DOWNWARD NOMINAL WAGE RIGIDITY, MONEY ILLUSION, AND IRREVERSIBILITY

Michael William Leamington Elsby

London School of Economics

Doctor of Philosophy in Economics

March 2005

Abstract

This thesis seeks to make three related contributions to our understanding of the causes and implications of downward nominal wage rigidity, the nature of money illusion on behalf of workers, and the theoretical treatment of irreversibility in factor demand and wage setting.

Chapter 1 seeks to contribute to the literature on downward nominal wage rigidity (DNWR) along two dimensions. First, I formulate and solve an explicit model of wage-setting in the presence of worker resistance to nominal wage cuts – something that has previously been considered intractable. In particular, I show that this resistance renders wage increases (partially) irreversible. Second, using this model, one can explain why previous estimates of the macroeconomic effects of DNWR have been so weak despite remarkably robust microeconomic evidence. In particular, one can show that previous studies have neglected the possibility that DNWR can lead to a compression of wage increases as well as decreases. Thus, the literature may have been overstating the costs of DNWR to firms. Using micro-data for the US and Great Britain, I find robust evidence in support of the predictions of the model. In the light of this evidence, Chapter 1 concludes that increased wage pressure due to DNWR may not be as large as previously envisaged, but that the behavioural implications of DNWR in respect of the reaction of workers to nominal wage cuts remain significant.

Chapter 2 then contrasts the implications of two proposed models of downward nominal wage rigidity — those based on the form of market contracts (MacLeod & Malcomson [1993]; Holden [1994]), and that based on money illusion explored in Chapter 1. In particular, I identify a method of distinguishing between these two foundations empirically, by observing how the distribution of wage changes varies with the rate of inflation. I find evidence that at least part of the observed rigidity cannot be easily explained by contract models, but can be explained in the context

of a model with money illusion.

Finally, Chapter 3 extends some of the theoretical developments of Chapter 1 with respect to models of irreversibility. In particular, Chapter 3 presents analytical results for models of dynamic factor demand in the presence of irreversibility in discrete time. It builds on previous work on irreversibility in the investment (Dixit & Pindyck [1994]) and labour demand (Bentolila & Bertola [1990]) literatures which use a continuous time, Brownian framework. I show that, whilst there are parallels between the discrete time models and their continuous time counterparts, the analytics in discrete time allow a more general treatment, principally by allowing the relaxation of the assumption of shocks following a unit root. I then explore the effects of relaxing this assumption on optimal factor demand.

Acknowledgements

I owe a great debt of gratitude to my supervisor, Alan Manning, for his perceptive comments, continued patience and encouragement, and for all the really excellent advice he gave me throughout my doctoral studies.

In addition I would like to thank Andrew Abel, Marianne Bertrand, Juan Dolado, Maarten Goos, Steinar Holden, Richard Layard, Jim Malcomson, Sendhil Mullainathan, Barbara Petrongolo, Steve Pischke, Danny Quah, Matthew Rabin, and participants at many seminars for valuable comments.

For help with accessing the data, I would like to thank Tanvi Desai.

Finally, I would like to thank my wife, Elizabeth, for convincing me, to her own cost, that I should give up my job and pursue graduate work in economics, and for her perseverance and support while I completed my studies.

Any errors are my own.

Contents

List of Figures

List of Tables

0.1 Introduction to Thesis

This thesis fuses insights from three distinct literatures: the empirical observation of downward rigidity in nominal wages, the psychological phenomenon of money illusion, and the theoretical treatment of irreversibility.

In particular, Chapter 1 formulates and solves an explicit model of wage-setting informed by recent survey-based evidence that workers resist nominal wage cuts (Bewley [1999], Shafir, Diamond, & Tversky [1997]). In particular, using this model we demonstrate that this resistance renders wage increases (partially) irreversible. In this way we establish the first link between the above literatures: that between money illusion in the form of a resistance to nominal loss, and irreversibility.

We then proceed to show that by using this model we can explain many of the properties of the observed downward rigidity of nominal wages found in developed economies (see Kramarz [2001]). Specifically, we obtain a novel finding that has not been addressed in the previous empirical literature on downward nominal wage rigidity (DNWR): that firms will optimally restrict nominal wage increases as a response to workers' resistance to wage cuts. Using micro-level data from the US and Great Britain, we find strong evidence that this is indeed the case.

This is a useful finding on a number of dimensions. First, it is a non-trivial implication of worker resistance to wage cuts that is testable using conventional data. This is useful because the previous literature providing evidence for aversion to nominal loss has typically just asked people about their reactions to wage cuts (Bewley [1999], Shafir et al. [1997]). Economists have instinctively avoided such data on the basis that there may be vast differences between what people say and what they do (Bertrand & Mullainathan [2001]). By providing additional evidence from more conventional sources, we obtain a more robust insight into the nature of wage setting, and more fundamentally, the form of workers' preferences.

Second, using this model, one can reconcile a tension in the empirical literature on downward nominal wage rigidity between the micro- and macro-level estimates. In particular, estimates of the macroeconomic effects of DNWR have been remarkably weak despite robust microeconomic evidence. We show that, by neglecting the possibility that downward nominal wage rigidity leads to the compression of wage increases as well as wage cuts, previous studies may have been overstating the costs of DNWR to firms. Our data from the US and Great Britain confirm that this is indeed the case: the estimated reduction in wage growth due to restricted wage increases offsets most of the estimated increase in wage growth due to restricted wage cuts in the data.

In the light of this evidence, Chapter 1 concludes that increased wage pressure due to DNWR may not be as large as previously envisaged, but that the behavioural implications of DNWR in respect of the reaction of workers to nominal wage cuts remain significant.

Chapter 2 then goes on to look more closely at the relationship between downward nominal wage rigidity and money illusion. In particular, it contrasts the implications of the model formulated in Chapter 1 with an alternative set of theories for DNWR based on the form of market contracts (MacLeod & Malcomson [1993]; Holden [1994]). Specifically, we identify a method of distinguishing between these two foundations empirically, by observing how the distribution of wage changes varies with the rate of inflation. We show that the model based on worker resistance to wage cuts of Chapter 1 predicts that the distribution of wage increases should vary with the rate of inflation; contract models, on the other hand, do not.

The intuition for this result is due to the fact that contract models are based on the existence of *real* frictions that render it costly for workers (firms) to switch firms (workers). Since in many countries there must be mutual consent of firm and worker to wage changes, the default contract in the event of disagreement is no change in wages. In particular, wages in these models are renegotiated up (down) only when the worker's (firm's) outside option (or equivalently strike (lockout) threat; Holden [1994]) becomes preferable. Thus, the existence of frictions drives a wedge between the firm's and the worker's outside option, and thus the nominal wage will remain constant for intervals of time. In this way such frictions can yield some fixity in wages. However, since wage rigidity in contract models is driven by *real* frictions, they cannot explain any changes in the compression of wage increases that are related to the rate of inflation. The model of money illusion in Chapter 1, on the other hand, predicts that wage increases will become compressed when inflation is low. This is because low inflation implies that the only way firms can reduce real labour costs in the future is by cutting the nominal wage, which is costly. Therefore, firms will restrict wage increases as a precaution against such future costs as inflation falls in the model of DNWR based on money illusion.

Taking these predictions to our micro-data for the US and Great Britain, we find evidence that firms actively compress wage increases when inflation is low. Thus, we conclude that at least part of the observed rigidity in nominal wages cannot be easily explained by contract models, but can be explained straightforwardly in the context of a model with money illusion. This reinforces the conclusion that downward nominal wage rigidity is, at least in part, symptomatic of a particular aversion to nominal wage cuts on behalf of workers.

Finally, Chapter 3 extends some of the theoretical developments of Chapter 1 with respect to models of irreversibility. In particular, Chapter 3 presents analytical results for models of dynamic factor demand in the presence of irreversibility in discrete time.

It builds on previous work on irreversibility in the investment (Dixit & Pindyck [1994]) and labour demand (Bentolila & Bertola [1990]) literatures which use a continuous time, Brownian framework. We show that, whilst there are parallels between the discrete time models and their continuous time counterparts, the analytics in discrete time allow a more general treatment, principally by allowing the relaxation of the assumption of shocks following a unit root.

Chapter 3 then explores the effects of relaxing this assumption on optimal factor demand. We first derive a general result that increased persistence of shocks leads to a greater response of factor demand to current shocks. The intuition for this is that, when shocks are more persistent, current shocks become more informative about future shocks. Thus, the firm has to worry less about any costly reversals of current factor demand decisions. Furthermore, we additionally find that reduced persistence in the form of i.i.d. shocks leads to greater rigidity in factor demand. Intuitively, i.i.d. shocks imply some mean-reverting aspect to shocks, which leads firms to worry more about having to reverse current factor demand decisions, at a cost, in the future. Finally, we examine the impact of distributional form by solving the model in the presence of exponential shocks. We find that the fatter tails implied by exponential shocks leads the firm to reduce demand more for bad shocks, and to increase demand more for good shocks.

Together, then, these chapters aim to enhance our understanding of money illusion and irreversibility both as distinct issues, but particularly through the lens of downward nominal wage rigidity in labour markets.

Chapter 1

Evaluating the Economic Significance of Downward Nominal Wage Rigidity

1.1 Introduction

The existence of rigidities in nominal wages (and prices) is a cornerstone of macroeconomic theory. Such rigidities act as the key theoretical motivation for the existence of a trade-off between inflation and unemployment in the form of the Phillips curve, and are thus of critical importance to the conduct and efficacy of macroeconomic policy.

A recent flurry of empirically oriented research has used micro-data to address the question of whether such nominal rigidity exists. In particular, this research details some striking characteristics of the distribution of nominal wage changes at the individual level. These include the existence of a mass point at zero nominal wage change and an asymmetry in the form of a deficit of nominal wage cuts, which are taken together as evidence for downward nominal wage rigidity (henceforth DNWR).

Such evidence has been found in numerous datasets spanning a vast number of developed economies (for a survey see Kramarz, 2001)[1]. However, a number of issues remain unresolved in the light of this research.

An important question relates to evidence for the expected macroeconomic effects of DNWR. In particular, a number of studies have shown that the above results predict the existence of a *convex, long run* Phillips curve (Akerlof, Dickens & Perry, 1996). Intuitively, low inflation implies that reductions in real labour costs can only be effected through nominal wage cuts. If firms are prevented from cutting nominal wages, then their only recourse is to layoff workers, leading to increased unemployment. Thus, when inflation is low, increased inflation can relax the constraint of DNWR on wage-setting for a significant fraction of firms, and thereby reduce unemployment. This result has been of particular interest in recent years due to the adoption of inflation targeting by many monetary authorities. In particular, the existence of a long-run Phillips curve implies that implementing a low inflation target could result in a persistent increase in unemployment.

Much of the research on DNWR addresses precisely this issue. A typical reference is the analysis of Card & Hyslop (1997) for the US. Their micro-level analysis finds strong evidence that nominal wage cuts are restricted when inflation is low, and they conclude that the existence of DNWR leads to an increase in average real wage growth of up to 1% per annum. Card & Hyslop then assess whether the predictions of this micro-level evidence are corroborated by evidence at a higher level of aggregation. In contrast to their micro-level results, Card & Hyslop's state-level results are much

[1]To get an impression of the sheer ubiquity of these observations, a non-exhaustive list of such studies is as follows. For the US: McLaughlin (1994), Kahn (1997), Card & Hyslop (1997), Altonji & Devereux (1999), Groshen & Schweitzer (1999), Lebow, Saks & Wilson (1999). For Canada: Farès & Hogan (2000), Christofides & Leung (2003). For Europe: Smith (2000), Nickell & Quintini (2003), Fehr & Goette (2003), Beissinger & Knoppik (2003), Decressin & Decressin (2003), Dessy (2002). For Australia: Dwyer & Leong (2003).

weaker. In particular, whilst they find some evidence for the existence of a Phillips curve trade-off, they obtain estimates that are too imprecise to conclude that this trade-off is stronger in periods of low inflation[2]. Moreover, informal observation of the recent incidence of low inflation together with low unemployment in the US and UK confirms the weakness of this prediction at the most basic level. Thus, there exists a puzzle: if the micro-level evidence for DNWR is so robust, why is the analogous macro-level evidence so fragile?

We argue that we can make progress in resolving this issue via a more careful consideration of the theoretical underpinnings of DNWR. In particular, we present a model of DNWR informed by recent evidence that wage-setters and negotiators are reluctant to cut the nominal wages of workers (see Bewley, 1995, 1998, 1999 and the survey in Howitt, 2002). In particular, by interviewing over 300 managers, pay professionals, labour leaders etc., Bewley finds that the most common explanation provided for this reluctance is the belief that nominal wage cuts damage worker morale. Moreover, there is additional evidence that agents are subject to money illusion (Shafir, Diamond & Tversky, 1997). In particular, these studies show that agents in different economic settings exhibit significant aversion to nominal losses – what we will term nominal loss aversion. A typical finding is that respondents believe it much more acceptable to receive a 5% nominal wage increase when inflation is 12%, than a 7% wage cut when there is no inflation (Kahneman, Knetsch & Thaler, 1986). This is corroborated by Genesove & Mayer (2001) who find evidence from real-estate data that condominium owners were reluctant to sell at a price below that they originally paid, even though they were typically moving locally, and hence were buying in the same market. Thus, nominal loss aversion applied to wage cuts can

[2] Weak macroeconomic effects have also been found by Lebow, Saks & Wilson (1999) for the US, and by Nickell & Quintini (2003) for the UK. Indeed, Lebow, Saks & Wilson coined the term "micro-macro puzzle" for the observed tension between micro- and macro-level estimates.

provide a key to explaining the existence of DNWR.

The need for an explicit model of wage-setting in the presence of worker resistance to wage cuts has been noted in the previous literature on money illusion, as well as by labour economists studying the distribution of wage changes:

> "Plausibly, the relationship [between wages and effort] is not continuous: there is a discontinuity coming from nominal wage cuts.... A central issue is how to model such a discontinuity." Shafir, Diamond & Tversky (1997), p.371.
>
> "[I]t is surprising to us that there is no rigorous treatment in the literature of how forward looking firms should set wages when it is costly to cut nominal wages." Altonji & Devereux (2000), p.423 note 7.

We address both these issues and show that a key insight into the implications of these behavioural models is that nominal wage increases become partially *irreversible*. In particular, consider a firm that raises the wage today, but reverses the wage increase by cutting the wage by an equal amount tomorrow. When workers resist wage cuts, the net effect on productivity will be negative: today's wage increase will raise productivity, but tomorrow's wage cut will reduce productivity by a greater amount. Thus, reversals of wage increases are costly to firms. In this sense we can think of there being an asymmetric adjustment cost to changing nominal wages.

Models of asymmetric adjustment costs have been widely studied in the investment (Dixit & Pindyck, 1994) and labour demand (Bentolila & Bertola, 1990) literatures, typically in the form of continuous time models with shocks following Brownian motions. In contrast, we formulate and solve our model of partial irreversibility in discrete time[3]. This is done for a number of reasons. First, since data are reported in discrete intervals, this method allows us to align theoretical and empirical concepts more naturally. Moreover, many wage contracts are renegotiated on an annual basis, which is more consistent with a discrete-time setup. Finally, when considering worker resistance to wage cuts, the time horizon over which workers evaluate a wage cut becomes important[4]. Plausibly, workers do not evaluate wage changes continually, but rather at discrete intervals, which again lends itself more to a discrete-time model. Whilst modelling these features in discrete time has to date been considered significantly less tractable than corresponding Brownian models, we develop a comparatively tractable solution method[5].

The solution to this "behavioural" model equips us with a number of predictions that can potentially reconcile the two strands of evidence mentioned above. We show that a key limitation in the previous empirical literature is that it assumes (implicitly or otherwise) that the existence of DNWR has no effect on the upper tail of the wage change distribution. In particular, this is a key identifying assumption in Card & Hyslop (1997), which leads them to use the observed upper tail of the distribution of wage changes to infer the properties of the lower tail in the absence of DNWR. The predictions of our model show that this may be misguided. In particular, the upper tail of wage changes will be compressed for two related reasons. First, we show that firms may actively reduce the nominal wage paid when they increase the

[3]Partial irreversibility of investment decisions has been studied in a continuous time Brownian framework by Abel & Eberly (1996).

[4]This point has been made by Benartzi & Thaler (1995) in the context of loss aversion over asset returns.

[5]An additional benefit of a discrete-time solution, pursued in Chapter 3, is that it allows one to relax assumptions on the distribution of shocks. In particular, it can be shown that a Brownian motion is the continuous-time analogue to a Gaussian random walk (see Dixit, 1993). A discrete-time framework allows one to use non-Gaussian shocks, as well as more generalised persistence assumptions in a more comfortable way.

wage relative to a "counterfactual" world without DNWR – what we will term "*active compression*". In the behavioural model, this results because raising the wage today increases the likelihood of having to cut the wage, at a cost, in the future. Second, the model shows that, even if firms do not actively compress nominal wage increases, the upper tail of the wage change distribution will still be compressed relative to the counterfactual with no DNWR. This is because DNWR raises the general level of wages in the economy, and thus firms do not have to raise wages as often or as much to obtain their desired wage level. In particular, we show that this process occurs as a result of a steady state requirement that average wages and productivity grow at the same constant rate in the long run. We refer to this process as "*latent compression*". Thus, by neglecting these effects, previous studies have potentially *overstated* the increase in wage growth due to DNWR. In this way, we can potentially reconcile the micro- and macro-level evidence on DNWR found in the previous empirical literature.

In the light of this, we seek evidence for these predictions using micro-data for the US and Great Britain. We find significant evidence that the upper tail of the wage change distribution exhibits a compression of wage increases that is related to DNWR. In particular, we find that this limits the estimated increase in real wage growth due to DNWR from around 1–1.5% to no more than 0.3%. We show that this is because firms can "save" at least 75% of the increase in wage growth due to restricted wage cuts by reducing nominal wage increases. This might go some way to explaining why the aggregate effects of DNWR are often found to be modest.

As an additional test of the implications of the model of DNWR presented, we show that the model also implies that increased rates of turnover should mitigate the necessity for firms to restrict wage increases. This occurs because higher turnover reduces the probability that a given worker will stay in the firm an additional period, and thus renders the firm more myopic. Thus firms do not need to compress wage increases as a precaution against future costly wage cuts to the same extent. We again find robust evidence for this hypothesis using the NESPD data, and to a lesser extent for the PSID also. This reinforces the claim that a model of DNWR based on worker resistance to nominal wage cuts is a useful way of understanding the empirical properties of wage setting.

In the light of this evidence, we conclude that the macro effects of DNWR may not be as large as previously envisaged, and thus may not provide such a strong argument against the adoption of a low inflation target. However, the behavioural implications of DNWR in respect of the reaction of workers to nominal wage cuts remain significant and we conclude that it does imply something fundamental about the nature of human behaviour.

The rest of this chapter is organised as follows. Section 2 presents an explicit behavioural model of wage-setting in the presence of worker resistance to nominal wage cuts; section 3 fleshes out some of the predictions of these models that we can take to the data; section 4 presents our empirical methodology and the results obtained; section 5 discusses some remaining issues for future work; and section 6 concludes. Where possible, we omit technical details from the main text, and relegate them to the appendix.

1.2 A Behavioural Model of DNWR

In this section we present an explicit model of downward nominal wage rigidity based on the observations detailed in the empirical literatures mentioned above. In particular, we study the optimal nominal wage policies of worker-firm pairs for whom the productivity of the worker (denoted e) depends upon the wage as follows:

$$e = \ln\left(\frac{\omega}{b}\right) + c\ln\left(\frac{W}{W_{-1}}\right)\mathbf{1}^{-} \tag{1.1}$$

where W is the nominal wage, W_{-1} the lagged nominal wage, $\mathbf{1}^{-}$ an indicator for a nominal wage cut, $\omega \equiv W/P$ the real wage, and b a measure of real unemployment benefits (which we assume to be constant over time). The parameter $c > 0$ varies the productivity cost to the firm of a nominal wage cut.

The motivation for this effort function is as follows. We assume that worker effort depends positively on the difference between the level of the real wage, ω, and real unemployment benefits, b. This captures the idea that, the higher the worker's real standard of living from being in work relative to unemployment, the harder that worker will work. In addition, we model the productivity loss due to nominal wage cuts by assuming that effort is falling in the *geometric* nominal wage cut. Our reasoning for this is that the most obvious alternative – that it is the absolute value of the cut in the nominal wage that reduces effort – is implausible in the following sense. It implies that a wage cut of a cent will cause the same loss in effort whether last period's nominal wage is \$1 or \$1,000,000. This is clearly extreme, so we employ the more sensible concept that it is the *percentage* cut in the nominal wage that affects effort.

The qualitative features of this effort function are illustrated in Figure 1-1. Clearly, there is a kink at $W = W_{-1}$ reflecting the existence of DNWR. In particular, the marginal productivity loss of a nominal wage cut exceeds the marginal productivity gain of a nominal wage increase:

$$\frac{\partial e/\partial W|_{W\uparrow W_{-1}}}{\partial e/\partial W|_{W\downarrow W_{-1}}} = 1 + c > 1 \tag{1.2}$$

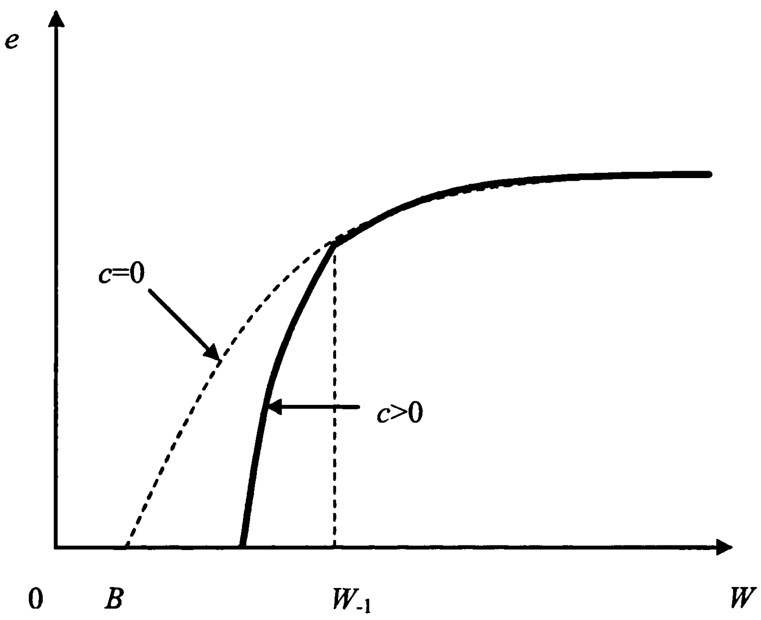


Figure 1-1: The Effort Function

This characteristic is what makes nominal wage increases (partially) irreversible – a nominal wage increase can only be reversed at an additional marginal cost of c. Clearly, the parameter c is what drives this feature of the model.

The effort function, (1.1), can be interpreted as a very simple way of capturing the basic essence of the motivations for DNWR mentioned in the literature. It is essentially a parametric form of effort functions in the spirit of the fair-wage effort hypothesis expounded by Solow (1979) and Akerlof & Yellen (1988), with an additional term reflecting the impact of nominal wage cuts on effort – as envisaged in the quote from Shafir, Diamond & Tversky (1997) in the introduction. Bewley (1999) also advocates such a characterisation[6]:

[6] However, such is the intricacy of Bewley's study, he would probably consider (1.1) a simplifica-

"The only one of the many theories of wage rigidity that seems reasonable is the morale theory of Solow..." Bewley (1999), p.423.

"The [Solow] theory...errs to the extent that it attaches importance to wage levels rather than to the negative impact of wage cuts." Bewley (1999), p.415.

In addition, an effort function with these properties can be derived from a compensating differentials model where worker utility exhibits nominal loss aversion[7]. The basic intuition for this is that, if workers dislike nominal losses and the firm wishes to cut the nominal wage, then the firm must compensate the worker in the form of lower on-the-job effort in order to prevent the worker from quitting. Thus, in this sense, (1.1) can be considered a reduced form of a model in which workers dislike nominal loss. The goal of this chapter is not to highlight the nuances of emphasis – which do indeed exist – between these behavioural foundations, but rather to show that they share a common, theoretically important, qualitative implication as to the nature of a firm's wage-setting choice. This is intended as a *start* towards richer models of these phenomena, and to this end aims to unify rather than to differentiate. We discuss the implications of alternative functional forms for the effort function in section 5.

The most comparable previous attempt at explicitly modelling the behavioural foundation to DNWR is that of Akerlof, Dickens & Perry (1996). However, Akerlof et al. present a model in which firms have no operational discretion over wage-setting – wages are given by a wage-setting relationship which firms take as exogenous, and which dictates that nominal wages can never fall. Thus, the implicit assumption in

tion, not least for its neglect of emphasis on morale as distinct from productivity, and of the internal wage structure of firms as a source of wage rigidity. We argue that it is a *useful* simplification as it provides key qualitative insights into the implied dynamics of wage-setting under more nuanced theories of morale.

[7] See the appendix for an explicit model of compensating differentials that yields the effort function (1.1) as its solution.

their model is that firms do not cut wages because, if they did, all of their workers would quit. In this way, their model effectively short-circuits any endogenous reaction on behalf of firms to the existence of DNWR. The model presented in this chapter differs critically in that firms do have a non-trivial wage-setting decision: firms can cut nominal wages if they wish, but it will have a strong adverse effect on productivity at the margin. We argue that this is a more desirable setup. In the first instance, it accords better with the evidence that firms restrict wage cuts due to concerns over *morale within the firm*, rather than because the external labour market dictates it (Bewley, 1999). Secondly, it also accords well with micro-data evidence that nominal wage cuts *do* occur, just with a lower frequency than might be expected[8].

The Wage Setting Problem

We consider a discrete-time, infinite-horizon model in which price-taking worker-firm pairs choose the nominal wage W_t at each date t to maximise the expected discounted value of profits. For simplicity, we assume that each worker-firm's production function is given by $a \cdot e$, where a is an *idiosyncratic* real technology shock which is observed contemporaneously, and acts as the source of uncertainty in the model. Thus, defining $\beta \in [0, 1)$ as the discount factor of the firm, the typical firm's decision problem is given by:

$$\begin{aligned} \max_{\{W_t\}} \quad & E_t \left[\sum_{s=t}^{\infty} \beta^{s-t} \left\{ a_s e_s - \omega_s \right\} \right] \\ \text{where} \quad & e_s = \ln\left(\frac{\omega_s}{b}\right) + c \ln\left(\frac{W_s}{W_{s-1}}\right) \mathbf{1}_s^- \end{aligned} \tag{1.3}$$

It turns out in what follows that it is convenient to re-express the firm's profit stream

[8] This is seen in the vast majority of micro-data studies. Moreover, even in samples *without* measurement error in which we might expect to see fewer wage cuts, one observes, if anything, more frequent nominal wage cuts (see Smith, 2000 and Nickell & Quintini, 2003).

in constant date t prices. To this end, we multiply through by P_t, which we define as the competitive price level at date t, and assume that it evolves according to $P_t = (1+\pi) P_{t-1}$, where π is the rate of inflation. Finally, defining the nominal counterparts, $A_t \equiv P_t a_t$ and $B_t \equiv P_t b$ and substituting for e_t, we obtain the following optimisation problem for the firm:

$$\max_{\{W_t\}} E_t \left[\sum_{s=t}^{\infty} \left(\frac{\beta}{1+\pi} \right)^{s-t} \left\{ A_s \left[\ln\left(\frac{W_s}{B_s} \right) + c \ln\left(\frac{W_s}{W_{s-1}} \right) \mathbf{1}_s^- \right] - W_s \right\} \right] \tag{1.4}$$

We assume that the nominal shock has support $[0, \infty)$ and that its evolution can be described by the cumulative density function $F(A'|A)$. Thus, rewriting the problem in recursive form[9] we have[10]:

$$v(W_{-1}, A) = \max_W \left\{ A \left[\ln\left(\frac{W}{B} \right) + c \ln\left(\frac{W}{W_{-1}} \right) \mathbf{1}^- \right] - W + \frac{\beta}{1+\pi} \int v(W, A') \, dF(A'|A) \right\} \tag{1.5}$$

(1.5) is the basic problem that we will attempt to solve in what follows[11]. Before we begin, though, we first present an intuitive outline of the type of results we might expect.

[9]We adopt the convention of denoting lagged values by a subscript, $_{-1}$, and forward values by a prime, $'$.

[10]In addition, we make the standard assumption that the measure $dF(A'|A)$ satisfies the Feller property, so that the mapping defined by (1.5) preserves continuity of the value function. A sufficient condition for this is that A is governed by the stochastic difference equation, $A' = g(A, \varepsilon')$, where g is a continuous function and ε' is an i.i.d. innovation (see Stokey & Lucas, 1989, pp.237, 261–262). We maintain this assumption throughout the paper.

[11]There is an issue that, for sufficiently low values of the wage, effort is potentially negative. However, accounting for such a non-negativity constraint significantly complicates the solution to the model without much gain in relevance. We maintain the assumption that the level of benefits is sufficiently low relative to wages as to allow almost all firms to ignore this constraint.

1.2.1 Some Intuition for the Behavioural Model

As the theory presented in the forthcoming sections can seem analytically complicated, in this section we present the economic intuition for each of the predictions of the model, which we deal with in turn.

First, the model predicts that there will be a spike at zero in the distribution of nominal wage changes across firms. This occurs because of the kink in the objective function at $W = W_{-1}$. In particular, this implies that for each firm there will be a range of values ("region of inaction") for the nominal shock, A, for which it is optimal not to change the nominal wage. Since A is distributed across firms, there will thus exist a positive fraction of firms each period whose realisation of A lies in their region of inaction that will in turn not change their nominal wage.

Second, in the event that a firm does decide to change the nominal wage, the wage change will be actively compressed relative to the case where there is no DNWR. That nominal wage cuts are attenuated is straightforward to explain – as wage cuts involve a discontinuous fall in productivity at the margin, the firm will be less willing to effect them. In particular, some small wage cuts that would have been implemented in the absence of DNWR will instead be implemented as wage freezes. Moreover, larger counterfactual wage cuts will be reduced in magnitude. It is slightly less obvious why nominal wage increases are also attenuated in this way. The reason is that, in an uncertain world, increasing the wage today increases the likelihood that you will have to cut the wage, at a cost, in the future.

A direct implication of this last prediction is that increases in the productivity cost of cutting the nominal wage, c, will accentuate all these effects. That is, a higher productivity cost due to nominal wage cuts will widen the region of inaction, thereby increasing the mass point at zero in the distribution of nominal wage changes, and will also render the active compression of nominal wage changes more acute.

The final prediction we want to emphasise at this stage is the effect of increased inflation on nominal wage increases. In particular, we find that the active compression of nominal wage increases becomes less pronounced as inflation rises. As explained earlier, this is because the only reason firms restrict wage increases in the model is the prospect of costly wage cuts in the future. Since higher inflation reduces the probability of this occurring, firms no longer need to worry as much about increasing the nominal wage.

1.2.2 Some Special Cases

In order to get a feeling for how the model works, we first solve some models that are special cases of the full dynamic model. In particular, we consider two cases: where nominal wage increases are fully reversible ($c = 0$), and the case where nominal increases are partially irreversible ($c > 0$), but where firms are myopic ($\beta = 0$). We will see that these will inform our approach to solving the full dynamic model given in (1.5).

The Case where $c = 0$

Note that the assumption that $c = 0$ removes any dynamic considerations from the firm's wage-setting choice by removing the dependence of effort on last period's wage. Thus the firm's problem is simply:

$$\max_{W} \left\{ A \ln\left(\frac{W}{B}\right) - W \right\} \tag{1.6}$$

The first-order condition for this problem is:

$$\frac{A}{W} - 1 = 0 \qquad (1.7)$$
$$\implies W = A, \quad \& \quad \Delta W = \Delta A$$

In this case, the distribution of nominal wage changes across firms will be exactly the same as the distribution of changes in the nominal shock. We term this result the *counterfactual solution.*

The Case where $\beta = 0$

In this case the firm's problem is given by:

$$\max_W \left\{ A \left[\ln\left(\frac{W}{B}\right) + c \ln\left(\frac{W}{W_{-1}}\right) \mathbf{1}^- \right] - W \right\} \qquad (1.8)$$

Here, the objective function is kinked at $W = W_{-1}$, and thus the derivative with respect to W is not well-defined at this point. To determine the optimal nominal wage policy, we take the first-order condition with respect to W, conditional on $\Delta W \neq 0$. Due to the concavity of the problem, this will determine when it is optimal to change the nominal wage (up or down) and by how much. It is thus trivial that whenever it is not optimal to change W, it is left unchanged. Thus, we have:

$$A \left[\frac{1}{W} + \frac{c}{W} \mathbf{1}^- \right] - 1 = 0, \quad \text{if } \Delta W \neq 0 \qquad (1.9)$$

which implies:

$$A = \begin{cases} W & \text{if} \quad \Delta W > 0 \\ \frac{W}{1+c} & \text{if} \quad \Delta W < 0 \end{cases} \qquad (1.10)$$

Moreover, from the concavity of the firm's objective, (1.8), it follows that:

$$\begin{aligned}
&\Delta W > 0 \quad \text{if} \quad A > A_u^S \\
&\Delta W < 0 \quad \text{if} \quad A < A_l^S \\
&\Delta W = 0 \qquad \text{otherwise}
\end{aligned} \tag{1.11}$$

where $A_u^S \geq A_l^S$ are trigger values for the firm's nominal shock that cause the firm to respectively raise or cut the nominal wage. Finally, we can relate (1.10) and (1.11) by noting that, due to the continuity and concavity of (1.8), the firm's optimal wage policy will be a continuous function of (A, W_{-1})[12]. Thus, it must be that:

$$\begin{aligned}
A_u^S &= W_{-1} \\
A_l^S &= \frac{W_{-1}}{1+c}
\end{aligned} \tag{1.12}$$

Thus, the solution implies a region of inaction for W at $W = W_{-1}$, and a trigger policy as follows:

$$\begin{array}{llll}
\text{If} & A > W_{-1} \equiv A_u^S, & \Delta W > 0 \text{ until} & W = A \\
\text{If} & A < \frac{W_{-1}}{1+c} \equiv A_l^S, & \Delta W < 0 \text{ until} & W = (1+c)\,A \\
\text{If} & A \in \left[A_l^S, A_u^S\right], & \Delta W = 0 \quad \text{or} & W = W_{-1}
\end{array} \tag{1.13}$$

To see this more clearly, consider Figure 1-2. The optimal policy function for the nominal wage in the non-rigid case ($c = 0$) is simply illustrated by the 45° line, whereas in this case, where $c > 0$ and $\beta = 0$, the bold line represents the optimal wage policy. By comparing this wage policy to the case where $c = 0$, we can see that the firm is taking counterfactual nominal wage cuts in the interval $\left[\frac{W_{-1}}{1+c}, W_{-1}\right]$ and is instead implementing them as wage freezes. Moreover, for all counterfactual wages

[12]This follows from the Theorem of the Maximum (see e.g. Stokey & Lucas, 1989, pp.62–63).

