the firm. For instance, a worker that receives a higher initial wage contract due to favourable economic conditions (high q_0) is more likely to be paid a higher wage in the subsequent period than an otherwise identical worker hired at a lower match productivity. Moreover, all else equal, the worker employed at a higher wage also faces a greater ex-ante probability of being laid off. This insight generates implications for labour market models that endogenize workers' job destruction rate through the derivation of a layoff reservation productivity similar to $\underset{\sim}{q}(r_1, \lambda)$

(e.g. Mortensen and Pissarides [1994], and most recently Eliaz and Spiegler [2013]). It also has implications for the empirical research that attempts to capture the persistent effects of labour market conditions at the time of hiring for the subsequent path of workers' wages during the employment relationship, and on their probability of being laid off (e.g. Beaudry and DiNardo [1991]; Schmieder and von Wachter [2010]). In our model this persistency depends on whether information about the state of the labour market is incorporated in the reference wage by the worker at the time of hiring.

Second, we have shown that the anticipation of the costs related to negative reciprocity and wage rigidity influence a farsighted firm's behavior at the time of hiring, putting upward pressure on the hiring reservation productivity $\underset{\sim}{\hat{q}}(r_0, \lambda)$ and downward pressure on the optimal initial wage contract. This prediction has implications for the understanding of the effects of wage rigidity for hiring behavior and wage dynamics. In the current macroeconomic literature concerned with labour market fluctuations, much attention has been devoted to the effect of wage rigidity of 'newly hired' workers on firms' job creation incentives (see for instance the discussion in Elsby et al. [2015a] and references therein). In contrast, in our model what matters for hiring behavior is the expected wage rigidity in the subsequent employment period once workers are employed. Thus our theory suggests that wage rigidity matters for job creation to the extent that it negatively influences the expected discounted value of profit that a firm gets from opening a vacancy and hiring.

6 Conclusion

Inspired by evidence from anthropological and experimental research on labor markets, in this paper we have advanced a theory of wage setting behavior based on contractual incompleteness, fairness, reciprocity, and reference dependence and loss aversion in the evaluation of wage contracts by workers. We identified loss aversion, rather than reciprocity *per se*, as the source of wage rigidity, and developed an understanding of the consequences of this for the wage setting and hiring behavior of firms in a simple dynamic framework. Our model provides a microeconomic account of many of the explanations of wage rigidity presented in the literature and, being simple, tractable, and based on first principles, establishes a clear link between the primitives of workers' evaluation of wage offers and the behavior of workers and firms.

Further exploring the insights of the model within a richer macroeconomic framework is firmly on our future research agenda. We believe that incorporating the asymmetric reference-dependent reciprocity we derived here in a macroeconomic model will act to suitably suppress the flexibility of wages to variation in economic conditions, that will convincingly contribute to the ongoing debate in the theory of labor market fluctuations (e.g. Shimer [2005]). Moreover since our model is clear in the distinction between newly hired and incumbent workers, it stands as a rich and tractable framework to analyze the consequences of the expected cost of wage rigidity in long-term employment relationships for hiring behavior and wage dynamics; an aspect that drew particular attention in light of recent cross-country labor market experiences in the aftermath of the Great Recession [Elsby et al., 2015b]. Therefore, a promising line of future research lies in developing a more complete understanding of the determinants and evolution of the reference wage for newly hired and incumbent workers, and the consequences for the dynamics of labor market outcomes, initial investigation of which is proving insightful.

Appendix

Proof of Theorem 1. We suppose Assumptions W1-W4 hold throughout. When $w = r$, $m(w) = m(r)$ and therefore $\Omega(e, w, r, \lambda) = b'(e) - c'(e)$. By assumption, $\Omega(0, r, r, \lambda) > 0$ and $\Omega^e(e, w, r, \lambda) < 0$, which implies $\tilde{e}(r, r, \lambda) \equiv \tilde{e}^n > 0$. Recalling the definition of $\mu(\cdot)$ in (3), when $w > r$, $\Omega(e, w, r, \lambda) = b'(e) - c'(e) + \zeta\eta(m(w) - m(r))$ with $m(w) - m(r) > 0$. As such $\Omega(\tilde{e}^n, w, r, \lambda) > \Omega(\tilde{e}^n, r, r, \lambda)$ and then the fact that $\Omega^e(e, w, r, \lambda) < 0$ implies $\tilde{e}(w, r)^+ > \tilde{e}^n$ for all $w > r$. When $w < r$, $\Omega(e, w, r, \lambda) = b'(e) - c'(e) + \zeta\lambda\eta(m(w) - m(r))$ with $m(w) - m(r) < 0$, so $\Omega(\tilde{e}^n, w, r, \lambda) < \Omega(\tilde{e}^n, r, r, \lambda)$ so $\Omega^e(e, w, r, \lambda) < 0$ implies $\tilde{e}(w, r, \lambda)^- < \tilde{e}^n$.

When $w < r$, the wage offered may be such that the worker would optimally choose $e < 0$ but cannot due to the constraint that $e \geq 0$. Define

$$\underline{w}(r, \lambda) = \max\{0, w : \Omega(0, w, r, \lambda) = 0\}.$$

Since $\Omega^w(e, w, r, \lambda) > 0$ this identifies the threshold wage at which the worker would choose $e = 0$ and below which they would like to choose $e < 0$ (since $\Omega(e, w, r, \lambda) < 0$ for all $e \geq 0$) but cannot, so we define $\tilde{e}(\underline{w}(r, \lambda), r, \lambda)^- \equiv 0$ for all $w \leq \underline{w}(r, \lambda)$.

For $w \neq r$, continuity of $\tilde{e}(w, r, \lambda)$ is readily established as $\Omega(e, w, r, \lambda)$ is continuous in all its arguments. Continuity at $w = r$ is established by noting that in the expression for $\Omega(e, w, r, \lambda)$, $\lim_{\epsilon \to 0} m(r + \epsilon) - m(r) = 0$ which implies $\tilde{e}(w, r)^+ \to \tilde{e}_n$ as $w \to r$ from above, and $\lim_{\epsilon\to 0} m(r-\epsilon) - m(r) = 0$ implying $\tilde{e}(w,r,\lambda)^- \to \tilde{e}_n$ as $w \to r$ from below.

When $w \neq r$ and $w > \underline{w}(r,\lambda)$ implicit differentiation of the first-order condition reveals

$$\tilde{e}^w(w,r,\lambda) = -\frac{\zeta\mu'(m(w)-m(r))m'(w)}{b''(e)-c''(e)} > 0$$

under our assumptions. Further differentiating this expression (recalling that $\mu(\cdot)$ is piecewise linear) we find that

$$\tilde{e}^{ww}(w,r,\lambda) = -\frac{\zeta\mu'(m(w)-m(r))m''(w)}{b''(e)-c''(e)} < 0,$$

establishing concavity of the optimal effort function.

Note from (3) that for $w > r$, $\mu'(\cdot) = \eta$ and when $w < r$, $\mu'(\cdot) = \lambda\eta$. To consider the response of effort to the wage above and below the reference wage we use the continuity of $\tilde{e}(w,r,\lambda)$ to establish that

$$\begin{aligned}\lim_{\epsilon\to 0}\tilde{e}^w(r-\epsilon,r,\lambda)^- &= -\lim_{\epsilon\to 0}\frac{\lambda\zeta\eta m'(r-\epsilon)}{b''(\tilde{e}(r-\epsilon,r,\lambda)^-)-c''(\tilde{e}(r-\epsilon,r,\lambda)^-)} = -\frac{\lambda\zeta\eta m'(r)}{b''(\tilde{e}_n)-c''(\tilde{e}_n)}\\ &= -\lambda\lim_{\epsilon\to 0}\frac{\zeta\eta m'(r+\epsilon)}{b''(\tilde{e}(r+\epsilon,r)^+)-c''(\tilde{e}(r+\epsilon,r)^+)} = \lambda\lim_{\epsilon\to 0}\tilde{e}^w(r+\epsilon,r)^+.\end{aligned}$$

Indeed, the fact that effort is increasing in the wage combined with concavity of $m(\cdot)$ imply that for any $w' < r < w''$ we have

$$\begin{aligned}\tilde{e}^w(w',r)^- &= -\frac{\lambda\zeta\eta m'(w')}{b''(\tilde{e}(w',r)^-)-c''(\tilde{e}(w',r)^-)}\\ &\qquad > -\frac{\lambda\zeta\eta m'(w'')}{b''(\tilde{e}(w'',r)^+)-c''(\tilde{e}(w'',r)^+)} = \lambda\tilde{e}^w(w'',r)^+.\end{aligned}$$

The relationship between $\tilde{e}(w,r,\lambda)$ and r is established by implicit differentiation:

$$\tilde{e}^r(w,r,\lambda) = \frac{\zeta\mu'(m(w)-m(r))m'(r)}{b''(e)-c''(e)} < 0.$$

Similarly, the effect of the degree of loss aversion on effort when $w < r$ is

$$\tilde{e}^\lambda(w,r,\lambda)^- = -\frac{\zeta\eta(m(w)-m(r))}{b''(e)-c''(e)} < 0.$$

Moreover, the effect of the degree of loss aversion on the effort response to the wage (for

$w < r$) is

$$\tilde{e}^{w\lambda}(w, r, \lambda) = -\frac{\zeta \eta m'(w)}{b''(e) - c''(e)} > 0.$$

$\square$

Proof of Theorem 2. Throughout the proof we assume the worker's productivity and reference wage are such that $q \geq \underset{\sim}{q}(r, \lambda)$ so the firm will be profitable if it hires the worker, and consider the properties of the threshold productivity at the end. We proceed by first stating some preliminaries, then considering the productivity thresholds, then demonstrating the nature of the optimal wage setting rule.

