SIRE DISCUSSION PAPER
SIRE-DP-2015-57

A Theory of Wage Setting Behavior

Marco Fongoni
Alex Dickson

UNIVERSITY OF STRATHCLYDE

www.sire.ac.uk
A Theory of Wage Setting Behavior∗

Marco Fongoni† and Alex Dickson‡

Department of Economics, University of Strathclyde

July 2015

Abstract

Concerns for fairness, workers’ morale and reciprocity influence firms’ wage setting policy. In this paper we formalize a theory of wage setting behavior in a simple and tractable model that explicitly considers these behavioral aspects. A worker is assumed to have reference-dependent preferences and displays loss aversion when evaluating the fairness of a wage contract. The theory establishes a wage-effort relationship that captures the worker’s reference-dependent reciprocity, which in turn influences the firm’s optimal wage policy. The paper makes two key contributions: it identifies loss aversion as an explanation for a worker’s asymmetric reciprocity; and it provides realistic and generalized microfoundation for downward wage rigidity. We further illustrate the implications of our theory for both wage setting and hiring behavior. Downward wage rigidity generates several implications for the outcome of the initial employment contract. The worker’s reference wage, his extent of negative reciprocity and the firms expectations are key drivers of the propositions derived.

Keywords: reference dependence; loss aversion; morale; reciprocity; employment contract; downward wage rigidity; wage setting behavior.

JEL Classification Numbers: 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. We would also like to thank the participants of the 2015 Stirling PhD Conference in Behavioural Sciences and, in particular, Leonhard Lades for valuable comments on the work.
†Corresponding author. E-mail: marco.fongoni@strath.ac.uk.
‡E-mail: alex.dickson@strath.ac.uk.
“No one would expect a workman who feels that his employer pays the bare minimum which economic conditions compel him to pay to reciprocate by doing his best to promote the interests of his employer.”

(Slichter, 1920)

1 Introduction

Wage cuts may be costly. The basic premise of this paper is that when making wage setting decisions employers take into account their workers’ perceptions of fairness, and are concerned about the potential impact of a wage offer on workers’ morale and productivity. Therefore, to understand wage setting behavior, and ultimately wage dynamics, it is crucial to explain the nature of these concerns, identify the factors that drive them and analyze their implications.

Building on an extensive body of research we propose a synthesis from which we draw our theory of wage setting behavior. The synthesis reconciles early theoretical insights of economists such as Slichter (1920), Hicks (1963), Okun (1981), Akerlof (1982) and Akerlof and Yellen (1990), with recent convergent findings of anthropological and experimental research, ranging from field surveys (Bewley, 2007) to laboratory and field experiments (Fehr et al., 2009). As such we believe our framework to be a first step toward a unified theory of wage setting behavior.

Workers’ morale is a key determinant of the psychological cost incurred by workers when they engage in productive activity, and it is strictly dependent on whether workers feel the firm is treating them fairly. Given the incomplete nature of the employment contract, that is, the quality of the workers’ performance is discretionary and not contractible, a firm is concerned about paying its workers a wage that does not violate their perceptions of fairness. The employment relationship is grounded on the mutual understanding of and compliance with the concept of reciprocity: as long as the firm will pay its workers a “fair” wage, the workers will perform productive activity at a “normal” standard. We discuss these aspects and the related literature in depth in section 2, where we also verbally expose our synthesis.

In Section 3 we formalize the theory into a general model of wage setting behavior featuring a worker-firm match. In the spirit of Kahneman and Tversky (1979) and Köszegi and Rabin (2006), a worker is assumed to have reference-dependent preferences and displays loss aversion when evaluating the fairness of a wage contract. The model also explicitly captures the role of the worker’s morale in relation with his perceptions of fairness through the formulation of a “morale function”. Wage changes, morale and the worker’s choice of effort are intrinsically linked through our assumptions about the worker’s preferences.
The worker’s perceptions of fairness are assumed to be captured by a given reference wage. As our interest is to elucidate the main psychological forces that drive the behavior of a worker and a firm engaged in wage negotiations, we do not specify the nature of the reference wage in the general formalization. This approach has the advantage of permitting the derivation of general theoretical implications for wage setting behavior that are valid for any given reference wage.

The model establishes a wage-effort relationship that characterizes a worker’s reference-dependent reciprocity, where effort is more responsive to wage changes below the reference wage than to wage changes above (Theorem 1). In turn, concerns about the worker’s reciprocity response influence the optimal wage setting policy of a firm and generate a discontinuous rate of adjustment of the wage with respect to changes in external market conditions, represented in our model by exogenous changes in the worker-firm match productivity (Theorem 2).

Among other implications, our theory makes two key contributions: i) it offers a psychological foundation for asymmetric reciprocity, identifying loss aversion as an explanation for negative reciprocity being a stronger force than positive reciprocity; and ii) it provides realistic microfoundations for downward wage rigidity, achieved by encompassing a number of key behavioral insights set out by other models (e.g. Akerlof (1982), Akerlof and Yellen (1990), Danthine and Kurmann (2008), Elsby (2009) and Eliaz and Spiegler (2013)).

The model is also useful to elucidate the nature and the drivers of reciprocal behavior in the employment relationship. In our model, the driver of reciprocity is the worker’s evaluation of the wage relative to a benchmark reference wage: a fair wage in excess of the reference wage is reciprocated with greater than normal effort; whilst unfair wages give rise to sub-normal effort. As such, reciprocity is intentions-based in the spirit of Rabin (1993), which distinguishes our model from other formalizations of reciprocal behavior in labor market relations based on motives of inequity aversion (e.g. Fehr and Schmidt (1999), Bolton and Ockenfels (2000) and Benjamin (2015)). Firms, conversely, are assumed to be purely self-interested, but of course respond to workers’ reciprocity when seeking to optimize their profits.

In Section 4 a parameterized version of the model is used to characterize an employment contract and to analyze renegotiation. Within a two-period dynamic stochastic optimization set-up we illustrate the implications of our theory for both wage setting and hiring behavior. Here we explore the hypothesis of “contracts as reference points” (Hart and Moore, 2008) and assume that an employed worker’s reference wage is endogenously determined by the initial wage contract signed with the firm. We find that wage contracts can be rigid and irreversible and show that such irreversibility generates several implications for the outcome of the initial employment contract. A forward-looking firm that anticipates the potential future cost of downward wage rigidity has an incentive to offer a
lower initial wage contract and requires a higher level of observed match productivity in order to hire a worker. The model also highlights the worker’s reference wage, his extent of negative reciprocity (linked to his degree of loss aversion) and the firm’s expectations as the key drivers of the propositions derived.

The theory proposed owes much to the different ideas and perspectives that have contributed to the emergent consolidation in our understanding of labor markets and the employment relationship. Ultimately, the contribution of our paper is to formalize this theory and provide a tractable benchmark model for the analysis of wage setting behavior. However, the theory also stresses the need for a better understanding of workers’ perceptions of what is a “fair” wage. Based on this conjecture, in the concluding section we make a case for a more systematic approach to the analysis of workers’ reference wage formation.

2 Morale, fairness and reciprocity in the employment relationship

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 underlie the observation of downward wage rigidity.

The main idea is that workers’ morale is an important determinant of their effort and productivity in the workplace. When workers perceive they are treated unfairly they withhold effort, as a consequence of low morale, and productivity is thereby adversely altered. Given the incomplete nature of the employment contract, firms’ managers have an incentive to treat workers fairly and refrain from cutting wages in the event of adverse economic conditions. This idea embeds the core of the consensus that has emerged from prominent behavioral theories and convergent findings in the recent anthropological and experimental research.

2.1 Early insights

That workers’ morale is strictly linked with their perceptions of fairness and with their productivity, and that employers are concerned about these issues when deciding upon wage policies, has long been acknowledged by economists, at least since the turn of the twentieth century. 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. Later on, Hicks (1963), Solow (1979) and Okun (1981) put forward similar arguments when discussing the possible
sources of the Keynesian wage floor, that is, the downward stickiness of money-wages. In support of the relative wage theory of Keynes (1936), they argued that resistance to cut nominal wages also comes from employers, concerned about the effects of wage cuts on workers’ morale, “ability” and “willingness to work” (Hicks, 1963, p. 94-95). Among these early insights, probably the most important contributions are the papers of Akerlof (1982) and Akerlof and Yellen (1990), which provide the first formal treatments of what would later be termed positive and negative reciprocity in the behavioral economics literature. Modelling the labor contract as a gift-exchange, Akerlof (1982) formalizes the idea that an employer is willing to pay a worker a higher wage relative to an assumed “fair” wage, with the expectations of receiving higher effort and productivity in exchange. On the other hand, drawing from psychology, sociology and empirical observation, Akerlof and Yellen (1990) propose the fair wage-effort hypothesis, according to which a worker’s effort is proportional only to wage changes below a specified “fair” wage. To maintain efficient levels of effort, firms offer fair wages, that are above the market clearing level. Though appealing, such theories remained untested since empirical analysis could only investigate the derived implications for wage determination and unemployment, and not the validity of their assumptions.

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 Galusca et al. (2012), Druant et al. (2012) and Du Caju et al. (2014), our understanding of the employment relationship has been greatly enhanced. By interviewing firms’ managers and labor leaders in several countries these studies provide a first falsification test of the behavioral assumptions put forward in the theoretical literature. The central finding is that firms’ managers are concerned about treating workers fairly, to avoid damaging their morale and productivity, and ultimately the firms’ profitability. 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). Moreover it has been found that wage rises could be used to induce good morale, resulting in improvements in effort and cooperation among workers. As such, these findings suggest the existence of a relationship between wage changes and workers’ effort.

However there are some exceptions. For instance Campbell and Kamlani (1997) find that effort responds more intensely to wage cuts than to wage raises: any positive effect of

---

1 Other anthropological studies include the papers of 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).
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

In support of the main findings coming from field surveys, there is an additional stream of evidence that comes from laboratory and field experiments. A key advantage of this approach is tight experimental control that allows for the collation of rigorous evidence and for the researcher to establish causation within employer and worker interactions (Fehr et al., 2009), albeit in a laboratory setting.

Overall the most important finding is confirmation of the existence of reciprocal behavior in the employment relationship: when people receive extra pay as a gift, 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 inefficient effort (negative reciprocity). However, although the evidence on negative reciprocity seems quite robust among a number of studies, field experiments document weaker evidence of positive reciprocity with respect to laboratory experiments (Malmendier et al., 2014). Cohn et al. (2014) try to address this inconsistency by combining results from the field into a subsequent laboratory experiment. They infer that positive reciprocity exists but may quickly disappear, which is consistent with previous research. 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. Experimental evidence reinforces the existence of a wage-effort relationship, suggesting a stronger response of effort to wage changes that are perceived as unfair—such as wage cuts—than to wage changes that are perceived as gifts.²

2.4 The proposed synthesis

Drawing from the literature streams discussed, we propose a synthesis of a theory of wage setting behavior. This synthesis captures the essential features of wage setting

²For more comprehensive surveys of this literature see Fehr et al. (2009), Cooper and Kagel (2013) and Malmendier et al. (2014).
and the employment relationship that emerge from several ideas and perspectives in the literature. 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 the productive activity. As Bewley (2007) puts it, 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. On the other hand when morale is low, workers tend to hold back cooperation and cease to identify themselves with the firm. We assume that workers’ morale is directly related to their willingness to exert effort: when morale is good, cooperation is enhanced and performing the productive activity becomes less psychologically costly to workers; when morale is low, workers are less motivated and the psychological costs of exerting effort increase.

Perceptions of Fairness. Changes in workers’ morale are strictly dependent on whether workers feel they are treated fairly by their employer. Although fairness in this sense could relate to many aspects of the job, we consider only the workers’ perceptions of fairness in terms of their remuneration. Following the standard approach in the literature we capture these perceptions within a reference “fair” wage. The fairness of a wage contract is evaluated by workers relative to a reference “fair” wage (reference wage henceforth): a wage below the reference wage is perceived as unfair, while a wage above is perceived as a gift. By incorporating the intrinsic psychological aspect of human decision making of loss aversion (Kahneman and Tversky, 1979) in the model—in individuals evaluate outcomes with respect to a reference point, and losses (due to deviations below this reference point) loom larger than equivalent-sized gains—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 workers’ morale than a same size wage raise (perceived as a gain). This assumption is consistent with the evidence reported by surveys and experiments as discussed above. 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”.

Reciprocity. Probably the result that stands out the most from the literature discussed here is the idea that the worker-firm relationship is based on a mutual understanding of workers’ reciprocal behavior. In our model, this concept arises from our assumptions

---

³The 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 “perceptions of entitlement” in Kahneman et al. (1986), “the reference frame of fairness judgements” in Fehr et al. (2009) and “feelings of entitlement” in Hart and Moore (2008).

⁴The mutual and tacit compliance with the norm of reciprocity has grounded foundations in psy-
about workers’ morale and perceptions of fairness. Suppose that a firm sets a wage contract that is considered unfair by the workforce. Workers start to feel a grievance against the firm and morale decreases which we capture by a greater psychological cost of exerting effort. As a consequence, workers negatively reciprocate the treatment they perceive as unfair by exerting less effort. The same response, but in the opposite direction, would arise if the firm sets a wage that is above the workers’ reference wage. Nevertheless, due to the assumptions that unfair behavior has a stronger impact on workers’ morale (loss aversion), effort is more responsive to wage changes that are considered unfair as opposed to wage changes considered gifts. This type of workers’ reciprocity differs conceptually from other formalizations that are based on inequity aversion (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000; Benjamin, 2015) or altruism (Levine, 1998). Since the workers’ reciprocity behavior strictly depends on their morale, a parallel can be drawn with the idea put forward by Cox et al. (2007), where reciprocity is an explicit function of the “emotional state”.

*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”. 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”) (Williamson, 1985, p. 262-63).

In the light of our theory, this feature of the employment contract involves important implications for firms’ wage setting behavior. As we have pointed out, workers’ effort is dependent on morale, which in turn is affected by the workers’ perceptions of being treated fairly. Due to the incomplete nature of the employment contract, a firm can only motivate its workers’ “consummate cooperation” by offering a wage contract that does not violate their standards of fairness. This insight is consistent with the consensus that firms’ managers recognize the importance of workers’ morale and are concerned about treating them fairly when setting wages. However, the reciprocal behavior of employers is not motivated by a general sense of fairness, equity or altruism, as has been commonly assumed by several theories of fairness and reciprocity in the literature (some of which

---

5See Malmendier et al. (2014) for a more exhaustive classification of reciprocity theories.
are cited above). In fact, the fairness concerns are induced by the interest of managers about workers’ productive effort due to the effect of this on the firm’s profitability.

Through using an example we can anticipate some insights that our theory can provide. Suppose that a firm is facing a persistent period of low productivity that adversely affects its profitability. In such an environment labor costs may be too high relative to productivity so the firm would benefit from reducing wages, absent other considerations. The key intuition of our theory is that there is a cost associated with such wage policy. If workers perceive the wage reductions to be unfair their effort and willingness to work will drop, exacerbating the decline in productivity and further influencing the firm’s profitability. Thus the firm’s managers may refrain from cutting wages, if the related cost in the form of negative reciprocity is greater than the benefit of paying lower wages. This insight is one of the major implications of our theory, which confirms the predictions of the early prominent hypotheses and offers a psychological foundation for it drawing from the recent anthropological and experimental research.

In the following two sections we formalize this theory in a framework consistent with utility and profit maximization. Subsequently we explore several theoretical implications for wage setting behavior and the employment contract.

3 A Model

We present a general formalization of our theory into a model of wage setting behavior. To simplify the analysis we consider an ongoing employment relationship between one employed worker and one operating firm. The purpose of this analysis is to elucidate the main behavioral forces underlying the worker’s and the firm’s decision making process and to assess how these forces shape the outcome of wage setting behavior.

We model this situation as a two-stage game with complete and perfect information: first, the firm sets the wage; then, after observing the wage, the worker chooses the level of productive effort. The firm seeks to maximize profits and the worker seeks to maximize utility; the timing along with the assumption of complete information means that the firm infers the worker’s best response to any wage offer and takes that into account when setting the optimal wage. This assumption reflects the evidence previously discussed that firms’ managers are aware of their workers’ perceptions of fairness and of the potential effects that a wage offer may have on their morale and effort.

Following the principle of backward induction we first derive the worker’s optimal effort response for any wage offer with respect to the reference wage. Then we derive the firm’s optimal wage policy taking into account the worker’s effort choice.
3.1 Employed Worker’s Preferences

We assume an employed worker’s overall utility $U(\cdot)$ to be a function of the wage $w$, his reference wage $r$ and the level of productive effort $e$. Consider the wage $w$ as the total compensation package that the worker negotiates with the firm. The reference wage $r$ captures the worker’s perceptions of fairness with respect to his remuneration, which for the purpose of this exposition is taken to be exogenously determined. The level of productive effort $e$ represents a combination of the worker’s effort and productivity, which translates into the worker’s performance quality as a measure of efficiency and cooperation. As such, the worker’s productive effort $e$ represents what Williamson (1985) defines “consummate cooperation”. For clarity of exposition we further assume that $U(\cdot)$ is additively separable in two components:

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

where $u(w|r)$ is the “wage utility”, representing the worker’s perceived utility from wage evaluations and $v(e; w, r)$ is the “effort utility”, representing the worker’s perceived net utility from productive effort given the wage in relation to the reference wage.