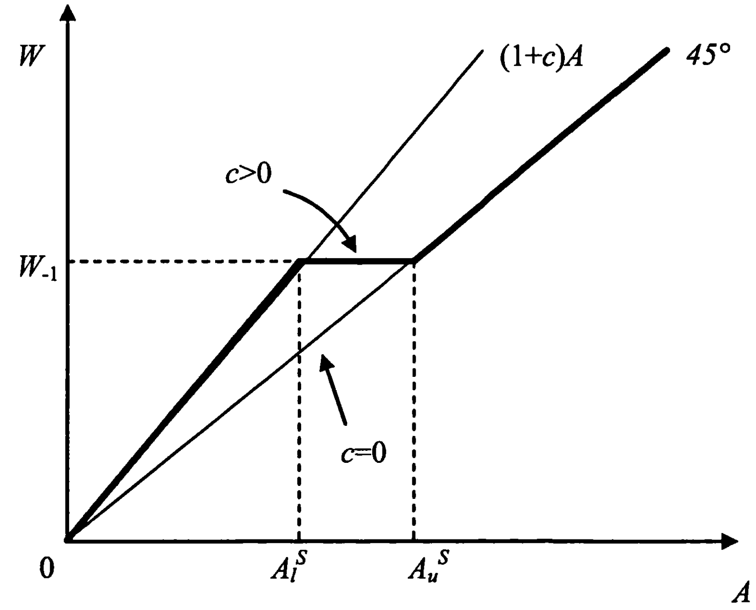


Figure 1-2: The Optimal Wage Policy in the Static Models

below $\frac{W_{-1}}{1+c}$, the firm is reducing the magnitude of wage cuts by a factor $\frac{1}{1+c}$. Thus nominal wage cuts are being actively compressed as a result of DNWR.

However, the same is *not* true for nominal wage increases. All counterfactual wage increases are being implemented without alteration. The reason for this is that $\beta = 0$ implies that the firm doesn't care about the future consequences of raising the nominal wage in the current period. We shall see that this is in stark contrast to the general case where we allow $\beta > 0$, to which we turn in the following section. However, it should be noted at this point that even in this simple case the Card & Hyslop (1997) method will be biased. Whilst this special case yields no *active* compression of wage increases by firms, there will still be some *latent* compression: since DNWR places upward pressure on the level of wages in the past, the firm does not have to raise wages as frequently to achieve their target wage today.

1.2.3 The Dynamic Model

In this case we must solve the full dynamic optimisation problem as stated above:

$$v(W_{-1}, A) = \max_{W} \left\{ A \left[\ln\left(\frac{W}{B}\right) + c\ln\left(\frac{W}{W_{-1}}\right) \mathbf{1}^{-} \right] - W + \frac{\beta}{1+\pi} \int v(W, A')\, dF(A'|A) \right\}$$

First we will present the general structure of the solution, and then we will obtain its specific form under additional assumptions as to the distribution of shocks $F(\cdot)$.

The basic structure of the solution to the full dynamic model is very similar to that of the model in which $\beta = 0$ above. As before, we solve the problem by first taking the first-order condition with respect to W, conditional on $\Delta W \neq 0$:

$$A\left[\frac{1}{W} + \frac{c}{W}\mathbf{1}^{-}\right] - 1 + \frac{\beta}{1+\pi}\int v_W(W, A')\, dF(A'|A) = 0, \quad \text{if } \Delta W \neq 0 \tag{1.14}$$

Clearly, a key step is obtaining an expression for $\int v_W(W, A')\, dF(A'|A)$, but we leave this for the moment and simply define it as the function, $D(W, A)$, so we can re-write our conditional first-order condition as:

$$\left(1 + c\mathbf{1}^{-}\right)\frac{A}{W} - 1 + \frac{\beta}{1+\pi}D(W, A) = 0, \quad \text{if } \Delta W \neq 0 \tag{1.15}$$

The following proposition confirms that the structure of the optimal nominal wage policy will be similar to that found in the case where $\beta = 0$:

Proposition 1 *The optimal nominal wage policy in the dynamic model is of the* *form:*

$$\begin{array}{lllll} \textit{If} & A > u(W_{-1}) \equiv A_u, & \Delta W > 0 & \textit{until} & W = u^{-1}(A) \\ \textit{If} & A < l(W_{-1}) \equiv A_l, & \Delta W < 0 & \textit{until} & W = l^{-1}(A) \\ \textit{If} & A \in [A_l, A_u], & \Delta W = 0 & \textit{or} & W = W_{-1} \end{array} \tag{1.16}$$

where the functions $u(\cdot)$ and $l(\cdot)$ satisfy:

$$\begin{aligned} \frac{u(W)}{W} - 1 + \frac{\beta}{1+\pi} D(W, u(W)) &\equiv 0 \\ (1+c)\frac{l(W)}{W} - 1 + \frac{\beta}{1+\pi} D(W, l(W)) &\equiv 0 \end{aligned} \tag{1.17}$$

Proof. See appendix. ■

The reasoning for this is very straightforward, and parallels that for the static example when $\beta = 0$. In particular, Proposition 1 uses the conditional first-order condition (1.15) to define the functions $u(\cdot)$ and $l(\cdot)$, as in (1.17). These functions determine the optimal relationship between the nominal wage, W, and the nominal shock, A, in the event that wages are adjusted up or down respectively. The rest of the result follows from the fact that, by virtue of the continuity and concavity of the firm's objective, (1.5), the optimal value of W must be a continuous function of A.

However, to complete our characterisation of the firm's optimal nominal wage policy, we need to establish the functions $u(\cdot)$, and $l(\cdot)$, to which we now turn. In particular, we can see from (1.17) that, in order to solve for these functions, we require knowledge of the functions $D(W, u(W))$ and $D(W, l(W))$. This is aided by Proposition 2:

Proposition 2 *The function $D(\cdot)$ satisfies:*

$$D(W, A) = \int_{l(W)}^{u(W)} \left[\frac{A'}{W} - 1\right] dF - \int_0^{l(W)} c\frac{A'}{W} dF + \frac{\beta}{1+\pi} \int_{l(W)}^{u(W)} D(W, A')\, dF \tag{1.18}$$

which is a contraction mapping in $D(\cdot)$ over the relevant range, and thus has a unique fixed point over this range.

Proof. See appendix. ■

The intuition for this result is as follows. The first term on the RHS of (1.18) represents tomorrow's expected within-period marginal benefit, given that W' is set equal to W. To see this, note that the firm will freeze tomorrow's wage if $A' \in [A'_l \equiv l(W), A'_u \equiv u(W)]$, and that in this event a wage level of W today will generate a within-period marginal benefit of $\left[\frac{A'}{W} - 1\right]$. Similarly, the second term on the RHS of (1.18) represents tomorrow's expected marginal cost, given that the firm cuts the nominal wage tomorrow. Finally, the last term on the RHS of (1.18) accounts for the fact that, in the event that tomorrow's wage is frozen, the marginal effects of W persist into the future in a recursive fashion. It is this recursive property that provides us with the key to determining the function $D(\cdot)$.

A digression at this point is worthwhile to avoid confusion. Recall that, when taking the derivative of the current revenue function with respect to W, we noted explicitly that this was not differentiable in W at $W = W_{-1}$. Surely, one might think, this would be the case for the continuation value as well. The key difference is that, from today's perspective, W_{-1} is predetermined and hence immutable, whereas both W and W' are yet to be determined, and are in that sense flexible. That is, there is not an immutable threshold value of W at which there exists a kink in the continuation value. Moreover, since we know that tomorrow's nominal wage cut, freeze, or increase regimes are determined by whether A' falls in each of three non-degenerate (for $c > 0$) intervals, we can take derivatives *conditional* upon being in each of these intervals, which are well-defined. This is the logic that underlies the above.

For the purposes of the present chapter, we use a specific form for $F(\cdot)$. In particular, we imagine that real shocks, a, evolve according to the following geometric random walk:

$$\begin{aligned} \ln a' &= \ln a - \frac{1}{2}\sigma^2 + \varepsilon' \\ \varepsilon' &\sim \mathcal{N}\left(0, \sigma^2\right) \end{aligned} \tag{1.19}$$

Given that prices are assumed to evolve according to $P' = (1+\pi) P$, we obtain the following process for nominal shocks, A:

$$\ln A' = \ln(1+\pi) + \ln A - \frac{1}{2}\sigma^2 + \varepsilon' \tag{1.20}$$

Note that this implies that $E(A'|A) = (1+\pi) A$. We can then use this information to determine the full solution as follows. First, we solve for the functions $D(W, u(W))$ and $D(W, l(W))$ using equation (1.18), via the method of undetermined coefficients. Then, given these, we obtain the solutions for $u(W)$ and $l(W)$ using the equations in (1.17). Following this method yields Proposition 3:

Proposition 3 *If nominal shocks evolve according to the geometric random walk, (1.20), the functions $u(\cdot)$ and $l(\cdot)$ are of the form:*

$$\begin{aligned} u(W) &= u \cdot W \\ l(W) &= l \cdot W \end{aligned}$$

where u and l are given constants that depend upon the parameters of the model, $\{c, \beta, \pi, \sigma\}$.

Proof. See appendix. ■

Thus, the optimal nominal wage policy takes the following piecewise linear form:

$$\begin{array}{llccc}\text{If} & A > u \cdot W_{-1} \equiv A_u, & \Delta W > 0 & \text{until} & W = A/u \\ \text{If} & A < l \cdot W_{-1} \equiv A_l, & \Delta W < 0 & \text{until} & W = A/l \\ \text{If} & A \in [A_l, A_u], & \Delta W = 0 & \text{or} & W = W_{-1}\end{array} \tag{1.21}$$

1.3 Predictions

This section seeks to bridge the gap between the theory presented above and the forthcoming empirical section by drawing out some testable predictions of the theory. Recall that we are interested in two potential forms of the compression of wage increases: *active* compression whereby firms actually reduce the wage paid when they increase the wage; and latent compression that arises because DNWR increases the general level of wages and thus lessens the need for firms to increase wages by as much in order to reach their desired level. We show that these predictions have a precise interpretation in the context of the model presented above.

The starting point to understanding the effects of DNWR is the following decomposition of the unconditional distribution of log nominal wage changes:

$$f(\Delta \ln W) = \int f(\Delta \ln W | W_{-1})\, dF(W_{-1}) \tag{1.22}$$

An important observation to note is that we would expect the existence of DNWR to affect *both* of the distributions on the RHS of this expression. That the conditional distribution, $f(\Delta \ln W | W_{-1})$, will be affected follows directly from the firms' optimal wage policies (1.21). However, we would also expect DNWR to affect the distribution of lagged nominal wages, $F(W_{-1})$. In particular, we would expect DNWR to raise the general level of nominal wages in the economy, as it restricts firms from cutting wages when they otherwise would have done. We will see that this simple decomposition is very important when it comes to deciding which wage change distribution is appropriate to analyse.

1.3.1 Active Compression

As can be seen from (1.21), active compression of wage changes can be related to the parameters u and l. Numerical simulations of the model establish that $u > 1 > l$ and that $1/l > u$.[13] This is precisely in accordance with our original intuition (section 2.1). Since $u > l$ there exists a region of inaction for the nominal shock variable in which it is optimal not to change the nominal wage. Moreover, because $l < 1$ there will be an active compression of nominal wage cuts. This follows directly from the discontinuous fall in effort following a wage cut at the margin. In addition, $u > 1$ means that nominal wage increases will also be actively compressed relative to the counterfactual solution. Recall that the intuition for this is that raising the nominal wage today raises the likelihood that the firm will wish to cut the wage, at a cost, in the future. Finally, the fact that $1/l > u$ implies that the active compression of wage increases will not be as strong as that for wage cuts. The reason for this is that the potential costs associated with wage increases are discounted in two ways. First, some discounting derives from the fact that raising the wage may only increase the costs of wage cuts *in the future.* But, in addition to this, the probability that these additional future costs will be realised is less than one, leading to further discounting. Figure 1-3 illustrates this result by superimposing the optimal policy implied by (1.21) on those obtained in the special cases where $c = 0$, or $\beta = 0$.

Recall that our main concern is with the characteristics of the nominal wage change distribution. Using (1.21) the following proposition derives the form of the log

[13] Unfortunately, due to the analytical complexity of the solution, a formal proof of this result has proved elusive.

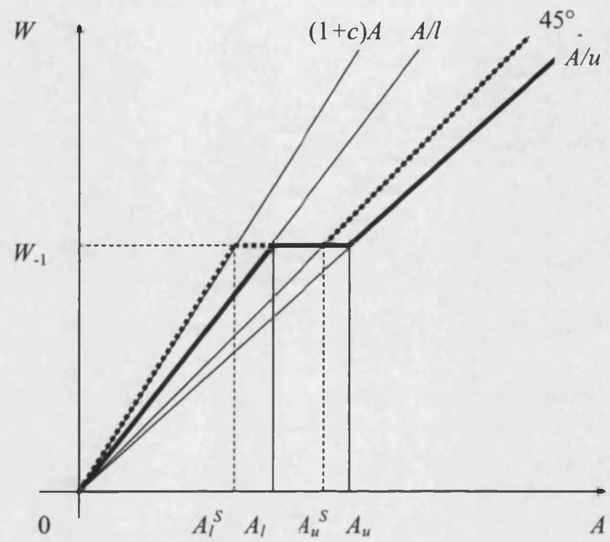


Figure 1-3: The Optimal Wage Policy in the Dynamic Model

nominal wage change distribution, conditional on the lagged wage, in the behavioural model:

Proposition 4 *The log nominal wage change density, conditional on the lagged wage, implied by the behavioural model is given by:*

$$f\left(\Delta \ln W|W_{-1}\right) = \begin{cases} \tilde{f}\left(\Delta \ln W + \ln u|W_{-1}\right) & \text{if } \Delta \ln W > 0 \\ \tilde{F}\left(\ln u|W_{-1}\right) - \tilde{F}\left(\ln l|W_{-1}\right) & \text{if } \Delta \ln W = 0 \\ \tilde{f}\left(\Delta \ln W + \ln l|W_{-1}\right) & \text{if } \Delta \ln W < 0 \end{cases} \tag{1.23}$$

where $\tilde{F}\left(\cdot|W_{-1}\right)$ and $\tilde{f}\left(\cdot|W_{-1}\right)$ are the c.d.f. and p.d.f. of the counterfactual (no DNWR) conditional log nominal wage change distribution.

Proof. See appendix. ■

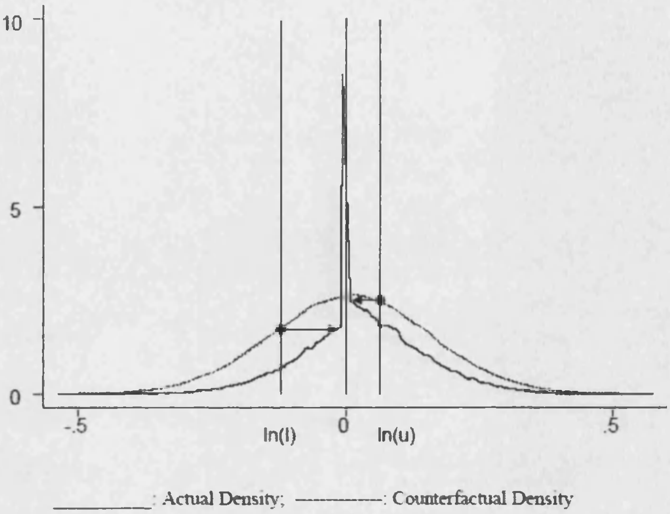


Figure 1-4: $f\,(\Delta \ln W | W_{-1})$ implied by Theory

Figure 1-4 illustrates this result. In particular, it shows that the distribution of log wage cuts is exactly the same as the counterfactual distribution below $\ln l < 0$, just shifted horizontally by an amount $-\ln l > 0$. A similar result obtains for wage increases. The residual density is "piled up" to a mass point at zero wage change. Thus, the effect of worker resistance to wage cuts is to yield a conditional log wage change distribution with dual censoring from above and below relative to the counterfactual[14].

The key prediction that we will test in our empirical work is the effect of the rate of inflation, π, on the compression of wage increases. To this end, figure 1-5 presents results for the effect of changes in the rate of inflation on the parameter u. It is

[14]This censoring result has interesting parallels in the previous empirical literature. Altonji & Devereux (2000) estimate an econometric model similar to (1.23) except that they neglect the possibility of compression of wage increases.

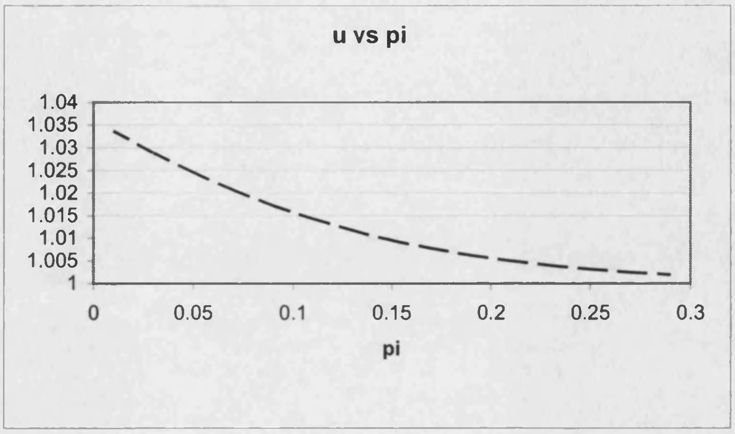


Figure 1-5: Properties of the Optimal Wage Policy Parameters

clear that the firm will reduce any active compression of wage increases as inflation rises since u falls as π rises. The intuition for this is that active compression of wage increases occurs only insofar as wage increases raise the likelihood of future costly nominal wage cuts. To see this, note that our special case in which the firm does not care about the future ($\beta = 0$) yielded no compression of wage increases ($u = 1$). Thus, since higher inflation reduces the likelihood of future costly nominal cuts, the firm no longer needs to worry about raising the nominal wage. A key related result that we want to emphasise is that, as inflation becomes large, $u \to 1$. That is, high inflation implies that wage increases cease to be compressed relative to a counterfactual world without DNWR. Thus, if the behavioural model is correct, we would expect to observe the upper tail of $f\left(\Delta \ln W | W_{-1}\right)$ becoming more dispersed as inflation rises. This is illustrated in Figure 1-6.

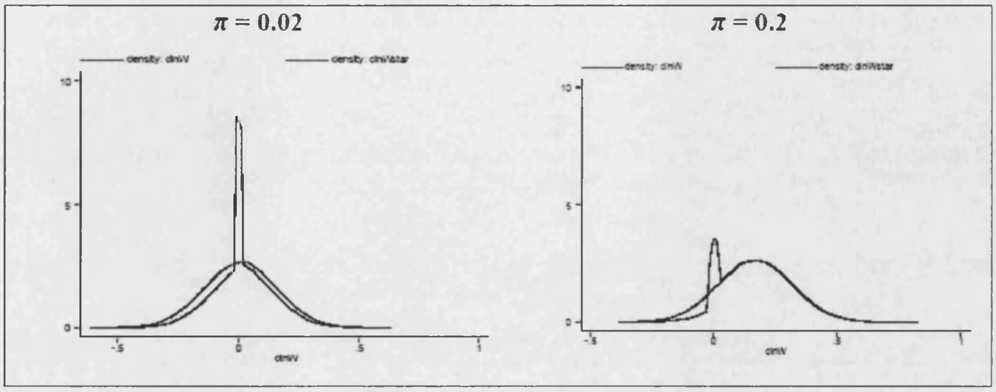


Figure 1-6: Theoretical $f\left(\Delta \ln W|W_{-1}\right)$ for Different Rates of Inflation

1.3.2 Latent Compression

All of the above discussion on active compression has been in terms of the nominal wage change distribution *conditional* on the lagged nominal wage. The reason for this is that the lagged wage is taken as given (is part of the state) at the time of setting the current wage, and so all theories will yield *direct* implications on the conditional distribution, $f\left(\Delta \ln W|W_{-1}\right)$. However, most of the previous empirical literature has concentrated on the properties of the unconditional distribution, $f\left(\Delta \ln W\right)$, typically by estimating some measure of the increase in average wage growth due to DNWR:

$$E\left(\Delta \ln W|DNWR\right) - E\left(\Delta \ln W|no\ DNWR\right) \tag{1.24}$$

to try to gain an impression of the effect of DNWR on the firms' real labour costs. The following proposition demonstrates that this emphasis in the previous literature may well be misleading:

Proposition 5 *DNWR has no effect on average wage growth in the long run for finite* $G \equiv u/l$.

Proof. See appendix. ■

This result can be interpreted in a number of ways. First, and closest to the form of the proof, note that the optimal wage policy (1.21) implies that the difference in the *levels* of the log wage with and without DNWR must be bounded (between $-\ln u < 0$ and $-\ln l > 0$). Thus, it follows that the rates of *growth* of actual and counterfactual log wages cannot be different in the long run, as it would necessarily imply a violation of these bounds.

An alternative interpretation for this result is that it is simply a requirement for the existence of a steady state in which average growth rates are equal. Since productivity shocks grow on average at a constant rate, so must wages grow at that same rate in the long run. Thus, even the model with DNWR must comply with this simple steady state condition in the long run.

How might this result come about? First, our results above indicate that firms may actively compress wage increases as a precaution against future costly wage cuts, thereby limiting the wage growth increasing effects of DNWR. However, this cannot be the whole story – we saw above that the active compression of wage increases will be less than that of wage cuts. In addition, we can find cases in which there will be no active compression of wage increases for which Proposition 5 still applies. So there must be an additional process at work.

Consider the case where $\beta = 0$. Recall that this is the case in which there is no active compression of wage increases as firms are myopic. Figure 1-7 shows a simulation of the unconditional wage change distribution implied by the behavioural model in this case. We can see from Figure 1-7 that, contrary to the assumption of previous studies, the upper tail of $f(\Delta \ln W)$ displays a compression in the presence

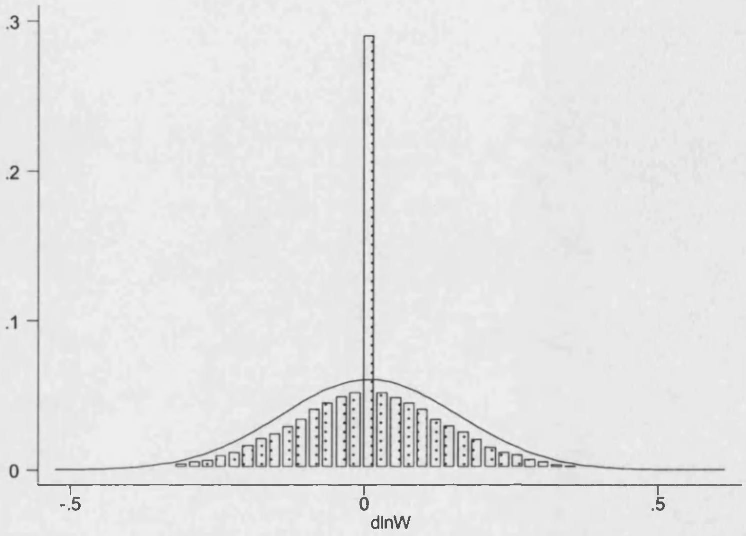


Figure 1-7: $f(\Delta \ln W)$ Implied by Theory when $\beta = 0$

of DNWR. Thus, the upper tail of the wage change distribution is still compressed, *even if firms do not actively compress wage increases.*

This provides an additional insight into the process by which this steady state requirement might be achieved in practice. If wage increases are not actively compressed, this means that when firms increase the wage, they increase it to the counterfactual level, A. However, recall that the existence of DNWR will tend to raise the general level of lagged wages in the economy, as firms will have been constrained in cutting wages in the past. Thus, when firms increase the wage, they do not have to increase it by as much or as often to reach the counterfactual wage level. Thus the upper tail of $f(\Delta \ln W)$ will indeed still be affected by the existence of DNWR – in particular, it will be less dispersed, as seen in Figure 1-7. We term this additional effect *"latent compression"*.

Thus, Proposition 5 has important implications with respect to the previous empirical literature. By not taking into account the compression of wage increases, previous empirical studies could have overstated the increase in wage growth due to DNWR. To see this, consider Figure 1-8. This shows three simulated wage change distributions derived from the model of section 2. The bold line shows the wage change distribution with DNWR ($c > 0$), whereas the thick dashed line illustrates the true counterfactual wage change density ($c = 0$). In addition, we include a "naive" counterfactual density that is derived by imposing symmetry in the upper tail of the distribution with DNWR (according to the method of Card & Hyslop, 1997). It can be clearly seen that, by using the naive counterfactual, we obtain an *overestimate* of the increase in average wage growth due to DNWR when there is a compression of the upper tail. By neglecting this compression, previous studies could have overstated the micro-effects of DNWR, which could go some way to explaining the observed tension between the micro- and macro- level evidence on DNWR. We will examine whether this is true in the ensuing empirical analysis.

1.3.3 Turnover Effects

In addition to the above, the model of section 2 can also provide predictions on the effect of turnover on the distribution of wage changes. To see this, imagine that there is now some exogenous probability that a worker will separate from the firm each period, $\delta < 1$. The effect of this is to reduce the firm's real discount factor from β to $\beta\delta$, since there is now a lower probability that the firm will survive until next period. As a result, sectors in which turnover is high (high δ) will act more myopically than sectors with low turnover. In other words, high turnover sectors should set wages more like the special case in which $\beta = 0$ (section 2.2), and low turnover sectors should act more like the forward looking firm of section 2.3. It

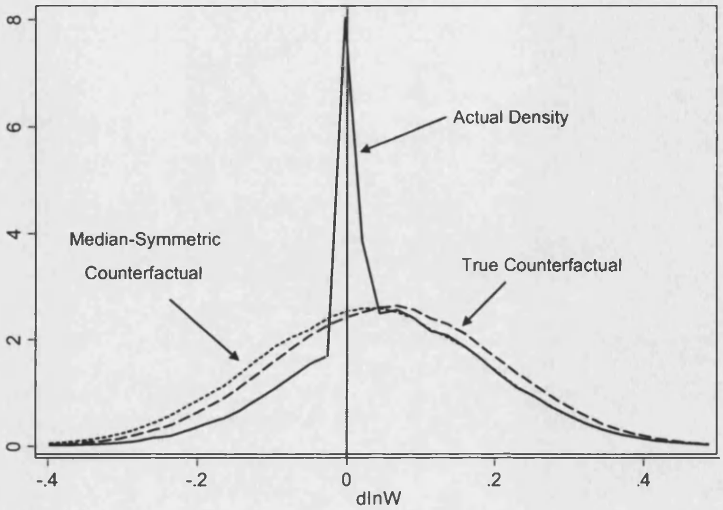


Figure 1-8: Overstatement of Costs of DNWR

follows that we should expect to see a greater active compression of wage increases in sectors with lower turnover. We will examine this claim in the forthcoming empirical section to which we now turn.

1.4 Empirical Implementation

The previous section has equipped us with a set of empirical predictions implied by the model of DNWR in section 2. In particular, we have shown that we would expect there to be a compression of wage increases as well as decreases that will reduce the increase in aggregate wage growth due to DNWR. In addition, we have shown that we would expect this compression of wage increases to be more pronounced in sectors with low rates of turnover. This section seeks to test these predictions using micro-data from the US and Great Britain.

1.4.1 Data

The data used in this analysis are taken from the Current Population Survey (CPS) and the Panel Study of Income Dynamics (PSID) for the US, and the New Earnings Survey Panel Dataset (NESPD) for Great Britain. For all datasets, the relevant wage measure used in this study is the basic hourly wage rate. Since the CPS and PSID are relatively well-known datasets, we only describe them briefly here.

The CPS samples are taken from the Monthly Outgoing Rotation Group (MORG) files from 1979 to 2002. We link respondents across consecutive years using a method similar to that advocated by Madrian & Lefgren (1999)[15]. This method yields approximately 25,000 individual annual wage changes each year from 1980–2002, although changes in sampling method yield lower sample sizes in 1985–86 and 1995–96 (see Table 1-1). Unfortunately, we cannot easily differentiate between job-stayers and changers using the CPS due to a lack of information on job characteristics and tenure[16]. Additional problems arise in the CPS resulting from the introduction of a computer-aided survey design (CAPI) in 1994. Figure 1-9 illustrates the dispersion of log wage changes in the CPS over the sample period, as measured by the standard deviation, and the 90-10 and 80-20 percentile differentials. One can clearly detect a significant rise in the dispersion of wage changes starting in 1994 with the introduction of CAPI. In our ensuing empirical analysis we attempt to control for this.

[15] In particular, first we match individuals according to their personal identifiers, as well as their month of interview. We then employ Madrian & Lefgren's "s|r|a" criterion – i.e. that matched observations must report the same sex and race across years, and that the difference in their age must lie in the interval $[0, 2]$.

[16] Card & Hyslop (1997) attempt to identify job-stayers in the CPS by restricting their analysis to those respondents who do not change occupation year-on-year. We do not make such an attempt as it is complicated by changes in the occupational classification over the period. However, we found that our sample displays very similar properties to that of Card & Hyslop.

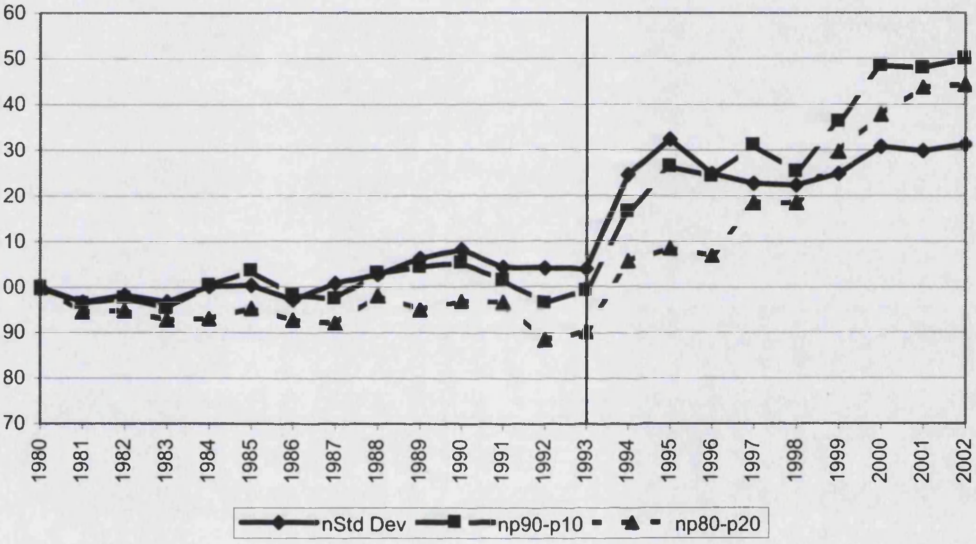


Figure 1-9: The Dramatic Increase in the Dispersion of Wage Changes after 1994 (CPS)

The PSID data are taken from the random (not poverty) samples for the years 1971 to 1992. We use data on regular hourly pay rates for household heads to construct individual annual wage changes. We concentrate on the wage changes of job-stayers by excluding workers with tenure of strictly less than 12 months[17], and additionally remove respondents who report that they live in a foreign country, and top-coded wage data. Our PSID sample provides us with much smaller samples than those from the CPS, with approximately 1,300–2,200 individual wage changes each year over the sample period (see Table 1-1).

Finally, the NESPD is an individual level panel which is collected in April of each year running from 1975 through to 2001 for Great Britain[18]. It is a 1% sample of British income tax-paying workers with a National Insurance (Social Security) number that ends in a given pair of digits, and in this sense is a random sample of the tax-paying population. The wage measure used is the gross hourly earnings, excluding overtime, of job-stayers whose pay is unaffected by absence. Table 1-1 provides summary statistics for the NESPD sample. An important observation to make is that the statistics for the level of real wage changes in 1977 are vastly lower than in all other periods. In particular, median real wage growth was -7.51% in 1977, but was never below -0.2% in any other year in the sample period. The reason for this is that the UK government of the time instituted an incomes policy in order to try to curb high inflation. In particular, these policies were remarkably successful in containing wage inflation in late 1976 to early 1977 as a result of the cooperation

[17] A selection issue arises when excluding job-changers. In particular, previous research has shown that "displaced" workers often accept significant reductions in earnings on re-employment (see Jacobson, LaLonde & Sullivan, 1993). Thus, by concentrating on job-stayers, our results might overstate the true extent of DNWR. However, it is also the case that much of the previous literature has focused on job-stayers, so our analysis will be comparable to that of other studies. We leave these empirical issues for future research.

[18] However, much of our analysis requires the use of consistent industry and occupation coding, which we have up to 1999 only.

of the unions (see Cairncross, 1995, pp. 220–221). Despite this, however, retail price inflation remained high, thereby leading to the significant real wage losses that we observe in our data. As a result of this, we treat the 1977 data as an outlier throughout the rest of our analysis.

The NESPD data for Great Britain have a number of key advantages for our purposes, especially in comparison with the CPS and PSID samples for the US. The first, and most obvious, is that the NESPD provides us with comparatively very large sample sizes: we obtain sample sizes of 60–80,000 wage change observations each year. This will help us to identify a more precise relationship between the distribution of wage changes and the rate of inflation, since we can be more confident that variation in the wage-change distribution is not driven by errors due to lower sample sizes.

The second advantage of the NESPD data is its sample period: from 1975–2001. This is particularly useful for our purposes given that we seek to use variation in the rate of inflation to gauge the impact of DNWR on wage changes, since the UK experienced significant variation in inflation over this period relative to the US. Figure 1-10 displays the time-series of the leading UK inflation indicator – the Retail Price Index (RPI) – and the CPI-U inflation rate for the US, over the relevant periods. It can be seen that the UK inflation rate varied substantially, with rates over 20% in the 1970s down to below 2% in the 1990s. Inflation in the US, on the other hand, displays much less variation, with rates no higher than 11%. Thus, again we can expect to be able to identify a more significant relationship, should one exist, between the wage-change distribution and inflation for the NESPD sample by virtue of this greater variation.