Preliminaries: First, note that under Assumption F2 and the results of Theorem 1, for $w \neq r$ we have that $\Psi^q(w; q, r, \lambda) = y^{qe}\tilde{e}^w > 0$; $\Psi^r(w; q, r, \lambda) = y^{ee}\tilde{e}^r\tilde{e}^w + y^e\tilde{e}^{wr} \geq 0$ (after noting that $\tilde{e}^{wr} = 0$); and $\Psi^w(w; q, r, \lambda) = y^{ee}(\tilde{e}^w)^2 + y^e\tilde{e}^{ww} - s''(w) < 0$. In addition, $\Psi^\lambda = y^{ee}\tilde{e}^\lambda\tilde{e}^w + y^e\tilde{e}^{w\lambda}$ so $\Psi^\lambda > 0$ if $w < r$ and $\Psi^\lambda = 0$ if $w > r$. These results also allow us to deduce that if $\lambda > 1$, $\Psi(w; q, r, \lambda)$ jumps down at the reference wage, since

$$\begin{aligned}
&\lim_{\epsilon \to 0} \Psi(r - \epsilon; q, r, \lambda) - \lim_{\epsilon \to 0} \Psi(r + \epsilon; q, r, \lambda) = \\
&\qquad y^e(q, \lim_{\epsilon \to 0} \tilde{e}(r - \epsilon, r, \lambda)^-) \lim_{\epsilon \to 0} \tilde{e}^w(r - \epsilon, r, \lambda)^- - s'(r - \epsilon) - \\
&\qquad y^e(q, \lim_{\epsilon \to 0} \tilde{e}(r + \epsilon, r)^+) \lim_{\epsilon \to 0} \tilde{e}^w(r + \epsilon, r)^+ - s'(r + \epsilon) \\
&\qquad = y^e(q, \tilde{e}_n)[\lim_{\epsilon \to 0} \tilde{e}^w(r - \epsilon, r, \lambda)^- - \lim_{\epsilon \to 0} \tilde{e}^w(r + \epsilon, r)^+] \\
&\qquad\qquad = y^e(q, \tilde{e}_n) \lim_{\epsilon \to 0} \tilde{e}^w(r + \epsilon, r)^+(\lambda - 1) \geq 0,
\end{aligned}$$

with a strict inequality if $\lambda > 1$. As such, $\Psi(w; q, r, \lambda)$ is everywhere decreasing in w, establishing concavity of the payoff function.

Productivity thresholds: As will be made clear in the remainder of the proof, the threshold $q_l(r, \lambda)$ identifies the critical match productivity below which the firm would want to set the wage below the reference wage, and $q_u(r)$ is the match productivity above which the firm would want to compensate the worker more than the reference wage. The former is the value of q below which profit is decreasing just below the reference wage; the latter is the value of q above which profit is increasing just above the reference wage. Since $\Psi(w, 0, r, \lambda) < 0$ when $w > 0$ and $\Psi^q > 0$ there will be a unique value of each productivity threshold.

We now want to establish some properties of the thresholds. Implicit differentiation allows us to deduce that

$$q_l^r(r,\lambda) = -\lim_{\epsilon\to 0}\frac{\frac{\mathrm{d}\Psi(r-\epsilon;q,r,\lambda)}{\mathrm{d}r}}{\Psi^q(r-\epsilon;q,r,\lambda)} \text{ and}$$
$$q_u'(r) = -\lim_{\epsilon\to 0}\frac{\frac{\mathrm{d}\Psi(r+\epsilon;q,r,\lambda)}{\mathrm{d}r}}{\Psi^q(r+\epsilon;q,r,\lambda)}.$$

Now,

$$\begin{aligned}\frac{\mathrm{d}\Psi(r\pm\epsilon;q,r,\lambda)}{\mathrm{d}r} &= \Psi^w(r\pm\epsilon;q,r,\lambda)+\Psi^r(r\pm\epsilon;q,r,\lambda)\\ &= y^{ee}(\tilde{e}^{w\pm})^2 + y^e\tilde{e}^{ww\pm} + y^{ee}\tilde{e}^{r\pm}\tilde{e}^{w\pm}\\ &= y^{ee}(\tilde{e}^{w\pm})(\tilde{e}^{w\pm}+\tilde{e}^{r\pm}) + y^e\tilde{e}^{ww\pm},\end{aligned}$$

As $\epsilon\to 0$ we can infer that $\tilde{e}^w(r\pm\epsilon,r,\lambda)^\pm + \tilde{e}^r(r\pm\epsilon,r,\lambda)^\pm \to 0$ (refer to the expressions of these objects in the proof of Theorem 1), implying $\lim_{\epsilon\to 0}\frac{\mathrm{d}\Psi(r\pm\epsilon;q,r,\lambda)}{\mathrm{d}r} = y^e\tilde{e}^{ww\pm} < 0$. This allows us to conclude that $q_l^r(r,\lambda) > 0$ and $q_u'(r) > 0$.

Turning next to investigate how the lower threshold depends on the degree of loss aversion, implicit differentiation gives

$$\begin{aligned}q_l^\lambda(r,\lambda) &= -\lim_{\epsilon\to 0}\frac{\Psi^\lambda(r-\epsilon,q,r,\lambda)}{\Psi^q(r-\epsilon,q,r,\lambda)}\\ &= -\frac{y^{ee}\lim_{\epsilon\to 0}\tilde{e}^\lambda(r-\epsilon,r,\lambda)^-\lim_{\epsilon\to 0}\tilde{e}^w(r-\epsilon,r,\lambda)^- + y^e\lim_{\epsilon\to 0}\tilde{e}^{w\lambda}(r-\epsilon,r,\lambda)^-}{\lim_{\epsilon\to 0}\Psi^q(r-\epsilon;q,r,\lambda)} < 0\end{aligned}$$

since we deduced in Theorem 1 that $\tilde{e}^{\lambda-} < 0$ and $\tilde{e}^{w\lambda-} > 0$.

The consideration of $\lim_{\epsilon\to 0}\Psi(r-\epsilon;q,r\lambda) - \lim_{\epsilon\to 0}\Psi(r+\epsilon;q,r,\lambda)$ in the preliminaries allows us to conclude that when $\lambda = 1$ these two objects are equal. This, combined with the observation that $\lim_{\epsilon\to 0}\tilde{e}(r-\epsilon,r,\lambda)^- = \tilde{e}_n = \lim_{\epsilon\to 0}\tilde{e}(r+\epsilon,r)$ (from Theorem 1) permits the conclusion that $q_l(r,1) = q_u(r)$. This, along with the fact that $q_l^\lambda(r,\lambda) < 0$ implies $q_l(r,\lambda) < q_u(r)$ for all $\lambda > 1$.

Optimal wage setting: We now turn to the optimal wage setting rule, which depends on the match productivity in relation to the productivity thresholds.

If $q \in [0, q_l(r,\lambda))$ then the definition of $q_l(r,\lambda)$ and fact that $\Psi^q > 0$ can be used to deduce that $\lim_{\epsilon\to 0}\Psi(r-\epsilon,q,r,\lambda) < 0$; since $\Psi(w;q,r,\lambda)$ is everywhere decreasing in w, the same is true for all $w \geq r$. As such, the optimizing wage must satisfy $w < r$ and will therefore be the solution to

$$y^e(q, \tilde{e}(w, r, \lambda)^-)\tilde{e}^w(w, r, \lambda)^- - s'(w) \leq 0,$$

with equality if $w > \underline{w}(r, \lambda)$ (recall from the proof of Theorem 1 that this is either zero or the wage below which effort takes the boundary value of zero). To account for the fact that the firm may pay the 'lowest feasible wage' for a range of match productivity, let $\breve{q}(r, \lambda) = \max\{0, q : \Psi(\underline{w}(r, \lambda); q, r, \lambda) = 0\}$ (at $\breve{q}(r, \lambda)$ the firm would want to pay $\underline{w}(r, \lambda)$ and since $\Psi^q > 0$ the same will be true for all $0 \leq q < \breve{q}(r, \lambda)$). For all $\breve{q}(r, \lambda) < q < q_l(r, \lambda)$ the optimal wage is given by the displayed first-order condition holding with equality, which is denoted by $\tilde{w}(r, q, \lambda)^-$. Implicit differentiation and our deductions in the preliminaries reveal

$$\begin{aligned}
\tilde{w}^q(r, q, \lambda)^- &= -\frac{\Psi^q}{\Psi^w} > 0, \\
\tilde{w}^r(r, q, \lambda)^- &= -\frac{\Psi^r}{\Psi^w} \geq 0, \text{ and} \\
\tilde{w}^\lambda(r, q, \lambda)^- &= -\frac{\Psi^\lambda}{\Psi^w} > 0.
\end{aligned}$$