**Wage Utility.** We assume that a worker’s utility from the wage is characterized by a standard utility function $m(w)$ augmented by a gain-loss function $n(w|r)$:

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

This formulation is inspired by Köszegi and Rabin (2006) and reflects the idea that the gain-loss perceptions relative to a reference wage are not all that workers care about. In fact, the function $m(w)$ captures the effects of absolute wage levels on the worker’s utility and it is assumed to be increasing in $w$ at a diminishing rate.

**A0.** $m(w)$ is increasing and continuously differentiable as many times as required, $m'(w) > 0$ and $m''(w) < 0$ for all $w$.

The function $n(w|r) \equiv \mu(m(w) - m(r))$ captures the net effect on the worker’s utility of wage deviations from the reference wage, where $\mu(\cdot)$ is a gain-loss function in the spirit of Kahneman and Tversky (1979) and Tversky and Kahneman (1991, 1992). For clarity of exposition we assume that $\mu(\cdot)$ is piecewise linear, implying that workers have constant, rather than diminishing as is sometimes assumed, sensitivity to gains and losses.

**A1.** $\mu(x)$ is continuous for all $x$, twice differentiable for $x \neq 0$ with $\mu(0) = 0$;

**A2.** $\mu(x)$ is strictly increasing and linear, that is, for all $x \neq 0$, $\mu'(x) > 0$ and $\mu''(x) = 0$;

**A3.** $\mu'_-(0)/\mu'_+(0) \equiv \lambda \geq 1$, where $\mu'_+(0) \equiv \lim_{x \to 0} \mu'(|x|)$ and $\mu'_-(0) \equiv \lim_{x \to 0} \mu'(-|x|)$.
Without loss of generality we can then write $\mu(x) = \eta x$ when $x \geq 0$ and $\mu(x) = \lambda \eta x$ when $x < 0$, where $\eta > 0$ is a scaling parameter that represents the worker’s subjective weight on gain-loss utility and $\lambda \geq 1$ represents the worker’s subjective degree of loss aversion.\(^6\) An employed worker is therefore assumed to have reference-dependent preferences when evaluating the fairness of a wage contract. Moreover, as argued in Section 2.4, our theory explicitly postulates that a given wage below the reference wage has a greater psychological impact on the worker’s utility than a wage the same amount above. We capture this feature by assuming that a worker is loss averse when evaluating the monetary outcome of the wage. In the reminder of the paper we therefore explore the properties of our model when $\lambda > 1$ and treat the possibility of $\lambda = 1$ as a special case.

**Effort Utility.** We assume that an employed worker’s preferences are also dependent on the utility he derives from performing the productive activity:

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

where $b(e)$ represents the worker’s intrinsic psychological benefits of being productive, $c(e)$ represents the worker’s intrinsic psychological and physical costs of being productive and $M(e; w, r) \equiv g(e) n(w | r)$ represents an additional psychological cost/benefit of productive effort which depends on the worker’s perceptions of fairness.\(^7\)

**B1.** $b(e)$ and $c(e)$ are continuously differentiable as many times as required, $b'(e) > 0$, $b''(e) \leq 0$ and $c'(e) > 0$, $c''(e) > 0$ for all $e$;

**B2.** $g(e)$ is continuous and linear, $g'(e) > 0$, $g''(e) = 0$ for all $e$.

Assumption B2 combined with A0-A3 define $M(\cdot)$ as a gain-loss function that is increasing in the level of productive effort $e$, increasing in the wage $w$ and decreasing in the reference wage $r$:

$$M(e; w, r) \equiv g(e) \cdot \mu(m(w) - m(r)). \quad (2)$$

For any level of productive effort $e$, whenever a worker is paid a wage above the reference wage $M(\cdot)$ is positive, representing additional psychological benefits of being productive; whenever a worker is paid a wage below the reference wage $M(\cdot)$ is negative, placing additional psychological costs of being productive. Moreover, as the worker is loss averse (A3 with $\lambda > 1$), these effects are asymmetric: the additional psychological cost of effort

\(^6\)Assumptions A1-A3 resemble closely the assumptions of Köszegi and Rabin (2006) over the properties of their “universal gain-loss function” without diminishing sensitivity. However we relax the strict inequality in the loss aversion coefficient $\lambda$, which allows us to explore the properties of the model even when a worker is not loss averse.

\(^7\)Although the assumption of including $b(e)$ in the worker’s utility is not standard in the theoretical literature, it is based on the idea that a worker also perceives a positive psychological satisfaction of performing the job (see for instance the discussion in the appendix of Altmann et al. (2014))
when a worker is paid a wage below the reference wage is larger than the additional psychological benefit of effort when a worker is paid a wage the same amount above the reference wage.

This function formalizes what happens to the worker’s morale when a wage offer deviates from what the worker perceives to be fair: if a wage is considered a gift, morale increases and the perceived net utility from productive effort increases; if a wage is considered unfair, morale decreases and the perceived net utility from productive effort decreases by more. Due to these features we refer to $M(\cdot)$ as the “morale function”.8

Our assumptions over the preferences of an employed worker place a particular emphasis on the worker’s perceptions of fairness and morale as two crucial determinants of overall utility $U(\cdot)$. Note that the gain-loss utility $\mu(\cdot)$ is present as an argument of both wage utility and effort utility. This specification is consistent with the idea that wage deviations from what is perceived as the “fair” wage affect a worker’s utility not only through influences on his standards of living/purchasing power (wage utility), but also in his willingness to exert effort and to cooperate with the organization (effort utility). Thus we explicitly formalize the idea that a worker’s satisfaction from the job does not depend on absolute wage levels only but also on the his perceptions of fairness and morale.

3.2 Reciprocity

Given a wage offer $w$ and a reference wage $r$, an employed worker seeks to maximize utility through his choice of effort $e$:

$$\max_{e \geq 0} U(e; w, r) = u(w|r) + v(e; w, r),$$

that is,

$$\max_{e \geq 0} U(e; w, r) = m(w) + \mu(m(w) - m(r)) + b(e) - c(e) + g(e) \cdot \mu(m(w) - m(r)),$$

Let $\tilde{e}(w, r)$ be the utility maximizing effort such that the following first order condition is satisfied:

$$\Omega(e, w, r) \equiv \frac{db}{de} - \frac{dc}{de} + \frac{dg}{de} \cdot \mu(m(w) - m(r)) \leq 0 \tag{3}$$

with equality if $\tilde{e}(w, r) > 0$.9 The following theorem defines the properties of the worker’s optimal effort function.

---

8From a modelling perspective, the functional form of $M(\cdot)$ is similar to the function used by Danthine and Kurmann (2008) to derive the workers’ reciprocal gift in terms of effort, in the spirit of Rabin (1993). However it is conceptually different as $M(\cdot)$ aims to capture the effect of the worker’s morale on his willingness to work, and not his preferences for reciprocal behavior.

9The second order condition that ensures the first order condition is both necessary and sufficient is $b''(e) - c''(e) < 0$, which is satisfied under assumption B1.
Theorem 1. Worker’s Reciprocity Behavior.

The worker’s optimal productive effort function can be expressed by the system:

\[ \tilde{e}(w, r) = \begin{cases} 
\tilde{e}(w, r)^+ & \text{if } w > r \\
\tilde{e}_n & \text{if } w = r \\
\tilde{e}(w, r)^- & \text{if } w < r. 
\end{cases} \]

For a given \( r \), the optimal effort \( \tilde{e}(w, r) \) is a continuous and increasing function of \( w \) with \( \frac{\partial \tilde{e}}{\partial w} > 0 \), \( \frac{\partial^2 \tilde{e}}{\partial w^2} < 0 \) for all \( w \neq r \). Moreover, if the worker is loss averse

\[ \frac{\partial \tilde{e}}{\partial w} \bigg|_{w<r} < \frac{\partial \tilde{e}}{\partial w} \bigg|_{w>r} = \lambda > 1. \]

For a given \( w \), the optimal effort \( \tilde{e}(w, r) \) is a continuous and decreasing function of \( r \) with \( \frac{\partial \tilde{e}}{\partial r} < 0 \) for all \( w \neq r \). Finally, for any given \( w \) and \( r \), whenever \( w = r \) the optimal effort \( \tilde{e}(w, r) \) is characterized by \( \tilde{e}_n \equiv \{ e : \frac{\partial^2 \tilde{e}}{\partial w^2} = \frac{\partial \tilde{e}}{\partial r} \} \).

Theorem 1 establishes a wage-effort relationship that is reference dependent and asymmetric: the worker’s effort \( \tilde{e}(w, r) \) is more responsive to wage changes below the reference wage than to wage changes above. As such, Theorem 1 provides a formal derivation of what can be interpreted as an employed worker’s reference-dependent reciprocity, where positive and negative reciprocity are captured by the partial derivatives of \( \tilde{e}(w, r) \) with respect to the wage \( w \), evaluated above or below the reference wage \( r \). The theorem also highlights that negative reciprocity is greater than positive reciprocity by a factor \( \lambda \), which is the worker’s degree of loss aversion. The established reference-dependent reciprocity is shown in Figure 1.\(^{10}\)

Corollary 1. For any given \( w < r \), \( \frac{\partial \tilde{e}(w, r)^-}{\partial \lambda} < 0 \) and \( \frac{\partial^2 \tilde{e}(w, r)^-}{\partial w \partial \lambda} > 0 \).

Corollary 1 is complementary to Theorem 1 and implies that when a worker perceives he is unfairly paid \( (w < r) \), the greater his degree of loss aversion the lower is the optimal effort \( \tilde{e}(w, r) \) exerted. Moreover, a greater loss aversion increases the worker’s negative reciprocity response, captured by the marginal effect of the wage on effort when \( w < r \). On the other hand when a worker perceives he has been treated fairly \( w = r \), the optimal choice of effort is purely determined by an intrinsic balance between his psychological and physical costs/benefits of productive activity. We refer to this as “normal” effort, denoted by \( \tilde{e}_n \): no matter how high is a wage, as long as the worker is paid the reference

---

\(^{10}\)Also note that a considerably low wage below the reference wage would result in negative productive effort. Although it could be interpreted as “output destruction” by the worker, with potential consequences for the firm’s profits, we abstract from such possibility and we focus only on positive values of effort.
wage he exerts the amount of effort that is determined by his intrinsic motivation to work (where the marginal benefit of effort $b'(e)$ equals the marginal cost $c'(e)$).

Since the properties of $\tilde{e}(w,r)$ are direct implications of our assumptions regarding the employed worker’s preferences (reference dependence, loss aversion and the morale function), our theory offers a psychological foundation for asymmetric reciprocity. Negative reciprocity is stronger than positive reciprocity because when a worker is loss averse, the evaluation of wage changes that are perceived to be unfair have a stronger impact on his morale (and therefore on the psychological costs of productive effort) than wage changes perceived as gifts. Hence our model identifies loss aversion as an explanation for negative reciprocity being stronger than positive reciprocity.\(^{12}\)

Theorem 1 is consistent with the main findings of anthropological and experimental evidence as reported in the previous section. Moreover it stands as a general microfoundation of the effort functions assumed by Akerlof (1982) in his gift exchange model (when $w > r$) and by Akerlof and Yellen (1990) in their fair wage-effort hypothesis (when $w < r$). It also provides a conceptual microfoundation of the reduced form effort function assumed by Elsby (2009) and the reduced form, reference-dependent output

\(^{11}\)This feature of the derived effort function addresses the critique put forward by Bewley (1999) to models in which effort is dependent on wage levels. It is true that our optimal effort function increases proportionally with the wage. However, it can also accommodate the theoretical prediction that, if a worker’s reference wage shifts to offset a wage raise above the reference wage, effort always gets back to “normal”. That is, if workers are characterized by the effort function as derived in Theorem 1 and always adapt their reference wage to the wage received, a firm cannot keep above-normal effort levels indefinitely, since workers would eventually adapt to the higher wages, getting back to “normal” effort. This implication is explored and discussed more in detail in the parameterized version of the model of Section 4.

\(^{12}\)This result supports the intuition put forward by other economists that workers’ effort is more responsive to unfair wage cuts due to psychological losses that could be explained by the concept of loss aversion (see for instance Campbell and Kamlani (1997), Bewley (2007) and Fehr et al. (2009)).
function assumed by Eliaz and Spiegler (2013).\footnote{Note that while Akerlof and Yellen (1990), Elsby (2009) and Eliaz and Spiegler (2013) assume away any positive reciprocity effect, our theory leaves the extent of positive reciprocity to be determined by the subjective weight that a worker puts on gain-loss utility, which is defined by the parameter $\eta$.}

For instance it is possible to draw a parallel with the reciprocal nature of gift giving proposed by Akerlof (1982). According to the wage-effort relationship established by Theorem 1 when $w > r$, a wage above the reference wage is reciprocated by the worker with higher than “normal” productive effort, $\tilde{e}(w, r)^+ > \tilde{e}_n$. As such in our model the gift exchange feature of the employment relationship is captured by the employer setting a wage above the worker’s standard of fairness and by the worker reciprocating the gift with higher than “normal” effort.

Note that the evaluation of the employer’s gift and the extent of the worker’s gift depend indirectly on the worker’s subjective evaluation of the gain-loss utility, captured by $\eta$: a greater $\eta$ means that the worker puts a greater weight on wage deviations from the reference wage and therefore responds more intensively to those deviations. Therefore, as $\eta$ may vary across individuals, this insight supports the experimental finding of individual heterogeneity for reciprocal preferences (see for instance Cohn et al. (2014)) and suggests a potential explanation: differences in workers’ weight on gain-loss utility $\eta$ may explain differences in their reciprocal preferences.

### 3.3 Firm’s Profit Function

A firm is assumed to be profit maximizing, where the per-worker profit $\pi(\cdot)$ is given by the difference between the per-worker value of output $y(q, e)$ and the respective per-worker cost of production $s(w)$:

$$\pi(w; q, e) = y(q, e) - s(w).$$

The per-worker cost of production is defined as an increasing function of the wage $w$ that the firm pays to the employed worker.

**C1.** $s(w)$ is continuously differentiable as many times as required, $s'(w) > 0$ and $s''(w) \geq 0$ for all $w$.

The value of output that an individual worker produces is captured by the production function $y(q, e)$, which is a combination of the worker’s level of productive effort $e$ and the level of productivity $q$ that characterizes the worker-firm match.

**C2.** $y(q, e)$ is continuously differentiable as many times as required in both its arguments, $y^e(q, e), y^q(q, e) > 0$, and $y^{ee}(q, e), y^{qq}(q, e) \leq 0$ and $y^{qe}(q, e) > 0$ for all $e$ and $q$.

We consider the match productivity $q$ as a comprehensive measure that embodies the level of technology available to the firm (or per-worker capital) together with the worker’s
intrinsic productivity that does not depend on monetary or other incentives. For any given level of productive effort, a higher match productivity increases the value of output \( y(\cdot) \) that a worker is able to produce. For the purpose of this analysis \( q \) is exogenously determined and it represents a link with the economic conditions that are external to the firm. Finally we assume \( q \in [\underline{q}, \overline{q}] \), where the lower bound of the interval is defined such that the profit will always be positive for any value of \( q \geq \underline{q} \). This assumption makes sure that the firm will always have the incentive to retain the worker and allows the analysis to focus on the implications of our theory for wage setting behavior only.