The final key advantage of the NESPD sample is that measurement error in these data is likely to be at a minimum for large scale datasets relative to individually

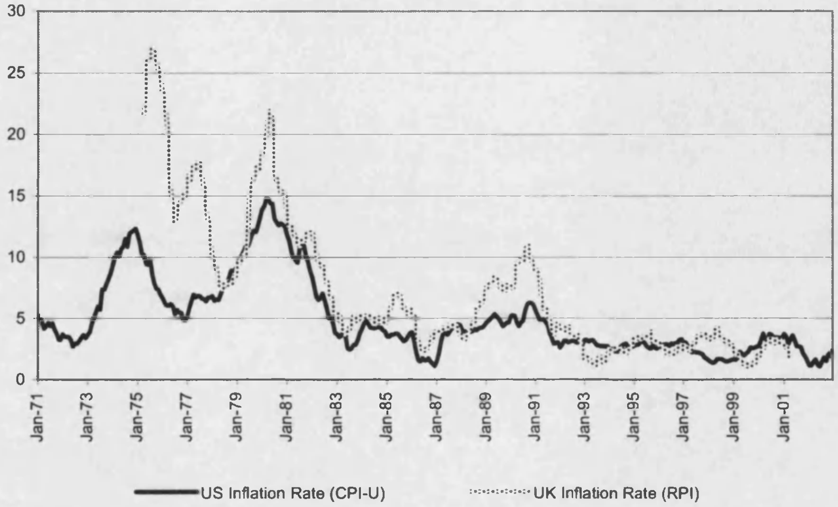


Figure 1-10: US & UK Inflation over the Sample Periods

reported data of the CPS and PSID samples. The reason for this is that the NESPD is collected from employers' payroll records, thereby leaving less scope for error due to imperfect memory etc. (see Nickell & Quintini, 2003, for more on this). Indeed validation studies of leading panel datasets have used matched data from employer surveys to assess the extent of measurement error in worker reported earnings data. In particular, Bound & Krueger (1990) and Card & Hyslop (1997) both seek to assess the importance of measurement error in the CPS via this method.

The existence of measurement error in hourly wages has been shown in previous empirical studies to act as a key impediment to inferring the extent of DNWR. As emphasised throughout this analysis, the existence of a spike at zero in the distribution of nominal wage changes is a key characteristic of DNWR. Classical measurement error in wages and hours data would yield an *understatement* of the extent of downward nominal wage rigidity. To see this, note that the addition of classical measurement error will render true wage freezes to be observed as (small) wage changes, thereby reducing the size of the observed spike (Akerlof, Dickens & Perry, 1996). Previous studies have also stressed that individuals may *round* their reported wages. This, in contrast to classical error, would yield an *overstatement* of the extent of DNWR as small true wage changes are reported as wage freezes (Smith, 2000).

Some existing studies have attempted to circumvent this problem in a number of ways. Altonji & Devereux (1999) developed an empirical model that allows for the existence of a Normally distributed classical measurement error. However, the main criticism of this is that no account is taken of potential rounding. In addition, a number of studies have augmented their analyses with data obtained from payroll records from individual establishments (see Altonji & Devereux, 1999, and Fehr & Goette, 2003). These, on the other hand, are subject to the criticism that any results are not representative. The relative accuracy of the NESPD allows us to avoid these difficulties, and is thus an important virtue in this context.

Since the descriptive properties of DNWR in all of these datasets have been well-explored in previous analyses – Card & Hyslop (1997) for the CPS, Kahn (1997) and Altonji & Devereux (2000) for the PSID, and Nickell & Quintini (2003) for the NESPD – we do not seek to provide a full descriptive account of DNWR. Rather, our aim is to assess the validity of the predictions of the model presented in section 2. To this end, we simply verify that all three of our datasets display the stylised features noted in the previous literature on DNWR: i.e. the existence of a spike at zero nominal wage change, and a relative deficit of nominal wage cuts. Figures 1-11, 1-12 and 1-13 display histograms of the observed log nominal wage change distributions for each of our samples, where we have differentiated between higher and lower inflation periods. In all histograms there is a clear spike in the distribution at zero nominal wage change, and a relative asymmetry in the form of a deficit of wage cuts. Moreover, it can be

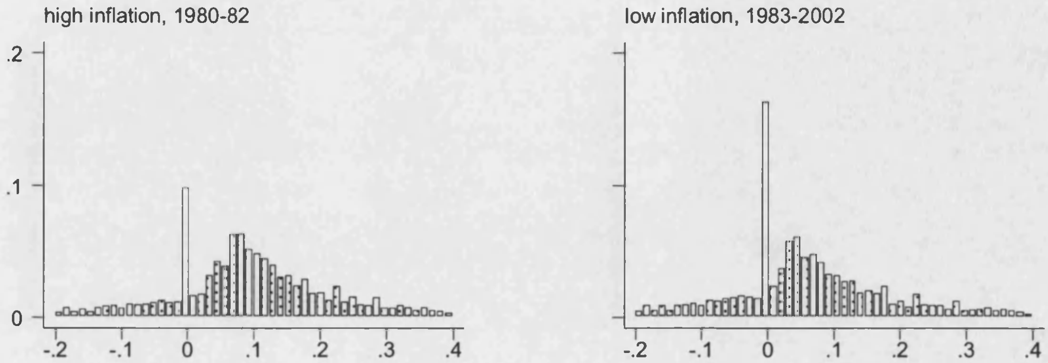


Figure 1-11: Log Nominal Wage Change Distributions in High and Low Inflation Periods (CPS, 1980 – 2002)

clearly seen that the higher the rate of inflation, the smaller is the spike at zero wage change. This is consistent with the logic that higher rates of inflation relax any DNWR constraint since firms are more able to reduce real labour costs without ever cutting the nominal wage.

A final note worth making in the context of our datasets is that inflation stayed at persistently low levels in the US and UK from 1992 onwards, with an average inflation rate of 2.56% for the US 1992–2002 and 2.69% for the UK 1992–2001. This is important, as a criticism levelled at previous studies of DNWR has been that individuals will get used to receiving nominal wage cuts when inflation has remained low for some time (Gordon, 1996, and Mankiw, 1996). Such a criticism becomes less compelling when the inflation rate has stayed low for the 9–10 years observed in our

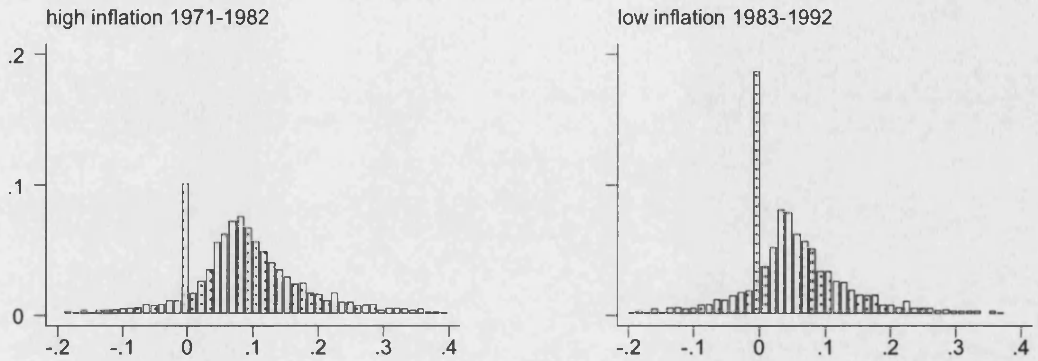


Figure 1-12: Log Nominal Wage Change Distributions in High and Low Inflation Periods (PSID, 1971 – 92)

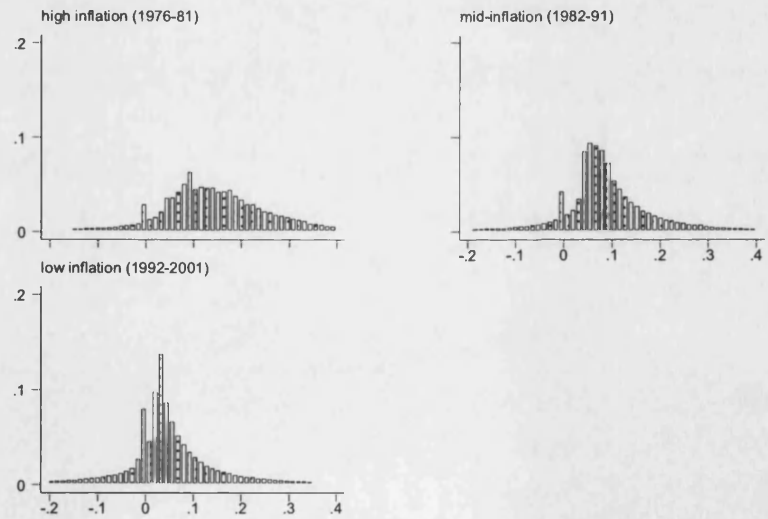


Figure 1-13: Log Nominal Wage Change Distributions in High, Mid, and Low Inflation Periods (NESPD, 1976 – 2001)

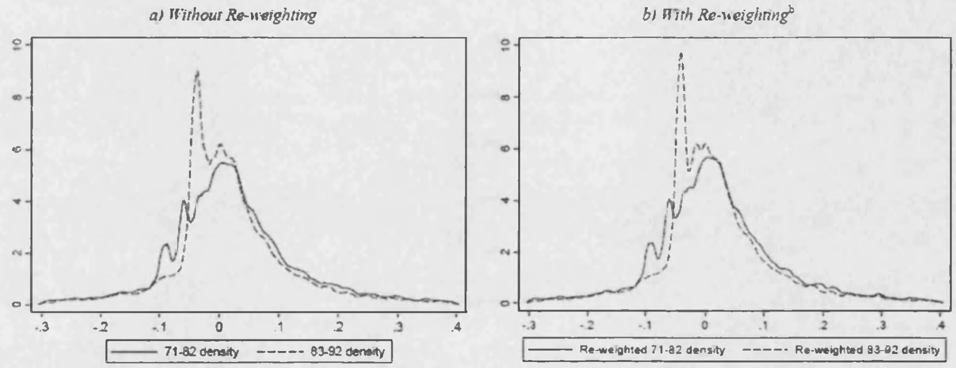
1.4.2 Does DNWR Increase Aggregate Wage Growth?

In order to test our hypotheses, we need a way of modelling empirically the wage change distributions, $f(\Delta \ln W|W_{-1})$ and $f(\Delta \ln W)$. In what follows, we will focus on the analogous real wage distribution counterparts to these. Note that this does not alter any substantive aspects of the analysis, since these are exactly the same shaped distributions, just shifted to the left by a constant (approximately equal to the rate of inflation)[19]. However, focusing on these does allow greater ease of comparison across years with different inflation rates.

The method we apply to our two empirical questions will turn out to be very similar. To start with, then, we motivate our preferred method in the context of trying to understand the impact of DNWR on the unconditional distribution of log wage changes, $f(\Delta \ln W)$. Let us begin by considering some naive approaches. First, we might think of simply looking at the differences between the wage change distributions in high inflation periods and low inflation periods to see if the predictions of section 4.1 are confirmed at this basic level. To this end, figures 1-14(a) and 1-15(a) present estimates of the density of log real wage changes for periods with different inflation rates using the PSID for the US, and the NESPD for Britain (the introduction of CAPI in the CPS renders this a less useful exercise for the CPS data). Notice that lower inflation leads to a compression of the lower and, more importantly for our purposes, the *upper* tail of the wage change distribution, precisely in accordance with the predictions of section 4.1[20].

[19]This follows because $\Delta \ln (W/P) = \Delta \ln W - \Delta \ln P \cong \Delta \ln W - \pi$ where π is the rate of inflation.

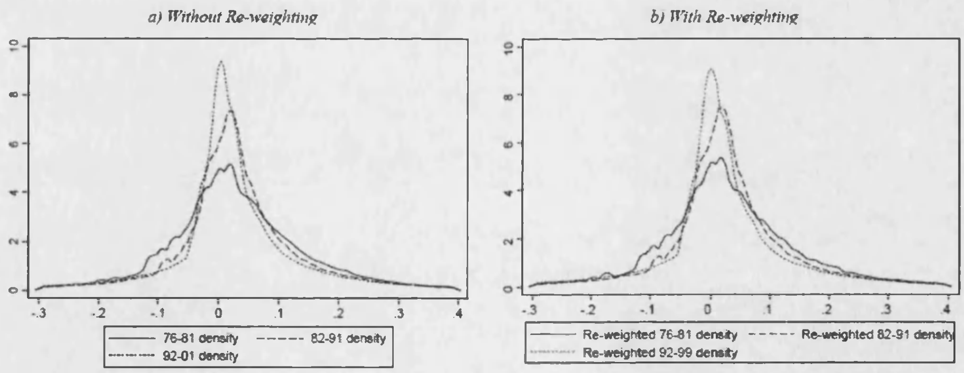
[20]It should be noted that the existence of the spike in the lower tail of the real wage change distribution (at approximately minus the rate of inflation) can lead to an overstatement of lower tail compression. However, our emphasis is on the effects on the upper tail, which are not subject to this problem.



Notes:

a. Kernel density estimates using an Epanechnikov kernel, over 250 data points, and a bandwidth of 0.005.
b. "Re-weighting" refers to the use of the DiNardo, Fortin & Lemieux (1996) re-weighting technique to control for changes in age, age^2, sex, education, 1-digit industry, 1-digit occupation, region, self employment, and tenure.

Figure 1-14: Density Estimates of Log Real Wage Change Distributions (PSID)



Notes:

a. Kernel density estimates using an Epanechnikov kernel, over 250 data points, and a bandwidth of 0.005.
b. "Re-weighting" refers to the use of the DiNardo, Fortin & Lemieux (1996) re-weighting technique to control for changes in age, age^2, sex, region (including London dummy), 2-digit industry, 2-digit occupation, and major union coverage.

Figure 1-15: Density Estimates of Log Real Wage Change Distributions (NESPD)

However, one could argue that at least some of the observed differences were due to changes in other variables that affect wage changes. For example, there have been changes in the industrial, age, gender, regional etc. compositions of the workforce in both the US and Britain over these time periods. So, we should control for factors such as these before attributing any differences to DNWR. To address this, we introduce a set of micro-level control variables for each dataset, summarised in Table 1-2. In particular, we control for changes in micro-level variables by re-weighting the observed wage change distributions according to the method of DiNardo, Fortin & Lemieux (1996) (henceforth DFL). To do this, we first define a "base year", T – for all datasets this will be the final sample year – and re-weight each year's observed wage change distribution to obtain an estimate of what the wage change distribution would have looked like if the distribution of micro-level characteristics were identical to that at date T. In particular, if we define the log wage change as Δw, micro-level characteristics as x, and the year of the relevant x distribution as t_x, we derive:

$$\begin{aligned} f\left(\Delta w_t; t_x = T\right) &= \int f\left(\Delta w|x\right) dF\left(x|t_x = T\right) \\ &= \int f\left(\Delta w|x\right) \cdot \psi \cdot dF\left(x|t_x = t\right) \end{aligned} \tag{1.25}$$

for all $t < T$. The key insight of DFL is that this is simply a re-weighted version of the observed date t wage change distribution, with weights ψ given by:

$$\psi = \frac{dF\left(x|t_x = T\right)}{dF\left(x|t_x = t\right)} = \frac{\Pr\left(t_x = T|x\right)}{\Pr\left(t_x = t|x\right)} \cdot \frac{\Pr\left(t_x = t\right)}{\Pr\left(t_x = T\right)} \tag{1.26}$$

where the second equality follows from Bayes' Rule. The conditional probabilities in (1.26) can then be estimated simply via a probit model.

Figures 1-14(b) and 1-15(b) displays density estimates of the DFL re-weighted distribution of log real wage changes for different inflation periods, again for the PSID and NESPD. Again, it can be seen clearly that lower rates of inflation are associated with a compression both of tails of the wage change distribution, in line with the predictions of section 3.2.

However, even having controlled for such factors, it is still not necessarily legitimate to attribute *all* the residual difference in the wage change distributions to DNWR. Thus we need a way of ensuring that only the variation in wage change distributions that varies *systematically* with DNWR is attributed. To do this, we estimate regressions of the form:

$$P_{n_{rt}} = \beta_{0n} + \beta_{1n} P_{50_{rt}} + \eta_n \pi_t + z'_{rt} \gamma_n + \varepsilon_{n_{rt}} \tag{1.27}$$

where $P_{n_{rt}}$ is the nth percentile of the real wage change distribution in region r at time t, π_t is the rate of inflation at time t and thereby measures the prominence of nominal zero in the distribution of log real wage changes, and z_{rt} is a vector of aggregate controls that could potentially affect the distribution of wage changes. $P_{50_{rt}}$ is included on the RHS of (1.27) in order to control for changes in the central tendency of the distribution of wage changes. That is, it "re-centres" the distributions over time in order to make them comparable. We estimate (1.27) by Least Squares, where we weight by the size of the region at each date[21].

The measure of inflation used will be the CPI-U-X1 series for the US, and the April to April log change in the Retail Price Index for Great Britain. The aggregate controls will be as follows. First, we control for any distortion to the wage change distributions caused by peculiarities of the datasets used. So, to control for the effects

[21]Formal quantile regression (Least Absolute Deviation) estimators were also tried with little difference in results. However, such is the computational intensity involved in estimating the correct standard errors for these estimators, we opted for simple OLS instead.

of the introduction of CAPI in 1994 in the CPS, we include a dummy variable that takes value one for all years from 1994 onwards when we estimate (1.27) for the CPS. In addition, to control for the incomes policies implemented in 1977 in the UK, we include a dummy that takes value one for the year 1977 in our NESPD regressions.

In addition, we control for the absolute change in the rate of inflation. This is motivated by the hypothesis that greater inflation volatility will yield greater dispersion in relative wages regardless of the existence of DNWR (see Groshen & Schweitzer, 1999). We also include both current and lagged regional unemployment rates. This is motivated by the idea that the existence of DNWR might lead to unemployment – indeed, as mentioned before, this is one of the principal reasons for interest in the topic. Since unemployment will lead to workers "leaving" the wage change distribution, it is important to control for any resulting distributional consequences. We also include lagged regional unemployment in accordance with the wage curve hypothesis of Blanchflower & Oswald (1994) that the *level* of wages is empirically associated with the level of unemployment. If this is true, then we would expect the change in unemployment to affect the change in wages, and so we include lagged regional unemployment to control for this possibility.

It should be noted that the empirical method described above is robust to a number of possible concerns. First, since we are exploiting variation in the *tails* of the distribution of wage changes, rather than the spike at zero, the above method is less subject to the measurement error concerns that much of the previous literature has suffered from. In particular, so long as measurement error is neither time-varying nor related to the rate of inflation, (1.27) will pick up the true effects of inflation on the wage change distribution[22]. Second, the specification is also robust to the existence

[22] One might be concerned that higher inflation leads to greater errors in reported wage data simply

of rigidity in *real* wages. The reason is that real wage rigidity, in its traditional form, will be invariant to inflation by definition. An exception to this is the argument put forward by Akerlof, Dickens & Perry (2001) that real wage rigidity is amplified as inflation rises because it becomes optimal for workers to direct their scarce attention to maintaining their real wage. However, if anything, such a possibility would work *against* the claim of the model in section 2, as it would predict that the upper tail of wage changes would become *more* compressed as inflation rises. If this were the case, any evidence we find for the predictions of section 3 could be interpreted as lower bounds on the true effects. A similar reasoning applies to any concerns one might have about the impact of skill-biased technical change (SBTC). Under SBTC, we might expect that workers obtaining high wage increases early in our samples will obtain even higher wage increases later on as technical change increasingly favours those in skilled sectors. However, since inflation is in practice declining over the sample periods of our data, SBTC would, if anything, work against the predictions of section 3. Thus, the above method is robust to a number of potential criticisms.

Clearly, the coefficients of interest in (1.27) for the purposes of estimating the effects of DNWR are η_n. In particular, the predictions of section 3.2 indicate that δ_n should be negative for low percentiles, and positive for high percentiles. The reasoning is that higher inflation should lead to an increased dispersion of wage changes, and thereby decrease negative percentiles, and increase positive ones.

We estimate (1.27) in three specifications. First, we simply include controls for the median wage change, P_{50}, and for any dataset peculiarities such as CAPI for the CPS and incomes policies of 1977 in the NESPD. We then include controls

because it becomes harder to keep track of exact changes in one's wage. However, our specification uses variation in *log* wage changes, and it is less clear that people would make greater *percentage* errors when inflation is high.

for the absolute change in the rate of inflation, and for regional current and lagged unemployment rates (where possible). Finally, we implement a specification with full controls that estimates (1.27) using percentiles of the DFL *re-weighted* wage change distributions so we can control for a full array of micro-level characteristics as well.

Recall that we would like to obtain an estimate of the increase in average wage growth due to DNWR:

$$\lambda \equiv E\left(\Delta w | DNWR\right) - E\left(\Delta w | no\ DNWR\right) \tag{1.28}$$

We show that such an estimate can be obtained using the estimates obtained from regressions of the form (1.27). In order to use this information to get an estimate of λ, we obtain an estimate of the predicted average wage change when inflation is very low (e.g. 1.3% in 1993 for Britain) and subtract the analogous average wage change when inflation is very high (e.g. 21.8% in 1980 for Britain)[23]:

$$\hat{\lambda} = \hat{E}\left(\Delta w | \pi = 1.3\%, x, z\right) - \hat{E}\left(\Delta w | \pi = 21.8\%, x, z\right) \tag{1.29}$$

To obtain these estimates, we use estimated percentiles from (1.27) to discretise the distribution of wage changes. In particular, if we estimate k equi-spaced percentiles of $f\left(\Delta w\right)$ then a best guess of the predicted average wage change is:

$$E\left(\Delta w | \pi, x, z\right) \approx \frac{1}{2\left(k-1\right)} \sum_{i=2}^{k} \left(\hat{P}_i + \hat{P}_{i-1}\right) \tag{1.30}$$

where i is an ascending index of the percentiles, with $i = 1$ indicating the lowest percentile, $i = 2$ the second lowest etc., and the $\hat{P}$s are the predicted values of these percentiles obtained from estimating equation (1.27). Thus, in this way, we can use

[23]Note that this involves out-of-sample predictions for the US data.

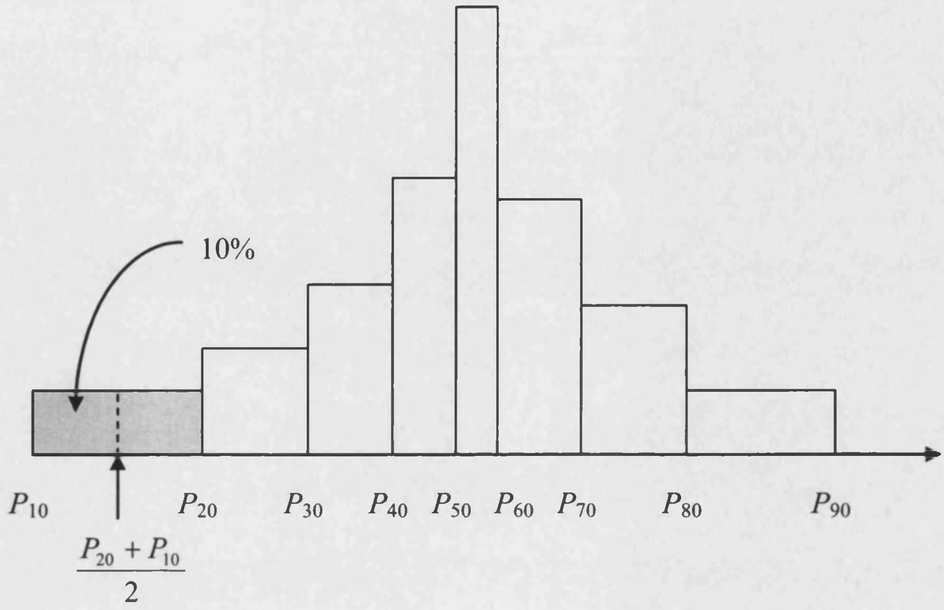


Figure 1-16: Discretising a Distribution using Percentiles: In this case, our best guess of $E(x)$ is given by $\sum_{i=2}^{9} \frac{1}{8}\left(\frac{P_i+P_{i-1}}{2}\right)$.

percentile regressions to obtain estimates of the increase in average wage growth due to DNWR. Figure 1-16 presents an intuitive treatment of this discretisation for the case of deciles ($k = 9$). Moreover, since our predicted percentiles allow us to sketch out a discretisation of the whole distribution of wage changes, we can decompose the increase in average wage growth due to DNWR into two components. The first is the increase in average wage growth due to compressed nominal wage cuts, which we refer to as "lower tail losses"; the second is the decrease in average wage growth [illegible]

compressed wage increases, "upper tail gains". In practice, we will perform due to

ocedure on 99 estimated wage change percentiles, $\hat{P}_1, \hat{P}_2, ..., \hat{P}_{99}$, for each of the this pr

pecifications detailed above. three s

Empirical Results

The results from estimating our three specifications of (1.27) for each dataset are reported in Tables 1-3–1-5. First, consider the results obtained for the CPS in Table 1-3. In all three specifications it can be seen that the estimated impact of inflation is negative for the 20th–30th percentiles, with strongest effects around the 30th percentile; and positive for the 40th–90th percentiles, with strong effects in the 70th and 90th percentiles. Thus, these results are in line with the hypothesis that higher inflation reduces the compression of both tails of the wage change distribution. Moreover, we see that the estimated effects of inflation at different points in the distribution are generally significant and fairly stable across all specifications. In addition, Table 1-3 presents estimates of the lower tail losses and upper tail gains due to DNWR. It can be seen that in all specifications there are substantial savings due to compressed wage increases, some of which even outweigh the costs from compressed wage cuts. In particular, our estimates broadly confirm the conclusion of Card & Hyslop (1997) that the increase in average wage growth due to compression in the lower tail is around 1%. However, this is offset by savings from compression of the upper tail of wage changes of around $1 - 1.5\%$.

Table 1-4 reports the analogous estimates for the PSID data. We can see that in all specifications the effect of inflation is negative for the 10th–20th percentiles, and positive for the 40th–90th percentiles. However, here the estimated effects are strongest in the 10th, and particularly the 20th, percentiles in the lower tail in contrast to the CPS results. The differences in the lower tail effects between the CPS and PSID results are likely to reflect the differences in the position of nominal zero in the respective wage change distributions, due to their different sample periods. In the CPS, nominal zero appears mostly between the 20th and 35th percentiles, whereas it appears at around the 10th–35th percentile in the PSID sample. Thus, the point at which DNWR binds differs across these two datasets.

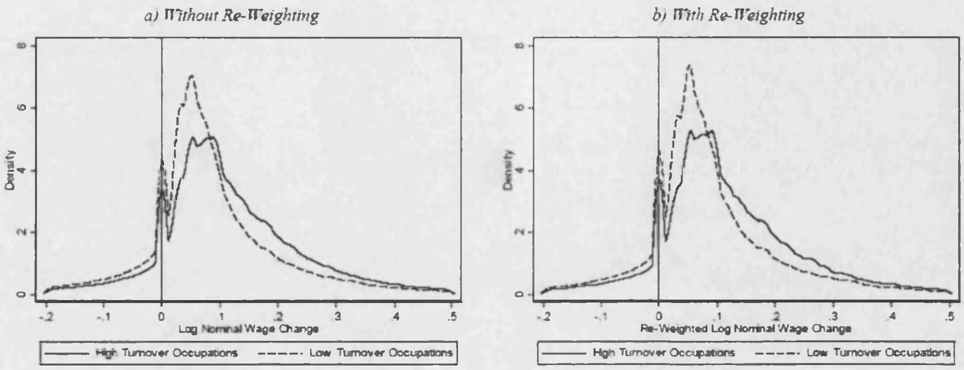
Whilst the PSID results are not as significant as those for the CPS, the upper tail effects remain significant, and are fairly stable across specifications. In addition, the coefficient estimates in the upper tail are comparable to those obtained in the CPS results, and we again observe that there are large savings from the compression of the upper tail, which again in some cases even outweigh the lower tail losses. In particular, we find an estimated increase in average wage growth due to lower tail losses of around $0.4 - 0.8\%$ which is offset by a reduction in average wage growth due to upper tail compression of $0.5 - 0.9\%$. It should however be noted that for the PSID, and to some extent the CPS data, these estimates are constructed from a number of regressions for which no significant inflation effect was detected. This is likely due to the relative lack of observations and inflation variation in the CPS and PSID compared to the NESPD. Thus, we do not want to place too much stock in the actual quantitative estimates obtained from this dataset. Rather, we consider our estimates of upper tail gains and lower tail losses for the PSID to be instructive of the fact that there is some significant compression of the upper tail of wage changes, and that this compression is of similar significance to the compression of the lower tail due to DNWR.

The results for the NESPD data are reported in Table 1-5. Again we observe that inflation has a negative impact on lower percentiles (10th–40th) and a positive impact on higher percentiles (60th–90th). Moreover, we obtain highly significant estimates for almost all percentiles and in all specifications. As mentioned above, this greater significance in comparison to the results for the PSID and the CPS is likely to be due to the superior quality and inflation variation of the NESPD. In addition, we again observe substantial upper tail gains due to compression of wage increases relative to lower tail losses, which are more consistent across specifications than those obtained for the CPS and the PSID. In particular, our results suggest that 77–95% of the lower tail losses due to DNWR is saved by restricting wage increases in the upper tail in the NESPD data, and that the increase in average real wage growth due to DNWR is of the order 0.05–0.3% – much lower than results obtained previously.

Together, these results provide strong evidence for the prediction that the upper tail of the wage change distribution will be less dispersed as a result of DNWR – in all specifications and for all datasets we see that wage increases become more restricted as inflation falls. As a result, by allowing both the upper and lower tails of the wage change distribution to be affected by DNWR, the estimated increase in average wage growth due to DNWR becomes much reduced and closer to zero – precisely in line with the predictions of section 3 and Proposition 5. In fact, we observe that estimates of the increase in average wage growth due to DNWR fall from around 1% to 0.3% at the most, and may even be negative. Thus, since previous studies have ignored the effects of DNWR on the upper tail of wage changes, they may well have vastly overstated the estimated "costs" due to DNWR.

1.4.3 Does Higher Turnover Reduce the Compression of Wage Increases?

In addition to the above, recall that section 3.3 established the claim that higher turnover sectors should act more myopically, and hence will feel more at liberty to raise nominal wages. We test this hypothesis in a manner similar to the above. First, we define a measure of "turnover" as the fraction of workers within an occupation group that are job changers each year. In a steady state this should closely match the fraction of workers who separate, and thus correspond to the parameter δ in section 3.3.



Notes:

a. Kernel density estimates using an Epanechnikov kernel, over 250 data points, and a bandwidth of 0.005.
b. "Re-weighting" refers to the use of the DiNardo, Fortin & Lemieux (1996) re-weighting technique to control for changes in adjusted lagged wages, age, age^2, sex, region (including London dummy), 2-digit industry, 2-digit occupation, and major union coverage.
c. High turnover refers to occupations for which the fraction of job changers in any given year exceeds the median.

Figure 1-17: Turnover Effects on the Distribution of Log Nominal Wage Changes for Job Stayers, NESPD

To gain an initial impression for whether such effects exist, Figure 1-17(a) plots density estimates of the nominal wage change distribution for job stayers in high and low turnover (respectively above and below median turnover) occupations using the NESPD data. It can be seen from this simple comparison that low turnover [redacted] any micro-level factors that may be driving the results, as well to focus on variation in the distribution of wage changes conditional on the lagged wage. Again, however, we see low turnover occupations compressing wage increases more than high turnover occupations, consistent with the predictions of section 3.3.

Figures 1-18(a) and 1-18(b) replicate the procedure for the PSID data. Again we can see that higher turnover sectors exhibit less compression of wage increases for both unweighted and re-weighted wage change distributions. However, it can be seen that the effects are not as strong as in the NESPD data. The main reason for this is that we only have an occupational classification at the 1-digit level in the PSID data, thereby limiting the variation in occupational level turnover we can identify and hence exploit. For this reason, we concentrate on using the NESPD data to identify turnover effects more formally.

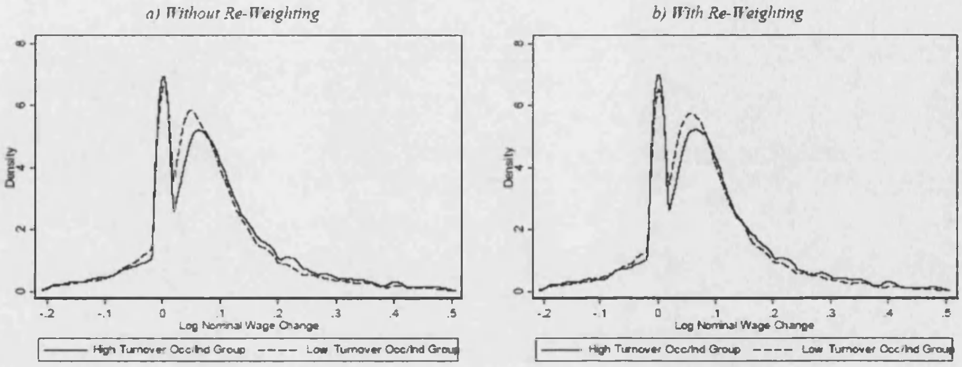
To do this, we run Least Squares regressions of the form:

$$P_{n_{ort}} = \alpha_{0n} + \alpha_{1n} P_{50_{ort}} + \theta_n \boldsymbol{\tau}_{ot} + z'_{rt}\psi_n + \varepsilon_{n_{ort}}$$

where $P_{n_{ort}}$ now refers to the nth percentile of nominal wage changes for job stayers, re-weighted for micro covariates and the lagged wage, in occupation o, region r, at time t. The variable of interest is $\boldsymbol{\tau}_{ot}$ which denotes the fraction of job changers in an occupation in a given year. Under the predictions of section 3.3, we would expect that the coefficients, θ_n will be positive for all positive percentiles of nominal wage changes.

Table 1-6 summarises the estimates of θ_n for the 60–90th percentiles. In the first

an additional control. This is legitimate provided that DNWR has no impact on either the price level or trend real wage growth. Given Proposition 5 and the results of section 4.2 this does not seem unreasonable.