If $q \in (q_u(r), \infty]$ then the definition of $q_u(r)$ and the fact that $\Psi^q > 0$ can be used to deduce that $\lim_{\epsilon \to 0} \Psi(r + \epsilon, q, r, \lambda) > 0$; since $\Psi(w; q, r, \lambda)$ is everywhere decreasing in w the same is true for all $w \leq r$ and, as such, the optimizing wage must exceed r and will therefore satisfy

$$y^e(q, \tilde{e}(w, r)^+)\tilde{e}^w(w, r)^+ - s'(w) = 0.$$

Letting $\tilde{w}(q, r)^+$ denote the solution (which is independent of λ), implicit differentiation gives

$$\begin{aligned}
&\tilde{w}^q(r, q)^+ > 0 \text{ and} \\
&\tilde{w}^r(r, q)^+ \geq 0.
\end{aligned}$$

If $q \in [q_l(r, \lambda), q_u(r)]$ then the fact that $\Psi^q > 0$ can be used to deduce that $\lim_{\epsilon \to 0} \Psi(r - \epsilon, q, r, \lambda) \geq 0$ and $\lim_{\epsilon \to 0} \Psi(r + \epsilon, q, r, \lambda) \leq 0$. That $\Psi^w < 0$ for all $w \neq r$ then implies $\Psi(w; q, r, \lambda) > 0$ for all $w < r$ and $\Psi(w; q, r, \lambda) < 0$ for all $w > r$, implying profit is maximized if and only if $w = r$.

Finally, if $q < \underset{\sim}{q}(r, \lambda)$ then then the employment relationship ends. Implicit differentiation of the zero profit condition defining the reservation productivity allows us to deduce that

$$\begin{aligned}\underset{\sim}{q}^r(r,\lambda) &= -\frac{\pi^w\tilde{w}^r + \pi^e(\tilde{e}^r + \tilde{e}^w\tilde{w}^r)}{\pi^w\tilde{w}^q + \pi^q + \pi^e\tilde{e}^w\tilde{w}^q} \\ &= -\frac{\tilde{w}^r(\pi^w + \pi^e\tilde{e}^w) + \pi^e\tilde{e}^r}{\tilde{w}^q(\pi^w + \pi^e\tilde{e}^w) + \pi^q} > 0\end{aligned}$$

since $\pi^w + \pi^e\tilde{e}^w = 0$ from the first-order condition, $\pi^q, \pi^e > 0$ by Assumption F2 and we found in Theorem 1 that $\tilde{e}^r < 0$. In addition,

$$\begin{aligned}\underset{\sim}{q}^\lambda(r,\lambda) &= -\frac{\pi^w\tilde{w}^\lambda + \pi^e(\tilde{e}^\lambda + \tilde{e}^w\tilde{w}^\lambda)}{\pi^w\tilde{w}^q + \pi^q + \pi^e\tilde{e}^w\tilde{w}^q} \\ &= -\frac{\tilde{w}^\lambda(\pi^w + \pi^e\tilde{e}^w) + \pi^e\tilde{e}^\lambda}{\tilde{w}^q(\pi^w + \pi^e\tilde{e}^w) + \pi^q}\end{aligned}$$

Again $\pi^w + \pi^e\tilde{e}^w = 0$ and $\pi^q, \pi^e > 0$, and we found in Theorem 1 that when $w > r$ $\tilde{e}$ is independent of λ, but when $w < r$, $\tilde{e}^\lambda < 0$. As such, if $\tilde{w}(r, \underset{\sim}{q}(r,\lambda)) > r$ then $\underset{\sim}{q}^\lambda(r,\lambda) = 0$ but when $\tilde{w}(r, \underset{\sim}{q}(r,\lambda)) < r$, $\underset{\sim}{q}^\lambda(r,\lambda) > 0$. □

The wage setting rule when the support of the match productivity distribution is $[\underline{q}, \bar{q}]$*.* For completeness, we elucidate the details of relaxing our simplifying assumption that the support of the distribution of match productivity is $[0,\infty]$, and rather suppose it is $[\underline{q}, \bar{q}]$. If this is the case, the definition of the reservation productivity becomes

$$\underset{\sim}{q}(r,\lambda) = \max\{\underline{q}, q : \pi(\tilde{w}(r,q,\lambda); q, \tilde{e}(\tilde{w}(r,q,\lambda), r, \lambda)) = 0\},$$

and the lower and upper thresholds of match productivity that characterize the wage setting rule are defined as

$$\begin{aligned}q_l(r,\lambda) &= \min\{\max\{\underline{q}, q : \lim_{\epsilon\to 0}\Psi(r-\epsilon; q, r, \lambda) = 0\}, \bar{q}\}, \text{ and} \\ q_u(r) &= \min\{\max\{\underline{q}, q : \lim_{\epsilon\to 0}\Psi(r+\epsilon; q, r, \lambda) = 0\}, \bar{q}\}.\end{aligned}$$

The form of these thresholds then depend on the worker's reference wage.

Consider first the lower threshold. Define $\bar{r}(\lambda) = \{r : \lim_{\epsilon\to 0}\Psi(r-\epsilon, \bar{q}, r, \lambda) = 0\}$ and $\underset{\sim}{r}(\lambda) = \{r : \lim_{\epsilon\to 0}\Psi(r-\epsilon, \underline{q}, r, \lambda) = 0\}$. If $r > \bar{r}(\lambda)$ then our deduction in the proof of Theorem 2 that $\lim_{\epsilon\to 0}\frac{\mathrm{d}\Psi(r\pm\epsilon;q,r,\lambda)}{\mathrm{d}r} = y^e\tilde{e}^{ww\pm} < 0$ and the fact that $\Psi^q > 0$ implies $\lim_{\epsilon\to 0}\Psi(r-\epsilon, q, r, \lambda) < 0$ for all $q < \bar{q}$ and we define $q_l(r,\lambda) = \bar{q}$ (the firm will always want to set the wage below the (high) reference wage). If $r < \underset{\sim}{r}(\lambda)$ then we can similarly deduce that $\lim_{\epsilon\to 0} \Psi(r-\epsilon, q, r, \lambda) > 0$ for all $q > \underline{q}$ and we define $q_l(r,\lambda) = \underline{q}$ (the firm will never want to set the wage below the (low) reference wage). Thus,

$$q_l(r,\lambda) = \begin{cases} \underline{q} & \text{if } r < \underset{\sim}{r}(\lambda) \\ \{q : \lim_{\epsilon\to 0} \Psi(r-\epsilon, q, r, \lambda) = 0\} & \text{if } \underset{\sim}{r}(\lambda) \leq r \leq \bar{r}(\lambda) \\ \bar{q} & \text{if } r > \bar{r}(\lambda) \end{cases}$$

Turning next to the upper threshold, define $\tilde{r} = \{r : \lim_{\epsilon\to 0} \Psi(r+\epsilon; \bar{q}, r, \lambda) = 0\}$ and $\underline{r} = \{r : \lim_{\epsilon\to 0} \Psi(r+\epsilon; \underline{q}, r, \lambda) = 0\}$. If $r > \tilde{r}$ then our previously cited monotonicity statements allow us to deduce that $\lim_{\epsilon\to 0} \Psi(r+\epsilon; q, r, \lambda) < 0$ for all $q < \bar{q}$ and so we define $q_u(r) = \bar{q}$. If $r < \underline{r}$, $\lim_{\epsilon\to 0} \Psi(r+\epsilon; q, r, \lambda) > 0$ for all $q > \underline{q}$ and we define $q_l(r) = \underline{q}$. As such,

$$q_u(r) = \begin{cases} \underline{q} & \text{if } r < \underline{r} \\ \{q : \lim_{\epsilon\to 0} \Psi(r+\epsilon, q, r, \lambda) = 0\} & \text{if } \underline{r} \leq r \leq \tilde{r} \\ \bar{q} & \text{if } r > \tilde{r} \end{cases}$$

The fact deduced in the preliminaries of the proof of Theorem 2 that $\lim_{\epsilon\to 0} \Psi(r-\epsilon; q, r, \lambda) > \lim_{\epsilon\to 0} \Psi(r+\epsilon; q, r, \lambda)$ when $\lambda > 1$, along with our observation that $\lim_{\epsilon\to 0} \frac{\mathrm{d}\Psi(r\pm\epsilon;q,r,\lambda)}{\mathrm{d}r} < 0$, allows us to conclude that $\underline{r} < \underset{\sim}{r}(\lambda)$ and $\tilde{r} < \bar{r}(\lambda)$.

FIGURE VI HERE.