3.4 The Wage Setting Rule

Given the level of match productivity \( q \), the firm seeks to set a wage \( w \) to maximize profit given the worker’s optimal choice of effort \( \hat{e}(w, r) \) after comparing this wage to his reference wage (characterized in Theorem 1):

\[
\max_w \pi(w; q, e) = y(q, e) - s(w) \\
\text{s.t. } e = \hat{e}(w, r),
\]

Let \( \hat{w}(r, q, \lambda) \) be the profit maximizing wage such that the following first order condition is satisfied, conditional on \( w \neq r \):

\[
\Psi(q, w, r) \equiv \frac{\partial y(q, e)}{\partial e} \frac{\partial \hat{e}(w, r)}{\partial w} - \frac{ds(w)}{dw} = 0 \tag{4}
\]

**Theorem 2. Wage Setting Rule.**

The optimal wage \( \hat{w}(r, q, \lambda) \) is a continuous function of \( q \) and \( r \) and it is characterized by the following system:

\[
\hat{w}(r, q, \lambda) = \begin{cases} 
\hat{w}(r, q)^+ > r & \text{if } q > q_u(r) \\
r & \text{if } q \in [q_l(r), q_u(r, \lambda)] \\
\hat{w}(r, q, \lambda)^- < r & \text{if } q < q_l(r, \lambda),
\end{cases}
\]

\[\text{\textsuperscript{14}}\text{Consider for instance two workers who are paid the same wage and perform the same job in a firm. One worker could be more productive than the other due to personal/intrinsic higher skills associated with the use of a specific technology or capital.}\]

\[\text{\textsuperscript{15}}\text{The match productivity could also capture the heterogeneity between different worker-firm matches, sectors, industries or countries depending on what measure of productivity is to be considered.}\]

\[\text{\textsuperscript{16}}\text{The second order condition that ensures the first order condition is both necessary and sufficient is } \frac{\partial^2 y}{\partial e^2} \left( \frac{\partial e}{\partial w} \right)^2 + \frac{\partial y}{\partial e} \frac{\partial^2 e}{\partial w^2} - \frac{\partial^2 s}{\partial w^2} < 0, \text{ which is satisfied under assumptions C1-C2 and Theorem 1.}\]
where the match productivity thresholds $q^u(r)$ and $q^l(r,\lambda)$ are defined as:

\[
q^u(r) \equiv \left\{ q : \frac{\partial y(q,e) \partial \tilde{e}(w,r)}{\partial e \partial w} \bigg|_{w \searrow r} = \frac{ds(r)}{dw} \right\},
\]

\[
q^l(r,\lambda) \equiv \left\{ q : \frac{\partial y(q,e) \partial \tilde{e}(w,r)}{\partial e \partial w} \bigg|_{w \nearrow r} = \frac{ds(r)}{dw} \right\}.
\]

For all $q > q^u$ and $q < q^l$, the optimal wage $\tilde{w}(r,q,\lambda)$ is an increasing function of $q$ with $\frac{\partial \tilde{w}}{\partial q} > 0$, while for all $q \in [q^l, q^u]$ the optimal wage $\tilde{w}(r,q,\lambda)$ is a non-decreasing function of $q$ with $\frac{\partial \tilde{w}}{\partial q} = 0$. Finally, for all $q \in [q^l, q^u]$ the optimal wage $\tilde{w}(r,q,\lambda)$ is increasing in $r$ with $\frac{\partial \tilde{w}}{\partial r} > 0$.

Theorem 2 characterizes the firm’s optimal wage policy for any given match productivity $q$ and reference wage $r$. Depending on whether $q$ is above, within or below the range $[q^l, q^u]$, the firm sets the optimal wage above, equal or below the employed worker’s reference wage, where the upper $q^u(r)$ and lower $q^l(r,\lambda)$ thresholds themselves depend on the reference wage. We refer to this optimal wage policy as a wage setting rule, the key determinants of which are the worker’s reference wage $r$ and the loss aversion parameter $\lambda$. Figure 2 plots the relationship between the optimal wage and the match productivity as established by Theorem 2.

![Figure 2: The wage setting rule](image)

For the values of $q$ outside the range $[q^l, q^u]$ the optimal wage $\tilde{w}(r,q,\lambda)$ is smoothly adjusted to the level of the match productivity $q$. However, for the values of $q \in [q^l, q^u]$

---

17Where the meaning is clear, we henceforth abbreviate $q^u(r)$ to $q^u$ and $q^l(r,\lambda)$ to $q^l$. Since $\tilde{e}(w,r)$ is not differentiable at the point $w = r$, we characterize the thresholds at the point where the wage approaches the reference wage either from above $w \searrow r$, or from below $w \nearrow r$. 

---

17
the optimal wage is not adjusted to \( q \) and is set equal to the worker’s reference wage. We therefore define \( [q', q''] \) as the range of rigidity. Finally, note that the optimal wage \( \hat{w}(r, q, \lambda) \) is always positively adjusted to changes in the reference wage \( r \): the greater the worker’s reference wage the higher is the optimal wage set by the firm, for any level of the match productivity.

The following corollaries help to define some additional implications of the wage setting rule. Corollary 2 considers how the range of rigidity changes with a change in the reference wage.

**Corollary 2.** The thresholds \( q^u(r) \) and \( q^l(r, \lambda) \) are continuous and increasing in \( r \), with \( \frac{\partial q^u(r)}{\partial r} > 0 \) and \( \frac{\partial q^l(r, \lambda)}{\partial r} > 0 \) for all \( r \).

A key insight is that an increase in the worker’s reference wage shifts the whole range of rigidity to the right. A resulting implication is that the range of \( q \in [q, q'] \) within which the firm optimally sets \( \hat{w} < r \) increases, that is, there are values of \( q < q' \) for which a worker with a high reference wage receives a wage below his reference wage, but a worker with a lower reference wage would not.

Next we consider how the degree of loss aversion influences the range of rigidity.

**Corollary 3.** The lower threshold \( q^l(r, \lambda) \) is decreasing in \( \lambda \), with \( \frac{\partial q^l(r, \lambda)}{\partial \lambda} < 0 \), and \( q^l(r, \lambda)|_{\lambda=1} = q^u(r) \).

Corollary 3 states that if the firm is employing a worker with a greater degree of loss aversion, for a given reference wage, the range of rigidity is larger due to the resultant reduction in the lower bound defined by \( q^l(r, \lambda) \). On the other hand if the employed worker is not loss averse (that is, if \( \lambda = 1 \)) the range of rigidity disappears since \( q^l = q^u \).

In such a case, the optimal wage setting rule is shown in figure 2 by the black, thick dotted curve. Hence when \( \lambda = 1 \) the optimal wage responds smoothly to changes in the match productivity for any \( q \in [q, q'] \).

**Corollary 4.** For all \( q < q' \), \( \frac{\partial \hat{w}(r, q, \lambda)}{\partial \lambda} > 0 \) and therefore \( \hat{w}(r, q, \lambda)|_{\lambda>1} > \hat{w}(r, q, \lambda)|_{\lambda=1} \).

Finally, according to Corollary 4, whenever a firm negotiates a wage below the worker’s reference wage, that wage is going to be higher the greater the worker’s degree of loss aversion. The optimal wage (thick black line) is always greater than the wage that the firm would have set (thick black dotted line) if it were employing a worker who is not loss averse.\(^{18}\)

---

18 This implication has been theoretically derived and empirically corroborated by Holden and Wulfsohn (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. For simplicity consider a wage below the worker’s reference wage as a wage cut. Negative reciprocity not only tempers the firm’s incentive to cut the wage (when \( \hat{w} = r \)); it also reduces the extent to which the wage is cut, that being the optimal policy (when \( \hat{w} < r \)).
The wage setting rule just formalized highlights a discontinuous rate of adjustment of the optimal wage $\tilde{w}(r,q,\lambda)$ to changes in the match productivity $q$. Such discontinuity is implied by a range of values of $q$ within which the firm optimally freezes the wage, setting it equal to the worker’s reference wage. The optimal wage is therefore rigid for the values of $q \in [q^l, q^u]$. This result is a direct implication of the potential worker’s reciprocity response that the firm will face, if it sets a wage below the reference wage when $q$ falls within the range of rigidity.

To illustrate why this is the case, consider an ongoing employment relationship where the worker-firm match is characterized by a wage contract $\tilde{w} = r$ and a match productivity $q = q^u$. Then suppose that the firm has an incentive to cut the wage (therefore the wage offer would fall below the worker’s reference wage), for instance due to a fall in the match productivity (e.g. $q < q^u$). The marginal effect on the firm’s profit from such a reduction in the wage is given by

$$\left. \frac{\partial y}{\partial \check{e}} \frac{\partial \check{e}}{\partial w} \right|_{w<r} - \left. \frac{ds}{dw} \right|_{w<r} = 0.$$  \hspace{1cm} (5)

The firm faces a tradeoff between the marginal benefit of a lower wage ($MB$) and the marginal cost of a lower value of output ($MC$) generated by the worker’s negative reciprocity in the form of a marginal decrease in productive effort. The key intuition is that as long as $MC > MB$, the marginal effect on the value of output generated by negative reciprocity is greater than the marginal effect of a lower wage. Thus the wage cut is not optimal and the firm is better off freezing the wage. This follows because the worker’s potential negative reciprocity implies a cost to the wage cut, and this cost is strictly increasing in the worker’s degree of loss aversion (Corollary 1). Therefore, the more a worker is loss averse the more costly it is for the firm to enact the wage cut. To compensate for the potential decrease in effort and to make it optimal for the firm to cut the wage the fall in $q$ has to be sufficiently large such that $q < q^l$. The behavioral mechanism illustrated here explains why the range of rigidity becomes larger as $\lambda$ increases (Corollary 3): the greater $\lambda$, the stronger the negative reciprocity and the larger is the fall in $q$ required to re-balance the tradeoff expressed in (5).

Throughout this illustration we have elucidated the interaction between Theorem 1 (the worker’s reciprocity behavior) and Theorem 2 (the wage setting rule) which form the two cornerstones of our theory. This conceptual framework gives a justification for rigidity in the adjustment of the wage to productivity and provides its theoretical foundations based on loss aversion, negative reciprocity, workers’ morale and their perceptions of fairness.$^{19}$

$^{19}$Someone could argue that the same asymmetric adjustment and range of rigidity could be generated
The reminder of the paper illustrates further the implications of our theory by exploring several hypothetical situations concerning a firm wage setting and hiring decisions and identifies some novel insights that arise from forward looking-behavior under uncertainty.

4 Illustration: the employment contract, renegotiation and downward wage rigidity

This section uses a parameterized version of the model to illustrate the implications of our theory for wage setting and hiring behavior. We use this framework to characterize the initial employment contract and to explore the conditions under which renegotiation takes place. To keep the focus on the behavioral forces underlying wage setting behavior, we abstract from time-contingent and legislative constraints that could potentially influence renegotiation. Moreover we do not consider the firm’s possibility of layoff and abstract from concerns about inflation (so only nominal wages matter). As such, our focus is on adjustments of the wage to changes in the match productivity. The parties of interest are a firm looking for a worker and an unemployed worker looking for a job. The firm is assumed to make a take-it-or-leave-it wage offer \( w_0 \) and the unemployed worker is assumed to accept only if he perceives the offer to be at least fair, that is, the worker accepts any \( w_0 \geq r_0 \).

**Time-line.** We consider a single employment period that starts with an initial wage contract at time 0 and ends with a potential renegotiation at time 1 (see Figure 3). These two dates are symbolic and do not necessarily delineate a fixed period of time. At time 0 the firm and the worker meet, match and contract; at time 1 the firm has the opportunity to rearrange the contract terms. The frequency of this opportunity may vary across countries, industries, sectors or firms depending in part on the relevant legislation in place.

**Uncertainty.** When the firm meets the worker it observes the value of productivity characterizing the match at time 0, namely \( q_0 \), which is drawn from a distribution \( h(Q) \). However it cannot predict how productivity will evolve during the employment period from time 0 to 1. We capture this uncertainty by assuming that \( q_1 \) is a non-negative random variable, drawn from a distribution \( f(Q) \). \( H(Q) \) and \( F(Q) \) are the corresponding cumulative distributions defined by \( \text{Prob}(q_t \leq Q) = H(Q) = F(Q) \) with support \([q, \bar{q}]\),

by the ad hoc assumption that wage cuts are more costly to the firm, based on the evidence previously discussed. However, such a model would not be able to explain and predict when such discontinuous rate of adjustment disappears and why (or when it is stronger and why). On the other hand our theory provides a realistic framework for analysing wage setting behavior in which the elements of asymmetric reciprocity, loss aversion \( \lambda \) and the worker’s perceptions of fairness \( r \) (incorporated in a morale function) are the main determinants of costly wage cuts and asymmetric adjustments. Therefore analyzing the extent and importance of these factors would not only allow us to draw predictions and hypotheses on whether wages will be rigid or not, but would also elucidate the behavioral reasons behind these predictions.
where $\bar{q} < \infty$, so that $H(q) = F(q) = 0$ and $H(\bar{q}) = F(\bar{q}) = 1$. The purpose of this assumption is to represent an hypothetical scenario in which changes in external economic conditions may affect the firm’s productivity (e.g. changes in technology or capital investments). Another interpretation is that the firm may have formed incorrect prior beliefs at the time of the initial contract, realizing the true value of the match productivity $q_1$ only during the employment period.

The purpose of this assumption is to represent an hypothetical scenario in which changes in external economic conditions may affect the firm’s productivity (e.g. changes in technology or capital investments). Another interpretation is that the firm may have formed incorrect prior beliefs at the time of the initial contract, realizing the true value of the match productivity $q_1$ only during the employment period.

20 These hypothetical changes in the match productivity are representative of two type of shocks that can potentially occur in the economy. The change in $q$ can represent the effect of a standard productivity shock (random, AR1 process or Poisson process), as it is commonly assumed by the macroeconomic literature that draws on the neoclassical tradition or by search and matching models of the labor market. The change in $q$ could also represent a structural change in the supply side of the economy due to a persistent shock of aggregate demand that affects a firm’s match productivity. In general we consider changes in $q$ as an event that alters the economic conditions and that gives an incentive to the firm to adjust labor costs.

Figure 3: Employment contract time-line

<table>
<thead>
<tr>
<th>Uncertainty 0</th>
<th>Employment</th>
<th>Resolution of Uncertainty 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$q_0$ is drawn</td>
<td>$\bar{w}_0$</td>
<td>$q_1$ is drawn</td>
</tr>
<tr>
<td>$r_0$ is given</td>
<td>$r_1 = \bar{w}_0$</td>
<td>$\bar{w}_1$</td>
</tr>
</tbody>
</table>

Wage contract as the reference “fair” wage. When the unemployed worker negotiates with the firm at time 0 he evaluates the fairness of the firm’s wage offer with respect to a given and unspecified reference wage $r_0$. As such $r_0$ is the unemployed worker’s reservation wage. However once employed the worker’s perceptions of fairness may change: during the employment period the worker gets used to the wage he receives and adopts it as the standard of fairness against which he evaluates future wage renegotiations. Thus, at time 1 the worker’s reference wage is assumed to be endogenously determined by the initial wage contract signed at time 0, that is, $r_1 = \bar{w}_0$.

Decision making process. At each negotiation date, the firm and the worker engage in a two stage game as described in Section 3. At time 0 the firm offers the wage contract $\bar{w}_0$ that maximizes the sum of its expected discounted profits, based on the observed $q_0$ and the expectations of what the match productivity might be during the employment period $q_1 \equiv \int_q q_1 dF(q_1)$. The firm is therefore assumed to be forward looking. This assumption will prove to be particularly relevant given the crucial role of the initial wage contract in the determination of the worker’s reference wage $r_1$. On the other hand the worker

21 For instance $r_0$ could be given by the national minimum wage collectively bargained in the labor market of interest. It could also be influenced by indicators of labor market conditions, such as the unemployment rate or by some other measure of the degree of labor market tightness. For the purpose of this illustration we leave the reference wage of an unemployed worker as exogenously determined and instead focus on the firm’s behavior for any given $r_0$. 

22
myopically chooses the optimal level of productive effort without considering the effects of this choice in the future. At each negotiation date, the worker’s response to any wage offer is entirely captured by the optimal productive effort function derived in Theorem 1.

Given these assumptions, we formalize the initial employment contract as a two-period dynamic stochastic maximization problem characterized as follows:

\[
J_0(r_0, q_0) = \max_{w_0} \pi_0(w_0; q_0, e_0) + \delta E_0[J_1(r_1, q_1)]
\]

s.t.
\[
e_t = \tilde{e}_t(w_t, r_t) \\
w_0 \geq r_0 \\
r_1 = w_0,
\]

where \(\delta\) is the firm’s discount factor. Note that the problem (6) is characterized by two state variables: the level of match productivity \(q_t\) and the worker’s reference wage \(r_t\). However, due to the assumption that the worker’s reference wage at time 1 is going to be determined by the firm’s optimal wage contract at time 0, a forward-looking firm is able to directly influence one of the state variables through its optimal wage policy. Since this feature is going to drive part of the key results in the following two sub-sections, it is sensible to discuss its behavioral nature.

The assumption of the wage contract determining the worker’s reference wage is consistent with several ideas in the literature. 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 proposed by Hart and Moore (2008) and further explored by Herweg and Schmidt (2015) to study the effects of \(ex \ ante\) contracts and loss aversion when the contracting parties engage in renegotiation. The laboratory experiments of Fehr et al. (2011, 2014) and Bartling and Schmidt (2014) provide strong support to 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 a direct evidence of this hypothesis and Stommel (2013) for a review of the literature).

The analysis that follows is divided in two parts: first we consider the case of a myopic firm \((\delta = 0)\) and study under which conditions wage renegotiation may or may not occur. Then we consider a forward-looking firm \((\delta > 0)\) and using the intuition drawn in the previous analysis we characterize the optimal employment contract that solves the dynamic problem in (6).
4.1 The case of a myopic firm

In this section we analyze the conditions under which an initial employment contract is signed and subsequently renegotiated when the firm is myopic ($\delta = 0$). Such digression helps to gain intuition and motivates the importance of analyzing a forward-looking firm.