Notes:

a. Kernel density estimates using an Epanechnikov kernel, over 250 data points, and a bandwidth of 0.008.
b. "Re-weighting" refers to the use of the DiNardo, Fortin & Lemieux (1996) re-weighting technique to control for changes in adjusted lagged wages, age, age^2, sex, education, 1-digit industry, 1-digit occupation, region, self employment, and tenure.
c. High turnover refers to occupation-industry groups for which the fraction of job changers in any given year exceeds the median.

Figure 1-18: Turnover Effects on the Distribution of Log Nominal Wage Changes for Job Stayers, PSID

specification (column 1), we include only basic controls for the median wage change and a dummy for 1977 to control for the incomes policies of that time. It can be clearly seen that turnover has a positive and highly significant impact on the 60–90th percentiles of nominal wage changes. The two additional columns address potential concerns one might have about the simple specification of column (1).

In particular, one concern might be that we would expect sectors with greater DNWR to have greater rates of turnover due to workers being made unemployed more often. In addition, we would also expect sectors with greater DNWR to exhibit a greater compression of wage increases, and thus create downward pressure on percentiles of wage increases. In this sense there may be an omitted variables bias to the estimates in column (1) – in particular, a *downward* bias to the estimates. Column (2) seeks to assess this possibility by including the current and lagged regional unemployment rates as controls. It can be seen, however, that this makes little difference to our estimated turnover effects, and, if anything, reduces the estimated coefficients. This suggests there is little evidence for an omitted variables problem of this type.

However, we may still be concerned that the measure of turnover is more generally cyclical, and thus correlated with the rate of inflation. Thus we may be worried that we are attributing to turnover the effects due to declining inflation. To address this concern, specification (3) includes time dummies. It can be seen that introducing these controls reduces the estimated effects of turnover, but that the coefficients remain positive and highly significant. We thus find robust evidence that increased turnover leads to an increased dispersion of wage increases, in line with the predictions of section 3.3.

1.5 Limitations and Future Directions

A number of issues remain in the light of the previous findings. First consider the theory presented in section 2. A particular assumption that one might be interested in relaxing is that of the form of adjustment costs in the effort function caused by nominal wage cuts. In particular, one might be interested in the implications of a fixed adjustment cost whereby effort falls dramatically for even very small wage cuts. This represents a more difficult theoretical challenge in the current framework, but has been considered in other applications. In particular, Degeorge, Patel & Zeckhauser (1999) study a simple two-period model of corporate earnings management with this fixed cost structure. They show that, for intermediate latent earnings losses, it is optimal for an executive to report no loss, but that for a sufficiently low latent earnings shock, it is optimal to "take a bath" – i.e. to report very low earnings now in order to reduce the chances of having to report losses in the future. This differs from the results of the model of section 2 in that we would expect to see a "hole" in the wage change distribution to the left of zero.

A related issue is that one may be interested in more general forms of convexity in wage cuts in the effort function, especially given the stylised evidence in favour for convexity of the utility function in losses from the literature on loss aversion (Kahneman & Tversky, 1979). Again, we might expect to observe firms "taking a bath" in a similar way to that described above. However, it should be noted that it is not immediately clear on theoretical grounds that such a convexity in losses should map from workers' utility functions into their effort functions. In particular it will depend on how workers' utility depends on effort as well as on wage changes. Whilst casual observation of the wage change distributions studied in this chapter does not seem to provide strong support to the claim that there is a hole to the left of zero in the distribution of nominal wage changes, further theoretical and empirical work may be worthwhile to assess this more formally.

Finally, our theory has neglected the possibility of DNWR motivated by factors other than worker resistance to wage cuts. Other models of DNWR have been formulated based on the legal requirement in many countries (notably excluding the US and UK) that wage contracts may only be renegotiated by mutual consent of the firm and the worker (MacLeod & Malcomson, 1993; Holden, 1994). The current chapter does not seek to deny the existence of such motivations, but merely to draw out and test the implications of behavioural foundations to DNWR. Indeed, as pointed out in Holden (2004), contract and behavioural motivations may even reinforce one another in explaining DNWR.

In addition to such theoretical issues, a number of questions arise from the empirical work of section 4. One such question is whether such findings can be explained by models of nominal rigidity other than DNWR. In particular, one may contend that a standard model of menu costs can explain the observed compression of wage increases in times of low inflation. In particular, Sheshinski & Weiss (1977) show that, in a deterministic price-setting model, increased inflation will result in more extreme price increases – as firms increase prices less often to avoid successive payment of menu costs in high inflation environments, when they do increase the price, they will increase it by more. However, if this were the correct model, we would again expect to see holes either side of zero in the distribution of nominal wage changes, and that these holes would *widen* as inflation rises. Whilst previous empirical work has found some evidence for menu cost effects, these effects have only a modest impact on wage changes around zero (Card & Hyslop, 1997), and certainly are not accentuated in times of high inflation. Moreover, in a deterministic setting with positive inflation, such as that of Sheshinski & Weiss (1977), firms always want to increase prices in the absence of menu costs. This no longer holds in an uncertain setting, as there will be situations in which the firm will wish to cut prices. In this case, it is no longer clear that the firm wants to set more extreme price increases under high inflation. The reason is that higher prices can increase the probability of wanting to cut the price in the future in an uncertain world, which is also costly in a menu cost setting. So it is by no means clear on *a priori* grounds that a menu cost model could explain the results presented in section 4.

More generally, there is a need in the literature on DNWR for an empirical model that can conform well with an explicit theory of wage setting as well as with the structure of the data. In particular, whilst the empirical methods of section 4 allow the data to speak more – by allowing different effects of inflation at different points in the wage change distribution, and by assuming nothing about the parametric form of counterfactual wage changes – they do not provide us with direct estimates that can be related back to a model of wage setting. However, the current chapter seeks to contribute to this process by showing how one can write down models of wage setting based on worker resistance to wage cuts, and by also providing empirical evidence that can inform future, more complex, models of DNWR. This will enable the formulation of more realistic structural models of DNWR that can be successfully estimated with meaningful parameter estimates.

1.6 Conclusions

This study seeks to make contributions on two outstanding issues in the literature on DNWR. In the first instance, it presents a fully explicit model of wage-setting in the presence of worker resistance to nominal wage cuts. We show that a key new insight in the context of such "behavioural" models is that nominal wage increases become partially irreversible. We then use this model to obtain testable predictions that allow us to address an outstanding issue in the literature on DNWR. In particular, we attempt to reconcile the remarkably robust micro-level evidence for DNWR across datasets and countries, with the weak evidence found for the expected macroeconomic effects. We show that the previous literature has neglected the fact that the upper tail of the distribution of wage changes will be compressed in such an environment for two related reasons. First, we find that the existence of DNWR can lead to the *active compression* of wage increases – i.e. firms pay a lower wage when they increase the wage relative to a world without DNWR. This occurs because increasing the nominal wage today raises the likelihood of having to cut the wage, at a cost, in the future. Second, we show that there will be a *latent compression* of the upper tail of the wage change distribution because the existence of DNWR raises the general level of wages in the economy, and thus firms do not have to increase the wage as often or as much in order to obtain their desired level. Thus, we argue that by neglecting these effects, previous studies could have overstated the increase in wage growth, and hence the expected macroeconomic effects, of DNWR.

Using panel data from the CPS and PSID for the US, and the NESPD for Great Britain, we find evidence that the upper tail of the wage change distribution is indeed compressed when we allow the entire distribution of wage changes to be affected by DNWR. In particular, we estimate that increased wage growth due to DNWR is no more than 0.3%, as opposed to figures of around 1% that are obtained by more naive estimates which only allow the lower tail to vary. Moreover, our results suggest that firms in practice can make significant savings relative to the costs of reduced wage cuts by compressing wage increases. In particular, we estimate that wage growth savings from compressed wage increases of at least 75% of the increase in costs due to restricted wage cuts. Thus, previous studies could have vastly overstated the wage costs of DNWR.

To further test the model of DNWR based on worker resistance to nominal wage cuts, we draw out additional predictions from the model in respect of the impact of turnover on the compression of the wage change distribution. In particular, we show that the model predicts higher turnover sectors should exhibit a reduced compression of wage increases as firms act more myopically. Again, we find robust evidence for this claim using the NESPD data, and, despite the limitations of the data, for the PSID also.

In the light of this evidence, we conclude that the increase in wage pressure due to the existence of DNWR may not be as large as previously envisaged. However, we have shown that the evidence on DNWR is consistent with a model in which workers resist nominal wage cuts along a number of dimensions. Hence, the behavioural implications of DNWR in respect of the reaction of workers to nominal wage cuts remain significant.

It should be noted, however, that it does not follow that DNWR is of little significance with respect to the workings of the macroeconomy. There may still remain important effects of nominal wage cuts on workers' productivity/effort at work. In this sense, low inflation may reduce productive efficiency in the economy. Moreover, the existence of kinked preferences implies first-order risk aversion on behalf of workers so that even small scale risk is welfare reducing (Rabin, 2000). It follows then that the welfare costs of business cycles may be much higher than previously claimed (Lucas, 1987). We leave these as topics for further research.

Tables for Chapter 1

Table 1-1: Descriptive Statistics of Wage Changes, CPS, PSID, NESPD

Year	US Data: CPI-U-X1	CPS N	CPS % Freeze	PSID N	PSID % Freeze	NES Data: RPI	N	% Freeze
1971	4.36			1,328	10.69			
1972	3.02			1,344	11.76			
1973	6.31			1,352	9.25			
1974	9.96			1,421	8.8			
1975	8.29			1,467	7.84			
1976	5.69			1,315	7.91	18.9	61,320	0.74
1977	6.4			1,321	9.16	17.5	65,802	1.5
1978	6.8			1,357	8.4	7.9	67,100	2.23
1979	9.63			1,406	7.25	10.1	66,490	2.38
1980	11.22	26,221	5.87	1,449	4.42	21.8	67,330	0.47
1981	9.48	28,981	5.95	1,534	8.02	12	71,068	2.67
1982	6.1	28,009	10.75	1,498	8.81	9.4	76,398	3.07
1983	4.18	27,131	13.11	1,456	14.9	4	78,477	2.1
1984	4.32	27,230	13.08	1,453	13.97	5.2	76,093	5.13
1985	3.56	13,386	12.51	1,519	11.92	6.9	75,737	1.73
1986	1.86	7,063	13.89	1,608	16.73	3	74,881	1.42

7	3.9	73,939	1.62
7	8	72,826	2.21
1	9.4	71,295	2.57
2	6.4	72,615	2.82
6	4.3	76,923	4.94
	1.3	78,758	7.02
	2.6	78,756	6.45
	3.3	80,281	5.65
	2.4	83,306	1.62
	2.4	81,018	1.79
	4	77,737	4.21
	1.6	77,967	4.49
	3	77,494	4.47
	1.8	80,427	0

1988	4.14	27,452	12.86	1,629	16.5
1989	4.82	27,351	12.27	1,614	15.3
1990	5.4	28,728	11.41	1,612	13.7
1991	4.21	29,392	11.87	2,240	14.8
1992	3.01	29,226	13.74	2,273	17.1
1993	2.99	29,157	13.53		
1994	2.56	27,257	12.21		
1995	2.83	10,474	12.55		
1996	2.95	8,686	11.67		
1997	2.29	26,003	10.92		
1998	1.56	25,905	10.52		
1999	2.21	26,168	10.09		
2000	3.36	25,898	9.88		
2001	2.85	25,244	9.54		
2002	1.58	27,301	10.61		

Table 1-2: Micro-Level Controls used in Addition to Age, Age^2, & Sex

CPS	PSID	NESPD
Education	Education	--
Industry (2-digit)	Industry (1-digit)	Industry (2-digit)
--	Occupation (1-digit)	Occupation (2-digit)
Region (50+ metropolitan dummies)	Region (6 dummies)	Region (10+ London dummies)
Non-white	--	--
--	--	Major Union Coverage
Public Sector	--	--
Self-employed	Self-employed	--
--	Tenure	--

Table 1-3: Regressions of Percentiles of Real Wage Changes on the Rate of Inflation and Controls (CPS, 1980 – 2002)

	Coefficient on Inflation Rate[a]					
Percentile	**No Controls**[b]		**Aggregate Controls**[c]		**Full Controls**[d]	
10th	-0.061	[0.134]	0.054	[0.143]	-0.054	[0.140]
20th	-0.257	[0.071]***	-0.165	[0.087]*	-0.167	[0.087]*
30th	-0.34	[0.073]***	-0.29	[0.073]***	-0.368	[0.067]***
40th	0.041	[0.044]	0.049	[0.048]	0.032	[0.042]
60th	0.05	[0.025]*	0.051	[0.027]*	0.044	[0.031]
70th	0.121	[0.050]**	0.121	[0.051]**	0.129	[0.055]**
80th	0.172	[0.091]*	0.191	[0.100]*	0.166	[0.104]
90th	0.148	[0.155]	0.247	[0.168]	0.349	[0.127]**
Lower Tail Losses	+1.03%		+0.42%		+1.01%	
Upper Tail Gains	-0.94%		-1.40%		-1.50%	
↑ in $\Delta\overline{w}$ due to DNWR[e]	+0.085%		-0.98%		-0.49%	

Notes:

a. Reports Least Squares estimates (weighted by region size) of real wage change percentiles on the rate of inflation and controls.
b. Includes a dummy for the years 1994 onwards to control for the increase in dispersion of real wage changes following introduction of CAPI.
c. As b, but includes additional controls for the absolute change in the rate of inflation, and the contemporaneous and lagged state unemployment rate.
d. As c, but uses real wage change percentiles re-weighted for changes in age, age^2, sex, race, region (including metropolitan dummy), 2-digit industry, education, public sector employment, and self-employment.
e. Predicted effect on real wage growth of a change in inflation from 22% (maximum NESPD sample inflation, 1980) down to 1.3% (minimum NESPD sample inflation, 1993). Computed from estimation of 99 percentile regressions of the form summarised in the Table using the method outlined in the main text.
f. Standard errors in brackets: robust to non-independence within years.
g. * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Table 1-4: Regressions Percentiles of Real Wage Changes on the Rate of Inflation and Controls (PSID, 1971 – 92)

	Coefficient on Inflation Rate[a]					
Percentile	**No Controls**		**Aggregate Controls[b]**		**Full Controls[c]**	
10th	-0.203	[0.075]**	-0.134	[0.056]**	-0.154	[0.068]**
20th	-0.617	[0.092]***	-0.579	[0.096]***	-0.586	[0.095]***
30th	0.012	[0.035]	0.008	[0.043]	-0.013	[0.041]
40th	0.024	[0.024]	0.033	[0.023]	0.027	[0.021]
60th	0.026	[0.024]	0.016	[0.024]	0.012	[0.026]
70th	0.1	[0.045]**	0.09	[0.037]**	0.037	[0.040]
80th	0.17	[0.073]**	0.137	[0.074]*	0.12	[0.066]*
90th	0.313	[0.115]**	0.333	[0.123]**	0.275	[0.132]*
Lower Tail Losses	+0.80%		+0.38%		+0.61%	
Upper Tail Gains	-0.87%		-0.73%		-0.54%	
↑ in $\Delta\overline{W}$ due to DNWR[a]	-0.063%		-0.35%		+0.069%	

Notes:

a. Reports Least Squares estimates (weighted by region size) of real wage change percentiles on the rate of inflation and controls. Observations from Alaska and Hawaii are dropped due to incomplete unemployment information before 1976.

b. Controls for the absolute change in the rate of inflation and the contemporaneous and lagged regional unemployment rate.

c. As b, but uses real wage change percentiles re-weighted for changes in age, age^2, sex, education, 1-digit industry, 1-digit occupation, region, self employment, and tenure.

d. Predicted effect on real wage growth of a change in inflation from 22% (maximum NESPD sample inflation, 1980) down to 1.3% (minimum NESPD sample inflation, 1993). Computed from estimation of 99 percentile regressions of the form summarised in the Table using the method outlined in the main text.

e. Standard errors in brackets: robust to non-independence within years.

f. * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Table 1-5: Regressions of Percentiles of Real Wage Changes on the Rate of Inflation and Controls (NESPD, 1976 – 2001)

	Coefficient on Inflation Rate[a]					
Percentile	No Controls[b]		Aggregate Controls[c]		Full Controls[d]	
10th	-0.137	[0.062]**	-0.057	[0.062]	-0.103	[0.070]
20th	-0.287	[0.033]***	-0.229	[0.018]***	-0.241	[0.020]***
30th	-0.197	[0.023]***	-0.149	[0.020]***	-0.144	[0.024]***
40th	-0.101	[0.016]***	-0.089	[0.016]***	-0.09	[0.011]***
60th	0.088	[0.008]***	0.077	[0.010]***	0.074	[0.008]***
70th	0.166	[0.016]***	0.147	[0.017]***	0.15	[0.014]***
80th	0.205	[0.028]***	0.175	[0.027]***	0.174	[0.023]***
90th	0.116	[0.053]**	0.065	[0.047]	0.1	[0.044]**
Lower Tail Losses	+1.31%		+0.74%		+1.11%	
Upper Tail Gains	-1.01%		-0.70%		-0.99%	
↑ in $\Delta\overline{W}$ due to DNWR[e]	+0.30%		+0.047%		+0.12%	

Notes:

a. Reports Least Squares estimates (weighted by region size) of real wage change percentiles on the rate of inflation and controls.
b. Includes a dummy for the year 1977 to control for the dramatic fall in real wage growth due to the incomes policies implemented in the UK at that time.
c. As b, but includes additional controls for the absolute change in the rate of inflation, and the contemporaneous and lagged regional unemployment rate.
d. As c, but uses real wage change percentiles re-weighted for changes in age, age^2, sex, region (including London dummy), 2-digit industry, 2-digit occupation, and major union coverage.
e. Predicted effect on real wage growth of a change in inflation from 22% (maximum NESPD sample inflation, 1980) down to 1.3% (minimum NESPD sample inflation, 1993). Computed from estimation of 99 percentile regressions of the form summarised in the Table using the method outlined in the main text.
f. Standard errors in brackets: robust to non-independence within years.
g. * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Table 1-6: Effect of Turnover on Percentiles of Nominal Wage Increases for Job Stayers, NESPD

Percentile	**Coefficient on Fraction of Job Changers in Occupation/Year Group**					
	(1)		**(2)**		**(3)**	
60th	0.057	[0.012]***	0.048	[0.010]***	0.026	[0.007]***
70th	0.094	[0.017]***	0.074	[0.017]***	0.043	[0.015]***
80th	0.151	[0.020]***	0.119	[0.019]***	0.091	[0.018]***
90th	0.257	[0.027]***	0.209	[0.027]***	0.222	[0.027]***
Controls	Median Wage Change 1977 Dummy		(1) + Current & Lagged Regional Unemployment Rates		(2) + Year Dummies	

Notes:

a. Report Least Squares estimates, weighted by occupation size.
b. Uses nominal wage change percentiles for job stayers, re-weighted for changes in adjusted lagged wage, age, age^2, sex, region (including London dummy), 2-digit industry, 2-digit occupation, and major union coverage.
c. Standard errors robust to heteroscedasticity.
d. * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Chapter 2

Is Downward Nominal Wage Rigidity driven by Money Illusion?

2.1 Introduction

There is now a burgeoning literature that provides evidence for the existence of downward nominal wage rigidity using micro data (for a survey, see Kramarz [2001]). In particular, much of the research identifies some intriguing properties of the cross-sectional distribution of wage changes. First, these distributions exhibit a mass-point at zero nominal wage change, indicating that there is some rigidity in nominal wages. Second, they display a relative lack of nominal wage cuts. Together, these findings have led many to conclude that nominal wages are downward rigid. These properties have been shown to be very robust, and have been replicated in a large number of studies for many countries, and across datasets within countries.

These results are of great interest to economists on a number of dimensions. First, nominal rigidities have been shown to be the key to the Phillips curve trade-off between inflation and unemployment in macroeconomics. This has led a number of researchers to study the implications of downward nominal wage rigidity (henceforth DNWR) for the conduct of macroeconomic policy (see Akerlof et al. [1996], Holden [2004], and Chapter 1 of this thesis). In this study, however, we address another key question in the light of the evidence for DNWR – what causes DNWR? In his recent book, Bewley [1999] set out to discover the answer to this question. In particular, he documents an array of interview evidence showing that wage-setters and negotiators are reluctant to cut the nominal wages of workers because they believe that wage cuts severely harm worker morale (see also the survey in Howitt [2002]). In addition to this evidence, there is a growing body of evidence that suggests people are subject to a particular form of money illusion in more general settings (Shafir, Diamond & Tversky [1997]). This research shows that agents are reluctant to accept nominal losses in a number of economic contexts. From a simple questionnaire study, Kahneman et al. [1986] report that most people prefer to receive a 5% nominal wage increase when inflation is 12%, than a 7% wage cut when there is no inflation. This is corroborated by recent evidence from an experimental pricing game (Fehr & Tyran [2001]) which demonstrates that participants are much more likely to raise their prices following a monetary expansion, than they are to cut their prices following a monetary contraction. Finally, in real-estate markets, Genesove & Mayer [2001] find evidence that condominium owners were reluctant to sell at a price below that they originally paid, *regardless of market prices.* Together, all these studies suggest a role for an asymmetric form of money illusion that we will term nominal loss aversion.

In the light of this evidence, it is tempting to conclude that the existence of downward nominal wage rigidity is simply a manifestation of nominal loss aversion in labour markets, along the lines of the evidence presented in Bewley [1999]. However, such a conclusion would be premature. An alternative set of theories has also been suggested to explain DNWR based on the form of market contracts in the presence of labour market frictions (MacLeod & Malcomson [1993]; Holden [1994]). In these models it is shown that institutional restrictions which stipulate that wage contracts may only be renegotiated by mutual consent of the firm and worker can result in fixed nominal wage contracts. In particular, it is shown that wages in these models are renegotiated up (down) only when the worker's (firm's) outside option (or equivalently strike (lockout) threat; Holden [1994]) becomes preferable. It follows that, if there exist some frictions that yield rents to the continuation of a current match, a wedge is driven between the firm's and the worker's outside option, and thus the nominal wage will remain constant for intervals of time. Moreover, the existence of inflation will mean that the worker's outside option will bind more often than the firm's, and that consequently nominal wages will be raised more often than cut. Thus, an alternative explanation based on the existence of labour market frictions can also explain the empirical evidence for DNWR. As a consequence, it is not immediately clear in practice that DNWR represents evidence of money illusion on behalf of workers. It is therefore of key importance to our understanding of both the nature of wage setting in labour markets, and more fundamentally of workers' preferences, to assess whether DNWR really is driven by money illusion.

In distinguishing between these theories, much of the discussion (if any) has come down to a fairly nuanced treatment of labour law in various countries (Malcomson [1997]). In particular, one can argue that significant evidence for DNWR has been found in countries in which renegotiation by mutual consent is not required by law (e.g. the US and Switzerland), thereby casting doubt on the relevance of contract models. However, even in the US, where employment is "at will", one can find cases where employer attempts to unilaterally alter a contract have not been upheld by courts (again, see Malcomson [1997]). Thus, it is not clear that such a discourse is likely to be fruitful in distinguishing between these theories.

In this chapter, we take a closer look at the implications of these two theories with respect to wage setting. In particular, we show that both theories predict that firms may actively reduce the nominal wage paid when they increase the wage relative to a "counterfactual" world without DNWR – what we will term "*active compression*" of wage increases. The intuition is as follows. In the "behavioural" model with money illusion, this results because raising the wage today increases the likelihood of having to cut the wage, at a cost, in the future (see Chapter 1 for an explicit model). Thus, it is optimal for firms to restrict wage increases as a precaution against future costly wage cuts. In contract models, on the other hand, such a compression can result if there exists a friction to workers in switching employers (or to striking), so that workers will accept a delay before bidding up their wage to outside levels.

We then demonstrate that the nature of this active compression of nominal wage increases in the behavioural and contract models provides us with a mechanism for differentiating between these two theories. In particular, we observe that the models' implications differ critically in the predicted impact of inflation on this upward compression. Specifically, the behavioural model implies that increased inflation reduces firms' desire to cut nominal wages, and hence relaxes the constraint of DNWR on the wage setting choice of the firm. This will in turn render the active compression of wage increases less pronounced when inflation is high in a model with money illusion. In contrast, contract models do not have this implication. The intuition for this is that, since workers' costs of switching employers are driven by *real* phenomena (e.g. search frictions), there is no reason to expect that workers' ability to bid up their wages would vary with inflation. Thus, by examining the impact of inflation on wage increases, we can determine which of these effects is likely to be present. It is this result that we argue is the key to distinguishing which theoretical foundation is driving the existence of DNWR.

In the light of this, we seek evidence for these predictions using micro-data for the US and Great Britain. We find robust evidence that the existence of DNWR leads firms to actively compress both nominal wage cuts and increases, and that the active compression of wage increases is diminished as inflation rises – precisely along the lines of a model with asymmetric money illusion. Thus, in the light of the above discussion, we argue that this represents important evidence for the behavioural motivation for DNWR, and in this sense tells us something fundamental about the nature of preferences.

However, we argue that this may not imply that contract-based models are irrelevant to the existence of DNWR. In particular, we go on to assess whether inflation-related compression of wage increases is diminished in unionised contexts. We do this because it seems more likely that the structure of contract models, with its emphasis on bargaining threats and the requirement of mutual consent to wage changes, is relevant for union wage setting. We find some modest evidence that inflation-related compression is reduced for union workers, but it is not fully offset by the existence of unions. That is, there still remains some residual compression of wage increases as inflation falls even in union contexts.

The rest of this chapter is organised as follows. Section 2 surveys results from the model of wage-setting in the presence of worker resistance to nominal wage cuts analysed in Chapter 1; section 3 then analyses the implications of contract models of DNWR; section 4 fleshes out some of the predictions of these models that we can take to the data; section 5 presents our empirical methodology and the results obtained; and section 6 concludes.

2.2 A Model of DNWR based on Money Illusion

This section briefly summarises some of the results obtained from the explicit model of downward nominal wage rigidity based on money illusion formulated in Chapter 1. In particular, in this model worker effort, e, is given by:

$$e = \ln\left(\frac{\omega}{b}\right) + c\ln\left(\frac{W}{W_{-1}}\right)\mathbf{1}^{-} \tag{2.1}$$

where W is the nominal wage, W_{-1} the lagged nominal wage, $\mathbf{1}^{-}$ an indicator for a nominal wage cut, $\omega \equiv W/P$ the real wage, and b a measure of real unemployment benefits. The parameter $c > 0$ varies the productivity cost to the firm of a nominal wage cut.

According to (2.1), workers increase their effort the higher is the real wage in work relative to real benefits out of work. This property is standard in fair-wage-effort models of the labour market (see Solow [1979], and Akerlof & Yellen [1988]). However, the key addition that captures workers' aversion to nominal wage cuts lies in the second term in (2.1). This says that workers reduce their effort discontinuously at the margin following a nominal wage cut, and that the subsequent loss in effort varies in proportion to the percentage cut in the nominal wage. This additional term is thus informed by the evidence for money illusion found in the previous literature.

Together, the properties of the effort function (2.1) can be seen in Figure 1-1 of Chapter 1. In particular, note the existence of a kink in the effort function at $W = W_{-1}$ which reflects the existence of DNWR. Another way of thinking about this is to note that the marginal effort loss of a nominal wage cut exceeds the marginal effort gain of a nominal wage increase. Chapter 1 shows that this interpretation is particularly useful as we can think of nominal wage increases as being (partially) irreversible – a nominal wage increase can only be reversed at an additional marginal cost of c. In this sense, nominal loss aversion acts an asymmetric adjustment cost on the firm's wage setting decision. Since models of asymmetric adjustment costs have been widely studied in the previous investment (Dixit & Pindyck [1994]) and labour demand (Bentolila & Bertola [1990]) literatures, much of the intuition from these literatures maps across to the current model.

Along these lines, the optimal wage setting policy of worker-firm pairs facing the effort function (2.1) turns out to take a convenient and familiar form in the presence of shocks that evolve according to a random walk:

Result 1 *If nominal shocks evolve according to a geometric random walk the optimal nominal wage policy takes the following piecewise linear form:*

$$W = \begin{cases} W^*/u_B & \text{if} \quad W^* > u_B \cdot W_{-1} \\ W_{-1} & \quad \text{otherwise} \\ W^*/l_B & \text{if} \quad W^* < l_B \cdot W_{-1} \end{cases} \tag{2.2}$$

where W^* is the nominal wage that would be paid in the absence of DNWR, and u_B, l_B are constants such that $u_B > 1 > l_B$ and $1/l_B > u_B$.

The intuition for this result is quite straightforward. First, since $l_B < 1$, it follows that firms are setting higher wages when they cut the wage than in a world without DNWR. That is, firms are actively compressing wage cuts. Clearly, this is a direct result of the fact that effort falls discontinuously at the margin following a nominal wage cut in a world with DNWR. More interestingly, the fact that $u_B > 1$ implies that firms are setting lower wages in the event that they increase the wage – i.e. they are actively compressing wage increases as well. The intuition for this is that increasing the wage today increases the probability that the firm will want to cut the wage, at a cost, in the future. Thus, firms restrict wage increases as a precaution against future costly wage cuts. Third, since $u_B > 1 > l_B$, there is a region of values for the counterfactual wage, $W^* \in [l_B \cdot W_{-1}, u_B \cdot W_{-1}]$, such that it is optimal for the firm not to change the nominal wage at all. This occurs because of the kink in the effort function at $W = W_{-1}$, which yields a familiar "region of inaction". Clearly, this region of inaction allows the model to explain the existence of the spike at zero in the distribution of nominal wage changes. Finally, the fact that $1/l_B > u_B$ implies that the compression of wage cuts will be stronger than that of wage increases. This occurs because any costs due to increasing the wage are *discounted*: first because they yield costs in the future, and second because the probability that the firm will indeed cut the wage in the future is less than one.

To see that this result can explain the observed properties of the distribution of wage changes, note that the typical firm's optimal wage setting policy (2.2) implies the following form of the log nominal wage change distribution in the behavioural model:

Result 2 *The log nominal wage change density, conditional on the lagged wage, implied by the behavioural model is given by:*

$$f(\Delta \ln W | W_{-1}) = \begin{cases} \tilde{f}(\Delta \ln W + \ln u_B | W_{-1}) & \text{if} \quad \Delta \ln W > 0 \\ \tilde{F}(\ln u_B | W_{-1}) - \tilde{F}(\ln l_B | W_{-1}) & \text{if} \quad \Delta \ln W = 0 \\ \tilde{f}(\Delta \ln W + \ln l_B | W_{-1}) & \text{if} \quad \Delta \ln W < 0 \end{cases} \tag{2.3}$$

where $\tilde{F}(\cdot | W_{-1})$ and $\tilde{f}(\cdot | W_{-1})$ are the c.d.f. and p.d.f. of the counterfactual (no DNWR) conditional log nominal wage change distribution.

Result 2 establishes clearly that the behavioural model implied that there will be a mass point at $\Delta \ln W = 0$ in the distribution of wage changes. In addition, it shows that the existence of the compression of wage cuts (through l_B) and wage increases (through u_B) will also alter the form of the wage change distribution (as illustrated in Figure 1-4 in Chapter 1). In particular, the compression of wage cuts will cause some counterfactual wage cuts to be "swept up" onto the spike at zero, and that the remainder of the left tail of wage changes will be shifted rightwards. Conversely, the compression of wage increases will lead to some counterfactual wage increases to be "swept back" onto the spike at zero, and that the remainder of the upper tail will be shifted leftwards.

The key comparative static result that we will use to differentiate this model with those based on contracts is the impact of the rate of inflation on the upper tail of the wage change distribution. In particular, a greater rate of inflation leads to a decline in the value of the parameter u_B in the behavioural model. The intuition for this is that higher inflation relaxes the constraint of DNWR on wage setting because firms can effect real reductions in labour costs without having to resort to nominal wage cuts. Thus, firms no longer need to be cautious about increasing the wage when inflation is high. We will see that this is in stark contrast to the implications of models based on contracts. The next section introduces such contract models.

2.3 The Contract-Based Approach to DNWR

A number of related theories based on the form of market contracts have been developed that can potentially explain the existence of DNWR (see MacLeod & Malcomson [1993]; Malcomson [1997]; Holden [1994], [1999], [2003]). These theories come in several forms, but share a key common implication, and so we review only one of them here: the model of general investments and switching costs of MacLeod & Malcomson [1993] (henceforth MM). The results of the model are driven by the following key assumptions:

1. Previous wage contracts may only be renegotiated by *mutual consent* of the firm and the worker.

2. The party responsible for a breakdown in trade is not verifiable.

3. There exist *frictions* that yield rents to the continuation of a current worker-firm match.

4. Firms and workers bargain over the nominal wage.

MM then show that a simple contract can ensure efficient general investments by the firm and the worker. This specifies a fixed wage such that both the firm and worker prefer to trade rather than not trade in all states of the world, which is renegotiated up (down) if and only if the worker's (firm's) outside option constraint binds.

To understand the intuition behind this result, three preliminary insights must be made. First, if the firm and the worker prefer to trade at the contract wage, then no renegotiation will occur. This is because at least one party must lose from renegotiation, and since renegotiation is by mutual consent, that party will not consent to a change. Second, if either the firm or the worker prefers not to trade *ex post*, the contract will be renegotiated according to surplus-sharing. The reason for this is that, since the identity of the cause of a breakdown in trade is not verifiable, the contract cannot specify penalties to a refusal to trade. Thus, refusal to trade can act as a credible threat during renegotiation, leading to surplus-sharing. Finally, if either party's outside option becomes preferable to the inside payoff, the wage is then renegotiated to match the outside option. Whilst this seems intuitive, one might expect surplus-sharing to obtain – as in the case with a refusal to trade. The reason this doesn't occur is that exercising an outside option leads to the immediate termination of bargaining, whereas a refusal to trade merely extends bargaining to another round. Thus, outside options do not act as a threat point during bargaining, but rather as a constraint on each party's payoff[1].

In the light of these insights, the intuition for the MM contract is as follows. First, the contract must ensure that both parties prefer to trade *ex post*, as otherwise renegotiation will result in surplus-sharing. Since surplus-sharing yields payoffs that do not reflect each party's investment, a positive probability of surplus-sharing leads to inefficient *ex ante* investments. Second, renegotiation of the wage when a party's outside option binds preserves investment incentives since the general nature of investments will ensure that outside option payoffs reflect the full marginal return on investment.

