école d'économie d'aix-marseille
aix-marseille school of economics

Working Papers / Documents de travail

Costly Wage Cuts, Relative Wage Comparisons, and Unemployment Hysteresis

Marco Fongoni

WP 2025 - Nr 40

Costly Wage Cuts, Relative Wage Comparisons, and Unemployment Hysteresis*

Marco Fongoni†

December 2025
Download the last version HERE

Abstract

This paper advances a theory of unemployment hysteresis—transitory shocks leave *permanent* effects—based on a model of endogenous path-dependent wage rigidity under incomplete employment contracts. Workers' relative wage comparisons—incumbents' aversion to wage cuts and new hires' concern with pay inequality—imply wage increases are partially irreversible, generating path dependence and asymmetry in wage adjustments. During recessions, hiring wages fail to adjust fully downward, depressing job creation and producing hysteresis effects and large unemployment fluctuations. A quantitative assessment shows that these effects can be significant under plausible calibrations of the cost of wage cuts and the sensitivity of workers to relative wages. A 1% transitory shock can generate a permanent increase in unemployment of about 0.5% to 15%, with benchmark values around 1.5–5.5%. The paper concludes by discussing the implications of the theory for the effectiveness of monetary policy and the empirical research on hysteresis effects, suggesting promising directions for future research.

Keywords: incomplete contracts; wage rigidity; irreversibility; hysteresis; unemployment.

JEL classification: E32; E52; E71; J23; J30; J60.

*This paper is dedicated to the memory of Rod Cross, a brilliant scholar and valued colleague whose friendship and pioneering contributions to the "Hysteresis Hypothesis" in macroeconomics were a profound source of inspiration. I would like to thank Georgios Angelis, Alexandros Loukas, Céline Poilly, Francesco Saverio Gaudio, Daniel Schaefer, Carl Singleton, and the participants to the JKU Economic Research Seminar (December, 2025) for very constructive discussion and feedback. I acknowledge financial support from the French government under the "France 2030" investment plan managed by the French National Research Agency Grant ANR-17-EURE-0020, and by the Excellence Initiative of Aix-Marseille University - A*MIDEX.

†Aix-Marseille Univ, CNRS, Aix-Marseille School of Economics (AMSE), Marseille, France. E-mail: marco.fongoni@univ-amu.fr.

1 Introduction

An enduring question in macroeconomics has been to explain the large, asymmetric, and persistent fluctuations in unemployment. The recent experiences of the global financial crisis and COVID-19 recession have revived macroeconomists' interest into a specific form of persistence known as "*hysteresis*" (see, for instance, the recent survey by Cerra, Fatás, and Saxena (2023)). A dynamic system is said to exhibit hysteresis whenever transitory disturbances have permanent effects. Macroeconomic models featuring hysteresis effects can therefore contribute meaningfully to our understanding of economic fluctuations, with crucial implications for the theory of unemployment—traditionally based on the "natural rate" hypothesis—and for the effectiveness of stabilisation policy (Cross, 2014; Blanchard, 2018). Models based on the natural rate are not hysteretic: as temporary shocks wear off, unemployment returns to the initial and unique steady state (the natural rate). Macroeconomic models modelling persistence through linear difference equations exhibiting a unit-root (zero-root) are also not hysteretic: a first shock followed by a second one of the same magnitude, but opposite direction, will bring the system back to the initial state. In contrast, dynamic systems exhibiting hysteresis do not return to the initial state after such a temporary disturbance (Cross, 1993).[1]

This paper advances a theory of unemployment hysteresis that is based on several inherent and well-established features of labour markets. Employment contracts are incomplete, meaning that while employers have monopsony power over wage setting, workers' productivity is in part discretionary and not contractible. Workers' relative wage considerations and aversion to "unfair" wages generate a cost of cutting the wage of existing workers, which spills over to the hiring wage due to pay equality concerns of new hires. This cost results in endogenous *downward wage rigidity* in response to moderate negative shocks and *muted wage cuts* during large shocks. Wage increases are therefore *partially irreversible*, leading to *path-dependence* and *asymmetry* in wage adjustments. One key implication is that, following a positive, fully transitory shock, hiring wages fail to adjust fully downwards, depressing hiring and permanently increasing unemployment—a phenomenon known as *remanence*. The contribution of this paper is to provide both a rigorous theoretical foundation and a first quantitative assessment of this form of hysteresis.