So that we can derive explicit solutions, consider an unemployed worker with wage utility:

$$u_0(w_0|r_0) = w_0^\alpha + \mu(x_0)$$

and effort utility:

$$v_0(e_0; w_0, r_0) = e_0 - e_0^2/2 + e_0\mu(x_0),$$

where $\mu(x_0) = (w_0^\alpha - r_0^\alpha)$ when $w_0 \geq r_0$ and $\mu(x_0) = -\lambda(r_0^\alpha - w_0^\alpha)$ when $w_0 < r_0$. When contracting with the firm at time 0, the worker observes a wage offer $w_0$ and chooses the optimal level of productive effort $\tilde{e}_0(w_0, r_0)$ that maximizes overall utility $U_0(e_0; w_0, r_0) = u_0(w_0|r_0) + v_0(e_0; w_0, r_0)$ given his reservation wage $r_0$. The resulting optimal effort function takes the following form:

$$\tilde{e}_0(w_0, r_0) = 1 + \mu(x_0) = \begin{cases} 
1 + w_0^\alpha - r_0^\alpha & \text{if } w_0 \geq r_0 \\
1 - \lambda(r_0^\alpha - w_0^\alpha) & \text{if } w_0 < r_0,
\end{cases}$$

where “normal” effort is $\tilde{e}_{n,0} = 1$ when $w_0 = r_0$. Taking this choice into account, the firm maximizes per-worker profit

$$\pi_0(w_0; q_0, e_0) = q_0 e_0 - w_0,$$

given the worker’s reservation wage $r_0$ and the observed match productivity $q_0$.

**Proposition 1.** If the firm is myopic $\delta = 0$, at time 0 the initial wage contract $\tilde{w}_0(r_0, q_0)$ is characterized as follows:

$$\tilde{w}_0(r_0, q_0) = \begin{cases} 
(\alpha q_0)^{1/\alpha} & \text{if } q_0 > q_0^u \\
r_0 & \text{if } q_0 = q_0^u \\
\text{no contract} & \text{otherwise},
\end{cases}$$

where the match productivity threshold $q_0^u(r_0)$ is defined as

$$q_0^u(r_0) \equiv \{ q_0 : \alpha q_0 r_0^{\alpha - 1} = 1 \} = \frac{r_0^{1-\alpha}}{\alpha},$$

and represents the myopic firm’s reservation productivity, below which no contract is offered.

Following the time-line set out in Figure 3 we illustrate the case of a hypothetical initial employment contract and a subsequent renegotiation.
Initial Contract $t=0$. Depending on the observed value of $q_0$, the firm either hires the worker and offers a wage contract $\tilde{w}_0 \geq r_0$ (if $q_0 \geq q_u^0$) or leaves and no contract is offered (if $q_0 < q_u^0$). The firm’s reservation productivity $q_u^0(r_0)$ is endogenously determined and it is an increasing function of the worker’s reservation wage. Thus we can infer that a higher $r_0$ reduces the range $[q_u^0, \overline{q}]$ within which hiring takes place.

**Corollary 5.** The higher the worker’s reservation wage $r_0$, the smaller is the range of the match productivity $q_0 \in [q_u^0(r_0), \overline{q}]$, within which the firm makes a wage offer to hire the worker.

A firm that is about to pay a worker a very high initial wage would intuitively require a high level of productivity from the match; therefore it is a sensible implication, under our assumption of complete information, that a worker with a greater reservation wage would find it harder to receive a “fair” wage offer, in a given range $[q, \overline{q}]$ of possible match productivity.

Now suppose that a worker-firm match occurs at time 0 and the observed match productivity is greater than the firm’s reservation productivity: $q_0 > q_u^0$. As shown in Figure 4 the optimal wage contract $\tilde{w}_0 > r_0$ is offered, the worker accepts and positively reciprocates with optimal productive effort $\tilde{e}_0 > 1$.

![Figure 4: Initial contract $t = 0$](image)

*Employment Period* $0 \rightarrow 1$. According to our set up, the employment period is characterized by two major events. The first relates to the worker’s reference wage: by entering the status of employed, the worker adjusts his feelings of entitlement and adopts the initial wage contract as the new standard of fairness, $r_1 = \tilde{w}_0$. Moreover, as a consequence of this adaptation, the worker’s level of productive effort goes back to “normal” for the time being. Thus the model identifies this form of reference wage
adaptation mechanism as an explanation for why positive reciprocity is only temporary.²²

The second is about a change in the match productivity, that is, \( q_1 \) is drawn from \( f(Q) \). For the purpose of this illustration, we consider a value of \( q_1 < q_0 \) lower than the match productivity that characterized the initial employment contract. As such we capture an hypothetical situation in which a change in the economic conditions, that are external to the firm, adversely impact productivity. Note that this event is observed either directly (through a change in the value of output), or indirectly (through a change in profit) and gives an incentive to the firm to renegotiate the wage at time 1.

Renegotiation \( t = 1 \). The firm would intuitively want to renegotiate the wage contract by setting a lower wage \( w_1 < \tilde{w}_0 \) since \( q_1 < q_0 \). However such downward wage adjustment may carry a cost: since the worker has adapted his reference wage to the optimal wage contract signed at time 0, any downward wage renegotiation would be perceived as unfair, triggering negative reciprocity in the form of subnormal effort. In fact, recalling Theorem 2, if \( q_1 \) is not sufficiently low, that is, if \( q_1 \in [q^l_1, q^u_1] \), the firm is better off not renegotiating the contract and freezing the wage. Moreover, since \( r_1 = \tilde{w}_0 \), the upper \( q^u_1(\tilde{w}_0) \) and lower \( q^l_1(\tilde{w}_0, \lambda) \) thresholds are now a function of the initial wage contract, and as the initial wage contract is in turn a function of \( q_0 \), we derive the following result:

**Proposition 2.** At time 1, if \( r_1 = \tilde{w}_0 \), the outcome of an optimal wage renegotiation \( \tilde{w}_1(r_1, q_1, \lambda) \) is characterized by the following wage setting rule:

\[
\tilde{w}_1(r_1, q_1, \lambda) = \begin{cases} 
(\alpha q_1)^{\frac{1}{1-\alpha}} > \tilde{w}_0 & \text{if } q_1 > q^u_1 \\
\tilde{w}_0 & \text{if } q_1 \in [q^l_1; q^u_1] \\
(\lambda \alpha q_1)^{\frac{1}{1-\alpha}} < \tilde{w}_0 & \text{if } q_1 < q^l_1,
\end{cases}
\]

where

\[
q^u_1(\tilde{w}_0) \equiv \frac{\tilde{w}_0^{1-\alpha}}{\alpha} = q_0 \\
q^l_1(\tilde{w}_0, \lambda) \equiv \frac{\tilde{w}_0^{1-\alpha}}{\lambda \alpha} = \frac{q_0}{\lambda}
\]

are the upper and lower thresholds of the range of rigidity at time 1.

Hence, as \( q_0 \) now defines the upper productivity threshold \( q^u_1 \), even though \( q_1 < q_0 \), the firm will wish to renegotiate and adjust the wage downwards (enact a wage cut)

---

²²This implication of the model is consistent with the evidence reported by field research (experiments and surveys) as opposed to laboratory experiments, 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 support the interpretation of this evidence according to which positive reciprocity quickly disappear 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)).
only if the observed value of the match productivity is substantially lower than the one characterizing the contract at time 0, that is, if and only if $q_1 < \frac{q_0}{\lambda}$.

If the fall in productivity is not sufficiently large, $\frac{q_0}{\lambda} \leq q_1 \leq q_0$, the firm’s optimal wage policy is to freeze the wage: $\tilde{w}_1 = r_1 = \tilde{w}_0$ (see the solid dot in Figure 5). Note that within this range the firm is incurring some “losses at the margin” due to a marginal decrease in the value of output per worker $y_1(q_1, e_1)$. However the firm’s incentive to compensate this marginal decrease with a downward adjustment of the wage is tempered, since to do so would yield an even greater decrease in the value of output due to the induced negative reciprocity. Finally, confirming the prediction of the general model, a greater extent of the worker’s loss aversion $\lambda$ enlarges the range of rigidity to the left. A greater $\lambda$ implies a stronger negative reciprocity response if the firm cuts the wage, therefore lowering the match productivity level at which it is optimal to do so, namely $q^*_1 = \frac{q_0}{\lambda}$.

This analysis shows that once an employment contract is signed, the firm may effectively be constrained in its ability to renegotiate the wage contract. Along with the potential cost of negative reciprocity represented by $\lambda$, this result is also driven by the assumed adaptation of the worker’s reference wage.

To see this, consider the counterfactual case of no adaptation.

**Corollary 6.** At time 1, if $r_1 = r_0$:

$$\tilde{w}_1(r_0, q_1) < \tilde{w}_1(r_0, q_0) \quad \forall \ q_1 < q_0 \in [\underline{q}, \overline{q}].$$

If the worker still considers his reservation wage to be the reference “fair” wage $r_1 = r_0$, a wage cut is not perceived as unfair. The firm can adjust the wage downwards without incurring the marginal cost of negative reciprocity (the unfilled dot in Figure 5 is now achievable). In fact Corollary 6 states that with no adaptation, the renegotiated wage is
always going to be lower than the initial wage contract if the match productivity decreases below $q_0$.\footnote{There is an additional implication concerning the outcome of renegotiation that becomes clearer after analyzing the counterfactual of no adaptation. If we consider the case of an increase in the match productivity (that is $\forall q_1 > q_0 \in [\underline{q}, \bar{q}]$) the renegotiated wage will always be greater than the initial wage contract. From this implication we can infer that if a worker always adopts the initial wage contract as the reference wage and the match productivity increases, in the following renegotiation the firm would not have to raise the wage as much as if there were no adaptation. This is a form of wage compression that is implicitly generated by the worker’s adaptation of the reference wage.}

Throughout this illustration our theory has shown that the initial wage contract may be rigid and irreversible due to the negative reciprocity effects that would be induced if a wage renegotiation is perceived to be unfair. The example has also analytically formalized the idea of Kahneman et al. (1986), that apart from legislative and budgetary constraints (for instance time-contingent contracts or wage floors), fairness is an additional constraint to the adjustment of market outcomes: the worker’s concerns about being paid the “fair” wage and the firm’s concerns about the worker’s morale and negative reciprocity limit the flexibility of the employment contract and generate a discontinuous rate of adjustment of the wage to productivity, namely downward wage rigidity when economic conditions worsen.

Under our assumptions, the greater is the initial wage contract $\tilde{w}_0$ the higher will be the worker’s reference wage $r_1$ at the time of renegotiation. In fact, the range $[\underline{q}, q^u_1(r_1)]$ within which the firm freezes the wage or enacts a wage cut, increases with the reference wage $r_1$ and therefore with the initial wage contract $\tilde{w}_0$. This implication generates a non-trivial prediction: if a firm could anticipate the costs associated with a higher initial wage contract, it would eventually negotiate a lower wage at time 0. As the value of $q_1$ is uncertain, by setting a lower initial wage the firm could avoid, or at least reduce the probability, of having to enact a costly wage freeze/cut in the event of a lower match productivity in the future.\footnote{This intuition and its implications have been analyzed in an infinite-horizon dynamic model by Elsby (2009), motivated by a puzzle of downward nominal wage rigidity revealed in in the empirical literature (micro-macro puzzle). Our theory, as it is shown later, encompasses Elsby’s model, confirms its main theoretical result and provides sound theoretical foundations to his assumptions over the worker’s effort function.} This follows because, for a given support $[\underline{q}, \bar{q}]$, a lower $\tilde{w}_0$ decreases the range $[\underline{q}, q^u_1(r_1)]$ within which these wage policies are optimal. A way to interpret this intuition is that a forward-looking firm that takes into account the potential future costs of a wage freeze/cut has an incentive to manipulate the employed worker’s reference wage: by offering a lower initial wage contract the firm’s managers keep the worker feelings of entitlement and perceptions of fairness relatively low.

Motivated by this prediction the following section characterizes an employment contract when a firm is forward looking.
4.2 Forward-looking behavior and uncertainty

When a match occurs at time 0, the firm offers the employment contract that solves the following two-period dynamic stochastic optimization problem:

\[
J_0(r_0, q_0) = \max_{w_0} \quad q_0 e_0 - w_0 + \delta \int_{q_1}^{q_l} J_1(r_1, q_1) \, dF(q_1)
\]

s.t. \( e_t = \bar{e}_t(w_t, r_t) \)

\( w_0 \geq r_0 \)

\( r_1 = w_0 \)

where \( J_1(r_1, q_1) = \max_{w_1} \quad q_1 e_1 - w_1. \)  

(7)

Hence, if a firm is forward looking, it also takes into account the expected future maximized profits for any possible realization of \( q_1 \), represented by \( \int_{q_1}^{q_l} J_1(r_1, q_1) \, dF(q_1) \). From Proposition 2 we know that depending on the value of the match productivity drawn between time 0 and time 1, the firm would either cut, freeze or raise the wage at the first opportunity of renegotiation. Moreover note that since \( r_1 = w_0 \), we can write the future maximized profit as function of \( w_0 \): \( J_1(w_0, q_1) \). As such it is convenient to express the future expected maximized profit in the following form:\(^25\)

\[
\int_{q_1}^{q_l} J_1(w_0, q_1) \, dF(q_1) = \int_{q_1}^{q_l} J_1(w_0, q_1)^- \, dF + \int_{q_1}^{q_l} J_1(w_0, q_1)^0 \, dF + \int_{q_1}^{q_l} J_1(w_0, q_1)^+ \, dF.
\]

(8)

This expression defines each possible future profit realization that the firm is facing when setting the initial wage contract, given the probability of drawing a match productivity \( q_1 \) that is below, within or above the range of rigidity derived in Proposition 2. However, since the thresholds of this range, \( q_1^-(w_0, \lambda) \) and \( q_1^+(w_0) \), are functions of \( w_0 \), by setting the initial wage contract the firm can also influence the range of the distribution, that is, the probability over which the wage \( w_1 \) will be cut, frozen or raised.\(^26\)

Let the initial wage contract \( \tilde{w}_0 \) be the solution to the optimization problem in (7) that satisfies the following first order condition for \( w \neq r \):

\[
q_0 \alpha w_0^{\alpha - 1} - 1 + \delta \theta(w_0, \lambda) = 0,
\]

(9)

\(^{25}\)Where +/-0/- indicate the event of a wage raise/freeze/cut.

\(^{26}\)Note that drawing from the implications of the general model set out in Section 3, a greater extent of the worker’s loss aversion increases the probability of having to enact a costly wage freeze in the future (see Corollary 3).
where the function
\[ \theta(w_0, \lambda) \equiv \int_{\frac{q}{2}}^{\bar{q}} J'_1(w_0, q_1) dF(q_1) \]
captures the marginal effect of the initial wage contract \( w_0 \) on the firm’s expected future maximized profit.\textsuperscript{27}

**Proposition 3.** The function \( \theta(w_0, \lambda) < 0 \) captures the marginal cost of the initial wage contract \( w_0 \) on the firm’s expected future maximized profit and it is characterized by the following expression:
\[
\theta(w_0, \lambda) = \int_{\frac{q}{2}}^{q_1^l (w_0, \lambda)} J'_1(w_0, q_1)^- dF + \int_{q_1^l (w_0, \lambda)}^{q_1^u (w_0)} J'_1(w_0, q_1)^0 dF + \int_{q_1^u (w_0)}^{\bar{q}} J'_1(w_0, q_1)^+ dF,
\]
where \( J'_1(w_0, q_1) < 0 \) for all \( q_1 \in [\frac{q}{2}, \bar{q}] \).

From Proposition 3 we infer that since \( \theta(w_0, \lambda) < 0 \) the marginal effect of \( w_0 \) on the expected future profit is negative where the term
\[
\int_{\frac{q}{2}}^{q_1^l} J'_1(w_0, q_1)^- dF
\]
represents the expected future marginal cost of \( w_0 \) in the event of a future wage cut \( w_1 < w_0 \), given a realization of \( q_1 \in [\frac{q}{2}, q_1^l (w_0, \lambda)] \); the term
\[
\int_{q_1^l}^{q_1^u} J'_1(w_0, q_1)^0 dF
\]
represents the expected future marginal cost of \( w_0 \) in the event of a future wage freeze \( w_1 = w_0 \), given a realization of \( q_1 \in [q_1^l (w_0, \lambda), q_1^u (w_0)] \); and the term
\[
\int_{q_1^u}^{\bar{q}} J'_1(w_0, q_1)^+ dF
\]
represents the expected future marginal cost of \( w_0 \) in the event of a future wage raise \( w_1 > w_0 \), given a realization of \( q_1 \in (q_1^u (w_0), \bar{q}] \). Therefore, the expected future maximized profit \( \int_{\frac{q}{2}}^{\bar{q}} J_1(w_0, q_1) dF(q_1) \) is a decreasing function of the initial wage contract \( w_0 \) for any possible realization of \( q_1 \in [\frac{q}{2}, \bar{q}] \). Proposition 3 reinforces the argument behind our prediction: a lower initial wage would decrease the probability of the firm having to enact a wage cut or a wage freeze in the future, by reducing the range of the distribution \([\frac{q}{2}, q_1^u (w_0)]\) within which these wage policies are optimal.