Thus, defining W and W^* as the current inside and outside nominal wage respectively, W_{-1} as the lagged nominal wage, c_W and c_F respectively as the worker's and the firm's real switching cost, and P as the current price level, a simple formalisation of this result would be:

Result 3 *An efficient contract in the context of Assumptions 1-4 is of the form:*

$$W = \begin{cases} (1) \quad W^* - P \cdot c_W & \text{if} \quad W^* - P \cdot c_W > W_{-1} \\ (2) \quad W_{-1} & \qquad \textit{otherwise} \\ (3) \quad W^* + P \cdot c_F & \text{if} \quad W^* + P \cdot c_F < W_{-1} \end{cases} \tag{2.4}$$

where c_W, c_F are real frictions and W^* is the nominal wage that would be paid in the absence of frictions.

[1]This is the result established by Binmore, Shaked & Sutton (1989).

That such a wage policy results in DNWR is typically justified as follows. First, the existence of frictions, $c_F, c_W > 0$, drives a wedge between the minimum wage the worker will accept, and the maximum wage the firm will pay. Thus, there will be intervals of time in which neither outside option will bind, and the nominal wage will remain constant in those intervals. Second, in an inflationary environment, the outside nominal wage, W^*, will increase on average over time, rendering it more likely that the worker's outside option constraint (régime (1)) binds than the firm's. Thus, nominal wages are bid up more often than they are bid down, providing a potential explanation for DNWR.

It should be noted that wage policies of the form of (2.4) have been derived via a number of characterisations for $c_F, c_W > 0$. In particular, such frictions can be justified by the existence of search costs (MacLeod & Malcomson [1993]), costly strike and lockout threats in union contexts, efficiency wages and employment protection in non-unionised contexts, and risk aversion with respect to uncertain payoffs from strikes or lockouts (Holden [1994], [2003], and [1999] respectively).

Recall that we are interested in comparing the impact of the rate of inflation on wage increases in contract models such as these with the behavioural model outlined above. In particular, section 2 showed that, in the behavioural model, firms actively compress wage increases, and that this compression declines as inflation rises. We can see from (2.4) that there may also be some upward compression of nominal wages in contract models if there exists a friction to the worker (i.e. if $c_W > 0$) so that workers wait before bidding up their nominal wage. How would we expect this compression to change as inflation rises?

To compare more directly the implications of the wage policy summarised in (2.4) with that derived from the behavioural model (2.2), let us further assume that the frictions c_F and c_W are given by[2]:

$$c_F = \gamma_F \cdot \frac{W^*}{P}; \ \ c_W = \gamma_W \cdot \frac{W^*}{P} \tag{2.5}$$

Such an assumption can be motivated by the idea that higher outside real wages reflect more productive matches for the employer or employee, which require greater search costs. Under this assumption, (2.4) can then be written as:

$$W = \begin{cases} (1) \ \ W^*/u_C & \text{if} \ \ W^* > u_C \cdot W_{-1} \\ (2) \ \ W_{-1} & \text{otherwise} \\ (3) \ \ W^*/l_C & \text{if} \ \ W^* < l_C \cdot W_{-1} \end{cases} \tag{2.6}$$

where $u_C \equiv \frac{1}{1-\gamma_W} > 1 > \frac{1}{1+\gamma_F} \equiv l_C$. Notice that this is of the exact same structure as (2.2), and thereby allows greater comparison across the two models. In particular, note that Result 2 will hold again here, yielding a similar "dual censoring" implication for the distribution of log wage changes. However, a key difference that we shall emphasise between the two solutions is that the compression parameters in the contract model, u_C and l_C, are predicted to be *unrelated* to the rate of inflation.

The key intuitive insight at this point is that, in these contract models, the only way a worker can obtain a wage increase is via exercising a credible threat, either to permanently leave the firm (i.e. taking an outside option), or to impose costs on the firm by striking etc. Thus, for increases in the rate of inflation to reduce the compression of wage increases, u_C, one would have to argue that workers' threats become more credible as inflation rises. In the model summarised in Result 3, there is no sense in which this can happen as the worker-side friction, c_W, reflects the *real*

[2]This assumption does not affect the subsequent qualitative statements, and is employed only in order to sharpen comparison with the behavioural model.

costs of search and work disruption, and these costs are therefore invariant to the rate of inflation. Therefore active compression of wage increases is not predicted to fall as inflation rises in these contract models. Note that this is in direct opposition to the implications of the model of DNWR informed by money illusion detailed earlier.

Moreover, this result is unlikely to be altered by the introduction of forward looking dimensions to these contract models. In particular, MacLeod & Malcomson [1993] (Proposition 10) show that a wage policy analogous to (2.4) will also ensure efficient investments in a multi-period version of their model, except that wage increases (decreases) are instead determined by the present discounted value of the worker's (firm's) outside options. By a completely analogous logic to that above, it can be seen that, since workers' outside options are unrelated to inflation, the compression of wage increases will also remain invariant to the rate of inflation.

2.4 Predictions

In this section we draw out precise predictions from the two models reviewed above that can be used to verify empirically whether a model of money illusion can explain DNWR.

An important first point to note is that, in order to differentiate the effects of money illusion on wage setting, we need to concentrate on the properties of the distribution of nominal wage changes *conditional* on the lagged nominal wage. The reason for this is that the lagged wage is known at the time of setting the current wage (it is a state variable), and hence all wage policies will be functions of the lagged nominal wage. To see this in practice, simply observe that the wage policies in both behavioural and contract models, (2.2) and (2.4) respectively, depend explicitly on the lagged wage. Since all previous empirical studies of DNWR concentrate on the unconditional distribution, this is an important note to bear in mind.

The previous sections argue that we can distinguish between contract and behavioural models of DNWR by observing the properties of the *upper tail* of this conditional distribution, $f(\Delta \ln W|W_{-1})$. In particular, we have shown that the key difference between these models lies in the impact of increased inflation on the active compression of wage increases relative to a counterfactual world with no DNWR. In the behavioural model, increased inflation reduces this compression because it reduces the likelihood that firms will want to cut nominal wages, and so the firm doesn't have to worry about the effect of current wage increases on the future costs of wage cuts. Thus, if the behavioural model is correct, we would expect to observe the upper tail of $f(\Delta \ln W|W_{-1})$ becoming more dispersed as inflation rises. However, we have shown in section 3 that contract models do not have this implication because workers' costs of switching employers are unrelated to inflation.

In the forthcoming empirical analysis, we will look at percentiles of the log wage change distribution to get an impression of whether there is an active compression of wage increases that is related to the rate of inflation. We can use our derivations in (2.3) to obtain precise predictions as to what we might expect to observe in the behavioural and contract cases respectively. In particular, note that the conditional c.d.f. of nominal wage increases is given by:

$$F(\Delta \ln W|W_{-1}, \Delta W > 0) = \tilde{F}(\Delta \ln W + \ln u_i|W_{-1}), \text{ for } i = \{B, C\} \tag{2.7}$$

and that the nth percentile in the domain of nominal wage increases, $P_n(\Delta \ln W|W_{-1})$, solves:

$$\tilde{F}[P_n(\Delta \ln W|W_{-1}) + \ln u_i|W_{-1}] = \frac{n}{100} \tag{2.8}$$

Thus:

$$P_n\left(\Delta \ln W | W_{-1}\right) = -\ln u_i + \tilde{F}^{-1}\left(\frac{n}{100}\right) = -\ln u_i + \tilde{P}_n \tag{2.9}$$

Thus, for all wage change percentiles in the domain of wage increases, both models imply that there will be a decrease in the percentile by a constant $\ln u_i$ relative to the counterfactual percentile, $\tilde{P}_n$. Thus, from the above discussion, we would expect these positive percentiles to increase as inflation rises in the behavioural model. In contract models, on the other hand, we would expect to see if anything the opposite. We will explore whether this prediction is verified in our empirical section.

2.5 Empirical Implementation

In this section, we assess the evidence for the predictions made in section 4 using data from the US and Great Britain. In particular, we use the Current Population Survey (CPS) and Panel Study of Income Dynamics (PSID) data for the US, and the New Earnings Survey Panel Dataset (NESPD) for Great Britain. The samples used are identical to those used in Chapter 1, so we review them only briefly here.

The CPS data are taken from the Monthly Outgoing Rotation Group samples for 1979–2002. We longitudinally match respondents using a method similar to that advocated by Madrian & Lefgren [1999] to yield a sample of around 25,000 wage change observations per year. Problems arise in the CPS data due to a change to a computer-aided survey design in 1994 (known as CAPI). This led to a sharp increase in the dispersion of wage changes from 1994 onwards. We control for this by including a dummy variable equal to one for all years from 1994 onwards into our econometric specifications. The PSID data are simply extracted from the random (not poverty) samples from 1970–1992. Finally, the NESPD data used run from 1975–2001 and provide around 70,000 wage change observations for each sample year.

Since much of the NESPD data are drawn from employer payroll records, the data are less subject to measurement error concerns (see Nickell & Quintini [2003] for more on this). A problem arises with the NESPD data for 1977 as a result of the wages policies instituted by the government in the UK at the time. In particular, this led to median real wage growth of -7.5% in 1977 when median real wage growth in all other sample periods was almost never negative. We control for this by including a dummy variable for 1977 in all empirical specifications for this data.

Empirical Method

Recall that section 4 provided us with predictions about the impact of higher rates of inflation on the dispersion of the upper tail of the wage change distribution *conditional* on the lagged wage. Thus an important empirical step is to identify variation in this conditional distribution. Clearly it will not be possible to obtain an exact empirical counterpart to this distribution, as in all probability the workers in our samples will have been paid different lagged wages. However, we can get close to this distribution if we can control for changes in the distribution of lagged wages. To see this, note that we can decompose the observed wage change distribution according to:

$$f(\Delta \ln W) = \int f(\Delta \ln W | W_{-1})\, dF(W_{-1}) \tag{2.10}$$

By controlling for changes in the distribution of lagged wages, $F(W_{-1})$, we can infer that any variation in the resulting distribution is due to variation in the conditional distribution, $f(\Delta \ln W | W_{-1})$. This is exactly the variation we require to differentiate between contract and behavioural foundations for DNWR. We identify this variation by employing the technique of DiNardo, Fortin & Lemieux [1996] (henceforth DFL).

In particular, we define a set of micro-level covariates (to be described shortly) that we would like to control for in estimating the impact of inflation on the wage change distribution, x, which includes a measure of lagged wages. DFL then show that one can obtain the distribution that would have resulted if the distribution of x had not changed over time by simply re-weighting the observed distribution. Specifically, we can re-weight all distributions so that the relevant time distribution of the xs, t_x, is equal to that in some comparison period, T (this will be the final sample date in our specifications). Thus, defining the log wage change as Δw, we derive:

$$f(\Delta w; t_x = T) = \int f(\Delta w|x)\, dF(x|t_x = T) = \int f(\Delta w|x) \cdot \psi \cdot dF(x|t_x = t) \quad (2.11)$$

The weights ψ are given by:

$$\psi = \frac{dF(x|t_x = T)}{dF(x|t_x = t)} = \frac{\Pr(t_x = T|x)}{\Pr(t_x = t|x)} \cdot \frac{\Pr(t_x = t)}{\Pr(t_x = T)} \quad (2.12)$$

where the second equality follows from Bayes' Rule. The conditional probabilities on the RHS of (??) can then be simply estimated using a Probit model.

Following this procedure, along the lines of (2.9), we test the predictions of section 4 by running regressions of the form:

$$P_{n_{rt}} = \beta_{0n} + \beta_{1n} P_{50_{rt}} + \eta_n \pi_t + z'_{rt} \gamma_n + \varepsilon_{rt} \quad (2.13)$$

$P_{n_{rt}}$ is the nth percentile of the *re-weighted* distribution of wage changes in region r at time t. π_t is the rate of inflation at each date. This is measured by the CPI-U-X1 index for the US, and the Retail Price Index for the UK. Finally z_{rt} is a set of aggregate control variables, to be defined shortly.

An issue arises when implementing this procedure, however, because the level of lagged wages has increased dramatically over the sample period with inflation and productivity growth, rendering lagged nominal wages at different dates potentially non-comparable. We address this issue by first adjusting the lagged nominal wage in each date t for price inflation to obtain equivalent lagged wages in our base year (date T) dollars/pounds. We then make a further adjustment by taking out a time trend[3]. Since all we want to control for is any changes in the distribution of lagged wages that may be due to the existence of DNWR, this is legitimate if we are willing to believe that DNWR has no effect on either the price level or trend real wage growth. Whilst the former seems reasonable, one might worry about the latter claim. However, Chapter 1 shows that, as both a theoretical and an empirical issue, DNWR has little effect on aggregate wage growth. This suggests that assuming DNWR has no effect on trend real wage growth is not an unrealistic assumption.

Controls

As mentioned previously, there are two levels at which we control for additional covariates in our empirical method. First, we control for micro-varying variables using the DFL technique. These controls are summarised for our three datatsets in Table 1-1 of Chapter 1, and we provide some motivation for these choices here.

First, we control for a quadratic in age and gender. This is simply to take into account demographic trends in the labour markets of the US and Britain. We also include, where possible, measures of education. This is motivated by the idea that more educated workers have become increasingly in demand with processes such as skill-biased technical change in the labour market. Thus the education composition of the workforce might affect the distribution of wage changes as demand polarises across skills, leading to, for instance, more extreme wage increases for the high-skilled. We

[3]In particular, we use residuals from an OLS regression of lagged log real wages on a quadratic time trend. Alternative methods of detrending, including the use of data on GDP per hour, and linear detrending were tried with little difference in results.

also control for industry, occupation, and region to purge any effects due to changing sectoral composition of the workforce, or regional migration over time.

In addition to these, we control for a set of aggregated controls by adding regressors to the RHS of (2.13). In particular, we include current and lagged regional unemployment rates to control for any unemployment effects of DNWR, leading to observations "leaving" the wage distribution. Moreover, we include a measure of the absolute change in the rate of inflation. This is to control for any increase in the dispersion of wage changes due to increased volatility of inflation (as suggested by Groshen & Schweitzer [1999]).

Results

Given these, we then perform regressions of the form in (2.13) using percentiles of the wage change distributions, re-weighted for the controls listed in Table 1-1, as well as adjusted lagged wages, using the DFL technique. The results are summarised in Tables 2-1–2-3. Table 2-1 presents the results for the CPS. It can be seen that inflation has a negative impact in the 20th and 30th percentiles, and a positive impact thereafter. Moreover, these effects are highly significant for the 20th and 30th percentiles in the lower tail, and the 70th–90th percentiles in the upper tail. Thus, lower wage change percentiles which represent negative wage changes become less negative as inflation falls, indicating a compression of wage cuts. In addition, however, it also implies that positive wage change percentiles become less positive as inflation falls. Thus, lower inflation leads to an active compression of the upper tail of the wage change distribution. Recall from the discussion of section 4 that this is predicted by a model of DNWR based on money illusion, but not from a model based on contracts.

Similar, if a little less significant, results are obtained for the PSID data in Table

2-2. Here we observe negative inflation effects in the 10th and 20th percentiles in the lower tail, and positive inflation effects in the 80th and 90th percentiles of the upper tail. In particular, the lower tail effect is very significant at the 20th percentile, in contrast to the CPS results. This is most likely due to the fact that a zero nominal wage change occurs further into the lower tail for the PSID as a result of the relatively higher rates of inflation experienced in the PSID sample period. However, the upper tail effects are again significant, and comparable to those found in the CPS. Again, this is in line with the predictions of the behavioural model, outlined in section 4, suggesting that at least some of the observed rigidity reflects money illusion.

Finally, Table 2-3 presents the results for the NESPD data. These represent the most significant estimates out of all the three datasets used in this study. This is likely to be a result of the relatively larger sample sizes, reduced levels of measurement error, and greater variation in the rate of inflation over the sample period for the NESPD compared to the CPS and PSID. Again, we see negative and significant inflation effects for the 20th–40th percentiles of wage changes, and positive and highly significant inflation effects in the 60th–90th percentiles.

Altogether, these results provide strong evidence from a range of datasets, with a number of controls, that lower inflation leads to an *active compression* of wage *increases* as well as decreases. This is precisely in line with the predictions of the behavioural model based on money illusion, and thus provides support to the hypothesis that DNWR is motivated by worker resistance to nominal wage cuts.

2.5.1 Is Money Illusion Important in Union Contexts?

The above results provide robust evidence that can be explained in the context of a model of money illusion for the wage change distribution of all workers. In this section, we delve a little deeper to see if our conclusions change when we concentrate on workers covered by union contracts. The reason we concentrate on unionised workers is because one might expect that the contract based models provide a more accurate description of wage setting in union contexts, with their emphasis on bargaining threats, and mutual consent to the alteration of existing wage contracts. Indeed, Holden [2004] argues that DNWR due to contract based considerations is likely to be more prominent in unionised sectors because firms need only replace one worker to achieve a reduction in that worker's wage in a decentralised context; in a union context, the firm would have to replace all workers, which is arguably much more costly.

To this end, Figure 2-1 illustrates density estimates of log nominal wage changes for union vs. non-union workers using the NESPD data. In particular, we divide the sample period into 3 sub-periods based on the rate of inflation: a high inflation period from 1976–81, a mid-inflation period from 1982–91, and a low inflation period from 1992–99. If the predictions of contract models are correct, we would expect that non-union wage increases will become compressed as inflation falls, but that union wage increases should be unchanged or less compressed. However, it can be seen from Figure 2-1 that wage increases become compressed in a very similar manner even for workers covered by union contracts. Moreover, it appears to be the case that, if anything, DNWR is less prominent in the union sector, in contrast to the predictions of the contract models. Thus, in this simple comparison, there seems to be evidence for money illusion in the union sector.

To address this question more formally, we adapt our percentile regression equation (2.13) to see if the active compression of wage increases is mitigated as unionisation increases. In particular, we add an interaction term on the RHS of (2.13) that

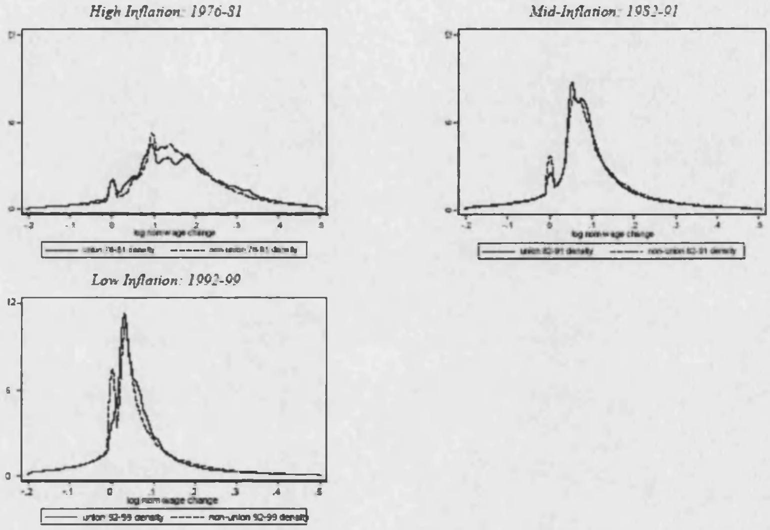


Figure 2-1: Union vs. Non-Union Distribution of Wage Changes, NESPD.

measures how the impact of inflation varies as unionisation increases:

$$P_{n_{rt}} = \beta_{0n} + \beta_{1n} P_{50_{rt}} + \eta_n \pi_t + \iota_n (\boldsymbol{\pi}_t \cdot \mathbf{u}_{rt}) + z'_{rt}\gamma_n + \varepsilon_{rt} \tag{2.14}$$

Here, $\mathbf{u}_{rt}$ is the fraction of union workers in region r at date t; all other variables are as before. Thus, if unionised workers experience reduced compression of wage increases as inflation falls, we would expect the coefficient on the interaction term, ι_n, to be negative.

Table 2-4 reports the results of this estimation. It can be seen that we find some evidence that greater unionisation leads to a diminished compression of wage increases as inflation falls, as indicated by negative parameter estimates for ι_n. Thus, more formal estimates provide evidence that suggests that contract based models may have greater predictive power in union contexts. However, in general these effects are not significant, except for at the 70th percentile. Moreover, comparing these estimates with the coefficients on the rate of inflation suggests that the magnitudes of the coefficients ι_n are not sufficient to cancel out the overall effect of inflation. This implies that there still remains some compression of wage increases in union sectors that is related to inflation.

2.6 Conclusions

By looking closely at the implications of the two proposed explanations of DNWR, this chapter has presented a method for differentiating between the behavioural theory for DNWR, and its alternative theories based on the form of market contracts (MacLeod & Malcomson [1993]). In particular, we have shown that the behavioural theory implies that the active compression of nominal wage increases will fade as the rate of inflation rises: the greater the rate of inflation, the less firms have to worry about the costs of wage cuts in the future, and the more they will feel at liberty to raise the wage today. This is in contrast to the predictions of contract models. In these models, workers can only bid up their wage by making a credible and costly threat to change employers or to strike etc. However, in these models, the costs of switching employers or striking are, critically, driven by *real* economic forces.. Thus, we would not expect a compression of wage increases related to inflation in these contract models, in direct contrast to the predictions of the behavioural model.

Using micro data for the US and Great Britain, we provide strong evidence that firms do indeed actively compress wage increases, and that this compression is reduced as inflation rises. Moreover, this is robust to controls for a number of characteristics, as well as to alternative theories for the effect of inflation on the distribution of wage changes. It follows that an important aspect of the observed properties of the distribution of wage changes can be explained simply in the context of a model with money illusion, and cannot be easily explained by contract models. In this way, it suggests that the subjective evidence presented by Shafir, Diamond & Tversky [1997] and Kahneman, Knetsch & Thaler [1986], the experimental evidence of Fehr & Tyran [2001], and the real estate evidence of Genesove & Mayer [2001] on nominal loss aversion are all corroborated by the evidence on DNWR. Thus, the evidence presented here provides yet another dimension to the evidence on money illusion, and reinforces a view of preferences in which nominal losses loom large. Thus, we conclude that behavioural concerns must play an important role in wage-setting.

However, it does not necessarily follow that contract models are completely irrelevant. In particular, we find some weak evidence that the compression of wage increases related to inflation is less prominent in unionised contexts. It is arguable that wage setting in unionised sectors is more closely related to the bargaining models that underlie contract models of DNWR. Thus, these findings present suggestive evidence that contract models may be relevant where unions are pervasive. Moreover, it may well be the case that money illusion and contract considerations are mutually reinforcing, as suggested by Holden [2004].

Thus, the results of this chapter should not be taken as a direct refutation of contract models, but rather as a statement that the motivations of the behavioural model have an appreciable bearing on the existence of DNWR. In this sense, the evidence presented in this chapter adds to the growing body of evidence for money illusion in the form of nominal loss aversion, and thus tells us something fundamental about the form of preferences.

Tables for Chapter 2

Table 2-1: Percentile Regressions of Real Wage Changes, including Controls for Adjusted Lagged Wages (CPS, 1980 – 2002).

Percentile	10th	20th	30th	40th	60th	70th	80th	90th
Inflation Rate	0.066 [0.127]	-0.141 [0.078]*	-0.307 [0.071]***	0.038 [0.050]	0.046 [0.028]	0.146 [0.060]**	0.269 [0.110]**	0.447 [0.154]***
Median	0.829 [0.329]**	0.75 [0.177]***	0.439 [0.096]***	0.775 [0.033]***	1.094 [0.032]***	1.311 [0.053]***	1.637 [0.110]***	2.108 [0.220]***
1(year>=1994)	-0.11 [0.009]***	-0.047 [0.005]***	-0.003 [0.002]	0 [0.001]	0.004 [0.002]**	0.012 [0.004]***	0.032 [0.007]***	0.067 [0.013]***
State U/E Rate	-0.01 [0.004]**	-0.006 [0.002]**	-0.001 [0.001]	-0.001 [0.000]***	0 [0.000]	0.001 [0.001]	0.003 [0.002]	0.005 [0.005]
State U/E Rate$_{-1}$	0.008 [0.004]**	0.005 [0.002]**	0.001 [0.001]	0.001 [0.000]	0 [0.000]	-0.001 [0.001]	-0.002 [0.002]	-0.003 [0.004]
|Change in CPI|	0.725 [0.492]	0.235 [0.283]	-0.217 [0.199]	0.077 [0.060]	0.033 [0.069]	-0.086 [0.187]	-0.365 [0.382]	-0.83 [0.587]
Constant	-0.232 [0.007]***	-0.082 [0.004]***	-0.031 [0.005]***	-0.026 [0.003]***	0.027 [0.002]***	0.064 [0.005]***	0.125 [0.008]***	0.261 [0.014]***
Observations	1107	1107	1107	1107	1107	1107	1107	1107
R-squared	0.36	0.26	0.42	0.65	0.8	0.58	0.42	0.32

Notes:

a. Reports Least Squares estimates (weighted by region size) of real wage change percentiles on the rate of inflation and controls.
b. Uses real wage change percentiles re-weighted for changes in **adjusted lagged wage**, age, age^2, sex, race, region (including metropolitan dummy), 2-digit industry, education, public sector employment, and self-employment.
c. Standard errors in brackets: robust to non-independence within years.
d. * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Table 2-2: Percentile Regressions of Real Wage Changes, including Controls for Adjusted Lagged Wages (PSID, 1971 – 92).

Percentile	10th	20th	30th	40th	60th	70th	80th	90th
Inflation Rate	-0.063 [0.065]	-0.569 [0.098]***	0.012 [0.045]	0.037 [0.021]*	0.032 [0.023]	0.072 [0.042]	0.176 [0.072]**	0.402 [0.157]**
Median	0.879 [0.150]***	0.278 [0.074]***	0.879 [0.046]***	1.001 [0.017]***	1.05 [0.025]***	1.113 [0.038]***	1.185 [0.062]***	1.279 [0.151]***
State U/E Rate	-0.465 [0.197]**	0 [0.120]	-0.135 [0.087]	-0.114 [0.039]***	0.121 [0.069]*	0.102 [0.123]	0.034 [0.160]	-0.495 [0.375]
State U/E Rate$_{-1}$	0.648 [0.193]***	0.044 [0.114]	0.052 [0.074]	0.057 [0.050]	-0.138 [0.057]**	-0.239 [0.111]**	-0.365 [0.153]**	-0.005 [0.343]
|Change in CPI|	0.071 [0.210]	-0.078 [0.154]	0.17 [0.099]	0.021 [0.049]	-0.073 [0.067]	-0.145 [0.117]	-0.002 [0.182]	0.195 [0.452]
Constant	-0.104 [0.012]***	-0.021 [0.005]***	-0.035 [0.006]***	-0.016 [0.002]***	0.017 [0.002]***	0.048 [0.005]***	0.09 [0.008]***	0.158 [0.020]***
Observations	88	88	88	88	88	88	88	88
R-squared	0.47	0.73	0.77	0.94	0.95	0.81	0.62	0.31

Notes:

a. Reports Least Squares estimates (weighted by region size) of real wage change percentiles on the rate of inflation and controls. Observations from Alaska and Hawaii are dropped due to incomplete unemployment information before 1976.
b. Uses real wage change percentiles re-weighted for changes in **adjusted lagged wage**, age, age^2, sex, education, 1-digit industry, 1-digit occupation, region, self employment, and tenure.
c. Standard errors in brackets: robust to non-independence within years.
d. * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Table 2-3: Percentile Regressions of Real Wage Changes, including Controls for Adjusted Lagged Wages (NESPD, 1976 – 1999).

Percentile	10th	20th	30th	40th	60th	70th	80th	90th
Inflation Rate	-0.111	-0.228	-0.162	-0.103	0.084	0.158	0.247	0.383
	[0.069]	[0.020]***	[0.021]***	[0.009]***	[0.008]***	[0.014]***	[0.033]***	[0.098]***
Median	0.983	0.756	0.793	0.866	1.09	1.183	1.266	1.516
	[0.162]***	[0.069]***	[0.046]***	[0.021]***	[0.024]***	[0.045]***	[0.081]***	[0.357]***
1(year=1977)	0.028	0.003	-0.01	-0.008	0.007	0.014	0.006	0.012
	[0.015]*	[0.007]	[0.004]***	[0.002]***	[0.002]***	[0.004]***	[0.006]	[0.021]
Region U/E Rate	0	-0.003	0	0	-0.001	-0.002	-0.004	-0.007
	[0.002]	[0.001]**	[0.001]	[0.001]	[0.000]*	[0.001]**	[0.001]***	[0.002]***
Region U/E Rate$_{-1}$	0.003	0.004	0.001	0	0.001	0.001	0.002	0.003
	[0.002]	[0.001]***	[0.001]	[0.001]	[0.000]	[0.001]	[0.001]	[0.002]
|Change in RPI|	0.06	0.041	-0.05	0.004	0.03	0.019	-0.062	-0.43
	[0.138]	[0.047]	[0.041]	[0.027]	[0.018]	[0.036]	[0.070]	[0.158]**
Constant	-0.121	-0.038	-0.013	-0.004	0.012	0.034	0.077	0.172
	[0.009]***	[0.003]***	[0.002]***	[0.001]***	[0.001]***	[0.002]***	[0.004]***	[0.011]***
Observations	240	240	240	240	240	240	240	240
R-squared	0.64	0.91	0.96	0.98	0.99	0.95	0.87	0.47

Notes:

a. Reports Least Squares estimates (weighted by region size) of real wage change percentiles on the rate of inflation and controls.
b. Uses real wage change percentiles re-weighted for changes in **adjusted lagged wage**, age, age^2, sex, region (including London dummy), 2-digit industry, 2-digit occupation, and major union coverage.
c. Standard errors in brackets: robust to non-independence within years.
d. * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Table 2-4: Percentile Regressions including Interaction between Inflation & Unionisation (NESPD, 1976 – 1999).

Percentile	60th	70th	80th	90th
Inflation Rate	0.121 [0.037]***	0.302 [0.062]***	0.39 [0.122]***	0.717 [0.271]**
Inflation*Union	-0.062 [0.060]	-0.241 [0.091]**	-0.24 [0.215]	-0.56 [0.545]
Median	1.097 [0.024]***	1.211 [0.045]***	1.294 [0.075]***	1.581 [0.337]***
1(year=1977)	0.008 [0.002]***	0.017 [0.004]***	0.009 [0.006]	0.02 [0.018]
Region U/E Rate	-0.001 [0.000]*	-0.002 [0.001]**	-0.004 [0.001]***	-0.006 [0.002]**
Region U/E Rate$_{-1}$	0.001 [0.000]	0.001 [0.001]	0.002 [0.001]	0.004 [0.002]
|Change in RPI|	0.029 [0.019]	0.017 [0.034]	-0.065 [0.071]	-0.437 [0.163]**
Constant	0.011 [0.002]***	0.03 [0.003]***	0.074 [0.005]***	0.163 [0.012]***
Observations	240	240	240	240
R-squared	0.99	0.96	0.87	0.47

Notes:

a. Reports Least Squares estimates (weighted by region size) of real wage change percentiles on the rate of inflation and controls.
b. Uses real wage change percentiles re-weighted for changes in **adjusted lagged wage**, age, age^2, sex, region (including London dummy), 2-digit industry, 2-digit occupation, and major union coverage.
c. Standard errors in brackets: robust to non-independence within years.
d. * significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level.

Chapter 3

Dynamic Factor Demand with Irreversibility: A Discrete Time Solution

3.1 Introduction

It has long been recognised in economics that the accumulation of capital and labour is subject to adjustment costs (see Oi [1962] for an early study in labour markets, and Arrow [1968] for a treatment of the investment case). In particular, much interest has focussed on the idea that investment and labour demand are subject to asymmetric adjustment costs that render factor demand *irreversible* to some degree. In the case of labour markets, these occur if there exist firing costs that make it difficult to reverse employment decisions. In the investment literature emphasis has been on the existence of installation and moving costs of machinery, capital specificity, and potential "lemons" problems in the market for second hand capital. It is argued that these costs drive a wedge between the purchase and resale prices of capital, and thereby render investment to some extent irreversible.

These observations have led to the development of a class of theoretical models which seek to characterise the effects of such irreversibility on optimal factor demand. To date, the vast majority of such models has adopted a particular specification in which time advances continuously and the shocks facing a firm evolve according to a Brownian motion (see Dixit & Pindyck [1994] for the investment case, and Bentolila & Bertola, [1990] for the labour demand case). This assumption has been used primarily for reasons of analytical tractability, but might be considered to be quite restrictive. In particular, it can be shown that a Brownian motion is the continuous time analogue to a discrete time random walk (Dixit [1993]).

This chapter seeks make a simple contribution to our understanding of such models by analysing the impact on factor demand of alternative assumptions about the structure of shocks facing a firm. To do this, we first formulate and solve a general model of partial irreversibility in discrete time. We find that we can characterise the solution in a manner analogous to, but more general than, existing Brownian models. In particular, the key analytical challenge in all such models is to characterise the marginal impact of current factor demand decisions on the future profits of a firm ("marginal q" in investment models). We show that this marginal value has a recursive structure which can be expressed in the form of a contraction mapping in discrete time. This contraction is analogous to the familiar differential equations obtained for the marginal value function in previous models that utilise Brownian motion. However, these differential equations may only be derived once one has assumed a structure for the shocks facing a firm; the contraction property derived in this study holds for a wider class of shocks. This greater generality thus allows a more comprehensive look at the implications of different shock processes.

Factor irreversibility in discrete time has been considered previously by, among others, Lucas & Prescott [1971]. The closest in spirit to the current chapter, however, is Sargent [1980], which considers the implications of *complete* investment irreversibility, in which it is infinitely costly to disinvest. In contrast, the model presented in this chapter addresses the more general case of *partial* irreversibility, in which disinvestment (or firing in the labour case) is costly, but not infinitely so. More fundamentally, though, the contraction property of the marginal value function characterised in this chapter has not been established previously in the literature on factor irreversibility, and significantly simplifies analysis of the discrete time case. Moreover, we show that this contraction property is analytically useful in the context of specific assumptions about the form of shocks facing the firm.