The nature of the productivity thresholds that define the range of rigidity depend on the worker's reference wage, which are illustrated in Figure VI, along with a typical reservation productivity $\underset{\sim}{q}(r,\lambda)$ that also depends on the reference wage. For workers for whom $\lambda > 1$, if $r > \bar{r}(\lambda)$ (defined above as the reference wage where $q_l(r,\lambda) = \bar{q}$) then the firm will always want to reduce the wage below the reference wage; and if $r < \underline{r}$ (defined as the reference wage where $q_u(r) = \underline{q}$) the firm will always want to increase the wage above the reference wage. If the reference wage is either particularly low or particularly high there is no range of rigidity. However, for $\underline{r} < r < \bar{r}(\lambda)$ there will be a non-empty range of match productivity in which the wage paid will be the same, so long as offering a contract that pays r to a worker whose productivity is $q_u(r)$ gives strictly positive profit (otherwise all contracts within the range of rigidity are unprofitable).

Consequently, in this more general case there is a non-empty range of reference wages in which the optimal wage contract will exhibit wage rigidity for loss averse workers, but outside this range wage rigidity is not a feature of the wage setting rule.

Corollary 1. *If $\lambda > 1$, $\underline{r} < r < \bar{r}(\lambda)$ and $\underset{\sim}{q}(r, \lambda) < q_u(r)$ there is a non-empty range of match productivity within which otherwise identical workers are paid the same wage, equal to their reference wage. The more loss averse a worker is the greater the range of rigidity and the higher will be their wage in the case of a wage cut, but the range of match productivity over which the employment relationship continues shrinks.*

Proof of Proposition 1. If in the first period the match productivity falls short of the reservation productivity with the updated reference wage, i.e. $q_1 < \underset{\sim}{q}(w_0, \lambda)$, then contracting with the worker for the final period would be unprofitable and the contract terminated; the employment relationship only has value for $q_1 \geq \underset{\sim}{q}(w_0, \lambda)$. As noted,

$$\begin{aligned}\mathbb{E}_0[J_1(w_0, q_1)] = &\int_{\underset{\sim}{q}(w_0,\lambda)}^{\max\{\underset{\sim}{q}(w_0,\lambda), q_l(w_0,\lambda)\}} J_1(w_0, q_1)^- \mathrm{d}F + \int_{\max\{\underset{\sim}{q}(w_0,\lambda), q_l(w_0,\lambda)\}}^{\max\{\underset{\sim}{q}(w_0,\lambda), q_u(w_0)\}} J_1(w_0, q_1)^= \mathrm{d}F \\ &+ \int_{\max\{\underset{\sim}{q}(w_0,\lambda), q_u(w_0)\}}^{\infty} J_1(w_0, q_1)^+ \mathrm{d}F\end{aligned}$$

where $J_1(w_0, q_1)^{-;=;+}$ represents the value of the contract if $w_1 < w_0; w_1 = w_0; w_1 > w_0$, in which effort is given by $\tilde{e}(w_1, w_0, \lambda)^-; \tilde{e}_n; \tilde{e}(w_1, w_0)^+$. (In subsequent proofs we are less pedantic and assume $\underset{\sim}{q}(w_0, \lambda) < q_l(w_0, \lambda)$.)

Let $\mathbb{I}^-$ be an indicator variable that takes the value 1 if $\underset{\sim}{q}(w_0, \lambda) < q_l(w_0, \lambda)$ and is otherwise zero; $\mathbb{I}^=$ an indicator variable that takes the value 1 if $q_l(w_0, \lambda) \leq \underset{\sim}{q}(w_0, \lambda) \leq q_u(w_0)$ and is otherwise zero; and $\mathbb{I}^+$ an indicator variable that takes the value 1 if $\underset{\sim}{q}(w_0, \lambda) > q_u(w_0)$ and is otherwise zero. Note that one, and only one, of $\mathbb{I}^-$, $\mathbb{I}^=$ and $\mathbb{I}^+$ is equal to 1. The marginal change in the value of the first-period employment contract (dropping the arguments of productivity thresholds) is

$$\begin{aligned}\theta(w_0, \lambda) = &\int_{\underset{\sim}{q}}^{\max\{\underset{\sim}{q}, q_l\}} J_1^{r_1}(w_0, q_1)^- \mathrm{d}F \\ &+ \mathbb{I}^-[q_l^r \lim_{\epsilon \to 0} J_1(w_0, q_l - \epsilon)^- f(q_l) - \underset{\sim}{q}^r J_1(w_0, \underset{\sim}{q}) f(\underset{\sim}{q})] \\ &+ \int_{\max\{\underset{\sim}{q}, q_l\}}^{\max\{\underset{\sim}{q}, q_u\}} J_1^{r_1}(w_0, q_1)^= \mathrm{d}F \\ &+ \mathbb{I}^=[q_u' J_1(w_0, q_u)^= f(q_u) - \underset{\sim}{q}^r J_1(w_0, \underset{\sim}{q}) f(\underset{\sim}{q})] + \mathbb{I}^-[q_u' J_1(w_0, q_u)^= f(q_u) - q_l^r J_1(w_0, q_l)^= f(q_l)] \\ &+ \int_{\max\{\underset{\sim}{q}, q_u\}}^{\infty} J_1^{r_1}(w_0, q_1)^+ \mathrm{d}F \\ &- \mathbb{I}^+ \underset{\sim}{q}^r J_1(w_0, \underset{\sim}{q}) f(\underset{\sim}{q}) - (1 - \mathbb{I}^+) q_u' \lim_{\epsilon \to 0} J_1(w_0, q_u + \epsilon)^+ f(q_u).\end{aligned}$$

By definition, $J_1(w_0, \underset{\sim}{q}) = 0$, and the continuity of the optimal effort function and wage setting rule imply $\lim_{\epsilon \to 0} J_1(w_0, q_l - \epsilon)^- = J_1(w_0, q_l)^=$ and $\lim_{\epsilon \to 0} J_1(w_0, q_u + \epsilon)^+ = J_1(w_0, q_u)^=$. It then follows for each of the three scenarios ($\mathbb{I}^- = 1; \mathbb{I}^= = 1; \mathbb{I}^+ = 1$), after cancelling terms, that

$$\theta(w_0, \lambda) = \int_{\underset{\sim}{q}}^{\infty} J_1^{r_1}(w_0, q_1) \mathrm{d}F.$$

Now, for $q_1 \in [\underset{\sim}{q}(w_0, \lambda), \infty] \setminus [q_l(w_0, \lambda), q_u(w_0)]$ (i.e. where the wage is not equal to the reference wage)

$$J_1^{r_1}(r_1, q_1) = \pi^w \tilde{w}^r + \pi^e(\tilde{e}^w \tilde{w}^r + \tilde{e}^r) \tag{9}$$

$$= \pi^e \tilde{e}^r < 0, \tag{10}$$

since from the first-order condition $\pi^w + \pi^e \tilde{e}^w = 0$ (assuming $\tilde{w}_1 > \underline{w}(w_0, \lambda)$ so the solution is interior) and we deduced in Theorem 1 that $\tilde{e}^r < 0$. For $q_1 \in [q_l(w_0, \lambda), q_u(w_0)]$ the wage is equal to the reference wage and effort is constant and equal to $\tilde{e}_n$, so $J_1^{r_1}(r_1, q_1) = \pi^w < 0$. As such,

$$\theta(w_0, \lambda) = \int_{\underset{\sim}{q}}^{q_l} \pi^e \tilde{e}^r \mathrm{d}F + \int_{q_l}^{q_u} \pi^w \mathrm{d}F + \int_{q_u}^{\infty} \pi^e \tilde{e}^r \mathrm{d}F < 0. \tag{11}$$

□

Proof of Theorem 3. The proof is qualitatively similar to the proof of Theorem 2, so the details are largely omitted. Let us, however, dwell on the assumption that $\Psi^w + \delta\theta^{w_0} < 0$. We know from the proof of Theorem 2 that $\Psi(w; q_0, r_0, \lambda)$ is decreasing in w as $\Psi^w < 0$ for $w \neq r$ and at $w = r$ there is a jump down. Recalling the expression for $\theta(w_0, \lambda)$ in (11) and recognising that both the integrand (except in the case of $q_1 \in [q_l(w_0, \lambda), q_u(w_0)]$) and the limits of integration depend on w_0, we deduce that

$$\begin{aligned}\theta^{w_0} = &\int_{\underset{\sim}{q}}^{q_l} \frac{\mathrm{d}}{\mathrm{d}r}\{\pi^e \tilde{e}^r\} \mathrm{d}F + q_l^r \pi^e \lim_{\epsilon \to 0} \tilde{e}^r(w_0 - \epsilon, w_0, \lambda) f(q_l) - \underset{\sim}{q}^r \pi^e \tilde{e}^r f(\underset{\sim}{q}) \\ &+ q_u' \pi^w f(q_u) - q_l^r \pi^w f(q_l) \\ &+ \int_{q_u}^{\infty} \frac{\mathrm{d}}{\mathrm{d}r}\{\pi^e \tilde{e}^r\} \mathrm{d}F - q_u' \pi^e \lim_{\epsilon \to 0} \tilde{e}^r(w_0 + \epsilon, w_0, \lambda) f(q_u).\end{aligned}$$