\textsuperscript{27}The prime ‘\( \prime \)’ is used to denote the first derivative of the value function \( J_1 \) with respect to the choice variable \( w_0 \).
Thus, we now turn into the characterization of the optimal employment contract at \( t = 0 \).

**Proposition 4.** If the firm is forward looking \( \delta > 0 \), at time 0 the initial wage contract \( \tilde{w}_0(r_0, q_0, q^e_1, \lambda) \) is characterized as follows:

\[
\tilde{w}_0(r_0, q_0, q^e_1, \lambda) = \begin{cases} 
\tilde{w}_0(r_0, q_0, q^e_1, \lambda) & \text{if } q_0 > q^u_0 \\
q_0 & \text{if } q_0 = q^u_0 \\
\text{no contract} & \text{otherwise},
\end{cases}
\]

where the match productivity threshold \( q^u_0(r_0, q^e_1, \lambda) \) is defined as

\[ q^u_0(r_0, q^e_1, \lambda) \equiv \{ q_0 : \alpha q_0 r_0^{\alpha-1} + \delta \theta(r_0, \lambda) = 1 \}, \]

and represents the forward-looking firm’s reservation productivity, below which no contract is offered.

Although we do not provide a closed form (analytical) solution for the result of Proposition 4, it is possible to draw out some implications based on the properties of our general model set out in Section 3.

First note that if the firm is forward looking, both the initial wage contract \( \tilde{w}_0 \) the reservation productivity \( q^u_0 \) are functions of the expectations about \( q_1 \) and the worker’s loss aversion parameter \( \lambda \). By comparing the results obtained for the case of a myopic firm and a forward-looking firm, we discuss two key implications that our theory generates.

**Corollary 7.** For a given \( r_0 \) and \( q_0 \), if a firm is forward looking \( \delta > 0 \), the initial wage offer \( \tilde{w}_0 \) is always lower than if the firm is myopic \( \delta = 0 \), that is:

\[ \tilde{w}_0(r_0, q_0, q^e_1, \lambda) < \tilde{w}_0(r_0, q_0) \quad \forall q_1 \in [q, q^l_0]. \]

This result confirms our prediction: since the initial wage contract at time 0 determines the employed worker’s reference wage at time 1, the firm has an incentive to offer a lower initial wage. Figure 6 shows the *ex ante* probabilities faced by a myopic and a forward-looking firm of drawing \( q_1 \) below, within or above the range of rigidity for a given distribution function of productivity. Note that while such *ex ante* probabilities are known if the firm is forward looking, they are rather unknown if the firm is myopic. Thus if we consider the properties of \( \theta(\cdot) \) and the characterization of the upper and lower thresholds \( q^u_1 \) and \( q^l_1 \) we can infer an additional insight: by offering a lower initial wage contract \( \tilde{w}_0 \) a forward-looking firm (Figure 6b) reduces the probability of being in the situation to enact a wage cut, \( q_1 \in [q, q^l_1] \) (shaded area), and increases the probability of
implementing a wage rise, \( q_1 \in [q_1^u, \bar{q}] \).\(^{28}\)

(a) Myopic firm

\[
f(q_1)
\]

(b) Forward-looking firm

\[
f(q_1)
\]

Figure 6: Ex ante probability of the renegotiation outcome

The implications of Corollary 7 are the result of our theory of wage setting behavior combined with forward-looking behavior and uncertainty. As such they can be considered a general result, which would hold for subsequent time periods and for a longer time horizon. Elsby (2009)’s seminal paper uses an infinite-horizon dynamic model to derive “compression of wage increases” 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 confirms the firm’s incentive to compress wage increases (in our case, entry wages) in the presence of downward wage rigidity. It also extends this implication to the initial employment contract and provides a model that elucidates the behavioral forces at play.\(^{29}\)

Corollary 8. For a given \( r_0 \) and \( q_0 \), if a firm is forward looking \( \delta > 0 \), the reservation productivity \( q_0^u \) is always greater than if the firm is myopic \( \delta = 0 \), that is:

\[
q_0^u(r_0, q_1^c, \lambda) > q_0^u(r_0) \quad \forall q_t \in [q, \bar{q}].
\]

This insight implies that uncertainty over the future unfolding of the match productivity \( q_1 \) affects the firm’s hiring behavior as well. If a firm is forward looking it chooses to offer a wage contract only on a smaller segment of the range \([q, \bar{q}]\) than if it was myopic. This results because the firm would require a higher match productivity at time

\(^{28}\)Note that this implication holds for any given distribution \( f(Q) \) with support \([q, \bar{q}]\).

\(^{29}\)This result also provides an explanation for firms preferring to implement policies that favour increasing wages over time instead of setting an initial high wage rate followed by a subsequent decrease. According to our theory this follows because firms understand that increasing wage profiles are a good device to keep workers’ morale and effort at “normal” levels while exploiting temporary positive reciprocity, and importantly not inducing negative reciprocity. This discussion draws a parallel with the seminal work of Loewenstein and Sicherman (1991), on workers’ preferences for increasing wage profiles.
0 to hire a worker in order to compensate for a potential future decrease in the match productivity around which it may not be able to adjust the wage. For a given $h(Q)$, the ex ante probability of hiring a worker is lower if the firm takes into account the effect of the initial employment contract in its future expected maximized profits. This result is shown in Figure 7 below.

(a) Myopic firm

(b) Forward-looking firm

Figure 7: Ex ante probability of hiring

The key implications established in Corollaries 7 and 8 derive directly from the fact that when the firm offers the wage contract at time 0, such decision is based on its expectations of the match productivity at time 1, namely $q_1$, as well as the worker’s degree of loss aversion $\lambda$. Regardless of the firm’s prior belief about the future realization of $q_1$ at time 0 (that is, regardless of the shape of $F(Q)$), and regardless of the worker’s subjective degree of loss aversion $\lambda$, if a firm is forward looking it will always have an incentive to set a lower initial wage contract and to require a higher match productivity to hire a worker for a given reservation wage $r_0$.

In what follows we analyze how the initial employment contract is influenced by changes in these two parameters.

**Corollary 9.** The function $\theta(w_0, \lambda)$ is decreasing in the loss aversion parameter $\lambda$, that is, $\frac{\partial \theta}{\partial \lambda} < 0$.

The intuition behind Corollary 9 is better understood if we recall the positive correlation between loss aversion, negative reciprocity and the marginal cost that a firm faces in the tradeoff between implementing a wage cut or enacting a wage freeze. A greater extent of loss aversion implies that a wage cut will be more costly into the future, due

---

30 As such this illustration also highlights that a worker’s initial wage not only depends on the current market conditions (i.e. $q_0$) but also on the expected path of the unfolding of the match productivity $q_1$ at the time of hiring.
to a stronger negative reciprocity. Moreover it also increases the probability of being in a situation to enact a wage freeze if the match productivity falls. Building on these insights we analyze the implications of loss aversion for the initial employment contract. Consider an optimal initial wage contract of a forward-looking firm that is facing either an unemployed worker \( i \) characterized by a loss aversion parameter \( \lambda_i \) or an unemployed worker \( j \) characterized by a loss aversion parameter \( \lambda_j \), where \( \lambda_i > \lambda_j \).

**Proposition 5.** For any given \( r_{0,j} = r_{0,i} \) and \( q_{0,j} = q_{0,i} \), the initial wage contract offered to worker \( i \) is always lower than the initial wage contract offered to worker \( j \):

\[
\tilde{w}_{0,i}(r_0, q_0, q_e^i, \lambda_i) < \tilde{w}_{0,j}(r_0, q_0, q_e^j, \lambda_j) \quad \forall q_t \in [q_1, q_e].
\]

Moreover, the reservation productivity if the firm is matched with worker \( i \) is always greater than the reservation productivity if the firm is matched with worker \( j \):

\[
q_{0,i}(r_0, q_e^i, \lambda_i) > q_{0,j}(r_0, q_e^j, \lambda_j) \quad \forall q_t \in [q_1, q_e].
\]

Thus, according to Proposition 5, the implications established by Corollaries 7 and 8 are even more enhanced. A firm that is facing a more loss averse worker has a greater incentive to set a lower initial wage contract and requires an even higher level of the initial match productivity in order to hire the worker.

For what concerns \( q_e^1 \), we pursue a theoretical exercise that gives to the firm’s expectations a central role in driving the outcome of an initial employment contract. We do so by treating the firm’s expectations about \( q_1 \) as a structural parameter and we compare two hypothetical scenarios in which a firm has two very different beliefs about how the future unfolds.\(^{31}\) Suppose that depending on the firm’s prior at time 0, \( q_1 \) can possibly be drawn from either \( F_a(Q) \) or \( F_b(Q) \). These two cumulative distribution functions are assumed to have the same support \([q_1, q_e]\) and are characterized by the following properties:

**D1.** \( F_a \) first order stochastically dominates \( F_b \), that is \( F_a(q_1) \leq F_b(q_1) \) for all \( q_1 \in [q_1, q_e] \). Moreover, for any given \( q_1^l \) and \( q_1^u \), \( f_a(q_1^l) > f_b(q_1^l) \), \( f_a(q_1^u) < f_b(q_1^u) \) and \( f_a(q_1^u) - f_a(q_1^l) = f_b(q_1^u) - f_b(q_1^l) \).

The first part of assumption D1 means that the probability of drawing a higher level of \( q_1 \) within the whole support is greater if the firm has its expectations represented by \( f_a(Q) \) rather than by \( f_b(Q) \). As such we define a firm “optimistic” if has expectations shaped by \( f_a(Q) \) and “pessimistic” if it has expectations shaped by \( f_b(Q) \). The second part of

---

\(^{31}\)This analysis is inspired by the “autonomous expectations” approach discussed in Frydman and Phelps (2013, p.28), which “treats participants’ expectations as a structural parameter […]and recognizes that the future is inherently open, and that how it unfolds depends crucially on participants’ revisions of their expectations”. 

33
D1 is useful to make the intuition of the following analysis clearer. If a firm is optimistic it has a greater *ex ante* probability of drawing a match productivity $q_1 > q_1^u$ and a lower *ex ante* probability of drawing $q_1 < q_1^l$ than if the firm is pessimistic. Furthermore note that the *ex ante* probability of drawing $q_1 \in [q_1^l, q_1^u]$ within the range of rigidity is the same. Thus an optimistic firm puts less weight to the probability of being in a situation to cut the wage at time 1; define these expectations as

$$q_1^a \equiv \int_{q}^{q_1} q_1 \, dF_a(q_1).$$

On the other hand a pessimistic firm puts a greater weight to the probability of being in a situation to enact a costly wage cut; define these expectations as

$$q_1^b \equiv \int_{q}^{q_1} q_1 \, dF_b(q_1).$$

Figure 8 plots the density distributions of $F_a(Q)$ and $F_b(Q)$, where the shaded area below $f_a(q_1)$ corresponds to $q_1^a$ and the shaded area below $f_b(q_1)$ corresponds to $q_1^b$.

![Figure 8: "Optimistic" vs. "pessimistic" firm](image)

**Proposition 6.** For a given $r_0$, $q_0$ and $\lambda$ the initial wage contract offered by a "pessimistic" firm $\tilde{w}_0^b$ is always lower than the initial wage contract offered by an "optimistic" firm $\tilde{w}_0^a$.

---

32Note that the figure aims to capture the different firm’s expectations before profit maximization takes place, that is, before the wage contract is offered. In fact depending on the initial wage contract $q_1^a$ and $q_1^b$ are going to change.
firm $\tilde{w}_0^a$:

$$\tilde{w}_0^b(r_0, q_0, q_1^r, \lambda) < \tilde{w}_0^a(r_0, q_0, q_1^r, \lambda) \quad \forall q_t \in [q, \overline{q}].$$

Moreover, the reservation productivity of a “pessimistic” firm $q_0^{u,b}$ is always greater than the reservation productivity of an “optimistic” firm $q_0^{u,a}$:

$$q_0^{u,b}(r_0, q_1^r, \lambda) > q_0^{u,a}(r_0, q_1^r, \lambda) \quad \forall q_t \in [q, \overline{q}].$$

If a firm has pessimistic expectations it would not only offer a less generous initial wage contract, but it would also require a greater initial match productivity to start the employment relationship. In fact, since the match productivity is expected to be lower in the future, the firm is more likely to be in the situation to enact a costly wage cut. The cost of the initial wage contract into the future firm’s profit is expected to be greater.

The reasoning behind Proposition 6 can be understood more clearly through the analysis of the function $\theta(\cdot)$. Recall that $\theta(\cdot)$ captures the marginal cost of the initial wage contract $w_0$ on the firm’s expected future maximized profit. As such the initial wage contract is always costly to the firm. However, since a wage cut is more costly than a wage raise due to the stronger negative reciprocity response from the worker, and since $q_1^b > q_1^a$, if a firm is pessimistic it expects the overall marginal costs of the initial wage contract to be greater than an optimistic firm, that is, $\theta(w_0, \lambda, F_b) > \theta(w_0, \lambda, F_a)$. By looking at Figure 8 it now is clear that a firm’s incentive to minimize the area below $q_1^b$ by setting a lower $w_0$ is greater if $q_1$ is drawn from $F_b(Q)$ than if it is drawn from $F_a(Q)$. Hence a pessimistic firm will offer a lower initial wage contract $\tilde{w}_0$, and by the definition of the reservation productivity, it will be less willing to hire a worker that has a relatively low match productivity.

4.3 Discussion

The illustration presented above has been useful to investigate several implications of the theory set out in Sections 2 and 3. The current section is dedicated to a looser discussion about additional insights and predictions that the theory generates.

Time horizon and upward wage rigidity. In the parameterized version of the model we implemented a single employment period time-line (two dates $t = 0, 1$) to study the conditions under which downward wage rigidity may arise. If the match productivity substantially declines ($q_1 < \frac{q_0}{2}$) the firm enacts a wage cut; whilst if the decline is moderate ($\frac{q_0}{2} \leq q_1 \leq q_0$) the firm optimally freezes the wage at the reference wage. Now suppose we introduce an additional employment period and an additional hypothetical opportunity for renegotiation at time $t = 2$. The most straightforward implication is that, if the firm is forward looking, both the initial wage contract $\tilde{w}_0$ and the renegotiated wage contract $\tilde{w}_1$ will be lower (wage compression). The marginal cost of the initial wage contract,
for any possible realization of the match productivity, will now bind for an additional period influencing the firm wage policy at time 0 and time 1.\textsuperscript{33} Another implication is that the model can generate upward wage rigidity; this can arise under two conditions. First, suppose that at time 1 the firm freezes the wage $\tilde{w}_1 = r_1 = \tilde{w}_0$ due to a realization of $q_1$ within the range of rigidity $[q_{l1}, q_{u1}]$. Even if in the following period the level of match productivity increases $q_2 > q_1$, as long as the realization of $q_2$ is within the range of rigidity $[q_{l2}, q_{u2}]$, the firm would not raise the wage and set $\tilde{w}_2 = r_2 = \tilde{w}_1 = r_1 = \tilde{w}_0$.

Secondly, suppose instead that at time 1 the firm enacts a wage cut $\tilde{w}_1 < r_1 = \tilde{w}_0$ due to a sufficiently low realization of the match productivity $q_1 < q_{l1}$. If the worker adapts to the new wage received, his reference wage $r_2$ will now be lower, shifting the range of rigidity and the wage setting curve down and to the left in the diagram. Although we do not provide a formal proof here, it is easy to demonstrate that the level of $q_1$ that induced a wage cut at time 1 will now characterize the lower threshold of the range of rigidity for the following wage renegotiation at time 2. As such, if $q_2$ is greater than $q_1$, but is not sufficiently high to go beyond the range of rigidity, again it is optimal for the firm to freeze the worker’s wage and to pay the reference wage $\tilde{w}_2 = r_2 = \tilde{w}_1$.\textsuperscript{34}

These two cases suggest that after the event of a wage cut or a wage freeze (under certain circumstances) the firm would not always raise the wage in response to an increase in the match productivity. The firm would optimally wait for the match productivity to get back to a sufficiently high level (above the range of rigidity) before raising the wage again. Extending this implication a step further and for a longer time horizon allows us to generate the following prediction: after a downturn, periods of persistently low productivity growth will be characterized by periods of relatively flat wage growth. Firms would prefer to wait for productivity growth to pick up again at a certain level before starting to raise wages.\textsuperscript{35}

Asymmetries in reciprocity. Anthropological and experimental research highlight systematic asymmetries with respect to workers’ reciprocity responses to wage changes: negative reciprocity is stronger and normally more persistent while positive reciprocity is weaker and believed to be temporary. A question that arises from the literature is whether the weaker positive reciprocity is the result of a lower weight on the worker’s

\textsuperscript{33}Note that extending the time horizon to infinite will not change the main qualitative implications of the model. Intuitively a longer time horizon implies larger costs of the initial wage contract and of each renegotiated contract, therefore generating an even larger wage compression like in Elsby (2009).