To this end, we apply this general contraction result to illustrate the impact of some simple alternative specifications for the evolution shocks. In particular, we first present the solution to the discrete time model in the face of a geometric Gaussian random walk. As explained above, this is the analogue to the common assumption of geometric Brownian motion in the previous literature, and can thus be used as a benchmark to assess the impact of alternative assumptions. We then analyse in turn the impact of simple alternative assumptions about the persistence of shocks, and the functional form of the distribution of shocks. We first derive a general result that increased persistence of shocks leads to a greater response of factor demand to current shocks. The intuition for this is that, when shocks are more persistent, current shocks become more informative about future shocks. Thus, the firm has to worry less about any costly reversals of current factor demand decisions. We then go on to verify that this is the case in practice by solving the model with i.i.d. Gaussian shocks. From this, we additionally find that reduced persistence in the form of i.i.d. shocks leads to greater rigidity in factor demand. Intuitively, i.i.d. shocks imply some mean-reverting aspect to shocks, which leads firms to worry more about having to reverse current factor demand decisions, at a cost, in the future. Finally, we examine the impact of distributional form by solving the model in the presence of exponential shocks. We find that the fatter tails implied by exponential shocks leads the firm to reduce demand more for bad shocks, and to increase demand more for good shocks.

It is also worth noting, however, that these findings are also useful for a number of other reasons. In particular, more generally the model provides guidance to the solution of any optimisation problem with a kinked objective function. This has been shown to have important applications in models of loss aversion as applied to downward nominal wage rigidity (see Chapters 1 and 2 of this thesis). Finally, by adopting the familiar techniques and findings of discrete time dynamic programming (Stokey & Lucas [1989]), providing a discrete time solution may open up the research agenda on adjustment costs to more researchers by no longer requiring investment into the concepts and methods of stochastic calculus.

The remainder of this chapter is organised as follows. Section 2 presents a general result on the structure of optimal factor demand in the face of partial irreversibility. Section 3 then proceeds to apply this general result to different assumptions about the evolution of firm shocks, in particular by considering different persistence and distributional assumptions. Section 4 concludes.

3.2 A General Result

We assume that the firm uses a single factor of production, x, to produce output, y, according to the production function:

$$y = g(x, A) \tag{3.1}$$

where $g_A > 0$, $g_x > 0$, and $g_{xx} < 0$. In addition, we assume that $g_{xA} > 0$ so that the demand for x is increasing in A. A is a shock variable that can be thought of as a combination of demand and/or technology shocks and acts as the source of uncertainty in the model. We make the standard assumption that A follows a Markov process, with c.d.f. given by $F(A'|A)$.

x can be thought of as either capital or labour. Accumulation of x comes at a cost $b_u > 0$ per unit; decumulation provides a return b_l which may be greater or less than zero depending on the case in hand. We assume that accumulation of x is irreversible, so that $b_u > b_l$. In the case where x is capital, b_u and b_l can be thought of as the purchase and resale prices of capital respectively, where b_l is typically positive. That the purchase price may exceed the resale price of capital has typically been justified by the existence of a lemons problem in the market for used capital (see Arrow [1968], for the original intuition, and Abel & Eberly [1996], for a Brownian model with this property). In the case where x is labour, b_u and $-b_l$ can be thought of as the per worker costs of hiring and firing respectively, where b_l is now assumed to be negative.

In addition to these adjustment costs, we allow for the existence of a per unit cost of maintaining and running x, which we denote w. Clearly, in the labour demand case this has the obvious interpretation of wage costs. Finally, we assume that a proportion δ of x depreciates each period, so that we can describe the evolution of x by:

$$\Delta x = \Delta X - \delta x_{-1} \tag{3.2}$$

where ΔX represents the change in x gross of depreciation. In the capital case, ΔX is just the total purchases less sales of capital, and δ is simply the fraction of capital that depreciates each period. In the labour case, ΔX is the number of hires less fires, and δ can be interpreted as the fraction of workers that quits each period.

The firm seeks to maximise the expected present discounted value of profits. Thus, defining the firm's discount factor, $\beta \equiv \frac{1}{1+r}$, where r is the real rate of interest, $\mathbf{1}^{+}$ and $\mathbf{1}^{-}$ as indicators for $\Delta X > 0$, respectively $\Delta X < 0$, we can write down the firm's objective as:

$$v(x_{-1}, A) \equiv \max_{x} \left\{ \begin{array}{c} g(x, A) - wx - b_u \Delta X \mathbf{1}^{+} - b_l \Delta X \mathbf{1}^{-} \\ +\beta \int v(x, A') \, dF(A'|A) \end{array} \right\} \tag{3.3}$$

To complete our characterisation of the firm's problem, we make the standard assumption that the measure $dF(A'|A)$ satisfies the Feller property, so that the functional mapping defined in (3.3) preserves continuity of the value function[1]. To solve this problem, we proceed by taking the first-order condition with respect to x, conditional on $\Delta X \neq 0$:

$$g_x(x, A) - w - b_u \mathbf{1}^{+} - b_l \mathbf{1}^{-} + \beta D(x, A) = 0 \tag{3.4}$$

where we define $D(x, A) \equiv \int v_x(x, A') \, dF(A'|A)$. Clearly, a key step to obtaining a solution is determining the function $D(\cdot)$. As a first step, however, we need to characterise the general form of the firm's optimal demand function, to which we now turn.

First, note that equation (3.4) depends only on A and x. Thus, in principle this can be solved for A as a function of x for the cases in which the firm accumulates x

[1] A sufficient condition for this is that the evolution of A be governed by the stochastic difference equation, $A' = h(A, \varepsilon)$, where ε is an i.i.d. random variable, and $h(\cdot)$ is a continuous function. See Stokey & Lucas [1989] for more details.

$(\mathbf{1}^{+}=1)$ and the firm decumulates x $(\mathbf{1}^{-}=1)$. We define these relationships as:

$$\begin{aligned} A &= u(x) \quad \text{if} \quad \Delta X > 0 \\ A &= l(x) \quad \text{if} \quad \Delta X < 0 \end{aligned} \tag{3.5}$$

Now notice that the firm's objective function, (3.3), is concave in x. Thus it follows that the firm will only accumulate (decumulate) x if it receives a sufficiently positive (respectively negative) shock:

$$\begin{aligned} &\Delta X > 0 \quad \text{if} \quad A > A_u \\ &\Delta X < 0 \quad \text{if} \quad A < A_l \\ &\Delta X = 0 \qquad\quad \text{otherwise} \end{aligned} \tag{3.6}$$

where $A_u \geq A_l$. To complete our characterisation of the firm's demand policy, we need to specify how these "trigger" values for A are determined. To do this, note that, since the firm's objective (3.3) is continuous and concave, the optimal value of x will be a continuous function of the state variables (x_{-1}, A)[2]. Thus, it must be the case that:

$$\begin{aligned} A_u &= u(\chi_{-1}) \\ A_l &= l(\chi_{-1}) \end{aligned} \tag{3.7}$$

where $\chi_{-1} \equiv (1-\delta)\, x_{-1}$. Piecing all these findings together, we establish Proposition 1:

[2]This follows directly from the Theorem of the Maximum – see Stokey & Lucas [1989], pp. 62–63.

Proposition 6 *The optimal demand policy for x is of the form:*

$$x = \begin{cases} u^{-1}(A) & \text{if } A > u(\chi_{-1}) \\ \chi_{-1} & \text{otherwise} \\ l^{-1}(A) & \text{if } A < l(\chi_{-1}) \end{cases} \tag{3.8}$$

where $\chi_{-1} \equiv (1-\delta)\,x_{-1}$ and the functions $u(\cdot)$ and $l(\cdot)$ are defined by:

$$\begin{aligned} g_x(x, u(x)) - w - b_u + \beta D(x, u(x)) &\equiv 0 \\ g_x(x, l(x)) - w - b_l + \beta D(x, l(x)) &\equiv 0 \end{aligned} \tag{3.9}$$

Clearly, to complete our solution for the functions $u(\cdot)$ and $l(\cdot)$, we need to characterise the functions $D(x, l(x))$ and $D(x, u(x))$. We approach this as follows. First, note that, due to the recursive nature of the problem, the trigger values next period will be given by:

$$\begin{aligned} A_l = l(\chi_{-1}) &\implies A'_l = l(\chi) \\ A_u = u(\chi_{-1}) &\implies A'_u = u(\chi) \end{aligned} \tag{3.10}$$

Moreover, if we define $v^{-/0/+}$ as the value functions conditional on each of the three possible continuation regimes in which $\Delta X' \lesseqgtr 0$ respectively, we can re-write $D(\cdot)$ as:

$$\begin{aligned} D(x, A) = & \int_0^{l(\chi)} v_x^-(x, A')\,dF + \int_{l(\chi)}^{u(\chi)} v_x^0(x, A')\,dF + \int_{u(\chi)}^{\infty} v_x^+(x, A')\,dF \\ & + (1-\delta)\,l'(\chi)\left\{v^-(x, l(\chi)) - v^0(x, l(\chi))\right\} \\ & + (1-\delta)\,u'(\chi)\left\{v^0(x, u(\chi)) - v^+(x, u(\chi))\right\} \end{aligned} \tag{3.11}$$

Since the value function is continuous[3], it follows that the last two terms above equal zero. These "value-matching" conditions therefore yield:

$$D(x,A) = \int_0^{l(\chi)} v_x^-(x,A')\,dF + \int_{l(\chi)}^{u(\chi)} v_x^0(x,A')\,dF + \int_{u(\chi)}^{\infty} v_x^+(x,A')\,dF \qquad (3.12)$$

Finally, straightforward application of the envelope theorem allows us to pin down two of the forward derivatives in (3.12):

$$\begin{aligned} v_x^-(x,A') &= (1-\delta)\,b_l \\ v_x^+(x,A') &= (1-\delta)\,b_u \end{aligned} \qquad (3.13)$$

To determine the forward derivative in the "do nothing" ($\Delta X' = 0$) regime, note that:

$$\begin{aligned} v_x^0(x,A') &= \frac{\partial}{\partial x}\left\{g(\chi,A') - w\chi + \beta\int v(\chi,A'')\,dF(A''|A')\right\} \\ &= (1-\delta)\left\{g_x(\chi,A') - w + \beta D(\chi,A')\right\} \end{aligned} \qquad (3.14)$$

Thus, by substituting these envelope-type properties into (3.12), we obtain Proposition 2:

Proposition 7 *The function $D(\cdot)$ satisfies:*

$$D(x,A) = (1-\delta)\left\{\begin{array}{c} \int_0^{l(\chi)} b_l dF + \int_{l(\chi)}^{u(\chi)} [g_x(\chi,A') - w]\,dF \\ + \int_{u(\chi)}^{\infty} b_u dF + \beta\int_{l(\chi)}^{u(\chi)} D(\chi,A')\,dF \end{array}\right\} \qquad (3.15)$$

where $\chi \equiv (1-\delta)x$. This is a contraction mapping in $D(\cdot)$ and hence has a unique fixed point.

[3] Recall that this follows from our assumption that the distribution of the shock variable A satisfies the Feller property.

Proof. For proof of contraction property, see appendix. ■

This result provides us with a relation that we can use to pin down $D(\cdot)$ and thereby derive a solution for the optimal demand for x using (3.4).

Proposition 2 is worthy of discussion from the perspective of its contribution to our understanding of the analytics of factor irreversibility. In particular, it builds on previous work of Sargent [1980], which deals with the case of complete irreversibility – i.e. the case in which $b_l \to -\infty$, so that factor demand reversals are infinitely costly. In particular, Sargent goes to some pains to establish the existence and uniqueness of the marginal value function. Proposition 2 allows us to establish these results more simply. Existence and uniqueness of the marginal value follow directly from the contraction property of (3.15). Thus, Proposition 2 provides a more parsimonious representation of the solution to models of irreversibility in discrete time than existing solutions.

3.2.1 Comparison with Brownian Solutions

In addition, however, Proposition 2 also provides a more general solution to a model of partial irreversibility (Abel & Eberly [1996] and Bentolila & Bertola [1990] are the closest Brownian analogues). To see this, a comparison with the structure of the existing Brownian models is instructive. In particular, the analogous representation to the objective function (3.3) in continuous time is given by the following Bellman equation:

$$\begin{aligned} rV(x,A) &\equiv \max_x \left\{ g(x,A) - wx - b_u dX \mathbf{1}^+ - b_l dX \mathbf{1}^- + \frac{E[dV(x,A)]}{dt} \right\} \quad (3.16) \\ \text{s.t } dx &= dX - \delta x \cdot dt \end{aligned}$$

Note that, in order to make any progress in solving such a model, we must specify a process for the shock variable A so that we may establish the structure of the expected capital gain, $E\left[dV\left(x,A\right)\right]/dt$. In particular, as mentioned above, the vast majority of studies employ the convenient assumption that shocks follow a geometric Brownian motion of the form:

$$dA = \mu A dt + \sigma A dz \tag{3.17}$$

where $dz = \varepsilon\sqrt{dt}$ is a standard Wiener process, and $\varepsilon \sim N\left(0,1\right)$. Under this assumption, application of Ito's Lemma provides us with the following representation for the future expected capital gain:

$$\frac{E\left[dV\left(x,A\right)\right]}{dt} = -\delta x V_x + \mu A V_A + \frac{1}{2}\sigma^2 A^2 V_{AA} \tag{3.18}$$

Only now can we begin to consider taking the derivative of the Bellman equation and thereby find the optimal demand for x. Doing so yields the following differential equation for the marginal value function[4]:

$$\begin{aligned} rV_x\left(x,A\right) &= g_x\left(x,A\right) - w - \delta V_x\left(x,A\right) - \delta x V_{xx}\left(x,A\right) \\ &\quad + \mu A V_{Ax}\left(x,A\right) + \frac{1}{2}\sigma^2 A^2 V_{AAx}\left(x,A\right) \end{aligned} \tag{3.19}$$

Given particular assumptions about the form of the production technology $g\left(\cdot\right)$, one can then use this equation to characterise the demand for x (see Abel & Eberly [1996] for details).

Proposition 2, on the other hand, characterises the solution to the model (3.3) for

[4]Note that the terms reflecting accumulation costs cancel out in (3.19). The reason is that an additional unit of x increases the value of the firm by V_x and costs b_u, thereby yielding a net contribution of $dX^+\left[V_x - b_u\right]$. This is always equal to zero since if it is optimal for the firm to accumulate x then the firm will set $V_x = b_u$; otherwise the firm does not accumulate and $dX^+ = 0$. A symmetric logic applies to decumulation of x.

a wide class shock processes $F(\cdot)$. In this sense, (3.15) represents a more general solution to models of irreversibility than those that have employed Brownian shocks reviewed above. In particular, (3.15) is the discrete time analogue to the differential equation obtained for the marginal value function in (3.19). However, recall that the only way to derive these differential equations in continuous time is to specify beforehand the evolution of shocks, typically given by a geometric Brownian motion such as (3.17). No such assumption had to be made to derive (3.15).

It is also worth noting that the contraction property of the mapping defined in (3.15) admits further potential in respect of deriving numerical solutions to more complicated technology and shock structures. In particular, it implies that successive recursions on (3.15) from an initial starting point will converge to the fixed point function $D(\cdot)$ defined implicitly by (3.15) in Proposition 2[5]. Thus, this also provides an insight into potential methods of solving adjustment cost models numerically.

In the remainder of this chapter, we consider just one of these developments – the impact of different assumptions about the distribution of shocks facing the firm on optimal factor demand. To focus ideas, we analyse the simple Cobb-Douglas production function:

$$g(x, A) = Ax^{\alpha} \tag{3.20}$$

with $\alpha \in (0, 1)$. We now turn to the implications of different assumptions about the evolution of A.

[5]This follows from a simple application of the "N–Stage Contraction Theorem" – see Stokey & Lucas (1989), p. 53.

3.3 Assessing the Impact of Different Forms of Uncertainty using Simple Examples

This section seeks to present some simple examples of how different assumptions about the structure of shocks facing the firm can alter optimal factor demand policies. We start with a treatment of the analogue to existing Brownian models, where shocks are assumed to follow a random walk. We then use this as a benchmark to examine the impact of first changes in persistence of shocks, and then changes in the distribution of shocks.

3.3.1 Geometric Gaussian Random Walk

As a first example, then, consider the case that is most closely linked to the Brownian motion assumption of the previous literature – that of a geometric Gaussian random walk. In particular, we assume that A evolves according to:

$$\begin{aligned} A' &= (1+\mu) A \exp\left(\varepsilon' - \frac{1}{2}\sigma^2\right) \\ \varepsilon' &\sim N\left(0,\sigma^2\right) \end{aligned} \tag{3.21}$$

Given this assumption, we can attempt to solve for $D(\cdot)$ using (3.15). To do this, we employ the method of undetermined coefficients to obtain solutions for $D(\cdot)$ and thereby for the functions $u(\cdot)$ and $l(\cdot)$ also. Doing so yields:

Proposition 8 *If a firm's shocks evolve according to the geometric random walk, (), then the functions $u(\cdot)$ and $l(\cdot)$ are of the form:*

$$\begin{aligned} u(x) &= u \cdot \frac{x^{1-\alpha}}{\alpha} \\ l(x) &= l \cdot \frac{x^{1-\alpha}}{\alpha} \end{aligned} \tag{3.22}$$

where u and l are given constants which depend upon the parameters of the model, $\{\alpha, \beta, \delta, \mu, \sigma, Q, R\}$ *and where* $Q \equiv w/b_u$, $R \equiv b_u/b_l$.

Proof. See appendix. ■

To get an idea of the implications of the adjustment costs b_l and b_u in the model, Figure 3-1 plots the firm's optimal demand policy in the case with fully reversible factor accumulation ($b_l = b_u = b$), and the case with partial irreversibility ($b_l < b_u$). In particular, we observe three departures from the reversible policy. First, the firm attenuates the extent to which it decumulates x. To see this, notice that for all values of A below A_l, the firm's demand for x is now greater than in the reversible case. This follows from the fact that the firm no longer receives the same return on decumulation of x. Recall that in the capital case this is due to the resale price of capital being lower than the purchase price. In the labour demand case, this results because it may be costly to fire workers. Second, the firm also attenuates the extent to which it accumulates x. This is because firms now realise that accumulation of x is irreversible to some extent, and as a result they limit accumulation as a precaution against being unable to decumulate as profitably in the future. Third, there is a range of values for the firm's shock, A, for which there is no change in the firm's optimal factor holdings gross of depreciation. This "region of inaction" arises as a result of the kink in the firm's objective function created by the existence of asymmetric kinked adjustment costs. We can now use this case a benchmark with which to compare the implications of other forms of shocks on optimal factor demand.

3.3.2 The Impact of the Persistence of Shocks

A question that immediately arises in the light of the results for the random walk case is that of the effect of the persistence of shocks on the optimal demand for x. In

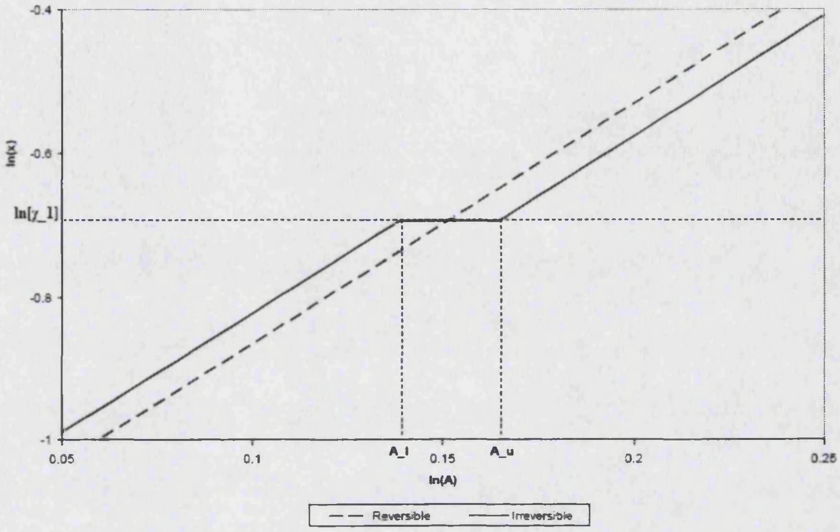


Figure 3-1: The Impact of Irreversibility on Optimal Factor Demand

particular, we have noted that the assumption of Brownian shocks in continuous time in the previous literature is analogous to assuming shocks follow a random walk in discrete time (Dixit [1993]). Thus, it is of interest to understand what might change as the persistence of shocks declines from a unit root. To this end, we characterise first a general result on the effect of shock persistence and then consider the specific case of i.i.d. Gaussian shocks as a concrete comparison with the Gaussian random walk results presented above.

First, assume that the evolution of shocks is governed by the stochastic difference equation:

$$A' = h\left(A, \varepsilon'\right) \tag{3.23}$$

where ε' is assumed to be an i.i.d. innovation. We can think of changes in the persistence of shocks facing the firm as changes in the response of A' to A: i.e.

$h_A(A, \varepsilon')$. Given this, we can then establish the following result:

Proposition 9 *An increase in the persistence of shocks, h_A, renders optimal factor demand more responsive to changes in A.*

Proof. See appendix. ■

The intuition behind this result is quite straightforward. As shocks become less persistent, or h_A falls, shocks will have more of a tendency to revert towards a mean in the future. Thus, low values of A today are more likely to be followed by higher values in the future. As a result, firms will not feel as compelled to reduce their holdings of factor x following a bad shock relative to a case in which shocks are very persistent. Similarly, when shocks are less persistent, we would expect good shocks to lead to a lesser increase in factor demand, as lower values of A will be expected in the future. Thus, reduced persistence pivots the firms optimal demand for x as a function of A clockwise.

I.i.d. Geometric Gaussian Shocks

We now examine how the general result of Proposition 4 might obtain in practice by returning to our case of Cobb-Douglas technology. In particular, we illustrate the impact of changes in persistence by comparing the results of section 3.1 with those from a model with i.i.d. Gaussian shocks:

$$\begin{aligned} A' &= \bar{A} \cdot \exp\left(\varepsilon' - \frac{1}{2}\sigma^2\right) \\ \varepsilon' &\sim N\left(0, \sigma^2\right) \end{aligned} \tag{3.24}$$

Note that this implies $E(A') = \bar{A}$ is the stationary mean value of the shock in every period[6]. In this case it is clear that, since today's shocks, A, have no impact on future

[6]For simplicity, we do not incorporate growth into this specification. Thus, all subsequent comparisons with the random walk case impose $\mu = 0$ in (3.21).

shocks, A', the marginal effects of today's factor demand on future profits must also be independent of A:

$$D(x,A) \equiv D(x) \tag{3.25}$$

Thus, taking this into account, and solving out the integrals in (3.15)[7], we obtain:

$$D(x) - \beta(1-\delta)(\Phi_2 - \Phi_1) D(\chi) = (1-\delta)\left\{ \begin{array}{c} b_l \Phi_1 + b_u (1-\Phi_2) - w(\Phi_2 - \Phi_1) \\ +\alpha \chi^{\alpha-1}(\Phi_3 - \Phi_4) \end{array} \right\} \tag{3.26}$$

where:

$$\begin{array}{ll} \Phi_1 = \Phi\left[\frac{1}{\sigma}\left(\ln l(\chi) - \ln \bar{A} + \frac{1}{2}\sigma^2\right)\right]; & \Phi_2 = \Phi\left[\frac{1}{\sigma}\left(\ln u(\chi) - \ln \bar{A} + \frac{1}{2}\sigma^2\right)\right] \\ \Phi_3 = \Phi\left[\frac{1}{\sigma}\left(\ln u(\chi) - \ln \bar{A} - \frac{1}{2}\sigma^2\right)\right]; & \Phi_4 = \Phi\left[\frac{1}{\sigma}\left(\ln l(\chi) - \ln \bar{A} - \frac{1}{2}\sigma^2\right)\right] \end{array} \tag{3.27}$$

Finding an analytic solution for $D(\cdot)$ using (3.26) is not as straightforward as in the random walk case, principally because of the non-linear terms Φ_i, $i = 1,..,4$, which depend on x. However, we can obtain further analytical results for the case in which there is no depreciation, $\delta = 0$, so that $\chi = x$. In this case, we solve (3.26) for $D(x)$ given by:

$$D(x) = \frac{1}{1-\beta(\Phi_2 - \Phi_1)}\left\{ \begin{array}{c} b_l \Phi_1 + b_u (1-\Phi_2) - w(\Phi_2 - \Phi_1) \\ +\alpha x^{\alpha-1}(\Phi_3 - \Phi_4) \end{array} \right\} \tag{3.28}$$

[7]Lemma 1 in the appendix establishes that, for a random variable $\ln x \sim N(\mu, \sigma^2)$ then $\int_{\underline{x}}^{\overline{x}} x dF(x) = \exp\left[\mu + \frac{1}{2}\sigma^2\right]\left\{\Phi\left[\frac{\ln \overline{x} - \mu}{\sigma} - \sigma\right] - \Phi\left[\frac{\ln \underline{x} - \mu}{\sigma} - \sigma\right]\right\}$, where $\Phi(\cdot)$ is the standard Normal c.d.f.

Substituting back into the first order conditions, (3.9), we obtain:

$$\begin{aligned} u(x) &= [w + b_u - \beta D(u(x), l(x))] \cdot \frac{x^{1-\alpha}}{\alpha} \\ l(x) &= [w + b_l - \beta D(u(x), l(x))] \cdot \frac{x^{1-\alpha}}{\alpha} \end{aligned} \tag{3.29}$$

where we have used a slight change of notation to emphasise the dependence of $D(\cdot)$ on x through the values of $u(\cdot)$ and $l(\cdot)$. These provide us with a system of two non-linear equations that can be solved for the functions $u(x)$ and $l(x)$, and thus allow us to identify the optimal demand for x.

Figures 3-2 and 3-3 illustrate the optimal demand functions for x with i.i.d. shocks obtained by solving the system (3.29), and compare them with the benchmark case of a geometric random walk. It can be seen that the existence of i.i.d. shocks leads to a clockwise tilting of the firm's optimal demand as a function of the shock, A, exactly along the lines of Proposition 4. In particular, Figures 3-2 and 3-3 illustrate the solution for two different values for $\bar{A}$ – the stationary mean in the i.i.d. case. We observe that, for values of the shock variable A below $\bar{A}$, the demand for x in the i.i.d. model exceeds that for the random walk model. Conversely, for values of A above $\bar{A}$, the demand for x is higher in the random walk model. Why might this be the case? The intuition is that, when shocks are i.i.d., there is a mean-reverting aspect to the firm's profitability. That is, the firm expects A to revert back to its mean, $\bar{A}$, in all future periods. Thus, when the current shock is below $\bar{A}$, the firm expects its profitability to get better in the future, and vice versa. In the random walk model, however, the firm expects future values of A to be equal to the current value for all future periods. Thus, for current values of A equal to $\bar{A}$, we obtain identical results for the i.i.d. and random walk cases.

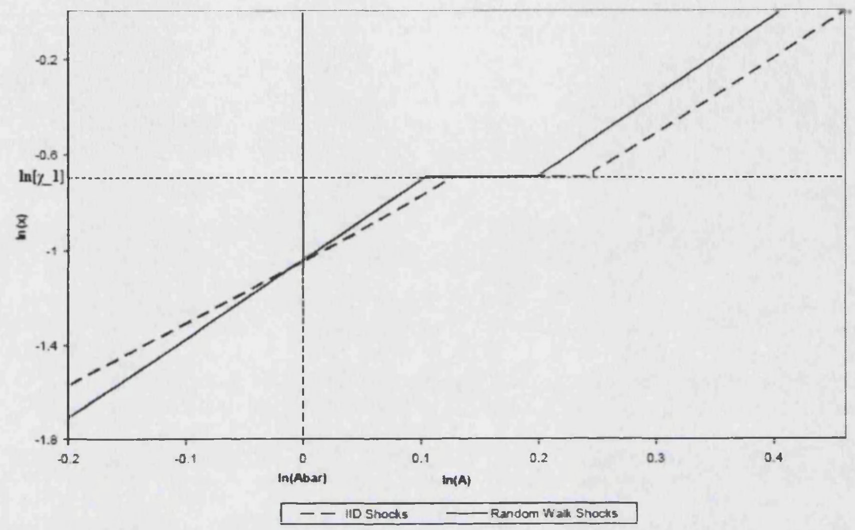


Figure 3-2: Geometric Random Walk vs. i.i.d. Gaussian Shocks: $\bar{A} = 1$

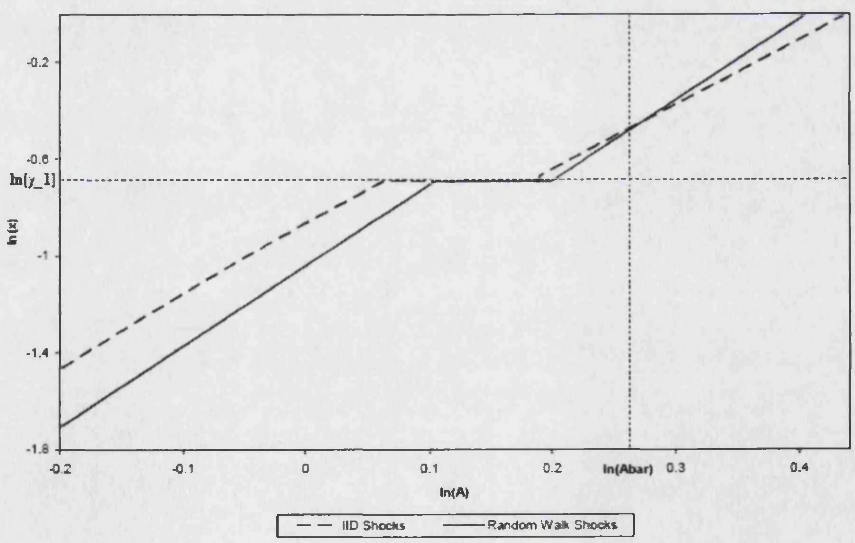


Figure 3-3: Random Walk vs. i.i.d. Gaussian Shocks: $\bar{A} = 1.3$

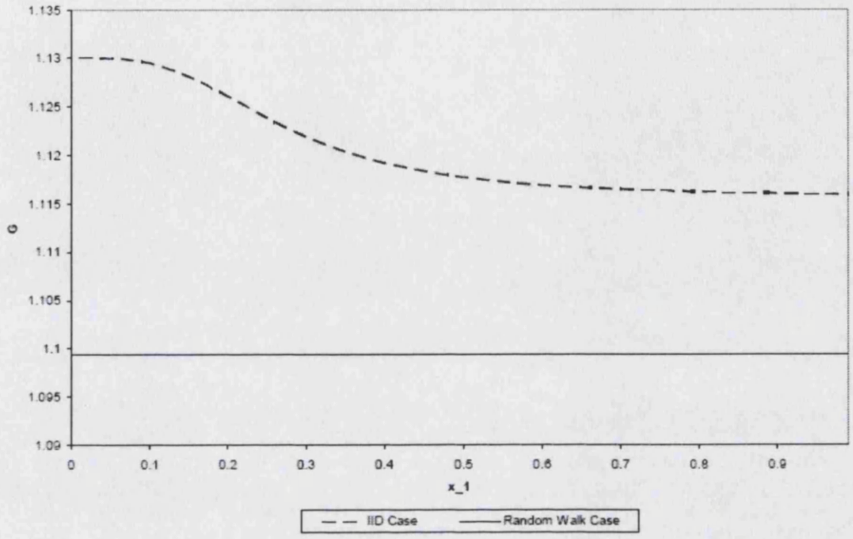


Figure 3-4: Extent of Inaction as a Function of x_{-1} - Random Walk vs. i.i.d. Gaussian

Figure 3-4 then illustrates the impact of assuming i.i.d. shocks on the extent of inaction. We measure this inaction by finding the geometric distance between the upper and lower triggers for accumulating and decumulating x: $G \equiv \frac{u(\chi_{-1})}{l(\chi_{-1})}$. In particular, Figure 3-4 plots G as a function of χ_{-1} in each case. We see that assuming i.i.d. shocks leads in general to an *increase* in the degree of inaction. The reason for

[redacted]

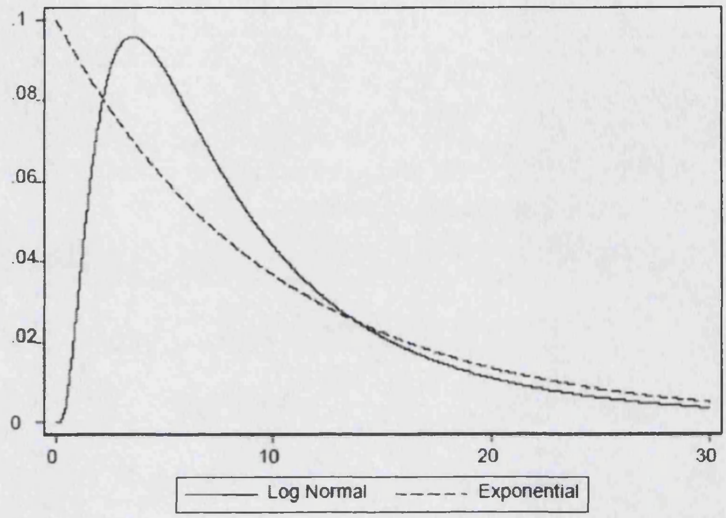


Figure 3-5: Exponential vs. Log-Normal Distribution

3.3.3 The Impact of the Distribution of Shocks

In this section, we consider the impact of departing from the assumption of log-Normal shocks by instead looking at shocks drawn from an exponential distribution. To see the implications of this, Figure 3-5 compares the exponential distribution to a log-Normal with identical mean and variance. It can be seen that the exponential distribution has fatter tails relative to the log-Normal, particularly in the left hand tail. Thus, the exponential distribution yields more extreme values more often than the log-Normal.

In what follows, we assume shocks A are i.i.d. exponential, and that their density and distribution functions are therefore respectively given by[8]:

$$\begin{aligned} f(A') &= \frac{1}{s}\exp(-A'/s) \\ F(A') &= 1-\exp(-A'/s) \end{aligned} \tag{3.30}$$

In this case, we can solve the integrals in (3.15) to obtain the analogue to (3.26) in the i.i.d. Gaussian case:

$$D(x) - \beta(1-\delta)(F_2 - F_1) D(\chi) = (1-\delta)\left\{ \begin{array}{c} b_l F_1 + b_u(1-F_2) - w(F_2 - F_1) \\ +\alpha\chi^{\alpha-1}\left[\begin{array}{c} [s+u(\chi)]F_2 - [s+l(\chi)]F_1 \\ -[u(\chi)-l(\chi)] \end{array} \right] \end{array} \right\} \tag{3.31}$$

where $F_1 = F[l(\chi)]$ and $F_2 = F[u(\chi)]$. As before in the i.i.d. Gaussian case, we concentrate on the analytically more tractable case in which there is no depreciation, $\delta = 0$. It follows that we can then solve for $D(\cdot)$ given simply by:

$$D(x) = \frac{1}{1-\beta(F_2 - F_1)}\left\{ \begin{array}{c} b_l F_1 + b_u(1-F_2) - w(F_2 - F_1) \\ +\alpha x^{\alpha-1}\left[\begin{array}{c} [s+u(x)]F_2 - [s+l(x)]F_1 \\ -[u(x)-l(x)] \end{array} \right] \end{array} \right\} \tag{3.32}$$

Once again, together with (3.29), this provides us with a set of non linear equations that can be solved for $u(\cdot)$ and $l(\cdot)$.