Now, from the expressions for $\tilde{e}^w$ and $\tilde{e}^r$ in the proof of Theorem 1 it follows that $\lim_{\epsilon \to 0} \tilde{e}^r(w_0 \pm \epsilon, w_0, \lambda) = -\lim_{\epsilon \to 0} \tilde{e}^w(w_0 \pm \epsilon, w_0, \lambda)$. Moreover, when $w \neq r$ the first-order condition holds with equality, which implies that $\pi^e \tilde{e}^w = -\pi^w$. These statements together give us $\pi^e \lim_{\epsilon\to 0} \tilde{e}^r(w_0 \pm \epsilon, w_0, \lambda) = \pi^w$, which allows several terms to cancel in the above expression. Noting that $\frac{\mathrm{d}}{\mathrm{d}r}\{\pi^e \tilde{e}^r\} = \pi^e \tilde{e}^{rr} + \pi^{ee}(\tilde{e}^r)^2$ then allows us to conclude that

$$\theta^{w_0} = \int_{\underset{\sim}{q}}^{q_l} (\pi^e \tilde{e}^{rr} + \pi^{ee}(\tilde{e}^r)^2)\mathrm{d}F + \int_{q_u}^{\infty} (\pi^e \tilde{e}^{rr} + \pi^{ee}(\tilde{e}^r)^2)\mathrm{d}F - \underset{\sim}{q}^r \pi^e \tilde{e}^r f(\underset{\sim}{q}).$$

We know from Theorem 1 that $\tilde{e}^r < 0$ and $\tilde{e}^{rr} > 0$, and from Theorem 2 that $\underset{\sim}{q}^r > 0$. As such, the sign of the second derivative of the expected future profit function with respect to the initial wage remains undetermined. To proceed, it is not necessary to assume that $\theta^{w_0} < 0$, but we do need to assume that $\Psi^w + \delta\theta^{w_0} < 0$, which we find reasonable. Since $\Psi^w < 0$ the inequality will be true for a sufficiently impatient firm, but our conjecture is $|\Psi^w| > |\theta^{w_0}|$, since we believe the direct current effect of a change in the wage on this period will be larger than the expected effect in the future that comes indirectly through this period's wage becoming the future reference wage (which is also discounted), and therefore even if $\theta^{w_0} > 0$ we will have $\Psi^w + \delta\theta^{w_0} < 0$.

Under this assumption the proof of the nature of the wage setting rule follows the same steps as the proof of Theorem 2 where Ψ is replaced with $\Psi + \delta\theta$. □

Proof of Proposition 2. The proof relies on investigation of the first-order condition of the two optimization problems, noting from Proposition 1 that $\theta(w_0, \lambda) < 0$. First we show that $\hat{q}_l(r, \lambda, \delta) > \tilde{q}_l(r, \lambda)$. Suppose, by contradiction, that $\hat{q}_l \leq \tilde{q}_l$, then the fact that $\Psi^q > 0$ (see the preliminaries in the proof of Theorem 2) implies

$$0 \equiv \lim_{\epsilon\to 0} \Psi(r - \epsilon; \tilde{q}_l, r, \lambda) \geq \lim_{\epsilon\to 0} \Psi(r - \epsilon, \hat{q}_l, r, \lambda),$$

but then since $\theta(w, \lambda) < 0$ we have that

$$\lim_{\epsilon\to 0} \Psi(r - \epsilon, \hat{q}_l, r, \lambda) > \lim_{\epsilon\to 0} \Psi(r - \epsilon, \hat{q}_l, r, \lambda) + \delta\theta(w, \lambda) \equiv 0,$$

yielding a contradiction. That $\hat{q}_u(r, \lambda, \delta) > \tilde{q}_u(r)$ is similarly proved.

We now want to compare $\hat{w}(r, q, \lambda, \delta)^{-;+}$ with $\tilde{w}(r, q, \lambda)^{-;+}$ where both functions are defined. We demonstrate that $\hat{w}(r, q, \lambda, \delta)^{-} < \tilde{w}(r, q, \lambda)^{-}$ for all $q < \tilde{q}_l(r, \lambda)$. Suppose, by contradiction, that $\hat{w}^{-} \geq \tilde{w}^{-}$. Then the fact that $\Psi^w < 0$ (see the preliminaries in the proof of Theorem 2) implies

$$0 \equiv \Psi(\tilde{w}^-; q, r, \lambda) \geq \Psi(\hat{w}^-; q, r, \lambda),$$

but then $\theta(w_0, \lambda) < 0$ implies

$$\Psi(\hat{w}^-; q, r, \lambda) > \Psi(\hat{w}^-; q, r, \lambda) + \delta\theta(r, \lambda) \equiv 0,$$

yielding a contradiction. The proof that $\hat{w}(r, q, \lambda, \delta)^+ < \tilde{w}(r, q, \lambda)^+$ for all $q > \hat{q}_u(r, \lambda, \delta)$ is similar and so omitted. $\square$

Proof of Proposition 3. First consider how the optimal wage changes with the degree of loss aversion. Implicit differentiation of the wage setting rule gives

$$\hat{w}^\lambda = -\frac{\Psi^\lambda + \delta\theta^\lambda}{\Psi^w + \delta\theta^w}.$$

We assume the denominator is negative, and we know from the preliminaries in the proof of Theorem 2 that $\Psi^\lambda = 0$ if $w \geq r$ and $\Psi^\lambda > 0$ if $w < r$. Recalling the definition of $\theta(w_0, \lambda)$ in (11) and noting that π^w and q_u are independent of λ, we deduce that

$$\begin{aligned}\theta^\lambda = &\int_{\underset{\sim}{q}}^{q_l} \frac{\mathrm{d}}{\mathrm{d}\lambda}\{\pi^e \tilde{e}^r\}\mathrm{d}F + q_l^\lambda \pi^e \lim_{\epsilon \to 0} \tilde{e}^r(w_0 - \epsilon, w_0, \lambda) f(q_l) - \underset{\sim}{q}^\lambda \pi^e \tilde{e}^r f(\underset{\sim}{q}) \\ &- q_l^\lambda \pi^w f(q_l) \\ &+ \int_{q_u}^{\infty} \frac{\mathrm{d}}{\mathrm{d}\lambda}\{\pi^e \tilde{e}^r\}\mathrm{d}F\end{aligned}$$

Now, $q_l^\lambda \pi^e \lim_{\epsilon \to 0} \tilde{e}^r(w_0 - \epsilon, w_0, \lambda) - q_l^\lambda \pi^w = 0$ since as we deduced previously for $w \approx r$, $\pi^e \tilde{e}^r \approx \pi^w$. Moreover, $\frac{\mathrm{d}}{\mathrm{d}\lambda}\{\pi^e \tilde{e}^r\} = \pi^{ee} \tilde{e}^\lambda \tilde{e}^r + \pi^e \tilde{e}^{r\lambda}$ which, according to our deductions in Theorem 1, is equal to zero for wages exceeding the reference wage. As such,

$$\theta^\lambda = \int_{\underset{\sim}{q}}^{q_l} (\pi^{ee} \tilde{e}^\lambda \tilde{e}^r + \pi^e \tilde{e}^{r\lambda})\mathrm{d}F - \underset{\sim}{q}^\lambda \pi^e \tilde{e}^r f(\underset{\sim}{q}).$$

In Theorem 1 we concluded that $\tilde{e}^\lambda, \tilde{e}^r, \tilde{e}^{r\lambda} < 0$. As such, if the layoff reservation productivity doesn't increase too much, i.e. $\underset{\sim}{q}^\lambda > 0$ is sufficiently small, then $\theta^\lambda < 0$, which allows us to conclude the statements regarding the wage in the proposition.

Next consider the reservation productivity in the subsequent employment contract governing layoff behavior, which is

$$\underset{\sim}{q}(\hat{w}(r_0, q_0, \lambda, \delta), \lambda) = \max\{0, q_1 : J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1) = 0\}.$$

Noting that

$$J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1) = \pi(\tilde{w}(\hat{w}(r_0, q_0, \lambda, \delta), q_1, \lambda); q_1, \tilde{e}(\tilde{w}(\hat{w}(r_0, q_0, \lambda, \delta), q_1, \lambda), \hat{w}(r_0, q_0, \lambda, \delta), \lambda)),$$

implicit differentiation reveals

$$\begin{aligned}
\frac{\mathrm{d}\underset{\sim}{q}}{\mathrm{d}\lambda} &= -\frac{\pi^w(\tilde{w}^r \hat{w}^\lambda + \tilde{w}^\lambda) + \pi^e(\tilde{e}^w(\tilde{w}^r \hat{w}^\lambda + \tilde{w}^\lambda) + \tilde{e}^r \hat{w}^\lambda + \tilde{e}^\lambda)}{\pi^w \tilde{w}^q + \pi^q + \pi^e \tilde{e}^w \tilde{w}^q} \\
&= -\frac{(\pi^w + \pi^e \tilde{e}^w)(\tilde{w}^r \hat{w}^\lambda + \tilde{w}^\lambda) + \pi^e(\tilde{e}^r \hat{w}^\lambda + \tilde{e}^\lambda)}{(\pi^w + \pi^e \tilde{e}^w)\tilde{w}^q + \pi^q} \\
&= -\frac{\pi^e(\tilde{e}^r \hat{w}^\lambda + \tilde{e}^\lambda)}{\pi^q}
\end{aligned}$$

after utilizing the fact that in the subsequent employment contract $\pi^w + \pi^e \tilde{e}^w = 0$. We know from Theorem 1 that $\tilde{e}^r < 0$ and $\tilde{e}^\lambda < 0$ when $w < r$ (which is the case here since we are supposing $\underset{\sim}{q}(w_0, \lambda) < q_l(w_0, \lambda)$), which allows us to conclude the statement in the proposition.