\textsuperscript{34}The prediction of upward wage rigidity along with downward wage rigidity is consistent with the evidence reported by Ball and Moffitt (2002). According to their investigation, in the short run, changes in productivity are not always couched with changes in real wages, allowing unemployment to adjust without inflationary pressure. In our model this discontinuous rate of adjustment of wages to productivity occurs within the range of rigidity. Although we do not discuss the implications for inflation and unemployment, it may be interesting to extend the model in that direction.

\textsuperscript{35}Exploring the model implications from this perspective could also provide an explanation for the current UK post-crisis experience.
reciprocal preferences or a shift in his reference point. The question is therefore concerned with the matters of intensity and persistence of reciprocal behavior. Our model allows to disentangle these two effects and to identify the relative sources of asymmetries. The intensity of reciprocity is captured by the worker’s weight on gain-loss utility \( \eta > 0 \) and by the loss aversion parameter \( \lambda \geq 1 \). If a worker is loss averse \( (\lambda > 1) \), a wage cut generates a stronger negative impact on morale than a same size wage raise, triggering stronger negative reciprocity response in the form of subnormal effort. In our model the asymmetry in intensity is explained by loss aversion.\(^{36}\) Negative reciprocity also seems to be more persistent than positive reciprocity. In the light of our theory the difference in persistence could potentially be explained by an asymmetric speed of adaptation of the reference wage. In the illustration the worker always adapts the reference wage to the wage raise (perceived as a gain), making positive reciprocity only last one period. Now suppose that the worker is going to receive a wage cut (perceived as a loss) in a following period from time 1 to time 2: would he adapt to it as fast as he did for the wage rise?

Experimental evidence on reference point adaptation has found that individuals adapt more rapidly to gains than to losses (Arkes et al., 2008, 2010). Bringing this insight into our theory allows us to conclude that a worker may adapt more rapidly to wage rises than to wage cuts. For this reason the positive reciprocity induced by a wage rise is going to be only temporary (one period) while the negative reciprocity induced by a wage cut is likely to be more persistent (more than one period). In our theory the greater persistence of negative reciprocity could be explained by a longer adaptation of the worker’s reference wage after a wage cut. This asymmetric adaptation would increase the firm’s cost of implementing a wage cut, reinforcing the propositions derived in the illustration of Section 4.

**Complete information.** One of the core assumptions of our theory is that the firm has complete information about the worker’s reciprocal behavior, captured by the optimal effort function established in Theorem 1. The firm knows the potential costs of a wage cut and negative reciprocity (of which the loss aversion parameter is a key determinant) and take them into account when making wage setting decisions to maximize profits. Although this assumption reflects evidence that firms’ managers have some information about workers’ responses to wage changes, it does not fully represent reality. In fact managers aware of the workers’ negative reciprocity may not be aware of how strong that response might be. That is, a firm may have incomplete information.

\(^{36}\)A possible methodology to test the validity of this implication (Theorem 1) would be to conduct an experiment divided in two treatments: the first with the objective to measure individuals’ degree of loss aversion; the second with the objective to measure the extent of their negative reciprocity behavior. If the two measures are positively correlated, that is, if individuals with a greater extent of loss aversion also exhibit stronger negative reciprocity, the theory is corroborated: loss aversion underlies workers’ negative reciprocity behavior.
Beliefs about its workers’ loss aversion may lead a firm to underestimate or overestimate the costs of wage cuts and downward wage rigidity, generating wage setting decisions that are sub-optimal. From a normative perspective we can conclude that if a firm can possibly infer its workers’ extent of loss aversion, this will become useful information to implement a more effective wage policy.

**The role of the firm’s expectations.** In the last part of the illustration we demonstrated that if a firm is “pessimistic” it will offer a lower initial wage contract and it will be less willing to hire a worker with a relatively low match productivity. By bringing the intuition of this analysis a step further we are able to draw some broader predictions.

For instance, consider a labor market in the aftermath of a recession characterized by a persistently low labor productivity. If firms believe that productivity will continue to stay low (if they are “pessimistic”), and if we assume that workers’ reservation wages are given and fixed, we would expect wages to grow at a slow pace with relatively fewer employment contracts offered. On the other hand, as soon as the firms’ expectations about future productivity start to improve (if they are “optimistic”), we would expect to observe higher wages and more employment contracts offered. This discussion sidesteps how firms’ expectations might be influenced, but nevertheless it recognizes their crucial role in driving the outcome of wage setting and hiring behavior. To conclude, note that these predictions are drawn by considering the workers’ reservation wages as given and constant, restricting the analysis to the demand-for-labor side of the story. Thus, whilst beyond the scope of the current paper, exploring the interaction between changes in workers’ reservation wages and firms’ expectations is certainly a promising route to pursue.

**The reference wage.** The general model set out in Section 3 considers the worker’s reference wage as exogenous and given, establishing the theorems for any given reference wage. Instead, by exploring the hypothesis of “contracts as reference points”, in the illustration of Section 4 we assumed that an employed worker’s reference wage is endogenously determined by the most recent employment contract signed with the firm (in our case the initial employment contract). This assumption reflects the shift in the workers’ feelings of entitlement when they enter the firm, as documented by the anthropological and experimental research discussed throughout the paper. Whilst this assumption is instructive in informing the theory, consideration of the formation of the reference wage could be more elaborate: our current assumption sidesteps the role of an unemployed worker’s reference wage when contracting with the firm at time 0 (i.e. his reservation wage) and the importance of other factors influencing an employed worker’s reference wage when renegotiating the contract with the firm at time 1.

For instance an unemployed worker’s reference wage may be influenced by factors such as the state of the labor market, his most recent employment contract with another firm or the wage of other workers performing the same job he is applying for in the
market. Combining these aspects with the model presented in this paper would provide a richer theoretical framework with which to study hiring behavior and wage setting for newly hired workers under different economic conditions. On the other hand an employed worker’s perceptions of what is the “fair” wage may be influenced by other factors that go beyond the most recent wage contract signed with the firm. For instance information about the firm’s profitability is a natural candidate to provide different implications than those derived in Section 4. Consider a worker who has information about the reasons behind the firm’s incentive to cut his wage and that consider this information relevant for his perceptions of fairness. The worker’s reference wage could now be written as a function that incorporates an indicator of the firm’s ability to pay as an additional argument. In periods of adverse financial conditions the worker may revise his perceptions of fairness and accept a lower wage contract from the firm. The worker’s reference wage would shift downwards and with it the whole wage setting rule, leaving a space for the firm to enact a wage cut without incurring the costs of negative reciprocity and low morale.

This hypothetical scenario reflects the opinions of managers discussed in Bewley (1999), that wage cuts without loss of morale are achievable when workers internalize the firm’s objectives and understand the reasons behind them. A similar discussion is included in Kahneman et al. (1986), who stress the importance of managers’ information disclosure in influencing workers’ perceptions of fairness. Obviously there may be many other influences of workers’ reference wages: labor market institutions, internal wage structures, external wage comparisons and labor market conditions are among those most frequently discussed in the literature. Incorporating these elements into our theory is likely to provide a broader set of implications for wage setting behavior and ultimately wage dynamics, on which we are currently making progress.

5 Conclusion

This paper has formalized a theory of wage setting behavior in a simple and tractable model that explicitly considers behavioral aspects of the employment relationship related to workers’ morale, fairness and reciprocity. The paper brings together developments in the theory of individuals’ decision making, i.e. reference-dependent preferences and loss aversion, with systematic behavior and findings documented by the recent anthropological and experimental research. This approach allowed us to provide a formal treatment of wage setting behavior that is grounded on a more in depth understanding of the employment relationship, and to identify the type of incentives and psychological forces that influence workers’ and firms’ decision making processes. Given its foundations that draw on the literature of labor markets, incomplete contracts and psychology, we believe this paper to be a first step toward a unified theory of wage setting behavior.

Workers’ reciprocity behavior in terms of effort is triggered by wage offers that deviate
from what they perceive to be a “fair” wage, namely their reference wage. Thus, a worker’s reciprocal behavior is reference dependent where negative reciprocity, induced by a wage cut below the worker’s reference wage, is stronger than positive reciprocity, induced by a wage raise above the worker’s reference wage. One of the main contributions of our theory is to identify loss aversion as the key psychological determinant of negative reciprocity being stronger than positive reciprocity: wage cuts (perceived as losses) have a stronger impact on a worker’s morale than wage raises (perceived as gains).

Concerns about paying the “fair” wage in turn influence the optimal wage policy of a firm. Due to the potential cost generated by negative reciprocity, a firm may refrain from enacting wage cuts even when facing adverse economic conditions. As such our theory provides realistic and generalized microfoundation for downward wage rigidity, encompassing other models in the literature such as Akerlof and Yellen (1990), Elsby (2009) and Eliaz and Spiegler (2013) as well as formalizing early insights of economists such as Slichter (1920), Hicks (1963) and Okun (1981). The illustration of Section 4 showed several implications of the theory in the context of an initial employment contract and renegotiation by exploring the hypothesis of contracts as reference points. The worker’s negative reciprocity represents a cost for a firm, influencing the outcome of renegotiation, and if the firm is forward looking, also the outcome of the initial employment contract. The core message delivered by our theory is that firms are undoubtedly concerned about paying workers “fair” wages, are aware of their reciprocal behavior and take these psychological aspects into account when deciding on their optimal wage policy to maximize profits.

In our general formalization the nature and the dynamics of the reference “fair” wage it is not specified and does not matter. This feature makes the two derived theorems general and valid for any given reference wage, enhancing the universality of our theory. Nevertheless our paper highlights the need for a better understanding of workers’ perceptions of what is the “fair” wage. As discussed throughout the paper, the reference wage is one of the key determinants of the major propositions derived: the worker’s optimal effort function and the firm’s wage setting rule are both explicit functions of the reference wage. In the light of our theory, understanding how workers form their perceptions of fairness, and how these perceptions may change over time or vary across type of workers, becomes a crucial element to fully understand a firm’s wage policy. Thus, in order to advance a step further towards formulating a more complete theory of wage setting behavior, it is essential to directly engage with a theory of workers’ reference wage formation. At this stage it would be useful to systematically follow the guidelines put forward by the research on individuals’ reference point formation to discriminate among the assumptions on reference wages investigated by the labor market literature. This approach seems to be promising, and is the subject of further research that is currently being pursued.
References


ttainen, H., editors, Behavioural Economics and Its Applications. Princeton University 
Press.


from a Long-Term Field Experiment. Working Paper, pages 1–44.

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.


Cox, J. C., Friedman, D., and Gjerstad, S. (2007). A tractable model of reciprocity and 


Firms’ price and wage adjustment in Europe: Survey evidence on nominal stickiness. 

avoid cutting wages: Survey evidence from European Firms. World Bank Group 


Appendix

Lemma 1. If assumptions A0-A3 hold, then \( \mu(\cdot) \) is increasing and concave in \( w \) and decreasing in \( r \), such that \( n^w(w|r) > 0 \), \( n^{wr}(w|r) < 0 \), \( n^r(w|r) < 0 \) and the cross partial \( n^{wr}(w|r) = 0 \) for all \( w \) and \( r \).

This lemma clarifies that the gain-loss function is increasing and concave with respect to the wage and decreasing with respect to the reference wage. It also highlights the properties of the cross partial derivative of the gain-loss function, which is useful to prove Theorem 2.

Proof of Lemma 1.

\[
n^w(w|r) = \frac{\partial \mu}{\partial w} = \mu'(m(w) - m(r))m'(w) > 0
\]

by A0 and A2;

\[
n^{wr}(w|r) = \frac{\partial^2 \mu}{\partial w \partial r} = \mu'(m(w) - m(r))m''(w) + \mu''(m(w) - m(r))m'(w) < 0
\]

since \( \mu''(\cdot) = 0 \), by A2 and \( m''(w) < 0 \), by A0;

\[
n^r(w|r) = \frac{\partial \mu}{\partial r} = \mu'(m(w) - m(r))m'(r) < 0,
\]

and the cross partial

\[
n^{wr}(w|r) = \frac{\partial^2 \mu}{\partial w \partial r} = \mu'(m(w) - m(r))m'^r(w) + \mu''(m(w) - m(r))m'(w)m'(r) = 0
\]

since \( \mu''(\cdot) = 0 \) by A2 and \( m'^r(w) \) is not defines since \( m'(w) \) is not a function of \( r \). \( \square \)

Proof of Theorem 1. We first prove continuity of \( \hat{e}(w,r) \) in both \( w \) and \( r \) and then, appealing to the implicit function theorem, deduce its properties.

Continuity. In the FOC given by equation (3) both \( b'(e) \) and \( c'(e) \) vary continuously with \( e \), by assumption B1. However, \( g'(e) \cdot \mu(x) \) does not as \( \mu(x) \) is not differentiable when \( x = 0 \). As such, for \( w \neq r \) it is obvious that \( \hat{e}(w,r) \) varies continuously with \( w \) and \( r \), and the only point of potential discontinuity is at \( w = r \). When \( w = r, \hat{e}(w,r) \) is given
by $b'(e) - c'(e) = 0$. Taking limits in the FOC given by equation (3) as $w \searrow r$ (i.e. $w$ approaches $r$ from above) implies $\tilde{e}(w, r) \to \tilde{e}(r, r)$, and the same is true for $w \nearrow r$ (i.e. $w$ approaches $r$ from below). As such, whilst $\tilde{e}(w, r)$ is not differentiable at $w = r$, it is a continuous function of $w$, and the same is true by similar arguments for $r$.

Properties of $\tilde{e}(w, r)$. For $w \neq r$, our assumptions imply that $\mu(x)$ is twice continuously differentiable, thus appealing to the implicit function theorem, we can deduce that
\[
\frac{\partial \tilde{e}}{\partial w} = -\frac{\partial \Omega/\partial w}{\partial \Omega/\partial e} = -\frac{g'(e) \cdot \mu'(m(w) - m(r))m'(w)}{b'(e) - c'(e)} > 0,
\]
where the denominator is negative by assumptions B1-B2 and the numerator is positive by assumptions A0-A2. Therefore $\tilde{e}(w, r)$ is increasing in $w$. Note that under the assumption of loss aversion (A3), which implies that $\mu'(m(w) - m(r))|_{w<r} \mu'(m(w) - m(r))|_{w>r} = \lambda > 1$,

we obtain the asymmetric relationship:
\[
\frac{\partial \tilde{e}/\partial w|_{w<r}}{\partial \tilde{e}/\partial w|_{w>r}} = \lambda > 1.
\]

Furthermore,
\[
\frac{\partial^2 \tilde{e}}{\partial w^2} = -\frac{(b''(e) - c''(e))(g'(e) \cdot \mu'(m(w) - m(r))m''(w))}{(b'(e) - c'(e))^2} < 0;
\]
since the numerator is positive due to $m''(w) < 0$, by A0, implicit differentiation proves that $\tilde{e}(w, r)$ is concave in $w$. In addition,
\[
\frac{\partial \tilde{e}}{\partial r} = -\frac{\partial \Omega/\partial r}{\partial \Omega/\partial e} = -\frac{g'(e) \cdot \mu'(m(w) - m(r))m'(r)}{b'(e) - c'(e)} < 0,
\]
where the numerator is negative as we are differentiating $\mu(\cdot)$ with respect to $r$. Therefore $\tilde{e}(w, r)$ is decreasing in $r$. Finally note that whenever $w < r$:
\[
\frac{\partial \tilde{e}}{\partial \lambda}|_{w<r} = -\frac{\partial \Omega/\partial \lambda}{\partial \Omega/\partial e} = -\frac{g'(e) \cdot \mu'(m(w) - m(r))}{b'(e) - c'(e)} < 0
\]
and
\[
\frac{\partial^2 \tilde{e}}{\partial w \partial \lambda}|_{w<r} = -\frac{(b''(e) - c''(e))(g'(e) \cdot \frac{\partial}{\partial \lambda} \{\mu'(m(w) - m(r))\})m'(w))}{(b'(e) - c'(e))^2} > 0,
\]
which, by assumptions A0-A3, together prove the statement of Corollary 1.

Proof of Theorem 2. We first prove continuity of $\tilde{w}(r, q, \lambda)$ in both $q$ and $r$. 

46
Subsequently we define the two match productivity thresholds \( q^u \) and \( q^l \) and we deduce the properties of \( \tilde{w}(r, q, \lambda) \) for the case of \( q > q^u \) and \( q < q^l \). We then turn to analyze the case of \( q \in [q^l, q^u] \) and prove that \( \tilde{w} = r \) characterizes the optimal (corner) solution when \( q \) falls within this range. Finally we deduce the properties of \( q^u \) and \( q^l \) which prove the statements of Corollaries 2 and 3.