The theory is developed within a simple general equilibrium model with search and matching frictions, fixed labour supply, and without capital, consisting of a large representative, risk-neutral household and a large representative, price-taking firm. In the baseline model,

[1]The word "hysteresis" is derived from the Greek "to be behind" and was coined by physicist-engineer James Alfred Ewing, referring to the behaviour of electromagnetic fields in ferric metals when exposed to magnetising cycles. For the history of hysteresis and its application to economics—including the use of the concept by Joseph Schumpeter, James Tobin and Edmund Phelps, among others—see, for instance, Cross and Allan (1988). Cross (1993), Amable, Henry, Lordon, and Topol (1995), and Göcke (2002) provide rigorous definitions of hysteresis and discuss the use—and misuse—of the term in economics. On the crucial distinction between hysteresis and persistence in unit-root (zero-root) processes, see Amable, Henry, Lordon, and Topol (1994). See also the discussion at the end of this section and the formal definition given in Section 3.

inflation is normalised to zero, thereby sidestepping the distinction between real and nominal variables (an extension including inflation and monetary policy is discussed in the last section). The only sources of disturbance are unanticipated, temporary, and persistent shocks to the firm's value of output. This approach enables a transparent and tractable analysis of the model's central results. First, it is formally established that a necessary and sufficient condition for the presence of hysteresis is the full procyclicality of wages in a given phase of the business cycle. It is precisely this property that renders wage changes partially irreversible when wage cuts are costly. A typical recession-induced unemployment hysteresis—due to the hiring wage being downward "rigid" (acyclical), or "sticky" (less-than-proportionally cyclical)–has its roots into a preceding expansion phase where wages were fully increased (i.e. with elasticity of at least one). Relatively lower wages at the start of the cycle warrant *fully procyclical* wage increases during the initial expansion. Then, workers' relative wage comparisons—incumbents' aversion to wage cuts and new hires' aversion to disadvantageous pay inequality—imply that labour is relatively more costly during the subsequent recession: the firm posts fewer vacancies and unemployment is permanently higher. Analogously, the analysis demonstrates that for an economy currently located at a high-wage, high-unemployment steady state (due to past hysteresis effects), a recession would warrant fully procyclical wage cuts, reducing workers' wage entitlements, and therefore wages. This enables the creation of relatively more vacancies during the subsequent expansion, resulting in permanently lower unemployment. Moreover, these asymmetric adjustments of the hiring wage naturally translate into asymmetric vacancy and unemployment fluctuations.

The mechanism at the heart of this form of unemployment hysteresis is a model of optimal wage setting under incomplete employment contracts, where the firm anticipates the asymmetric effects that wage changes can have on workers' effort. The partial irreversibility of wage changes implied by this model, which belongs to the same class of the model developed by Elsby (2009), and further extended by Dickson and Fongoni (2019) and Fongoni (2024a), results in an optimal wage policy of the "Ss" form: depending on the *size* and *history* of shocks, wages can be either fully procyclical, rigid, or sticky.[2] By embedding this model in a general equilibrium framework in which new hires' wage entitlements are influenced by the wage currently paid to existing workers, the paper establishes that the path-dependent rigidities in the wage of existing workers "spill over" to the hiring wage, with fundamental consequences for the dynamics of vacancies and unemployment. Essentially, pay equality concerns of new hires endogenously give rise to "internal pay equity".[3] Since the wage dynamics resulting from

[2]That optimal policies of the Ss form can generate path-dependence and hysteresis effects—both at the individual and aggregate level—was already understood in the literature of kinked adjustment costs and state-dependent pricing (see, e.g., Bertola and Caballero, 1990; Caplin and Leahy, 1991; Dixit, 1992).