The reservation productivity governing hiring behavior in the initial contract is

$$\underset{\sim}{\hat{q}}(r, \lambda, \delta) = \{q_0 : \pi(\hat{w}(r_0, q_0, \lambda, \delta); q_0, \tilde{e}(\hat{w}(r_0, q_0, \lambda, \delta), r_0, \lambda)) + \delta \mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)] = 0\}.$$

Implicit differentiation reveals

$$\begin{aligned}
\frac{\mathrm{d}\underset{\sim}{\hat{q}}}{\mathrm{d}\lambda} &= -\frac{\pi^w \hat{w}^\lambda + \pi^e(\tilde{e}^w \hat{w}^\lambda + \tilde{e}^\lambda) + \delta \frac{\mathrm{d}\mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)]}{\mathrm{d}\lambda}}{\pi^w \hat{w}^{q_0} + \pi^q + \pi^e \tilde{e}^w \hat{w}^{q_0} + \delta \frac{\mathrm{d}\mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)]}{\mathrm{d}q_0}} \\
&= -\frac{\pi^e \tilde{e}^\lambda + \hat{w}^\lambda(\pi^w + \pi^e \tilde{e}^w) + \delta \frac{\mathrm{d}\mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)]}{\mathrm{d}\lambda}}{\pi^q + \hat{w}^{q_0}(\pi^w + \pi^e \tilde{e}^w) + \delta \frac{\mathrm{d}\mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)]}{\mathrm{d}q_0}}
\end{aligned}$$

Letting $\pi^{-;=;+}$ be the profit function when $w_1 < w_0; w_1 = w_0; w_1 > w_0$, in which effort is given by $\tilde{e}(w_1, w_0, \lambda)^-; \tilde{e}_n; \tilde{e}(w_1, w_0)^+$, we have

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{E}_0[J_1]}{\mathrm{d}\lambda} &= \int_{\underset{\sim}{q}}^{q_l} \frac{\mathrm{d}\pi^-}{\mathrm{d}\lambda}\mathrm{d}F - \frac{\mathrm{d}\underset{\sim}{q}}{\mathrm{d}\lambda}\,\pi|_{q=\underset{\sim}{q}}\, f(\underset{\sim}{q}) + \frac{\mathrm{d}q_l}{\mathrm{d}\lambda}\pi^= f(q_l) \\
&\quad + \int_{q_l}^{q_u} \frac{\mathrm{d}\pi^=}{\mathrm{d}\lambda}\mathrm{d}F + \frac{\mathrm{d}q_u}{\mathrm{d}\lambda}\pi^= f(q_u) - \frac{\mathrm{d}q_l}{\mathrm{d}\lambda}\pi^= f(q_l) \\
&\quad + \int_{q_u}^{\infty} \frac{\mathrm{d}\pi^+}{\mathrm{d}\lambda}\mathrm{d}F - \frac{\mathrm{d}q_u}{\mathrm{d}\lambda}\pi^= f(q_u).
\end{aligned}$$

Noting that $\pi|_{q=\underset{\sim}{q}} \equiv 0$ and the other effects on the limits of integration cancel out, this reduces to

$$\frac{\mathrm{d}\mathbb{E}_0[J_1]}{\mathrm{d}\lambda} = \int_{\underset{\sim}{q}}^{q_l} \frac{\mathrm{d}\pi^-}{\mathrm{d}\lambda}\mathrm{d}F + \int_{q_l}^{q_u} \frac{\mathrm{d}\pi^=}{\mathrm{d}\lambda}\mathrm{d}F + \int_{q_u}^{\infty} \frac{\mathrm{d}\pi^+}{\mathrm{d}\lambda}\mathrm{d}F.$$

Now,

$$\begin{aligned}
\frac{\mathrm{d}\pi^{-;+}}{\mathrm{d}\lambda} &= \pi^w(\tilde{w}^r\hat{w}^\lambda + \tilde{w}^\lambda) + \pi^e(\tilde{e}^w(\tilde{w}^r\hat{w}^\lambda + \tilde{w}^\lambda) + \tilde{e}^r\hat{w}^\lambda + \tilde{e}^\lambda) \\
&= \pi^e(\tilde{e}^r\hat{w}^\lambda + \tilde{e}^\lambda)
\end{aligned}$$

since $\pi^w + \pi^e\tilde{e}^w = 0$. However, within the range of rigidity we have

$$\frac{\mathrm{d}\pi^=}{\mathrm{d}\lambda} = \pi^w\hat{w}^\lambda.$$

Note that $\hat{w}^\lambda$ doesn't depend on q_1 and $\tilde{e}^\lambda = 0$ when the wage exceeds the reference wage. Then, recalling the expression for $\theta(w_0, \lambda)$ in (11), we have that

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{E}_0[J_1]}{\mathrm{d}\lambda} &= \int_{\underset{\sim}{q}}^{q_l} \pi^e(\tilde{e}^r\hat{w}^\lambda + \tilde{e}^\lambda)\mathrm{d}F + \int_{q_l}^{q_u} \pi^w\hat{w}^\lambda\mathrm{d}F + \int_{q_u}^{\infty} \pi^e\tilde{e}^r\hat{w}^\lambda\mathrm{d}F \\
&= \hat{w}^\lambda\theta + \int_{\underset{\sim}{q}}^{q_l} \tilde{e}^\lambda\mathrm{d}F
\end{aligned}$$

This allows us to conclude that

$$\begin{aligned}
\pi^e\tilde{e}^\lambda + \hat{w}^\lambda(\pi^w + \pi^e\tilde{e}^w) + \delta\frac{\mathrm{d}\mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)]}{\mathrm{d}\lambda} &= \hat{w}^\lambda(\pi^w + \pi^e\tilde{e}^w + \delta\theta) + \pi^e\tilde{e}^\lambda + \delta\int_{\underset{\sim}{q}}^{q_l} \pi^e\tilde{e}^\lambda\mathrm{d}F \\
&= \pi^e\tilde{e}^\lambda + \delta\int_{\underset{\sim}{q}}^{q_l} \pi^e\tilde{e}^\lambda\mathrm{d}F
\end{aligned}$$

since the first-order condition for the initial employment contract implies $\pi^w + \pi^e\tilde{e}^w + \delta\theta = 0$.

Similar deductions allow us to conclude that $\frac{\mathrm{d}\mathbb{E}_0[J_1]}{\mathrm{d}q_0} = \hat{w}^{q_0}\theta$, and therefore that

$$\begin{aligned}\pi^q + \hat{w}^{q_0}(\pi^w + \pi^e \tilde{e}^w) + \delta \frac{\mathrm{d}\mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)]}{\mathrm{d}q_0} &= \hat{w}^{q_0}(\pi^w + \pi^e \tilde{e}^w + \delta\theta) + \pi^q \\ &= \pi^q.\end{aligned}$$

As such,

$$\frac{\mathrm{d}\hat{\underset{\sim}{q}}}{\mathrm{d}\lambda} = -\frac{\pi^e \tilde{e}^\lambda + \delta \int_{\underset{\sim}{q}}^{q_l} \pi^e \tilde{e}^\lambda \mathrm{d}F}{\pi^q} > 0$$

since we know from Theorem 1 that when the wage is below the reference wage $\tilde{e}^\lambda < 0$.

In summary, the hiring reservation productivity is unambiguously larger for a more loss averse worker; the layoff reservation productivity will also increase if either the firm pays a higher wage in the initial wage contract, or the direct effect of loss aversion on effort dominates the indirect effect. □

References

Agell, J. and Bennmarker, H. (2007). Wage Incentives and Wage Rigidity: A Representative View From Within. *Labour Economics*, 14(3):347–369.

Agell, J. and Lundborg, P. (1995). Theories of Pay and Unemployment: Survey Evidence from Swedish Manufacturing Firms. *Scandinavian Journal of Economics*, 97:295–307.

Agell, J. and Lundborg, P. (2003). Survey Evidence on Wage Rigidity and Unemployment: Sweden in the 1990s. *Scandinavian Journal of Economics*, 105(1):15–30.

Akerlof, G. A. (1982). Labor Contracts as Partial Gift Exchange. *Quarterly Journal of Economics*, 97(4):543–569.