**Continuity.** In the FOC given by equation (4) \( y'(\cdot) \) varies continuously with both \( q \) and \( e \), by assumption C2, \( \bar{e}'(\cdot) \) is not differentiable at \( w \neq r \) but varies continuously with both \( w \) and \( r \) as proved before and \( s'(w) \) varies continuously with \( w \), by assumption C1. As such, \( \tilde{w}(r, q, \lambda) \) is a continuous function of both \( q \) and \( r \).

**Properties of \( \tilde{w}(r, q, \lambda) \) for \( q > q^u \).** \( q^u \) is defined as the match productivity threshold characterized by:

\[
q^u(r) \equiv \left\{ q : \frac{\partial y(q, e) \partial \bar{e}(w, r)}{\partial w} \bigg|_{w \to r} = \frac{ds(r)}{dw} \right\}.
\]

Note that every function is evaluated at \( r \). To simplify notations, denote the left hand side with \( M^+ \equiv \frac{\partial y(q, e) \partial \bar{e}(w, r)}{\partial w} \bigg|_{w \to r} \). Hence, for all \( q > q^u \), \( M^+|_{q>q^u} > M^+|_{q=q^u} \) and since \( s'(w) \) is non-decreasing, there exists an optimal wage \( \tilde{w}(r, q, \lambda) > r \). Denote this wage as \( \tilde{w}(r, q)^+ \):

\[
\tilde{w}(r, q)^+ \equiv \left\{ w : \frac{\partial y(q, e) \partial \bar{e}(w, r)}{\partial w} \bigg|_{w \to r} = \frac{ds(w)}{dw} \right\}.
\]

Which means that whenever \( q > q^u \) the firm sets the profit maximizing wage \( \tilde{w}(r, q)^+ > r \) such that the tradeoff between the marginal value of output and the marginal labor cost is equalized.

![Margins Diagram](image)

To analyze how the optimal wage \( \tilde{w}(r, q)^+ \) adjusts to \( q \) and \( r \), let us first denote the FOC expressed in equation (4) when \( w > r \) as \( \Psi(q, e, r)^+ \equiv \Psi(q, e, r)|_{w>r} \). Thus,
appealing to the implicit function theorem we can deduce that
\[
\frac{\partial \tilde{w}^+}{\partial q} = -\frac{\partial \Psi^+ / \partial q}{\partial \Psi^+ / \partial w} = -\frac{\partial^2 y}{\partial e \partial q} \frac{\partial e}{\partial w} \bigg|_{w>r} > 0,
\]
where the denominator is negative and the numerator is positive by assumptions C1-C2 and Theorem 1. Moreover we deduce that
\[
\frac{\partial \tilde{w}^+}{\partial r} = -\frac{\partial \Psi^+ / \partial r}{\partial \Psi^+ / \partial w} = -\frac{\partial^2 e}{\partial e \partial r} \frac{\partial e}{\partial w} \bigg|_{w>r} + \frac{\partial y}{\partial e} \frac{\partial^2 e}{\partial w \partial r} \bigg|_{w>r} > 0,
\]
where by similar arguments the numerator is positive (with \( \frac{\partial^2 e}{\partial w \partial r} \bigg|_{w>r} = 0 \)) and the denominator is negative. Therefore we have proved that \( \tilde{w}(r, q)^+ \) is increasing in \( q \) and increasing in \( r \).

**Properties of \( \tilde{w}(r, q, \lambda) \) for \( q < q^l \).** Recall that \( q^l \) is defined as the match productivity threshold characterized by:
\[
q^l(r, \lambda) \equiv \left\{ q : \frac{\partial y(q, e)}{\partial e} \frac{\partial e(w, r)}{\partial w} \bigg|_{w>r} = \frac{ds(r)}{dw} \right\}.
\]
Note that every function is evaluated at \( r \). To simplify notations, denote the left hand side with \( M^- \equiv \frac{\partial y(q, e)}{\partial e} \frac{\partial e(w, r)}{\partial w} \bigg|_{w>r} \). Hence, for all \( q < q^l \), \( M^- \bigg|_{q<q^l} < M^- \bigg|_{q=q^l} \) and since \( s'(w) \) is non-decreasing, there exists an optimal wage \( \tilde{w}(r, q, \lambda) < r \). Denote this wage as \( \tilde{w}(r, q, \lambda)^- \):
\[
\tilde{w}(r, q, \lambda)^- \equiv \left\{ w : \frac{\partial y(q, e)}{\partial e} \frac{\partial e(w, r)}{\partial w} \bigg|_{w<r} = \frac{ds(w)}{dw} \right\}.
\]
Whenever \( q < q^l \) the firm sets the profit maximizing wage \( \tilde{w}(r, q, \lambda) < r \) such that the tradeoff between the marginal value of output and the marginal labor cost is equalized.

To analyze how the optimal wage \( \tilde{w}(r, q, \lambda)^- \) adjusts to \( q \) and \( r \), let us first denote
the FOC expressed in equation (4) when \( w < r \) as \( \Psi(q,e,r) \equiv \Psi(q,e,r)|_{w<r} \). Thus, appealing to the implicit function theorem we can deduce that

\[
\frac{\partial \tilde{w}^-}{\partial q} = \frac{\partial \Psi^- / \partial q}{\partial \Psi^- / \partial w} = -\frac{\partial^2 y}{\partial e \partial q} \bigg|_{w<r} \frac{\partial \Psi}{\partial w} \bigg|_{w<r} - \frac{d^2 s}{dw^2} > 0,
\]

and that

\[
\frac{\partial \tilde{w}^-}{\partial r} = \frac{\partial \Psi^- / \partial r}{\partial \Psi^- / \partial w} = -\frac{\partial^2 y}{\partial e \partial r} \bigg|_{w<r} \frac{\partial \Psi}{\partial w} \bigg|_{w<r} + \frac{\partial y}{\partial e} \frac{\partial^2 \tilde{e}}{\partial w \partial r} \bigg|_{w<r} > 0,
\]

which, by similar arguments as above, proves that \( \tilde{w}(r,q,\lambda)^- \) is increasing in \( q \) and increasing in \( r \). Moreover note that, as \( \frac{\partial^2 \tilde{e}}{\partial w \partial \lambda} \bigg|_{w<r} > 0 \) (see Corollary 1) we deduce that

\[
\frac{\partial \tilde{w}^-}{\partial \lambda} = \frac{\partial \Psi^- / \partial \lambda}{\partial \Psi^- / \partial w} = -\frac{\partial y}{\partial e} \frac{\partial^2 \tilde{e}}{\partial w \partial \lambda} \bigg|_{w<r} - \frac{d^2 s}{dw^2} > 0,
\]

which means that whenever \( w < r \), the optimal wage \( \tilde{w}(r,q,\lambda)^- \) is increasing in \( \lambda \), therefore proving the statement of Corollary 4.

**Properties of** \( \tilde{w}(r,q,\lambda) \) **for** \( q \in [q^l, q^u] \). First note that due to assumption A3, for \( \lambda > 1 \), \( q^l < q^u \) and when \( \lambda = 1 \) \( q^l = q^u \). This follows from the observation that \( q^l(r,1) = q^a(r) \) by definition; \( q^u(r) \) does not depend on \( \lambda \); and from Corollary 1 (which is proved below) \( \frac{\partial q^l(r,\lambda)}{\partial \lambda} < 0 \). As such, for loss averse workers (\( \lambda > 1 \)), the range of rigidity is non-empty.

We now seek to prove that for all \( q^l \leq q \leq q^u \) the optimal wage is characterized by \( \tilde{w}(r,q,\lambda)^- = r \). We do so by demonstrating that for the value of \( q \in [q^l, q^u] \), \( \frac{\partial \pi}{\partial w} < 0 \) for all \( w > r \) and \( \frac{\partial \pi}{\partial w} > 0 \) for all \( w < r \).

First, for any \( q \leq q^u \), \( \frac{\partial \pi}{\partial w} < 0 \) whenever \( w > r \). To see this note that at \( q = q^u \), \( \frac{\partial \pi}{\partial w} \bigg|_{w<r} = 0 \), and therefore for all \( q < q^u \), \( \frac{\partial \pi}{\partial w} \bigg|_{w<r} < 0 \), since \( \frac{\partial^2 y}{\partial e \partial q} > 0 \) by assumption C2. Moreover, by our assumptions and the result of Theorem 1, \( \frac{\partial \pi}{\partial w^2} < 0 \) for all \( w > r \), which implies that for any \( w > r \), \( \frac{\partial \pi}{\partial w} < 0 \). Similarly for any \( q \geq q^l \), \( \frac{\partial \pi}{\partial w} > 0 \) whenever \( w < r \): at
\[ q = q' \frac{\partial e}{\partial w} \big|_{w=r} = 0, \text{ and therefore } \frac{\partial e}{\partial w} \big|_{w=r} > 0 \text{ for all } q > q', \text{ since } \frac{\partial^2 e}{\partial q^2} > 0. \text{ Concavity of the profit function then implies that for all } w < r, \frac{\partial e}{\partial w} > 0. \text{ As such, for all } q' \leq q \leq q^u, \text{ profit is increasing in } w \text{ below } r \text{ and decreasing in } w \text{ above } r, \text{ therefore it is maximized at } \hat{w} = r. \]

Hence it is straightforward that for all \( q \in [q', q^u] \) the optimal wage characterized by \( \hat{w}(r, q, \lambda) = r \) is increasing in \( r \) and non-decreasing in \( q \).

**Properties of the thresholds \( q^u(r) \) and \( q^l(r, \lambda) \).** By similar arguments used to prove continuity of \( \hat{w}(r, q, \lambda) \), the threshold \( q^u(r) \) is a continuous function of \( r \) and \( q^l(r, \lambda) \) is a continuous function of both \( r \) and \( \lambda \). Thus appealing to the implicit function theorem we can deduce that

\[
\frac{\partial q^u}{\partial r} = -\frac{\partial \Psi / \partial r |_{w=r}}{\partial \Psi / \partial q |_{w=r}} = -\frac{\frac{\partial^2 y}{\partial e^2} \frac{\partial e}{\partial w} \big|_{w=r} \frac{\partial e}{\partial w} \big|_{w=r} + \frac{\partial y}{\partial e} \frac{\partial y}{\partial \lambda} \big|_{w=r}}{\frac{\partial^2 y}{\partial e^2} \frac{\partial e}{\partial w} \big|_{w=r}} > 0,
\]

and that

\[
\frac{\partial q^l}{\partial r} = -\frac{\partial \Psi / \partial r |_{w=r}}{\partial \Psi / \partial q |_{w=r}} = -\frac{\frac{\partial^2 y}{\partial e^2} \frac{\partial e}{\partial w} \big|_{w=r} \frac{\partial e}{\partial w} \big|_{w=r} + \frac{\partial y}{\partial e} \frac{\partial y}{\partial \lambda} \big|_{w=r}}{\frac{\partial^2 y}{\partial e^2} \frac{\partial e}{\partial w} \big|_{w=r}} > 0,
\]

where in both cases the denominator is positive by assumption C2 and Theorem 1 and the numerator is negative by assumption C1-C2 and by definition of \( q^u \) and \( q^l \). Note that when we are differentiating the productivity thresholds \( q^u \) and \( q^l \) with respect to \( r \), we are differentiating two functions that are evaluated at \( r \). As such, the partial derivative \( \frac{\partial e}{\partial w} \) which is normally negative when \( w \neq r \), is now equal to zero. In fact, when \( w \rightarrow r \) (from above or below) optimal effort is \( \hat{e}(w, r) \rightarrow \hat{e}(r, r) \), which approaches “normal” effort \( \hat{e}_n \) as defined in Theorem 1. Moreover, the expression \( \frac{\partial}{\partial r} \left( \frac{\partial e}{\partial w \big|_{w-r}} \right) \) means that we are taking the partial derivative with respect to the reference wage of the partial derivative of effort with respect to the wage \( \frac{\partial e}{\partial w} \), evaluated at \( r \). Therefore it is not a cross partial derivative of effort and the effect is decreasing with \( \frac{\partial}{\partial r} \left( \frac{\partial e}{\partial w \big|_{w-r}} \right) < 0 \). By similar arguments \( \frac{\partial}{\partial r} \left( \frac{\partial e}{\partial w \big|_{w-r}} \right) > 0 \) which proves that the numerator is negative. Therefore we can conclude that both \( q^u \) and \( q^l \) are increasing functions of \( r \) which proves the statement of Corollary 2.

Finally note that

\[
\frac{\partial q^l}{\partial \lambda} = -\frac{\partial \Psi / \partial \lambda |_{w=r}}{\partial \Psi / \partial q |_{w=r}} = -\frac{\frac{\partial y}{\partial e} \frac{\partial y}{\partial \lambda} \big|_{w=r}}{\frac{\partial^2 y}{\partial e^2} \frac{\partial e}{\partial w} \big|_{w=r}} < 0,
\]

which means that the lower threshold \( q^l \) is decreasing in \( \lambda \) and therefore proves the statement of Corollary 3. \(\square\)
Proof of Proposition 1. We follow the analytical structure of the proof of Theorem 2. However, since by assumption a worker will not accept a wage offer \( w_0 < r_0 \), we do not consider this case and characterize the optimal initial wage contract for \( w_0 \geq r_0 \) only.

Thus, we define \( q_0^u \) as the match productivity threshold:

\[
q_0^u(r_0) \equiv \{ q_0 : \alpha q_0 r_0^{\alpha-1} = 1 \},
\]

characterized by the firm’s FOC, evaluated at \( r_0 \), conditional on \( w_0 > r_0 \). Solving the equation for \( q_0 \) we find that \( q_0^u(r_0) = \frac{1}{\alpha} r_0^{1-\alpha} \). Hence, for all \( q_0 > q_0^u \) there exists an optimal wage \( \tilde{w}_0(r_0, q_0) > r_0 \) such that the tradeoff between the marginal value of output and the marginal labor cost (given by the FOC, conditional on \( w_0 > r_0 \)) is equalized:

\[
\tilde{w}_0(r_0, q_0) \equiv \{ w_0 : \alpha q_0 w_0^{\alpha-1} = 1 \}.
\]

Solving for \( w_0 \) we find that \( \tilde{w}_0(r_0, q_0) = (\alpha q_0)^{\frac{1}{1-\alpha}} \). Hence, as in the proof of Theorem 2, if \( q_0 = q_0^u \) the optimal wage is \( \tilde{w}_0(r_0, q_0) = r_0 \). Note that this result can also be obtained if we evaluate the optimal wage \( \tilde{w}_0 = (\alpha q_0)^{\frac{1}{1-\alpha}} \) for the value of \( q_0 = q_0^u = \frac{1}{\alpha} r_0^{1-\alpha} \).

Finally, since \( q_0^u(r_0) \) is an increasing function of \( r_0 \), as established by Theorem 2, it is straightforward to see that a higher \( r_0 \) makes the range \([q_0^u(r_0), \tilde{q}]\) smaller. This proves the statement of Corollary 5. \( \square \)

Proof of Proposition 2. We follow the analytical structure of the proof of Theorem 2. Thus we define \( q_1^u \) and \( q_1^l \) as the productivity thresholds that define the range of rigidity within which the optimal wage is \( \tilde{w}_1(r_1, q_1, \lambda) = r_1 \):

\[
q_1^u(r_1) \equiv \{ q_1 : \alpha q_1 r_1^{\alpha-1} = 1 \}
\]

\[
q_1^l(r_1, \lambda) \equiv \{ q_0 : \lambda \alpha q_1 r_1^{\alpha-1} = 1 \},
\]

characterized by the firm’s FOC, evaluated at \( r_1 \), respectively conditional on \( w_1 > r_1 \) and \( w_1 < r_1 \). Solving the equation with respect to \( q_1 \) we find that \( q_1^u = \frac{1}{\alpha} r_1^{1-\alpha} \) and \( q_1^l = \frac{1}{\lambda\alpha} r_1^{1-\alpha} \).

Hence, for all \( q_1 \geq q_1^l \) there exists an optimal wage \( \tilde{w}_1(r_1, q_1)^+ > r_1 \) which solves the firm’s FOC:

\[
\tilde{w}_1(r_1, q_1)^+ \equiv \{ w_1 : \alpha q_1 w_1^{\alpha-1} = 1 \}.
\]

Solving for \( w_1 \) we find that \( \tilde{w}_1(r_1, q_1)^+ = (\alpha q_1)^{\frac{1}{1-\alpha}} \). And for all \( q_1 < q_1^l \) there exists an optimal wage \( \tilde{w}_1(r_1, q_1, \lambda)^- < r_1 \) which solves the firm’s FOC:

\[
\tilde{w}_1(r_1, q_1, \lambda)^- \equiv \{ w_1 : \lambda \alpha q_1 w_1^{\alpha-1} = 1 \}.
\]

Solving for \( w_1 \) we find that \( \tilde{w}_1(r_1, q_1, \lambda)^- = (\lambda \alpha q_1)^{\frac{1}{1-\alpha}} \). Finally note that since \( r_1 = \tilde{w}_0 \)
we can write:

\[ q_u(\tilde{w}_0) = \frac{\tilde{w}_0^{1-\alpha}}{\alpha} \]

\[ q_l'(\tilde{w}_0, \lambda) = \frac{\tilde{w}_0^{1-\alpha}}{\lambda \alpha}. \]

Substituting the optimal wage derived in Proposition 1 \((\tilde{w}_0 = (\alpha q_0)^{1/\alpha})\) and rearranging the equations we obtain \(q_u^* = q_0\) and \(q_l^* = q_0/\lambda\), which proves Proposition 2.