Figure 3-6 now compares the solution to the exponential model with that obtained for the Gaussian model. We can see that the existence of exponential shocks with fatter tails leads to a reduction of the demand for x for low values of the firm's shock

[8]Note that the exponential is a one parameter distribution with mean and variance given respectively by s and s^2. In all comparisons with log-Normal results, we match these two moments so that they are identical.

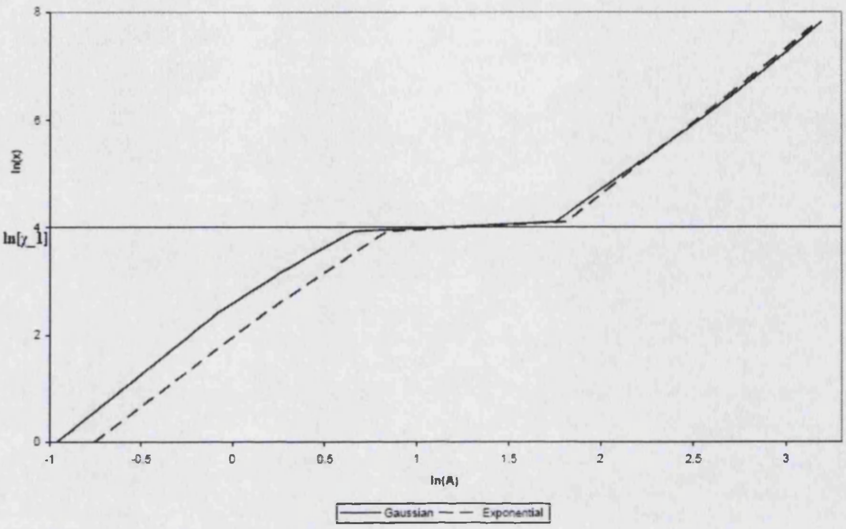


Figure 3-6: Optimal Factor Demand in the face of Gaussian vs. Exponential Shocks

A, and to higher demand for x for higher values of A. The intuition for this is that lower values of A are more common in the exponential relative to the Gaussian case (see Figure 3-5). Thus, firms are more likely to reduce their demand for x following a bad shock because they are less confident that the shock will return to a higher value in the future in the exponential model (although on average, by construction they expect the same value of A in all future periods). The converse is also true to a lesser extent for good shocks. In particular, the fatter upper tail in the exponential case leads the firm to be more confident of obtaining very good shocks in the future, and thus leads the firm to be more likely to accumulate x as well.

3.4 Conclusions

This chapter has provided new results on the analytics of factor accumulation in the presence of asymmetric adjustment costs in discrete time. In particular, we have established that the marginal value function obeys a contraction mapping, and that this property has a number of useful implications. First, we have shown that this result holds for a wider class of shock structures than previous solutions. Specifically, we have shown that previous continuous time models have had to specify the form of shocks – typically Brownian motions – before being able to characterise the marginal value function. No such assumption has to be made in the framework discussed in this chapter.

We then go on to show that the contraction property allows one to characterise the properties of optimal factor demand under alternative assumptions about first the persistence of shocks, and second the distribution of shocks. In particular, we establish a general result that an increased persistence of shocks raises the response of factor demand to current shocks. In addition, by solving a model with i.i.d. Gaussian shocks, we find that reduced persistence leads to greater rigidity in factor demand. In this sense, Brownian models may imply a lower extent of inaction than would be predicted by less persistent shocks. Finally, by solving a model with exponential rather than Gaussian shocks, we establish that fatter tails in the distribution of shocks lead to reduced demand under bad shocks, but higher demand under good shocks. In this way, we can begin to understand the impact of distributional assumptions on models of factor demand under irreversibility.

Chapter 4

Conclusion to Thesis

This thesis has sought to unify the concepts of downward nominal wage rigidity in labour markets, money illusion in the form of an aversion to nominal loss, and irreversibility in a novel way.

In particular, Chapter 1 shows that money illusion in the form of nominal loss aversion on behalf of workers implies that firms' wage decisions become irreversible to some extent. To see this, note that if a firm increases the wage today, and then reverses its decision by cutting the wage by an equal amount tomorrow, the net effect on productivity will be negative. This is because cutting the wage involves a larger reduction in productivity than an equal-sized wage increase raises productivity.

Chapter 1 then links this finding to the observation of DNWR in labour markets. In particular, it shows that we can obtain a better understanding of the nature of DNWR by noting that such a model has the novel that firms have an incentive to attenuate wage increases as well as decreases.

This finding has not been addressed in the previous literature on DNWR, and allows a potential explanation for why the expected macroeconomic effects of DNWR have not been observed, despite remarkably robust evidence for DNWR in micro-level data. In particular, by not accounting for this compression of wage increases, and the related cost savings to firms, Chapter 1 argues that the previous literature may have overstated the increase in wage pressure due to DNWR at the micro level. In this way, we can therefore potentially reconcile the macro and micro level results in the literature.

Chapter 1 then tests whether this hypothesis does indeed explain the data on DNWR using micro data from the US and Great Britain. We find strong evidence that firms do indeed compress wage increases, precisely along the lines of a model in which workers resist nominal wage cuts. This then augments and links the current evidence on money illusion and DNWR in the literature. Moreover, we also find that this compression of wage increases is of a sufficient magnitude in the data to offset much of the increase in aggregate wage growth due to restricted wage cuts. Thus, again along the lines of the theory, we can reconcile the micro and macro level estimates for DNWR.

Chapter 2 then seeks to reinforce the link between the observed DNWR in labour markets and the phenomenon of money illusion. It does this by drawing out more carefully the implications of the model of DNWR based on money illusion in Chapter 1, and comparing them to the implications of alternative models of DNWR based on wage contracting. Chapter 2 shows that the key intuitive distinction between these two sets of models is that these alternative "contract" models are based on *real* economic frictions, whereas, almost by definition, the model of money illusion is based on a *nominal* friction (i.e. a resistance to nominal wage cuts).

In particular, contract models explain wage rigidity by noting that in many countries (notably excluding the US and UK) there is a legal requirement of mutual consent of firm and worker to any wage change. It follows that the default option in any renegotiation is for the wage to remain unchanged. When combined with frictions (e.g. search costs) that make it costly for firms (workers) to switch workers (firms), and hence drive a wedge between the outside options of each party, it follows that wages will stay constant for periods of time. Moreover, in an inflationary environment, wages will be bid up more often than down, and hence wages will appear downward rigid. However, since this rigidity in contract models is intrinsically related to *real* frictions, it follows that these frictions should be unrelated to the rate of inflation. In the money illusion model of Chapter 1, on the other hand, lower inflation will lead to a greater compression of wage increases as firms worry more about their ability to cut nominal wages in future periods.

Chapter 2 then tests this additional prediction using micro data from the US and Great Britain. Specifically, it finds clear evidence that nominal wage increases become more compressed as inflation falls, exactly in line with the predictions of the model of money illusion. This suggests that at least part of the observed DNWR is driven by resistance to nominal wage cuts an behalf of workers, rather than the motives underlying contract models. Thus, the results of Chapter 2 confirm the link between money illusion and DNWR underlying the model in Chapter 1.

Finally, Chapter 3 extends the theoretical findings of Chapter 1 with respect to models of irreversibility. In particular, it analyses these models in their original context – that of dynamic factor demand – to obtain some insights into the effects of the process of shocks facing a firm on optimal factor demand.

Specifically, Chapter 3 presents a more general solution to models of irreversibility by developing a discrete time setup. It shows that previous models of irreversibility in continuous time require the assumption that the shocks facing a firm follow a Brownian motion – the continuous time analogue to a random walk. By adopting a discrete time approach, the solution in Chapter 3 allows a more general characterisation of optimal factor demand in the face of a much wider class of shocks.

Chapter 3 then goes on to summarise the effects of relaxing the assumption of random walk shocks. In particular, we establish a quite general result that reduced persistence of shocks renders optimal factor demand less responsive to shocks. The intuition for this is that greater persistence renders current shocks more informative about future shocks. It follows that firms are more willing to respond to current shocks because they are more certain that they will not have to reverse their decision in the future. In addition to this general finding, Chapter 3 also presents some results from specific assumptions about the process of shocks facing the firm. By assuming i.i.d. Gaussian shocks, we find that this creates greater rigidity in optimal factor demand. Intuitively, when shocks are i.i.d., the firm knows that they will revert back to some mean in the future. Thus, is makes less sense to sink resources into changing factor employment following a shock when the firm knows it will likely have to reverse its decision, at a cost, in the future.

Together, then, these chapters aim to enhance our understanding of money illusion and irreversibility both as distinct issues, but particularly through the lens of downward nominal wage rigidity in labour markets.

4.1 References

Abel, Andrew B., and Janice C. Eberly, "Optimal Investment with Costly Reversibility", *Review of Economic Studies*, Vol. 63, No. 4. (Oct., 1996), pp. 581-593.

Akerlof, George A., William T. Dickens, and George L. Perry, "The Macroeconomics of Low Inflation", *Brookings Papers on Economic Activity*, Vol. 1996, No. 1. (1996), pp. 1-76.

Akerlof, George A., William T. Dickens, and George L. Perry, "Near Rational Wage and Price Setting and the Long-Run Phillips Curve", *Brookings Papers on Economic Activity*, Vol. 2001, No. 1. (2001), pp. 1-44.

Akerlof, George A., and Janet L. Yellen, *Efficiency Wage Models of the Labor Market*, Cambridge University Press, 1986.

Altonji, Joseph & Paul Devereux, "The Extent and Consequences of Downward Nominal Wage Rigidity", *Research in Labor Economics*, Vol 19, Elsevier Science Inc. (2000): 383-431.

Arrow, Kenneth J., "Optimal Capital Policy with Irreversible Investment" in *Value, Capital and Growth, Papers in honour of Sir John Hicks*, J. N. Wolfe (ed.), Chicago: Aldine Publishing Company, 1968, pp. 1-19.

Beissinger, T. & Knoppik, K., "How Rigid Are Nominal Wages? Evidence and Implications for Germany", forthcoming in: *Scandinavian Journal of Economics*, 2003.

Benartzi, Shlomo, and Richard H. Thaler, "Myopic Loss Aversion and the Equity Premium Puzzle", *Quarterly Journal of Economics*, Vol. 110, No. 1. (Feb., 1995), pp. 73-92.

Bentolila, Samuel, and Giuseppe Bertola, "Firing Costs and Labour Demand: How Bad is Eurosclerosis?", *Review of Economic Studies*, Vol. 57, No. 3. (Jul., 1990), pp. 381-402.

Bertrand, Marianne, and Sendhil Mullainathan, "Do People Mean What They Say? Implications for Subjective Survey Data", *American Economic Review*, Vol. 91, No. 2, Papers and Proceedings of the 113th Annual Meeting of the American Economic Association. (May, 2001), pp. 67-72.

Bewley, Truman F., "A Depressed Labor Market as Explained by Participants", *American Economic Review*, Vol. 85, No. 2, Papers and Proceedings of the 107th Annual Meeting of the American Economic Association Washington, DC, January 6-8, 1995. (May, 1995), pp. 250-254.

Bewley, Truman F., "Why Not Cut Pay?", *European Economic Review.* May 1998; 42(3-5): 459-90.

Bewley, Truman F., *Why wages don't fall during a recession*, Cambridge and London: Harvard University Press, 1999.

Binmore, Ken, Avner Shaked, and John Sutton, "An Outside Option Experiment", *Quarterly Journal of Economics*, Vol. 104, No. 4. (Nov., 1989), pp. 753-770.

Blanchflower, David G., Andrew J. Oswald, *The Wage Curve*, Cambridge and London: MIT Press, 1994.

Cairncross, Alec, *The British Economy since 1945*, 2nd edition, Blackwell, 1995.

Card, David & Dean Hyslop "Does Inflation 'Grease the Wheels of the Labor Market'?" in Christina D. Romer and David H. Romer, editors, *Reducing Inflation: Motivation and Strategy.* University of Chicago Press, 1997.

Christofides Louis N., and Man Tuen Leung , "Nominal Wage Rigidity in Contract Data: A Parametric Approach", *Economica*, November 2003, vol. 70, no. 280, pp. 619-638(20).

Decressin, Anja, and Jorg W. Decressin, "On Sand and the Role of Grease in Labor Markets: How Does Germany Compare?" IMF Working Paper No. 02/164.

Degeorge, Francois, Jayendu Patel, and Richard Zeckhauser, "Earnings Management to Exceed Thresholds", *Journal of Business*, 1999, Vol. 72, No. 1.

Dessy, Orietta, "Nominal wage rigidity in the European Countries: evidence from the Europanel", working paper, 2002.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach", *Econometrica*, Vol. 64, No. 5. (Sep., 1996), pp. 1001-1044.

Dixit, Avinash, *The art of smooth pasting*, Fundamentals of Pure and Applied Economics, vol. 55: Stochastic Methods in Economic Analysis section. Harwood Academic, 1993.

Dixit, Avinash, and Robert Pindyck, *Investment under Uncertainty*, Princeton University Press, 1994.

Dwyer, Jacqueline, and Kenneth Leong, "Nominal Wage Rigidity in Australia", *Australian Journal of Labour Economics.* March 2003; 6(1): 5-24.

Farès, Jean, and Seamus Hogan, "The Employment Costs of Downward Nominal-Wage Rigidity." Bank of Canada Working Paper 2000–1.

Fehr, E. & L. Götte, "Robustness and Real Consequences of Nominal Wage Rigidity", Institute for Empirical Research in Economics, University of Zürich, Working Paper No. 44.

Genesove, David & Christopher Mayer, "Loss Aversion and Seller Behavior: Evidence from the Housing Market", *Quarterly Journal of Economics.* November 2001; 116(4): 1233-60.

Gordon, R., Comment on: Akerlof, George A., William T. Dickens, and George L. Perry, "The Macroeconomics of Low Inflation", *Brookings Papers on Economic Activity,* Vol. 1996, No. 1. (1996), pp. 1-76.

Groshen, Erica L.,and Mark E. Schweitzer, "Identifying Inflation's Grease and Sand Effects in the Labor Market", Feldstein, Martin, ed. *The costs and benefits of price stability.* NBER Conference Report series. Chicago and London: University of Chicago Press, 1999; 273-308.

Holden, Steinar, "Wage Bargaining and Nominal Rigidities", *European Economic Review.* May 1994; 38(5): 1021-39.

Holden, Steinar, "Wage Bargaining, Holdout, and Inflation", *Oxford Economic Papers.* April 1997; 49(2): 235-55.

Holden, Steinar, "Downward nominal wage rigidity — contracts or fairness considerations?", June 2002, working paper.

Holden, Steinar, "The costs of price stability - downward nominal wage rigidity in Europe", March 2003, forthcoming in *Economica.*

Howitt, Peter, "Looking Inside the Labor Market: A Review Article", *Journal of Economic Literature* 40 (March 2002): 125-38.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan, "Earnings Losses of Displaced Workers" *American Economic Review*, Vol. 83, No. 4. (Sep., 1993), pp. 685-709.

Kahn, Shulamit, "Evidence of Nominal Wage Stickiness from Microdata", *American Economic Review*, Vol. 87, No. 5. (Dec., 1997), pp. 993-1008.

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler, "*Fairness as a Constraint on Profit Seeking: Entitlements in the Market*", American Economic Review, Vol. 76, No. 4. (Sep., 1986), pp. 728-741.

Kahneman, Daniel, and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, Vol. 47, No. 2. (Mar., 1979), pp. 263-292.

Kramarz, F., "Rigid Wages: What Have we Learnt from Microeconometric Studies", in *Advances in Macroeconomic Theory*, J. Drèze ed., 194-216, Oxford University Press, Oxford, UK.

Lebow, David E., Raven E. Saks, Beth-Anne Wilson, "Downward Nominal Wage Rigidity: Evidence from the Employment Cost Index", Board of Governors of the Federal Reserve System, Finance and Economics Discussion Series: 99/31 July 1999.

Lucas, Robert E. Jr, and Edward C. Prescott, "Investment Under Uncertainty", *Econometrica*, Vol. 39, No. 5. (Sep., 1971), pp. 659-681.

Lucas, Robert E. Jr., *Models of Business Cycles.* Oxford: Basil Blackwell, 1987.

MacLeod, W. Bentley, and James M. Malcomson, "Investments, Holdup, and the Form of Market Contracts", *American Economic Review*, Vol. 83, No. 4. (Sep., 1993), pp. 811-837.

Madrian, Brigitte, and Lars J. Lefgren, "A Note on Longitudinally Matching Current Population Survey (CPS) Respondents" NBER Technical Working Paper 247, November 1999.

Mankiw, N. G., Comment on: Akerlof, George A., William T. Dickens, and George L. Perry, "The Macroeconomics of Low Inflation", *Brookings Papers on Economic Activity*, Vol. 1996, No. 1. (1996), pp. 1-76.

McLaughlin, Kenneth J., "Rigid Wages?", *Journal of Monetary Economics.* December 1994; 34(3): 383-414.

Nickell, S., and Glenda Quintini, "Nominal wage rigidity and the rate of inflation", *Economic Journal*, October 2003, vol. 113, no. 490, pp. 762-781(20).

Oi, Walter Y., "Labor as a Quasi-Fixed Factor", *Journal of Political Economy*, Vol. 70, No. 6. (Dec. 1962), pp. 538-555.

Rabin, Matthew, "Risk Aversion and Expected Utility Theory: A Calibration Theorem", *Econometrica*, September 2000, Vol. 68, No. 5, pp. 1281-1292.

Sargent, Thomas J., "Tobin's 'q' and the Rate of Investment in General Equilibrium," in *On The State of Macroeconomics*, (Carnegie-Rochester No. 12) ed. by K. Brunner and A. Meltzer, North-Holland, 1980, p. 107–54.

Shafir, Eldar, Peter Diamond, and Amos Tversky, "Money Illusion", *Quarterly Journal of Economics.* May 1997; 112(2): 341-74.

Sheshinski, Eytan, and Yoram Weiss, "Inflation and Costs of Price Adjustment", *Review of Economic Studies*, Vol. 44, No. 2. (Jun., 1977), pp. 287-303.

Smith, Jennifer, "Nominal wage rigidity in the United Kingdom", *Economic Journal* 110 (March 2000), 176-195.

Solow, Robert M., "Another Possible Source of Wage Stickiness", *Journal of Macroeconomics*, Vol. 1, pp.79–82, Winter 1979.

Stokey, Nancy L., and Robert E. Lucas Jr., *Recursive methods in economic dynamics*, with Edward C. Prescott Cambridge, Mass. and London: Harvard University Press, 1989.

Appendix A

Nominal Loss Aversion as a Foundation for DNWR

In this appendix we show that an effort function with the qualitative features of (1.1) can be derived from a stylised compensating differentials model with worker nominal loss aversion. The setup is as follows. The firm must choose an effort level (working conditions), e, as well as the nominal wage, W, to maximise the expected discounted value of profits in such a way as to maintain worker utility above their reservation utility. If we define R as the gross real rate of interest in the economy, the typical firm's decision problem is given by:

$$\begin{aligned} \max_{\{W_t, e_t\}} \quad & E_t \left[\sum_{s=t}^{\infty} R^{-(s-t)} \{a_s e_s - \omega_s\} \right] \qquad \text{(A.1)} \\ \text{s.t.} \quad & U_{EE_t} \geq U_{EU_t}, \text{ if worker-firm was employed in } t-1 \\ & U_{UE_t} \geq U_{UU_t}, \text{ if worker-firm was unemployed in } t-1 \end{aligned}$$

where U_{ij_t} denotes the utility of a worker who was in state i in period $t-1$, and in state j in period t where $i, j \in \{E, U\}$ and E and U denote employment and unemployment respectively. We adopt the following formulation of within-period utility, u[1]:

$$u = \frac{M}{P}\left(\frac{M}{M_{-1}}\right)^{c\mathbf{1}^-}\exp\left(-e\right) \tag{A.2}$$

$$\text{where,} \quad M = \begin{cases} W & \text{if worker is currently employed} \\ B & \text{if worker is currently unemployed} \end{cases}$$

or any positive monotonic transform thereof[2]. Thus, we can write the worker's lifetime utility as:

$$U = u + \frac{1}{R}E\left[U'\right] \tag{A.3}$$

The key characteristic of this specification of worker utility is that the marginal disutility of a nominal loss exceeds the marginal utility of a nominal gain. This property corresponds to:

$$\frac{\partial u/\partial M|_{M\uparrow M_{-1}}}{\partial u/\partial M|_{M\downarrow M_{-1}}} = 1 + c > 1 \tag{A.4}$$

This characteristic is what we term *nominal loss aversion.* Finally, we assume that an employed worker-firm becomes unemployed with probability δ, and that an unemployed worker-firm becomes employed with probability m. each period. Given this setup, we can establish the following result:

[1] From now on, for notational simplicity, we drop time subscripts, and denote lagged values by a subscript, $_{-1}$, and forward values by a prime, $'$.

[2] Note that here we define $\mathbf{1}^-$ more generally so that:

$$x^{\mathbf{1}^-} = \begin{cases} x & \text{if} \quad x < 1 \\ 1 & \text{if} \quad x \geq 1 \end{cases} \quad \text{and} \quad y\mathbf{1}^- = \begin{cases} y & \text{if} \quad y < 0 \\ 0 & \text{if} \quad y \geq 0 \end{cases}$$

Proposition 10 *If the worker's reservation utility is not subject to nominal loss aversion, then a firm solving (A.1) acts as if it were solving the wage-setting problem in the main text, (1.5).*

Proof. First, note that if a worker's utility from unemployment is unrelated to nominal earnings losses, it follows that without loss of generality, we can write:

$$\begin{aligned} U_{EE} &= \omega\left(\frac{W}{W_{-1}}\right)^{c1^{-}} \exp(-e) + \frac{1}{R}\left\{\delta EU'_{EU} + (1-\delta) EU'_{EE}\right\} \\ U_{EU} &= b + \frac{1}{R}\left\{mEU'_{UE} + (1-m) EU'_{UU}\right\} = U_{UU} \end{aligned} \tag{A.5}$$

A direct implication of the second line is that:

$$U_{EU} = \frac{b}{1-\frac{1}{R}} = U_{UU} \tag{A.6}$$

We now concentrate on worker-firms that were employed last period (so we can see the effect of a wage cut). Then note that, since firm profits are increasing in e and U_{EE} is decreasing in e, the firm will choose effort such that $U_{EE} = U_{EU}$, if the worker-firm was employed last period. It therefore follows that:

$$U_{EE} = \omega\left(\frac{W}{W_{-1}}\right)^{c1^{-}} \exp(-e) + \frac{1}{R}\left\{\frac{b}{1-\frac{1}{R}}\right\} = \frac{b}{1-\frac{1}{R}} = U_{EU} \tag{A.7}$$

Solving this equation for effort yields:

$$e = \ln\left(\frac{\omega}{b}\right) + c\ln\left(\frac{W}{W_{-1}}\right)1^{-} \tag{A.8}$$

Clearly, this is precisely the effort function stated in the main text, (1.1). ■

Thus, if workers do not treat nominal earnings losses from job loss as they do wage cuts, then we obtain the exact same wage-setting problem as in the main text.

The intuition for this is that, if the firm wishes to cut the nominal wage, then it has to compensate the worker in the form of lower on-the-job effort in order to ensure it is optimal for the worker to remain working at the firm.

However, one may consider this assumption unrealistic: job loss is typically associated with nominal earnings losses associated with losing one's wage and moving onto benefits. The next result allows for this, and establishes that the implicit effort function relevant to wage-setting in this case retains the same qualitative properties as those in (1.1).

Proposition 11 *If a worker's utility from unemployment is related to nominal income loss through job loss, then a firm solving (A.1) acts as if it were solving a wage-setting problem analogous to (1.5) in which effort is given by:*

$$e = \ln\left(\frac{\omega}{b}\right) + c\ln\left(\frac{W}{W_{-1}}\right)\mathbf{1}^{-} - \chi\left(W_{-1}, W\right)$$

where $\chi = \ln\left[\frac{1}{R} + \left(\frac{B}{W_{-1}}\right)^{c\mathbf{1}^{-}} - \frac{1}{R}\left(\frac{B'}{W}\right)^{c\mathbf{1}^{-}}\right]$*, and reflects the impact of nominal loss related to job loss in the present and future.*

Proof. For simplicity, we assume that real benefits are constant over time ($b' = b$), which implies that nominal benefits rise deterministically with inflation ($B' = (1+\pi)B$) and thus never fall[3]. In this case, the relevant value functions for the

[3]This assumption is entirely innocuous to the firm's wage choice, since the level and change in benefits over time is completely exogenous to the firm. Thus, none of the qualitative properties of the ensuing analysis depends on this assumption.

worker are given by:

$$
\begin{aligned}
U_{EE} &= \omega\left(\frac{W}{W_{-1}}\right)^{c1^-} \exp\left(-e\right)+\frac{1}{R}\left\{\delta EU'_{EU}+\left(1-\delta\right)EU'_{EE}\right\} \qquad (A.9)\\
U_{EU} &= b\left(\frac{B}{W_{-1}}\right)^{c1^-} +\frac{1}{R}\left\{mEU'_{UE}+\left(1-m\right)EU'_{UU}\right\}\\
U_{UE} &= \omega\left(\frac{W}{B_{-1}}\right)^{c1^-} \exp\left(-e\right)+\frac{1}{R}\left\{\delta EU'_{EU}+\left(1-\delta\right)EU'_{EE}\right\}\\
U_{UU} &= b+\frac{1}{R}\left\{mEU'_{UE}+\left(1-m\right)EU'_{UU}\right\}
\end{aligned}
$$

Again, note that, since firm profits are increasing in e and U_{EE}, U_{UE} are decreasing in e, the firm will choose effort such that $U_{EE}=U_{EU}$, if the worker-firm was employed last period, and $U_{UE}=U_{UU}$, if the worker-firm was unemployed last period. It then follows that:

$$
\begin{aligned}
U_{EE} &= \omega\left(\frac{W}{W_{-1}}\right)^{c1^-} \exp\left(-e_E\right)+\frac{1}{R}EU'_{EE}=b\left(\frac{B}{W_{-1}}\right)^{c1^-}+\frac{1}{R}EU'_{UU}=U_{EU} \qquad (A.10)\\
U_{UE} &= \omega\left(\frac{W}{B_{-1}}\right)^{c1^-} \exp\left(-e_U\right)+\frac{1}{R}EU'_{EE}=b+\frac{1}{R}EU'_{UU}=U_{UU} \qquad (A.11)
\end{aligned}
$$

Again, we concentrate on worker-firms that were employed last period (so we can see the effect of a wage cut). Using (A.10) we obtain:

$$
e_E=\ln\omega+c\ln\left(\frac{W}{W_{-1}}\right)1^- -\ln\left[b\left(\frac{B}{W_{-1}}\right)^{c1^-}+\frac{1}{R}E\left(U'_{UU}-U'_{EE}\right)\right] \qquad (A.12)
$$

Furthermore, subtracting (A.10) from (A.11) and forwarding one period, we obtain:

$$
U'_{UU}-U'_{EE}=b\left[1-\left(\frac{B'}{W}\right)^{c1^-}\right] \qquad (A.13)
$$

Substituting back into (A.12) above completes the proof. ∎

Note that this effort function is almost exactly the same as that in the main text, (1.1), except for the term χ. In particular, it retains a kink at $W = W_{-1}$, which implies that nominal wage increases are still partially irreversible in this model – the key qualitative characteristic this chapter seeks to analyse. The intuition for χ is as follows. The greater the nominal loss from job loss today (lower $(B/W_{-1})^{c1^{-}}$), the worse is the worker's outside alternative from quitting the job, and thus the greater the effort that the firm can extract from the worker. However, less clear is the effect of $(B'/W)^{c1^{-}}$. The intuition here is that, the greater the nominal loss from job loss tomorrow (lower $(B'/W)^{c1^{-}}$), the worse is the ability of the worker to obtain a high wage tomorrow. This then reduces the value of continued employment to the worker, and thus the worker works less hard. Crudely put, the greater the nominal loss from job loss tomorrow, the more the worker has to lose by staying in the firm.

Our final result in this section establishes that the dynamic structure of χ is such that it may well not have a significant impact on wage-setting:

Proposition 12 *If (i) the worker's reservation utility is subject to nominal loss aversion; (ii) the firm's shocks, a, evolve according to the random walk, (1.19); and (iii) the level of uncertainty is sufficiently low, then a firm solving (A.1) acts (approximately) as if it were solving the wage-setting problem in the main text, (1.5).*

Proof. First we establish the result for the deterministic problem ($\sigma = 0$). It therefore follows from continuity of the problem in σ that the result holds in a neighbourhood around $\sigma = 0$ – i.e. for sufficiently low levels of uncertainty.

In the deterministic case, $a' = a$ with probability one, and it follows that $\chi = \chi'$. The contribution of χ and χ' in the intertemporal maximand is given by:

$$\mathbf{I} \equiv -a\chi - \frac{a'\chi'}{R}$$

and their effect on the wage-setting choice is determined by:

$$\frac{\partial \mathrm{I}}{\partial W}=\frac{a}{R\cdot\exp\left(\chi\right)}\frac{\partial}{\partial W}\left(\frac{B'}{W}\right)^{\mathrm{c1}^{-}}-\frac{a'}{R\cdot\exp\left(\chi'\right)}\frac{\partial}{\partial W}\left(\frac{B'}{W}\right)^{\mathrm{c1}^{-}}=0$$

Thus, the existence of the χ terms has no effect on the wage-setting choice, and the firm acts as if effort were given simply by (1.1). ■

To understand the flavour of this result, note that raising today's wage, W, has two effects through increasing the nominal loss through job loss tomorrow ($(B'/W)^{\mathrm{c1}^{-}}$ falls). First, this reduces the worker's current valuation of continued employment at the firm, and thereby reduces effort today along the lines discussed above. However, by increasing the nominal loss through job loss tomorrow, a higher wage today also means that the firm can extract greater effort from the worker tomorrow. Thus, these two effects work against each other, and under the assumptions of Proposition 8, cancel out completely. Thus, there are good reasons to believe that a firm faced with a worker subject to nominal loss aversion will set wages similar to those derived from the problem in the main text, (1.5).

Appendix B

Technical Details of Chapter 1

B.1 Lemmas and Proofs

Proof of Proposition 1. Note first that equation (1.15) depends only on A and W. In principle this can be solved for A as a function of W when the nominal wage is increased ($1^- = 0$), and when it is decreased ($1^- = 1$). We define these relationships as:

$$\begin{aligned} A &= u(W) \quad \text{if} \quad \Delta W > 0 \\ A &= l(W) \quad \text{if} \quad \Delta W < 0 \end{aligned} \tag{B.1}$$

At this point, we assume that the problem is concave and verify this later[1]. It follows that:

$$\begin{aligned} \Delta W > 0 \quad &\text{if} \quad A > A_u \\ \Delta W < 0 \quad &\text{if} \quad A < A_l \\ \Delta W = 0 \quad &\quad \text{otherwise} \end{aligned} \tag{B.2}$$

where $A_u \geq A_l$. Moreover, we can relate (B.1) and (B.2) by noting that the continuity and concavity of v ensure, via the Theorem of the Maximum, that the

[1]Note that concavity of the problem is not trivial because the revenue function is convex in W_{-1}. We verify that the problem is indeed concave in W in Appendix C.

optimal value of W as a function of (A, W_{-1}) is continuous. It therefore follows that:

$$\begin{aligned} A_u &= u(W_{-1}) \\ A_l &= l(W_{-1}) \end{aligned} \tag{B.3}$$

Thus, the statement must hold. ■

Lemma 1 The value function defined in (1.5) has the following properties:

$$\begin{aligned} v_W^-(W, A') &= -c\frac{A'}{W} \\ v_W^0(W, A') &= \frac{A'}{W} - 1 + \frac{\beta}{1+\pi}D(W, A') \\ v_W^+(W, A') &= 0 \end{aligned} \tag{B.4}$$

Proof. This is simply an extended application of the Envelope Theorem. First define:

$$\Omega(W', W, A') \equiv A'\left[\ln\left(\frac{W'}{B'}\right) + c\ln\left(\frac{W'}{W}\right)\right] - W' + \frac{\beta}{1+\pi}\int v(W', A'')\,dF(A''|A') \tag{B.5}$$

as the ex ante objective function of the firm. Thus, by definition:

$$v(W, A') \equiv \max_{W'} \Omega(W', W, A') \tag{B.6}$$

and:

$$v_W(W, A') = \Omega_1\frac{dW'}{dW} + \Omega_2 \tag{B.7}$$

Now consider the value of this derivative in each of the three continuation regimes. Note first that, if W' is indeed changed, optimality dictates that it will be changed in such a way that $\Omega_1 = 0$. This is the exact same logic used in the Envelope Theorem.