Akerlof, G. A. and Yellen, J. L. (1990). The Fair Wage-Effort Hypothesis and Unemployment. *Quarterly Journal of Economics*, CV(2):255–283.

Altmann, S., Falk, A., Grunewald, A., and Huffman, D. (2014). Contractual incompleteness, unemployment, and labour market segmentation. *Review of Economic Studies*, 81(October 2013):30–56.

Arkes, H. R., Hirshleifer, D., Jiang, D., and Lim, S. S. (2008). Reference Point Adaptation:

Tests in the Domain of Security Trading. *Organizational Behavior and Human Decision Processes*, 105:67–81.

Arkes, H. R., Hirshleifer, D. A., Jiang, D., and Lim, S. S. (2010). A Cross-Cultural Study of Reference Point Adaptation: Evidence from China, Korea, and the US. *Organizational Behavior and Human Decision Processes*, 112:99–111.

Baker, G., Gibbs, M., and Holmstrom, B. (1994). The Wage Policy of a Firm. *Quarterly Journal of Economics*, 109(4):921–955.

Bartling, B. and Schmidt, K. M. (2015). Reference Points, Social Norms, and Fairness in Contract Renegotiations. *Journal of the European Economic Association*, 13(1):98–129.

Beaudry, P. and DiNardo, J. (1991). The Effect of Implicit Contracts on the Movement of Wages over the Business Cycle: Evidence from Micro Data. *Journal of Political Economy*, 99(4):665–688.

Benjamin, D. J. (2015). A Theory of Fairness in Labour Markets. *Japanese Economic Review*, 66(2):182–225.

Bewley, T. F. (1999). *Why Wages Don't Fall During a Recession.* Harvard University Press, London.

Bewley, T. F. (2007). Fairness, Reciprocity, and Wage Rigidity. In Diamond, P. and Vartiainen, H., editors, *Behavioural Economics and Its Applications*. Princeton University Press.

Bhaskar, V. (1990). Wage Relativities and the Natural Range of Unemployment. *Economic Journal*, 100(400):60–66.

Blinder, A. S. and Choi, D. H. (1990). A Shred of Evidence on Theories of Wage Stickiness. *Quarterly Journal of Economics*, 105:1003–1015.

Bolton, G. E. and Ockenfels, A. (2000). A Theory of Equity, Reciprocity, and Competition. *American Economic Review*, 90(1):166–193.

Campbell, C. M. and Kamlani, K. S. (1997). The Reasons for Wage Rigidity: Evidence from Survey of Firms. *Quarterly Journal of Economics*, 112:759–789.

Chemin, M. and Kurmann, A. (2012). Do Workers Feel Entitled to High Wages? Evidence from a Long-Term Field Experiment. *Working Paper.*

Cohn, A., Fehr, E., and Goette, L. (2014). Fair Wages and Effort Provision: Combining Evidence from a Choice Experiment and a Field Experiment. *Management Science*, 61(8):1777–1794.

Danthine, J.-P. and Kurmann, A. (2007). The Macroeconomic Consequences of Reciprocity in Labor Relations. *Scandinavian Journal of Economics*, 109(4):857–881.

Driscoll, J. C. and Holden, S. (2004). Fairness and Inflation Persistence. *Journal of the European Economic Association*, 2(2-3):240–251.

Druant, M., Fabiani, S., Kezdi, G., Lamo, A., Martins, F., and Sabbatini, R. (2012). Firms' Price and Wage Adjustment in Europe: Survey Evidence on Nominal Stickiness. *Labour Economics*, 19(5):772–782.

Du Caju, P., Kosma, T., Lawless, M., Messina, J., and Rõõm, T. (2014). Why firms avoid cutting wages: Survey evidence from European Firms. *World Bank Group Policy Research Working Paper 6976*, pages 1–39.

Dufwenberg, M. and Kirchsteiger, G. (2004). A Theory of Sequential Reciprocity. *Games and Economic Behavior*, 47(2):268–298.

Eliaz, K. and Spiegler, R. (2013). Reference Dependence and Labor-Market Fluctuations. *NBER Macroeconomics Annual*, 28:159–200.

Elsby, M. W. L. (2009). Evaluating the Economic Significance of Downward Nominal Wage Rigidity. *Journal of Monetary Economics*, 56(2):154–169.

Elsby, M. W. L., Michaels, R., and Ratner, D. (2015a). The Beveridge Curve: A Survey. *Journal of Economic Literature*, 53(3):571–630.

Elsby, M. W. L., Shin, D., and Solon, G. (2015b). Wage Adjustment in the Great Recession. In Mas, A. and Card, D., editors, *The Labour Market in the Aftermath of the Great Recession*. University of Chicago Press.

Falk, A. and Fischbacher, U. (2006). A Theory of Reciprocity. *Games and Economic Behavior*, 54(2):293–315.

Fehr, E., Goette, L., and Zehnder, C. (2009). A Behavioral Account of the Labor Market: The Role of Fairness Concerns. *Annual Review of Economics*, 1(1):355–384.

Fehr, E., Hart, O., and Zehnder, C. (2011). Contracts as Reference Points—Experimental Evidence. *American Economic Review*, 101(April):493–525.

Fehr, E., Hart, O., and Zehnder, C. (2014). How Do Informal Agreements and Revision Shape Contractual Reference Points? *Journal of the European Economic Association*, 13(1):1–28.

Fehr, E. and Schmidt, K. M. (1999). A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics*, 114(3):817–868.

Galuscak, K., Keeney, M., Nicolitsas, D., Smets, F., Strzelecki, P., and Vodopivec, M. (2012). The Determination of Wages of Newly Hired Employees: Survey Evidence on Internal Versus External Factors. *Labour Economics*, 19(5):802–812.

Gneezy, U. and List, J. A. (2006). Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments. *Econometrica*, 74(5):1365–1384.

Hart, O. and Moore, J. (2008). Contracts as Reference Points. *Quarterly Journal of Economics*, CXXIII(February):1–48.

Herweg, F. and Schmidt, K. M. (2015). Loss Aversion and Inefficient Renegotiation. *Review of Economic Studies*, 82:297–332.

Herz, H. and Taubinsky, D. (2014). What Makes a Price Fair? An Experimental Study of Market Experience and Endogenous Fairness Norms. *Working Paper*, pages 1–59.

Hicks, J. R. (1963). *The Theory of Wages*. MacMillan and Co., London, 2nd edition.

Holden, S. and Wulfsberg, F. (2014). Wage Rigidity, Inflation, and Institutions. *Scandinavian Journal of Economics*, 116(2):539–569.

Howitt, P. (2002). Looking Inside the Labor Market: A Review Article. *Journal of Economic Literature*, 40(1):125–138.

Kahneman, D., Knetsch, J. L., and Thaler, R. H. (1986). Fairness as a Constraint on Profit Seeking : Entitlements in the Market. *American Economic Review*, 76(4):728–741.

Kahneman, D. and Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2):263–292.

Kaufman, T. (1984). On Wage Stickiness in Britain's Competitive Sector. *British Journal of Industrial Relations*, 22:101–112.

Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money*. London: MacMillan and Cambridge University Press.

Kőszegi, B. and Rabin, M. (2006). A Model of Reference-Dependent Preferences. *Quarterly Journal of Economics*, CXXI(November):1133–1165.

Koch, C. (2014). Do Reference Points Lead to Wage Rigidity? Experimental Evidence. *Working Paper*, pages 1–55.

Koenig, F., Manning, A., and Petrongolo, B. (2014). Reservation Wages and the Wage Flexibility Puzzle. *CEP Discussion Paper No 1319*.

Lindbeck, A. and Snower, D. J. (1986). Wage Setting, Unemployment, and Insider-Outsider Relations. *American Economic Review*, 76(2):235–239.

Malmendier, U., Velde, V. L., and Weber, R. (2014). Rethinking Reciprocity. *Annual Review of Economics*, 6:849–974.

Marshall, A. (1890). *Principles of Economics*. MacMillan and Co., London, 8th edition.

Mas, A. (2006). Pay, Reference Points, and Police Performance. *Quarterly Journal of Economics*, CXXI(3):783–821.

Mortensen, D. T. and Pissarides, C. A. (1994). Job Creation and Job Destruction in the Theory of Unemployment. *Review of Economic Studies*, 61(3):397–415.

Okun, A. M. (1981). *Prices and Quantities: A Macroeconomic Analysis*. Basil Blackwell Publisher, Oxford.

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *American Economic Review*, 83(5):1281–1301.

Schmieder, J. F. and von Wachter, T. (2010). Does Wage Persistence Matter for Employment Fluctuations? Evidence from Displaced Workers. *American Economic Journal: Applied Economics*, 2(July):1–21.

Shapiro, C. and Stiglitz, J. E. (1984). Equilibrium Unemployment as a Worker Discipline Device. *American Economic Review*, 74(3):433–444.

Shimer, R. (2005). The Cyclical Behavior of Equilibrium Unemployment and Vacancies. *American Economic Review*, 95(1):25–49.