[3]As such, the present model provides a micro-foundation of existing models that exogenously impose internal pay equity—sometimes referred to as "equal treatment" (see, e.g., Snell and Thomas, 2010; Martins, Snell, and Thomas, 2010; Gertler and Trigari, 2009; Fukui, 2020). Further note that in Elsby (2009) and Fongoni (2024a) it is the wage of existing workers only that is downward rigid. The dynamics of hiring wages is not considered in Elsby (2009), whereas in Fongoni (2024a) it is fully procyclical due to new hires' low wage entitlements.

this model is *the* mechanism generating unemployment hysteresis in the present framework, the paper also discusses at length the empirical plausibility of the underlying assumptions, as well as that of the resulting implications for wage dynamics (see Section 5). In fact, the model builds on an extensive literature concerning labour markets and incomplete employment contracts, combining elements of efficiency-wage theory (Malcomson, 1981; Akerlof and Yellen, 1990) with well-documented aspects of employment relationships—work morale, fairness, reciprocity, and relative wage comparisons (Bewley, 2007; Fehr, Goette, and Zehnder, 2009). Moreover, the resulting path-dependent wage dynamics plausibly replicates the distribution of log-wage changes observed in the data (see, e.g., Elsby and Solon, 2019), it generates state-dependent wage rigidity (e.g. Grigsby, Hurst, and Yildirmaz, 2021), and implies similar asymmetries between the hiring and existing workers' wage (e.g. Hazell and Taska, 2025). Overall, these considerations provide strong support for path-dependent *hiring* wage rigidity and for the resulting theory of unemployment hysteresis.

Another contribution of the paper is a first step toward a quantitative assessment of the model's hysteresis mechanism. A comparative statics analysis identifies two parameters primarily affecting the size of hysteresis effects. One reflects workers' sensitivity to relative wage comparisons. This governs the elasticity of incumbent workers' effort to wage changes, which in turn affects the *level* of the optimal wage set by the firm. Importantly, while the effort of existing workers fluctuates procyclically whenever their wage is adjusted, that of new hires is constant: internal pay equity is such that new hires are paid their reference "fair" wage and exert "normal effort" (independent of the wage). The other parameter reflects workers' relative sensitivity to wages that are perceived as unfair (in effect, the firm's perception of this). This crucially governs the relative *cost* of cutting the wage of existing workers, which determines the partial irreversibility of wage increases. While this cost is calibrated to target empirically plausible frequencies of wage freezes, the calibration of the parameter governing the elasticity of effort poses a greater challenge. By definition, discretionary effort is hard to measure. To make progress, the analysis explores the implications of effort elasticities within an empirically plausible, but also wide range between 0 and 1 (see Fongoni (2024a) Section III.B. for a survey of the available evidence on the elasticity of workers' discretionary effort).

The quantitative implications of the model are studied via impulse response analyses to unanticipated, transitory, and persistent shocks to the firm's value of output. This exercise provides a clear picture of the parameter space (and implied empirical targets) within which hysteresis effects can be quantitatively significant. For instance, when the cost of wage cuts is calibrated to achieve a benchmark frequency of wage freezes of 28%, for elasticities of effort of 0.4 and 1, the model generates an unemployment remanence of 0.5% and 5.6% respectively. The latter figure implies a permanent increase in the unemployment rate of about half percentage point from an initial steady-state rate of 7%—which is significant, especially for a transitory shock of 1%. A quantitatively important result is obtained for a frequency of wage freezes of 40% and an elasticity of effort of 1. This calibration generates a remanence of 15.2%, that is, unemployment permanently increases from 7% to 8.1%. In terms of asymmetric fluctuations, the response of vacancies and unemployment to negative shocks is at least twice as large as their response to positive shocks in benchmark simulations—the effect being strongest in the presence of downward wage rigidity.

The paper concludes with a discussion of how the baseline model can be extended to allow for the role of inflation and monetary policy. If workers evaluate wages relative to a nominal reference wage—due to "money illusion" (Shafir, Diamond, and Tversky, 1997)—optimal wage setting implies path-dependent downward *nominal* wage rigidity. This provides a rationale for a central bank to increase the price level during downturns (by cutting the interest rate) to keep the path-dependent *real* wage aligned with the counterfactual frictionless wage. An interesting finding is that when a central bank successfully offsets hysteresis effects, the nominal interest rate remains permanently lower—i.e. the unemployment remanence is absorbed by the nominal interest rate.