It therefore follows that:

$$
\begin{aligned}
v_W^- (W, A') &= \Omega_2^- = -c\frac{A'}{W} \\
v_W^+ (W, A') &= \Omega_2^+ = 0
\end{aligned}
\tag{B.8}
$$

where $\Omega^{+/-}$ is defined analogously to $v^{+/-}$ to denote the ex ante objective conditional on W' being adjusted up or down respectively. It is only slightly less obvious what happens when $\Delta W' = 0$, i.e. when the wage is not adjusted. In this case, $W' = W$ and this implies that:

$$
\begin{aligned}
v^0 (W, A') &\equiv \Omega^0 (W, W, A') \\
&= A' \ln\left(\frac{W}{B'}\right) - W + \frac{\beta}{1+\pi}\int v (W, A'') \, dF (A''|A')
\end{aligned}
\tag{B.9}
$$

It therefore follows that:

$$
v_W^0 (W, A') = \frac{A'}{W} - 1 + \frac{\beta}{1+\pi}\int v_W (W, A'') \, dF (A''|A') \tag{B.10}
$$

Since, by definition $D (W, A') \equiv \int v_W (W, A'') \, dF (A''|A')$, the statement holds as required. ■

Proof of Proposition 2. First, note that we can re-write the continuation value conditional on each of the three possible continuation regimes:

$$
v (W, A') = \begin{cases} v^- (W, A') & \text{if} \quad A' < A'_l \\ v^0 (W, A') & \text{if} \quad A' \in [A'_l, A'_u] \\ v^+ (W, A') & \text{if} \quad A' > A'_u \end{cases} \tag{B.11}
$$

where superscripts $^{-/0/+}$ refer to whether the nominal wage is cut, frozen, or raised tomorrow. Note also that, due to the recursive nature of the problem:

$$\begin{aligned} A_l &\equiv l(W_{-1}); \qquad & A_u &\equiv u(W_{-1}) \\ \Longrightarrow A_l' &\equiv l(W); \qquad & A_u' &\equiv u(W) \end{aligned} \tag{B.12}$$

Thus we can write[2]:

$$\int v(W,A')\,dF(A'|A) = \int_0^{l(W)} v^-(W,A')\,dF + \int_{l(W)}^{u(W)} v^0(W,A')\,dF + \int_{u(W)}^{\infty} v^+(W,A')\,dF \tag{B.13}$$

Taking derivatives with respect to W and recalling the definition of $D(\cdot)$, we can write:

$$D(W,A) = \begin{cases} \int_0^{l(W)} v_W^-(W,A')\,dF + v^-(W,l(W))\,l'(W) \\ +\int_{l(W)}^{u(W)} v_W^0(W,A')\,dF + v^0(W,u(W))\,u'(W) - v^0(W,l(W))\,l'(W) \\ +\int_{u(W)}^{\infty} v_W^+(W,A')\,dF - v^+(W,u(W))\,u'(W) \end{cases} \tag{B.14}$$

Since $v(W,A')$ is continuous, it must be that $v^-(W,l(W)) = v^0(W,l(W))$ and $v^0(W,u(W)) = v^+(W,u(W))$. These "value matching" conditions allow us to write:

$$D(W,A) = \int_0^{l(W)} v_W^-(W,A')\,dF + \int_{l(W)}^{u(W)} v_W^0(W,A')\,dF + \int_{u(W)}^{\infty} v_W^+(W,A')\,dF \tag{B.15}$$

[2]Henceforth, "dF" without further elaboration is to be taken as "$dF(A'|A)$".

Finally, using the Envelope conditions in Lemma 1, and substituting into (B.15) we obtain (1.18) in the main text:

$$D(W,A) = \int_{l(W)}^{u(W)} \left[\frac{A'}{W} - 1\right] dF - \int_0^{l(W)} c\frac{A'}{W} dF + \frac{\beta}{1+\pi}\int_{l(W)}^{u(W)} D(W,A')\,dF \equiv (\mathbf{C}D)(W,A) \tag{B.16}$$

To verify that C is a contraction mapping over the "relevant range" (to be defined shortly), we confirm that Blackwell's sufficient conditions for a contraction hold here (see Stokey & Lucas, 1989, p.54). First, note that any values for (W,A) that render C unbounded cannot obtain under optimality, since they will necessarily violate the conditional first-order condition, (1.15). Thus, we can restrict our attention to a subset of values for (W,A) around the optimum for which C is bounded. This is what we define as the "relevant range". That C then maps the space of bounded functions into itself over this range holds by definition. Given this, monotonicity and discounting are straightforward to verify. To verify monotonicity, fix $(W,A) = (\bar{W},\bar{A})$, and take $\hat{D} \geq D$. Then note that:

$$\begin{aligned}\int_{l(\bar{W})}^{u(\bar{W})} \hat{D}(\bar{W},A')\,dF(A'|\bar{A}) - &\int_{l(\bar{W})}^{u(\bar{W})} D(\bar{W},A')\,dF(A'|\bar{A}) \\ &= \int_{l(\bar{W})}^{u(\bar{W})} \left[\hat{D}(\bar{W},A') - D(\bar{W},A')\right] dF(A'|\bar{A}) \geq 0\end{aligned} \tag{B.17}$$

Since $(\bar{W},\bar{A})$ were arbitrary, it thus follows that C is monotonic in D. To verify discounting, note that:

$$\begin{aligned}[\mathbf{C}(D+a)](W,A) &= (\mathbf{C}D)(W,A) + \frac{\beta}{1+\pi}a\left[F(u(W)|A) - F(l(W)|A)\right] \\ &\leq (\mathbf{C}D)(W,A) + \frac{\beta}{1+\pi}a\end{aligned} \tag{B.18}$$

Since we know that $\frac{\beta}{1+\pi} < 1$ it follows that C is a contraction over the relevant range. It therefore follows from the Contraction Mapping Theorem that C has a unique fixed point over the relevant range. ■

Proof of Proposition 4. Using (1.21), note that we can re-write the optimal nominal wage policy as:

$$\Delta \ln W = \begin{cases} (1) & \ln (A/W_{-1}) - \ln u \quad \text{if} \quad \ln (A/W_{-1}) > \ln u \\ (2) & 0 \qquad\qquad\qquad\qquad \text{otherwise} \\ (3) & \ln (A/W_{-1}) - \ln l \quad \text{if} \quad \ln (A/W_{-1}) < \ln l \end{cases}$$

To derive the implied log nominal wage change density, given W_{-1}, we take each régime in turn. The log nominal wage change c.d.f. in régime (1) is given by:

$$\begin{aligned} F(\Delta \ln W | W_{-1}, \Delta \ln W > 0) &= \Pr [\ln (A/W_{-1}) - \ln u \leq \Delta \ln W | W_{-1}] \\ &= \Pr [\ln (A/W_{-1}) \leq \Delta \ln W + \ln u | W_{-1}] \\ &= \tilde{F} (\Delta \ln W + \ln u | W_{-1}) \end{aligned}$$

where the last line follows from the fact that $W = A$ in the counterfactual (no DNWR) case. It follows that the log nominal wage change density, given W_{-1}, in régime (1) is simply $\tilde{f} (\Delta \ln W + \ln u | W_{-1})$ as stated. A completely analogous logic applies to régime (3). Finally, the density in régime (2) is given by:

$$\begin{aligned} f(\Delta \ln W | W_{-1}, \Delta \ln W = 0) &= \Pr \{\ln (A/W_{-1}) \in [\ln l, \ln u] \, | W_{-1}\} \\ &= \tilde{F} (\ln u | W_{-1}) - \tilde{F} (\ln l | W_{-1}) \end{aligned}$$

as stated. ■

Proof of Proposition 5. Denote the counterfactual nominal wage at time t as $W_t^* = A_t$. We seek the properties of the difference in average wage growth between the "actual" (with DNWR) and counterfactual cases, which we define as λ:

$$\begin{aligned} \lambda_T &\equiv \frac{1}{T}\sum_{s=t+1}^{t+T} \ln\left(\frac{W_s}{W_{s-1}}\right) - \frac{1}{T}\sum_{s=t+1}^{t+T} \ln\left(\frac{W_s^*}{W_{s-1}^*}\right) \\ &= \frac{1}{T}\left\{\ln\left(\frac{W_{t+T}}{W_{t+T}^*}\right) - \ln\left(\frac{W_t}{W_t^*}\right)\right\} \end{aligned} \tag{B.19}$$

Then note that, from the optimal wage policy of the firm, (1.21), it follows that the log–difference between the actual and counterfactual wages must be bounded:

$$\ln\left(\frac{W_t}{W_t^*}\right) \in [-\ln u, -\ln l] \tag{B.20}$$

Thus:

$$\begin{aligned} \sup \lambda_T &= \frac{1}{T}[\ln u - \ln l] = \frac{1}{T}\ln G \\ \inf \lambda_T &= \frac{1}{T}[\ln l - \ln u] = -\frac{1}{T}\ln G \end{aligned} \tag{B.21}$$

Therefore, for finite G:

$$\lim_{T\to\infty} \sup \lambda_T = 0 = \lim_{T\to\infty} \inf \lambda_T \tag{B.22}$$

■

B.2 Technical Details of Proposition 3

The following lemma will turn out to be useful in what follows:

Lemma 2 If $\ln x \sim N(\mu, \sigma^2)$ then it follows that:

$$\int_{\underline{x}}^{\overline{x}} x dF(x) = \exp\left[\mu + \frac{1}{2}\sigma^2\right]\left\{\Phi\left[\frac{\ln \overline{x} - \mu}{\sigma} - \sigma\right] - \Phi\left[\frac{\ln \underline{x} - \mu}{\sigma} - \sigma\right]\right\} \quad \text{(B.23)}$$

where $\Phi(\cdot)$ is the c.d.f. of the standard Normal.

Proof. Since x is log-Normally distributed, the p.d.f. of x is given by:

$$f(x) = \frac{1}{\sigma x}\phi\left(\frac{\ln x - \mu}{\sigma}\right) \quad \text{(B.24)}$$

where $\phi(\cdot)$ is the p.d.f. of the standard Normal. It follows that:

$$\int_{\underline{x}}^{\overline{x}} x dF(x) = \int_{\underline{x}}^{\overline{x}} x \frac{1}{\sigma x\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right] dx \quad \text{(B.25)}$$

Defining $z \equiv \ln x - \mu \Longrightarrow dx = \exp(\mu + z)\, dz$, we obtain:

$$\int_{\underline{x}}^{\overline{x}} x dF(x) = \int_{\ln \underline{x}-\mu}^{\ln \overline{x}-\mu} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\mu + z - \frac{1}{2\sigma^2}z^2\right] dz \quad \text{(B.26)}$$

Completing the square for the term in brackets:

$$\frac{1}{2\sigma^2}z^2 - z = \frac{1}{2\sigma^2}\left(z^2 - 2\sigma^2 z\right) = \frac{1}{2\sigma^2}\left(z - \sigma^2\right)^2 - \frac{1}{2}\sigma^2 \quad \text{(B.27)}$$

Substituting back into the former expression:

$$\begin{aligned}
\int_{\underline{x}}^{\overline{x}} x dF(x) &= \int_{\ln \underline{x}-\mu}^{\ln \overline{x}-\mu} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\mu + \frac{1}{2}\sigma^2 - \frac{1}{2}\left(\frac{z-\sigma^2}{\sigma}\right)^2\right] dz \qquad \text{(B.28)}\\
&= \exp\left[\mu + \frac{1}{2}\sigma^2\right]\int_{\ln \underline{x}-\mu}^{\ln \overline{x}-\mu} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{z}{\sigma} - \sigma\right)^2\right] dz\\
&= \exp\left[\mu + \frac{1}{2}\sigma^2\right]\left\{\Phi\left[\frac{\ln \overline{x} - \mu}{\sigma} - \sigma\right] - \Phi\left[\frac{\ln \underline{x} - \mu}{\sigma} - \sigma\right]\right\}
\end{aligned}$$

as required. ■

B.2.1 Obtaining the functions $D(W, u(W))$ and $D(W, l(W))$

For ease of reference we re-write (1.18) here as:

$$D(W,A) - \frac{\beta}{1+\pi}\int_{l(W)}^{u(W)} D(W,A')\,dF = \int_{l(W)}^{u(W)} \left[\frac{A'}{W} - 1\right] dF - \int_{0}^{l(W)} c\frac{A'}{W}dF \quad \text{(B.29)}$$

and proceed by using the method of undetermined coefficients. We conjecture that $D(W,A)$ is of the form:

$$D(W,A) = \alpha_1 \frac{A}{W} + \alpha_2 \quad \text{(B.30)}$$

and verify that this will indeed be the case for $A = u(W)$ or $l(W)$, using Lemma 2 to solve out the integrals in (B.29). Following this method yields[3]:

$$\begin{aligned} D(W,u(W)) &= (1+\pi)\frac{u(W)}{W}\left[\frac{\kappa_1 - (1+c)\Phi_1}{1-\beta(\kappa_1 - \Phi_1)}\right] - \frac{\kappa_2 - \Phi_2}{1 - \frac{\beta}{1+\pi}(\kappa_2 - \Phi_2)} \quad \text{(B.31)} \\ D(W,l(W)) &= (1+\pi)\frac{l(W)}{W}\left[\frac{\Phi_3 - (1+c)\kappa_1}{1-\beta(\Phi_3 - \kappa_1)}\right] - \frac{\Phi_4 - \kappa_2}{1 - \frac{\beta}{1+\pi}(\Phi_4 - \kappa_2)} \end{aligned}$$

where:

$$\begin{aligned} &\Phi_1 = \Phi\left[\tfrac{1}{\sigma}\left(-\ln G - \ln(1+\pi) + \tfrac{1}{2}\sigma^2\right) - \sigma\right] && \Phi_2 = \Phi\left[\tfrac{1}{\sigma}\left(-\ln G - \ln(1+\pi) + \tfrac{1}{2}\sigma^2\right)\right] \\ &\Phi_3 = \Phi\left[\tfrac{1}{\sigma}\left(\ln G - \ln(1+\pi) + \tfrac{1}{2}\sigma^2\right) - \sigma\right] && \Phi_4 = \Phi\left[\tfrac{1}{\sigma}\left(\ln G - \ln(1+\pi) + \tfrac{1}{2}\sigma^2\right)\right] \\ &\kappa_1 = \Phi\left[\tfrac{1}{\sigma}\left(-\ln(1+\pi) + \tfrac{1}{2}\sigma^2\right) - \sigma\right] && \kappa_2 = \Phi\left[\tfrac{1}{\sigma}\left(-\ln(1+\pi) + \tfrac{1}{2}\sigma^2\right)\right] \end{aligned} \quad \text{(B.32)}$$

and we define $G \equiv \frac{u(W)}{l(W)}$, the geometric gap between the two trigger values for A.

[3] Technical details of this derivation are available on request from the author.

B.2.2 Obtaining the functions $u(W)$ and $l(W)$

It is now straightforward to solve for the functions $u(W)$ and $l(W)$ by substituting the above (B.31) into the equations (1.17) to obtain after some algebra:

$$\left.\frac{\partial v}{\partial W}\right|_{\Delta W>0} = \frac{u(W)}{W}\left[\frac{1-c\beta\Phi_1}{1-\beta(\kappa_1-\Phi_1)}\right] - \frac{1}{1-\frac{\beta}{1+\pi}(\kappa_2-\Phi_2)} = 0 \quad \text{(B.33)}$$

$$\left.\frac{\partial v}{\partial W}\right|_{\Delta W>0} = \frac{l(W)}{W}\left[\frac{1+c-c\beta\Phi_3}{1-\beta(\Phi_3-\kappa_1)}\right] - \frac{1}{1-\frac{\beta}{1+\pi}(\Phi_4-\kappa_2)} = 0 \quad \text{(B.34)}$$

It is thus clear that the functions $u(W)$ and $l(W)$ are given by:

$$\begin{aligned} u(W) &= \left[\frac{1-\beta(\kappa_1-\Phi_1)}{1-\frac{\beta}{1+\pi}(\kappa_2-\Phi_2)}\cdot\frac{1}{1-c\beta\Phi_1}\right]\cdot W \\ l(W) &= \left[\frac{1-\beta(\Phi_3-\kappa_1)}{1-\frac{\beta}{1+\pi}(\Phi_4-\kappa_2)}\cdot\frac{1}{1+c-c\beta\Phi_3}\right]\cdot W \end{aligned} \quad \text{(B.35)}$$

These two equations clearly depend on $G \equiv \frac{u(W)}{l(W)}$, which is unknown so far. However, we can determine G using our expressions for $u(W)$ and $l(W)$ above:

$$\frac{u(W)}{l(W)} \equiv G = \frac{1+\pi-\beta(\Phi_4-\kappa_2)}{1+\pi-\beta(\kappa_2-\Phi_2)}\cdot\frac{1-\beta(\kappa_1-\Phi_1)}{1-\beta(\Phi_3-\kappa_1)}\cdot\frac{1+c-c\beta\Phi_3}{1-c\beta\Phi_1} \equiv T(G) \quad \text{(B.36)}$$

Note that all the terms on the RHS of this equation are functions of G, and not of W. Obtaining the relevant value of G requires solving for the fixed point(s) of the mapping defined by this equation. Given the relevant value of G, this implies that the Φ_is, $i = 1, ..., 4$, will be given constants, as will the coefficients on W in (B.35), and it follows that:

$$\begin{aligned} u(W) &= u \cdot W \\ l(W) &= l \cdot W \end{aligned} \quad \text{(B.37)}$$

as stated in the main text.

B.2.3 Properties of the Map $T(G)$

The completion of the solution requires finding the set of fixed points, $G^* = \{G : G = T(G)\}$. The set G^* represents the set of inaction regions that are internally consistent with the optimality conditions for W defined by the conditional first-order condition, (1.15). If G^* is a singleton, then the solution is complete. However, if there exist multiple solutions for G, then there exists an associated value of the firm for each $G \in G^*$. Simulations of the mapping $T(G)$ in (B.36) reveal that, whilst there always exists at least one fixed point for $T(G)$, there is not, in general, a unique fixed point. Thus, in the case where there exists more than one fixed point, we need a criterion for identifying which fixed point value of G maximises the value function, which is provided by the following proposition:

Proposition 13 *Where there exist multiple fixed points for the mapping $T(G)$, the wage policy that maximises the value function is that associated with $G^1 \equiv \min\{G : G = T(G)\}$.*

Proof. Take the case in which there exist 3 fixed points, and define these as $G^1 < G^2 < G^3$. The associated value functions are then given by:

$$v^i(W_{-1}, A) = \max_W \left\{ \begin{array}{c} A\left[\ln\left(\frac{W}{B}\right) + c\ln\left(\frac{W}{W_{-1}}\right)1^-\right] - W \\ +\frac{\beta}{1+\pi}\int v(W, A')\,dF(A'|A) \end{array} \right\}, \text{ s.t. } G = G^i \qquad \text{(B.38)}$$

for $i = 1, .., 3$. We claim that the following must be true:

$$v^1 \geq v^2 \geq v^3 \qquad \text{(B.39)}$$

To see this, note first that a higher value of G only serves to restrict the firm's choice of W by widening the region in which wages are not changed. In particular, under a lower value of G, the firm can always choose a W arbitrarily close to W_{-1}, and hence replicate the wage policy under a higher G, if it wishes. In general, though, the firm can do better than this under lower values of G. Thus the statement must hold. ■

B.2.4 Verifying Concavity of the Value Function in W

In order to verify that the solution obtained is indeed a maximum, and that the solution method employed in the main text is valid, we need to verify concavity of the value function in W, to which we now turn. Taking derivatives of equation (1.15) we obtain:

$$\frac{\partial^2 v\left(W_{-1}, A\right)}{\partial W^2} = -\left(1 + c\mathbf{1}^{-}\right)\frac{A}{W^2} + \frac{\beta}{1+\pi} D_W\left(W, A\right) \tag{B.40}$$

Then note that, from the definition of $D\left(W, A\right)$ in equation (B.29):

$$\begin{aligned} D_W\left(W, A\right) &= \int_0^{l(W)} c\frac{A'}{W^2}dF - \int_{l(W)}^{u(W)} \frac{A'}{W^2}dF + \frac{\beta}{1+\pi}\int_{l(W)}^{u(W)} D_W\left(W, A'\right)dF \\ &\quad - \left[c\frac{l\left(W\right)}{W}\right] l'\left(W\right) - \left[\frac{l\left(W\right)}{W} - 1 + \frac{\beta}{1+\pi} D\left(W, l\left(W\right)\right)\right] l'\left(W\right) \\ &\quad + \left[\frac{u\left(W\right)}{W} - 1 + \frac{\beta}{1+\pi} D\left(W, u\left(W\right)\right)\right] u'\left(W\right) \end{aligned} \tag{B.41}$$

Recalling the conditions for $u\left(W\right)$ and $l\left(W\right)$ in (1.17), the terms in square brackets cancel, and we obtain:

$$D_W\left(W, A\right) - \frac{\beta}{1+\pi}\int_{l(W)}^{u(W)} D_W\left(W, A'\right)dF = \int_0^{l(W)} c\frac{A'}{W^2}dF - \int_{l(W)}^{u(W)} \frac{A'}{W^2}dF \tag{B.42}$$

which is recursive in the function $D_W\left(W, \cdot\right)$. Thus, again we can use the method of undetermined coefficients to verify its form. We conjecture that this function is of the form:

$$D_W\left(W, A\right) = \alpha\frac{A}{W^2}$$

and verify that this will be the case for $A = u(W)$ or $l(W)$, again using the results in Lemma 2. Following this method in each case yields[4]:

$$\begin{aligned} D_W(W, u(W)) &= (1+\pi)\frac{u(W)}{W^2}\left[\frac{(1+c)\Phi_1 - \kappa_1}{1-\beta(\kappa_1 - \Phi_1)}\right] \\ D_W(W, l(W)) &= (1+\pi)\frac{l(W)}{W^2}\left[\frac{(1+c)\kappa_1 - \Phi_3}{1-\beta(\Phi_3 - \kappa_1)}\right] \end{aligned} \tag{B.43}$$

Substituting into (B.40) and using the solutions for the functions $u(W)$ and $l(W)$, we obtain:

$$\begin{aligned} \left.\frac{\partial^2 v(W_{-1}, A)}{\partial W^2}\right|_{\Delta W>0} &= -\frac{1}{W}\left[1 - \frac{\beta}{1+\pi}(\kappa_2 - \Phi_2)\right]^{-1} < 0 \\ \left.\frac{\partial^2 v(W_{-1}, A)}{\partial W^2}\right|_{\Delta W<0} &= -\frac{1}{W}\left[1 - \frac{\beta}{1+\pi}(\Phi_4 - \kappa_2)\right]^{-1} < 0 \end{aligned} \tag{B.44}$$

Since both of these expressions are strictly negative, it follows that the value function is concave in W at the optimum, and the solution obtained above is indeed a maximum.

[4] Again, technical details of these derivations are available on request from the author.

Appendix C

Technical Details of Chapter 3

Proof of Contraction Property in Proposition 2. First, we define the functional mapping in (3.15) as $\mathbf{C}$:

$$D(x,A) = (1-\delta)\left\{\begin{array}{c} \int_0^{l(\chi)} b_l dF + \int_{l(\chi)}^{u(\chi)} \left[g_x(\chi,A') - w\right] dF \\ + \int_{u(\chi)}^{\infty} b_u dF + \beta \int_{l(\chi)}^{u(\chi)} D(\chi,A')\, dF \end{array}\right\} \equiv (\mathbf{C}D)(x,A)$$

To verify that $\mathbf{C}$ is a contraction mapping we confirm that Blackwell's sufficient conditions for a contraction hold here (see Stokey & Lucas, 1989, p.54). First, note that any values for (x,A) that render $\mathbf{C}$ unbounded cannot obtain under optimality, since they will necessarily violate the conditional first-order condition, (3.4). Thus, we can restrict our attention to a subset of values for (x,A) around the optimum for which $\mathbf{C}$ is bounded. That $\mathbf{C}$ then maps the space of bounded functions into itself over this range holds by definition. Given this, monotonicity and discounting are straightforward to verify. To verify monotonicity, fix $(\chi,A) = (\bar{\chi},\bar{A})$, and take

$\hat{D} \geq D$. Then note that:

$$\int_{l(\bar{\chi})}^{u(\bar{\chi})} \hat{D}(\bar{\chi}, A')\, dF(A'|\bar{A}) - \int_{l(\bar{\chi})}^{u(\bar{\chi})} D(\bar{\chi}, A')\, dF(A'|\bar{A}) = \int_{l(\bar{\chi})}^{u(\bar{\chi})} \left[\hat{D}(\bar{\chi}, A') - D(\bar{\chi}, A')\right] dF(A'|\bar{A}) \geq 0 \tag{C.1}$$

Since $(\bar{\chi}, \bar{A})$ were arbitrary, it thus follows that $\mathbf{C}$ is monotonic in D. To verify discounting, note that:

$$\begin{aligned} [\mathbf{C}(D+a)](x, A) &= (\mathbf{C}D)(x, A) + \beta(1-\delta)\, a\, [F(u(\chi)|A) - F(l(\chi)|A)] \\ &\leq (\mathbf{C}D)(x, A) + \beta(1-\delta)\, a \end{aligned} \tag{C.2}$$

Since we know that $\beta(1-\delta) < 1$ it follows that $\mathbf{C}$ is a contraction over the relevant range. It therefore follows from the Contraction Mapping Theorem that $\mathbf{C}$ has a unique fixed point over this range. ■

Proof of Proposition 3. We conjecture that the functions $D(\cdot)$, $u(\cdot)$, and $l(\cdot)$ are of the forms:

$$\begin{aligned} D(x, A) &= \psi_1 A x^{\alpha-1} + \psi_2 \\ u(x) &= v x^{1-\alpha} \\ l(x) &= \lambda x^{1-\alpha} \end{aligned} \tag{C.3}$$

and verify that these are the case for $A = l(\chi)$or $A = u(\chi)$. To this end, first re-write (3.15) as:

$$D(x,A) - (1-\delta)\beta\int_{l(\chi)}^{u(\chi)} D(\chi, A')\,dF$$
$$= (1-\delta)\left\{\int_0^{l(\chi)} b_l dF + \int_{l(\chi)}^{u(\chi)} \left[A'\alpha\chi^{\alpha-1} - w\right] dF + \int_{u(\chi)}^{\infty} b_u dF\right\} \quad \text{(C.4)}$$

First consider the case where $A = u(x)$ or $\Delta X > 0$. In this case, we can use (C.3) and Lemma 1 to solve out the integrals in (C.4). Doing so and equating coefficients yields:

$$\begin{aligned} D(x,u(x)) &= u(x)\,\alpha x^{\alpha-1}\frac{(1-\delta)^{\alpha}(1+\mu)(\kappa_2-\Phi_2)}{1-\beta(1-\delta)^{\alpha}(1+\mu)(\kappa_2-\Phi_2)} \\ &\quad + (1-\delta)\frac{b_u(1-\kappa_1)+b_l\Phi_1 - w(\kappa_1-\Phi_1)}{1-\beta(1-\delta)(\kappa_1-\Phi_1)} \end{aligned} \quad \text{(C.5)}$$

Similarly, following the same procedure for the case where $A = l(x)$ or $\Delta X < 0$ we obtain:

$$\begin{aligned} D(x,l(x)) &= l(x)\,\alpha x^{\alpha-1}\frac{(1-\delta)^{\alpha}(1+\mu)(\Phi_4-\kappa_2)}{1-\beta(1-\delta)^{\alpha}(1+\mu)(\Phi_4-\kappa_2)} \\ &\quad + (1-\delta)\frac{b_l\kappa_1 + b_u(1-\Phi_3) - w(\Phi_3-\kappa_1)}{1-\beta(1-\delta)(\Phi_3-\kappa_1)} \end{aligned} \quad \text{(C.6)}$$

where:

$$\begin{aligned} &\Phi_1 = \Phi\left[\tfrac{1}{\sigma}\left(-\ln G + \ln\tfrac{(1-\delta)^{1-\alpha}}{1+\mu} + \tfrac{1}{2}\sigma^2\right)\right]; \quad \Phi_2 = \Phi\left[\tfrac{1}{\sigma}\left(-\ln G + \ln\tfrac{(1-\delta)^{1-\alpha}}{1+\mu} - \tfrac{1}{2}\sigma^2\right)\right] \\ &\Phi_3 = \Phi\left[\tfrac{1}{\sigma}\left(\ln G + \ln\tfrac{(1-\delta)^{1-\alpha}}{1+\mu} + \tfrac{1}{2}\sigma^2\right)\right]; \quad \Phi_4 = \Phi\left[\tfrac{1}{\sigma}\left(\ln G + \ln\tfrac{(1-\delta)^{1-\alpha}}{1+\mu} - \tfrac{1}{2}\sigma^2\right)\right] \\ &\kappa_1 = \Phi\left[\tfrac{1}{\sigma}\left(\ln\tfrac{(1-\delta)^{1-\alpha}}{1+\mu} + \tfrac{1}{2}\sigma^2\right)\right]; \quad \kappa_2 = \Phi\left[\tfrac{1}{\sigma}\left(\ln\tfrac{(1-\delta)^{1-\alpha}}{1+\mu} - \tfrac{1}{2}\sigma^2\right)\right] \end{aligned} \quad \text{(C.7)}$$

and $G \equiv \frac{u(x)}{l(x)}$ is the geometric gap between the upper and lower trigger values for changes in x. Substituting these back into the relevant first order conditions, and solving for $u(x)$ and $l(x)$ we obtain:

$$\begin{aligned} u(x) &= \left[w + b_u\left\{1-\beta(1-\delta)\left[1-\frac{R-1}{R}\Phi_1\right]\right\}\right]\frac{1-\beta(1-\delta)^{\alpha}(1+\mu)(\kappa_2-\Phi_2)}{1-\beta(1-\delta)(\kappa_1-\Phi_1)}\frac{x^{1-\alpha}}{\alpha} \\ l(x) &= \left[w + b_l\left\{1-\beta(1-\delta)\left[R-(R-1)\Phi_3\right]\right\}\right]\frac{1-\beta(1-\delta)^{\alpha}(1+\mu)(\Phi_4-\kappa_2)}{1-\beta(1-\delta)(\Phi_3-\kappa_1)}\frac{x^{1-\alpha}}{\alpha} \end{aligned} \tag{C.8}$$

where $R \equiv \frac{b_u}{b_l}$ is the geometric gap between the costs of gaining and losing the factor x. To complete the solution, we need a method for determining G, to which we now turn. First, if we define $Q \equiv \frac{w}{b_u}$, note that we can write:

$$\begin{aligned} G \equiv \frac{u(x)}{l(x)} &= \frac{QR + R\left\{1-\beta(1-\delta)\left[1-\frac{R-1}{R}\Phi_1\right]\right\}}{QR+1-\beta(1-\delta)\left[R-(R-1)\Phi_3\right]} \\ &\quad \times\frac{1-\beta(1-\delta)^{\alpha}(1+\mu)(\kappa_2-\Phi_2)}{1-\beta(1-\delta)^{\alpha}(1+\mu)(\Phi_4-\kappa_2)} \\ &\quad \times\frac{1-\beta(1-\delta)(\Phi_3-\kappa_1)}{1-\beta(1-\delta)(\kappa_1-\Phi_1)} \\ &\equiv T(G) \end{aligned} \tag{C.9}$$

To determine the relevant value of G, we must solve for the fixed point of the mapping $T(G)$. Given this fixed point, the Φ_i, $i = 1, ..., 4$ will be given constants, and the functions $u(\cdot)$ and $l(\cdot)$ will be of the form:

$$\begin{aligned} u(x) &= u\cdot\frac{x^{1-\alpha}}{\alpha} \\ l(x) &= l\cdot\frac{x^{1-\alpha}}{\alpha} \end{aligned} \tag{C.10}$$

as required. ■

Proof of Proposition 4. Note first that we can totally differentiate the first order condition, (3.4), to obtain:

$$\frac{dx}{dA} = -\frac{g_{xA} + \beta D_A}{g_{xx} + \beta D_x} > 0 \tag{C.11}$$

which is positive by virtue of the concavity of the value function ($g_{xx} + \beta D_x < 0$). The impact of the persistence of shocks, h_A, operates through the term D_A, to which we now turn. It will turn out to be useful to use the following change of variables result:

$$\int_a^b j(A')\, dF(A'|A) = \int_{k(a;A)}^{k(b;A)} j[h(A,\varepsilon')]\, d\tilde{F}(\varepsilon') \tag{C.12}$$

where $\varepsilon' = k(A';A)$ is the inverse of the A-section of h, and where $\tilde{F}$ is the distribution function of ε'. Using this we can write:

$$D(x,A) = (1-\delta)\left\{\begin{array}{c} \int_{k[0;A]}^{k[l(\chi);A]} b_l d\tilde{F} + \int_{k[l(\chi);A]}^{k[u(\chi);A]} g_x[\chi, h(A,\varepsilon')]\, d\tilde{F} - \int_{k[l(\chi);A]}^{k[u(\chi);A]} w d\tilde{F} \\ + \int_{k[u(\chi);A]}^{k[\infty;A]} b_u d\tilde{F} + \beta \int_{k[l(\chi);A]}^{k[u(\chi);A]} D[\chi, h(A,\varepsilon')]\, d\tilde{F} \end{array}\right\} \tag{C.13}$$

Thus, defining $k_l \equiv k[l(\chi);A]$ and $k_u \equiv k[u(\chi);A]$ for notational simplicity, and taking the derivative w.r.t. A we obtain[1]:

$$D_A(x,A) = (1-\delta)\left\{\begin{array}{c} \frac{\partial k_l}{\partial A}\left(b_l + w - g_x[\chi, h(A,k_l)] - \beta D[\chi, h(A,k_l)]\right) \\ +\frac{\partial k_u}{\partial A}\left(-b_u - w + g_x[\chi, h(A,k_u)] - \beta D[\chi, h(A,k_u)]\right) \\ + \int_{k_l}^{k_u} g_{xA}[\chi, h(A,\varepsilon')] \cdot h_A(A,\varepsilon')\, d\tilde{F} \\ +\beta \int_{k_l}^{k_u} D_A[\chi, h(A,\varepsilon')] \cdot h_A(A,\varepsilon')\, d\tilde{F} \end{array}\right\} \tag{C.14}$$

Now notice from the first order condition, (3.4), the first two lines of this expression

[1]We assume that the support of A', $[0,\infty)$, is invariant to A, so that $\partial k[0;A]/\partial A = 0 = \partial k[\infty;A]/\partial A$.

must equal zero. Thus:

$$D_A(x,A) = (1-\delta)\left\{\int_{k_l}^{k_u} g_{xA}\left[\chi, h\left(A,\varepsilon'\right)\right]\cdot h_A \cdot d\tilde{F} + \beta\int_{k_l}^{k_u} D_A\left[\chi, h\left(A,\varepsilon'\right)\right]\cdot h_A \cdot d\tilde{F}\right\} \tag{C.15}$$

It follows that D_A is increasing in the persistence of shocks, h_A.

Using a similar derivation, we can also establish that:

$$D_x(x,A) = (1-\delta)\left\{\int_{k_l}^{k_u} g_{xx}\left[\chi, h\left(A,\varepsilon'\right)\right] d\tilde{F} + \beta\int_{k_l}^{k_u} D_x\left[\chi, h\left(A,\varepsilon'\right)\right] d\tilde{F}\right\} \tag{C.16}$$

It is thus also possible that h_A will affect D_x through its effect on h. However, the size of h_A has a first-order effect on D_A, whereas its effect on D_x is higher order. Thus the effect through D_A must dominate in (C.11), and it follows that greater persistence (higher h_A) leads to a greater response of factor demand to shocks at the optimum, dx/dA, as required. ■