Slichter, S. H. (1920). Industrial Morale. *Quarterly Journal of Economics*, 35(1):36–60.

Smith, J. C. (2015). Pay Growth, Fairness and Job Satisfaction: Implications for Nominal and Real Wage Rigidity. *Scandinavian Journal of Economics*, 117(3):852–877.

Snell, A. and Thomas, J. P. (2010). Labor Contracts, Equal Treatment, and Wage-Unemployment Dynamics. *American Economic Journal: Macroeconomics*, 2(3):98–127.

Solow, R. M. (1979). Another Possible Source of Wage Stickiness. *Journal of Macroeconomics*, 1(1):79–82.

Stommel, E. (2013). *Reference-Dependent Preferences: A Theoretical and Experimental Investigation of Individual Reference-Point Formation*. Springer Gabler, 2013 edition.

Stüber, H. and Beissinger, T. (2012). Does Downward Nominal Wage Rigidity Dampen Wage Increases? *European Economic Review*, 56(4):870–887.

Summers, L. H. (1988). Relative Wages, Efficiency Wages, and Keynesian Unemployment. *American Economic Review*, 78(2):383–388.

Williamson, O. E. (1985). *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. The Free Press, New York.

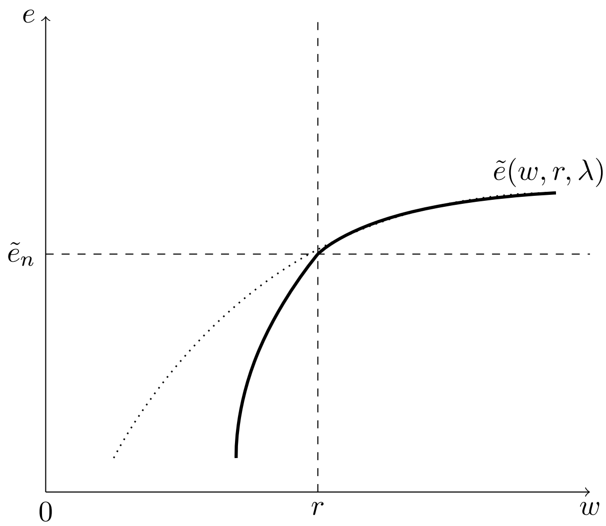

Figure I:
Asymmetric Reference-dependent Reciprocity

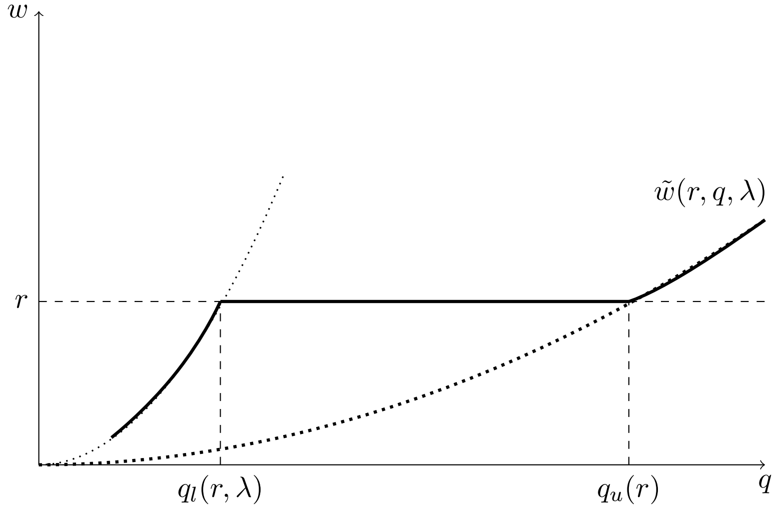

Figure II:
The Wage Setting Rule for a Loss Averse Worker whose Reference Wage is r. The thick dotted line represents $\tilde{w}(r,q)^{+}$ and the thin dotted line $\tilde{w}(r,q,\lambda)^{-}$.

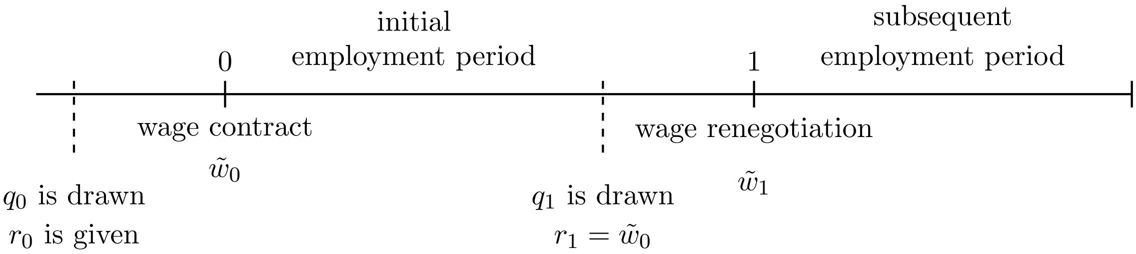

Figure III:
Employment Relationship Time-line.

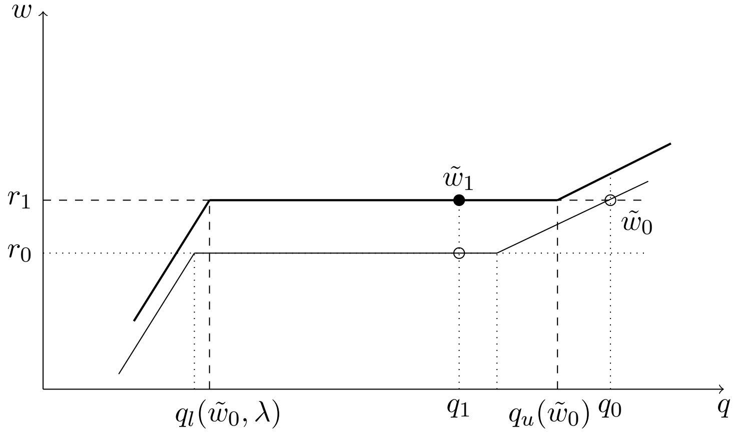

Figure IV:
Adaptation and Downward Wage Rigidity when a Worker is Initially Paid Above their Reference Wage.

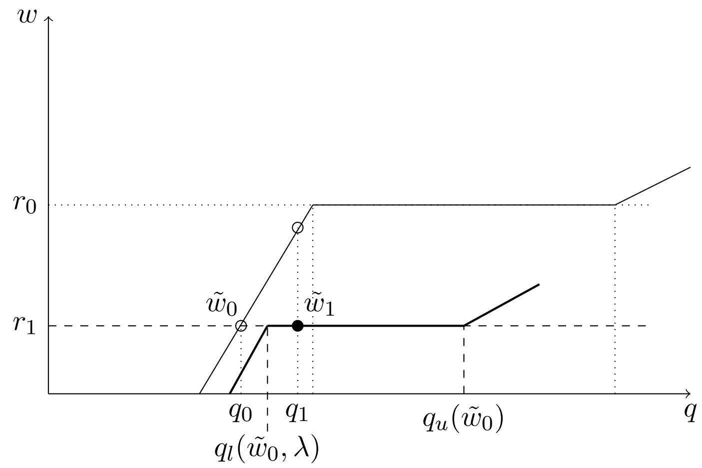

Figure V:
Adaptation and Upward Wage Rigidity when a Worker is Initially Paid Below their Reference Wage.

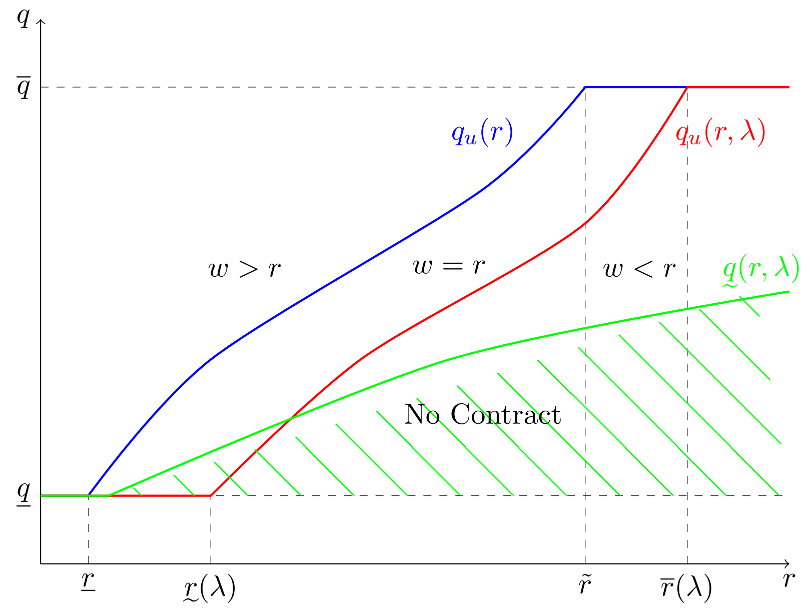

Figure VI:
The Range of Rigidity with a Truncated Support of the Match Productivity Distribution.