**Proof of Corollary 6.** Consider the case of no adaptation \(r_1 = r_0\) (hence \(q_u^* = q_0^*\)) and two different wage contracts characterized respectively by \(q_1 = q_0\) and \(q_1 < q_0\):

\[ \tilde{w}_1(r_0, q_0) = (\alpha q_0)^{1/\alpha} \]

\[ \tilde{w}_1(r_0, q_1) = (\alpha q_1)^{1/\alpha}. \]

It is straightforward to see that \((\alpha q_1)^{1/\alpha} < (\alpha q_0)^{1/\alpha}\) for all \(q_1 < q_0 \in [\underline{q}, \overline{q}]\).

**Proof of Proposition 3.** The function \(\theta(w_0, \lambda) \equiv \int_{\underline{q}}^{\overline{q}} J'_1(w_0, q_1) dF(q_1)\) is the first derivative with respect to \(w_0\) of the expected future maximized profit for any realization of \(q_1 \in [\underline{q}, \overline{q}]\). We are interested in the sign of \(\theta(w_0, \lambda)\).

First, consider the following envelope conditions:

\[ J'_1(w_0, q_1)^+ = -\alpha q_1(w_0)^{\alpha-1} \]

\[ J'_1(w_0, q_1)^0 = -1 \]

\[ J'_1(w_0, q_1)^- = -\lambda \alpha q_1(w_0)^{\alpha-1}, \]

which characterize the derivative of the value function \(J_1(w_0, q_1)\) respectively in the event of a wage raise, freeze or cut for a realization of the \(q_1\) above, within or below the range of rigidity. These conditions are obtained as follows: i) substitute the optimal wage derived in Proposition 2 into the firm’s profit function at time 1 for each possible realization of \(q_1\); ii) differentiate the obtained value of the maximized profit with respect to \(w_0\).

Therefore, recall equation (8):

\[
\int_{\underline{q}}^{\overline{q}} J_1(w_0, q_1) dF(q_1) = \int_{\underline{q}}^{q_1^*(w_0, \lambda)} J_1(w_0, q_1)^- dF + \int_{q_1^*(w_0)}^{q_1^*(w_0)} J_1(w_0, q_1)^0 dF + \int_{q_1^*(w_0)}^{\overline{q}} J_1(w_0, q_1)^+ dF, \tag{11}
\]
and differentiate with respect to $w_0$:

\[
\int_{q}^{\bar{q}} J'_1(w_0,q_1) dF(q_1) = \\
\int_{q}^{q_1^l(w_0,\lambda)} J'_1(w_0,q_1)^- dF + J_1(w_0,q_1)^- \frac{\partial q_1^l}{\partial w_0} \\
+ \int_{q_1^l(w_0,\lambda)}^{q_1^l(w_0)} J'_1(w_0,q_1)^0 dF + J_1(w_0,q_1)^0 \frac{\partial q_1^u}{\partial w_0} - J_1(w_0,q_1)^0 \frac{\partial q_1^l}{\partial w_0} \\
+ \int_{q_1^l(w_0)}^{\bar{q}} J'_1(w_0,q_1)^+ - J_1(w_0,q_1)^+ \frac{\partial q_1^u}{\partial w_0} dF. \tag{12}
\]

Note that all the derivatives with respect to the limits cancel out since

\[
J_1(w_0,q_1)^- = J_1(w_0,q_1)^0 \quad \text{and} \quad J_1(w_0,q_1)^0 = J_1(w_0,q_1)^+,
\]

that is, the value functions of the firm’s future profit evaluated respectively at $q_1 = q_1^l$ and $q_1 = q_1^u$ are the same. Thus we can rewrite equation (12) as:

\[
\int_{q}^{\bar{q}} J'_1(w_0,q_1) dF(q_1) = \\
\int_{q}^{q_1^l(w_0,\lambda)} J'_1(w_0,q_1)^- dF + \int_{q_1^l(w_0,\lambda)}^{q_1^l(w_0)} J'_1(w_0,q_1)^0 dF + \int_{q_1^l(w_0)}^{\bar{q}} J'_1(w_0,q_1)^+ dF, \tag{13}
\]

which corresponds to the function $\theta(w_0,\lambda)$ as stated in Proposition 3. Finally by substituting the envelope conditions derived in (10) we can write $\theta(w_0,\lambda)$ in its analytical form:

\[
\theta(w_0,\lambda) = \\
- \int_{q}^{q_1^l(w_0,\lambda)} \lambda \alpha q_1(w_0)^{\alpha-1} dF - \int_{q_1^l(w_0,\lambda)}^{q_1^l(w_0)} 1 dF - \int_{q_1^l(w_0)}^{\bar{q}} \alpha q_1(w_0)^{\alpha-1} dF.
\]

Hence $\theta(w_0,\lambda) < 0$  \(\square\)

**Proof of Proposition 4.** As when proving Proposition 1, we follow closely the analytical structure of the proof of Theorem 2.

Thus, define $q_0^u$ as the match productivity threshold:

\[
q_0^u(r_0,q_1^c,\lambda) \equiv \{q_0 : \alpha q_0 r_0^{\alpha-1} + \delta \theta(r_0,\lambda) = 1\},
\]

characterized by the firm’s FOC in (9), evaluated at $r_0$, conditional on $w_0 > r_0$. To simplify notations denote $T_0^+ \equiv \alpha q_0 r_0^{\alpha-1} + \delta \theta$. Hence, for all $q_0 > q_0^u$, $T_0^+|_{q=q^u} > T_0^+|_{q=q^u}$.
which means that there exists an optimal wage $\tilde{w}_0(r_0, q_0, q_1^e, \lambda) > r$. Denote this wage as $\tilde{w}_0(r_0, q_0, q_1^e, \lambda)^+$:

$$\tilde{w}_0(r_0, q_0, q_1^e, \lambda)^+ \equiv \{w_0 : \alpha q_0 w_0^{\alpha - 1} + \delta \theta(w_0, \lambda) = 1\}.$$ 

Which means that whenever $q_0 > q_0^*$ the firm sets the profit maximizing wage $\tilde{w}_0(r_0, q_0, q_1^e, \lambda)^+ > r$ such that the tradeoff between the marginal value of output at time 0 and the marginal labor cost at time 0, augmented by the discounted marginal cost of the initial wage contract into the future expected maximized profit of the firm at time 1, is equalized.

![Margins diagram](image)

The figure illustrates the optimal solution.

**Proof of Corollary 7.** When a firm is myopic $\delta = 0$ the optimal initial wage contract is characterized as:

$$\tilde{w}_0(r_0, q_0)^+ \equiv \{w_0 : \alpha q_0 w_0^{\alpha - 1} = 1\}.$$ 

Denote the left hand side by $M_0^+ \equiv \alpha q_0 w_0^{\alpha - 1}$. When a firm is forward looking ($\delta > 0$) the optimal initial wage contract is characterized as:

$$\tilde{w}_0(r_0, q_0, q_1^e, \lambda)^+ \equiv \{w_0 : \alpha q_0 w_0^{\alpha - 1} + \delta \theta(w_0, \lambda) = 1\}.$$ 

Where the left hand side is denoted with $T_0^+ \equiv \alpha q_0 w_0^{\alpha - 1} + \delta \theta$. It is straightforward to see that since $\theta < 0$, as established in Proposition 3, the margins $T_0^+ < M_0^+$ for all $q_t \in [\underline{q}, \bar{q}]$.

That is, a forward-looking firm faces an additional marginal cost associated with the initial wage contract and therefore optimally sets a lower profit maximizing wage in order to equalize the tradeoff defined by its first order condition.
Proof of Corollary 8. When a firm is myopic ($\delta = 0$) its reservation productivity is characterized as:

$$q^u_0(r_0) \equiv \{q_0 : \alpha q_0 r_0^{\alpha - 1} = 1\}.$$  

Solving for $q_0$ we found that $q^u_0(r_0) = \frac{r_0^{1-\alpha}}{\alpha}$. When a firm is forward looking ($\delta > 0$) its reservation productivity is characterized as:

$$q^u_0(r_0, q^e_1, \lambda) \equiv \{q_0 : \alpha q_0 r_0^{\alpha - 1} + \delta \theta(r_0, \lambda) = 1\}.$$  

Solving for $q_0$ we find that $q^u_0(r_0, q^e_1, \lambda) = \frac{(1-\delta \theta(r_0, \lambda)) r_0^{1-\alpha}}{\alpha}$.

Since $\theta < 0$,

$$(1 - \delta \theta(r_0, \lambda)) > 1$$

which implies that

$$\frac{(1 - \delta \theta(r_0, \lambda)) r_0^{1-\alpha}}{\alpha} > \frac{r_0^{1-\alpha}}{\alpha}.$$  

Hence, for all $q_1 \in [q, \bar{q}]$:

$$q^u_0(r_0, q^e_1, \lambda) > q^u_0(r_0).$$

The reservation productivity of a forward-looking firm is greater.

Proof of Corollary 9. We are interested in studying the effect of the loss aversion parameter $\lambda$ into the function $\theta(w_0, \lambda)$. Consider $\theta(w_0, \lambda)$ in its general form:

$$\theta(w_0, \lambda) = \int_{q^l(w_0, \lambda)} J'_1(w_0, q_1) \ dF + \int_{q^l(w_0, \lambda)} J'_0(w_0, q_1) \ dF + \int_{q^l(w_0, \lambda)} J'_0(w_0, q_1) \ dF.$$
Differentiate with respect to λ to obtain:

\[
\frac{\partial \theta}{\partial \lambda} = \frac{\partial}{\partial \lambda} \left\{ \int_2 q^l(w_0, \lambda) J'_1(w_0, q_1^-) dF \right\} + J'_1(w_0, q_1^a) \frac{\partial q_1^a}{\partial \lambda} - J'_1(w_0, q_1^a) \frac{\partial q_1^a}{\partial \lambda} < 0.
\]

That is because \(\frac{\partial}{\partial \lambda} \left\{ \int_2 q^l(w_0, \lambda) J'_1(w_0, q_1^-) dF \right\} < 0\) and the derivatives with respect to the limits cancel out since

\[
J'_1(w_0, q_1^a) = J'_1(w_0, q_1^a).
\]

that is, the value functions of the future profit evaluated at \(q_1 = q_1^a\) are the same. To see this write the function \(\theta\) in its analytical form using the envelope conditions derived in the proof of Proposition 3:

\[
\theta(w_0, \lambda) = - \left( \int_2 q^l(w_0, \lambda) \lambda q_1(w_0)^{\alpha-1} dF + \int_2 q^l(w_0) q_1(w_0) \alpha q_1(w_0)^{\alpha-1} dF + \int_2 q^l(w_0) = 1 \right).
\]

Then, differentiate with respect to λ to obtain:

\[
\frac{\partial \theta}{\partial \lambda} = - \left( \int_2 q^l(w_0, \lambda) \alpha q_1(w_0)^{\alpha-1} dF + \lambda \alpha q_1(w_0, \lambda) q_1(w_0) \alpha q_1(w_0)^{\alpha-1} \right) \frac{\partial q_1^l}{\partial \lambda} - \frac{\partial q_1^l}{\partial \lambda}.
\]

Substituting for \(q_1^l(w_0, \lambda) = \frac{w_0^{1-\alpha}}{\lambda \alpha}\) and rearranging the equation we obtain:

\[
\frac{\partial \theta}{\partial \lambda} = - \left( \int_2 q^l(w_0, \lambda) \alpha q_1(w_0)^{\alpha-1} dF + \lambda \alpha \frac{w_0^{1-\alpha}}{\lambda \alpha} q_1(w_0)^{\alpha-1} \right) \frac{\partial q_1^l}{\partial \lambda} - \frac{\partial q_1^l}{\partial \lambda},
\]

which implies that the last two terms on the left cancel out. Hence we obtain:

\[
\frac{\partial \theta}{\partial \lambda} = - \int_2 q^l(w_0, \lambda) \alpha q_1(w_0)^{\alpha-1} dF < 0.
\]

Therefore \(\theta(w_0, \lambda)\) is a decreasing function of the loss aversion parameter \(\lambda\). \(\square\)

**Proof of Proposition 5.** We divide the proof in two steps: first we provide a lemma that establishes the properties of the function \(\theta\) with respect to the firm’s initial wage contract and the firm’s reservation productivity; then we use these properties to prove the statement of Proposition 5.

Thus, consider the firm’s optimal wage contract at time 0:

\[
\hat{w}_0(r_0, q_0, q_1^e, \lambda)^* \equiv \left\{ w_0 : \alpha q_0 w_0^{\alpha-1} + \delta \theta(w_0, \lambda) = 1 \right\},
\]

56
where the left hand side is denoted with $T_0^+ \equiv \alpha q_0 w_0^{\alpha-1} + \delta \theta(w_0, \lambda)$. Then, consider the firm’s reservation productivity:

$$q^u_0(r_0, q^e_1, \lambda) \equiv \{ q_0 : \alpha q_0 r_0^{\alpha-1} + \delta \theta(r_0, \lambda) = 1 \}$$

$$\equiv \frac{(1 - \delta \theta(r_0, \lambda)) r_0^{1-\alpha}}{\alpha}.$$  

Lemma 2. The margins denoted by $T_0^+$ are an increasing function of $\theta$, that is, $\frac{\partial T_0^+}{\partial \theta} > 0$. Hence for any $\theta_i < \theta_j$ we have $T_0^+(\theta_i) < T_0^+(\theta_j)$. Moreover the firm’s reservation productivity $q^u_0$ is a decreasing function of $\theta$, that is, $\frac{\partial q^u_0}{\partial \theta} < 0$. Hence for any $\theta_i < \theta_j$ we have $q^u_0(\theta_i) > q^u_0(\theta_j)$.

Proof of Lemma 2. Differentiating $T_0^+$ with respect to $\theta$ yields:

$$\frac{\partial T_0^+}{\partial \theta} = \delta > 0.$$  

Differentiating $q_0^u$ with respect to $\theta$ yields:

$$\frac{\partial q_0^u}{\partial \theta} = \frac{-\delta r_0^{1-\alpha}}{\alpha} < 0.$$  

Now consider two unemployed workers $i$ and $j$ with loss aversion parameters $\lambda_i > \lambda_j$ as it is the case of Proposition 5. From Corollary 9 we infer that $\theta_i(w_0, \lambda_i) < \theta_j(w_0, \lambda_j)$, which implies that $T_0^+(\theta_i) < T_0^+(\theta_j)$ as established by Lemma 2. By plotting the margins that characterize the optimal initial contracts for workers $i$ and $j$ we infer that $\hat{w}_{0,i}(r_0, q_0, q^e_1, \lambda_i) < \hat{w}_{0,j}(r_0, q_0, q^e_1, \lambda_j)$ for all $q_t \in [\underline{q}, \overline{q}]$.

Moreover, since $\theta_i(w_0, \lambda_i) < \theta_j(w_0, \lambda_j)$, by Lemma 2 it is straightforward to see that $q^u_0(r_0, q^e_1, \lambda_i) > q^u_0(r_0, q^e_1, \lambda_j)$ for all $q_t \in [\underline{q}, \overline{q}].$  

\[57\]
Proof of Proposition 6. By assumption D1, $F_a \leq F_b$ and $f_a(q_1^t) > f_b(q_1^t)$, which implies that:

$$\int_q^{q_1^t} J'_1(w_0, q_1) dF_a < \int_q^{q_1^t} J'_1(w_0, q_1) dF_b,$$

since

$$\int_q^{q_1^t} q_1 dF_a < \int_q^{q_1^t} q_1 dF_b.$$

Hence, by the definition of $\theta \equiv \int_q^{q_1^t} J'_1(w_0, q_1) dF$ we infer that $\theta(F_a) < \theta(F_b)$. Finally, by appealing to Lemma 2 and for similar arguments used in the the proof of Proposition 5 we infer that $\tilde{w}_0^c(r_0, q_0, q_1^c, \lambda) < \tilde{w}_a^c(r_0, q_0, q_1^c, \lambda)$ and $q_0^{u,b}(r_0, q_1^c, \lambda) > q_0^{u,a}(r_0, q_1^c, \lambda)$ for all $q_1 \in [q, \overline{q}]$. \qed