The theory advanced in this paper carries novel theoretical and empirical implications for the study of unemployment and business cycles. It establishes a new channel for hysteresis effects rooted in the nature of labour markets, rather than in institutional factors—such as unionisation or welfare provisions—that are commonly linked to mechanisms of insider-outsider relations and skill deterioration of the unemployed (discussed below). A key insight is that hysteresis effects might be present in industries or sectors with a higher prevalence of incomplete employment contracts, and where workers are particularly sensitive to relative wage considerations. Moreover, the analysis establishes a link between periods of procyclical (irreversible) wage changes and subsequent hysteresis effects. This result can potentially guide empirical research, by pointing to periods of expansion (or recession) with highly procyclical hiring wages as a possible underlying cause of unemployment persistence. Nevertheless, the paper also underscores that empirically detecting hysteresis effects may be challenging when monetary policy actively offsets negative shocks in the presence of nominal rigidities. At the same time, it identifies the dynamics of the interest rate as a possible indicator: episodes of persistently low interest rates combined with stable inflation—such as in the aftermath of the Great Recession—may signal that the policy rate is absorbing part of the underlying hysteresis pressures. Accordingly, hysteresis effects are more likely to affect unemployment when monetary policy is constrained by the zero lower bound.

After a discussion of the related literature, the manuscript is organised as follows. Section 2 presents the baseline model. Section 3 characterises the equilibrium and establishes the central results of the paper. Section 4 explores the quantitative implications of the model's main mechanisms. Section 5 discusses the empirical plausibility of both assumptions and implications of the path-dependent rigid-wage model. Finally, Section 6 presents an extension of the model considering the role of inflation and monetary policy, it summarizes the key findings, and discusses directions for future research. Mathematical proofs, detailed derivations of the model, and a description of the computational approach are contained in an the Appendix.

Related literature. The theory advanced in this paper is related to the literature on unemployment hysteresis and to that of asymmetric unemployment fluctuations.

Hysteresis in unemployment. Hysteresis effects are those that persist after the initial cause generating the effects is removed. Hence, it may be misleading to associate unemployment hysteresis with theories of *persistent* effects of transitory shocks, or with models in which unemployment follows a unit-root (zero-root) process. In the theory of equilibrium unemployment, the term hysteresis was first disseminated by Phelps (1972), and later invoked by Blanchard and Summers (1986) to explain the persistently high unemployment in Europe in the 70s and 80s.[4] More recently, macroeconomists' interest into hysteresis effects is witnessing a revival (as a testimony of this, see the speech by Yellen (2016), the article by Blanchard (2018) in the *Journal of Economic Perspectives*, and the more recent survey by Cerra et al. (2023) in the *Journal of Economic Literature*). Unemployment hysteresis, broadly defined, has been associated with two main mechanisms in the literature: "insider-outsider" relations and "loss of skill" of the unemployed. Models of insider-outsider relations were originally proposed by Lindbeck and Snower (1987) and Blanchard and Summers (1986). The key mechanism can be summarised by a union that bargains the wage to achieve an employment target—a measure of the insiders—which evolves exogenously as a linear function of past employment. Unemployment persistence is governed by the coefficient on past employment, which may equal one (i.e. employment dynamics have a unit root). More recently, Galí (2022) incorporates this mechanism into a New Keynesian framework to study the implications for monetary policy. The loss of skill channel was first analysed by Pissarides (1992): after an adverse shock, the human capital of the unemployed deteriorates, resulting in fewer vacancies due to a lower average quality of unemployed workers. The negative feedback loop between the loss of skill and the decline in vacancies can result in permanently higher unemployment after adverse, temporary shocks. Acharya, Bengui, Dogra, and Wee (2022) explore the qualitative implications of this mechanism for optimal monetary policy into a general equilibrium search model with nominal wage rigidities and a zero lower bound constraint.[5] To date, theories based on the deterioration of skills are the most strongly supported by the available evidence (see the recent evidence in Furlanetto, Lepetit, Robstad, Rubio-Ramírez, and Ulvedal (2025), and the one discussed in Blanchard (2018) and Cerra et al. (2023)).[6] The present paper contributes to this literature by advancing a novel, plausible mechanism for hysteresis effects, based on several inherent features of employment relationships and the dynamics of wages.

[4]For the literature emphasising output hysteresis, such as endogenous growth models, see the survey by Cerra et al. (2023).

[5]Ljungqvist and Sargent (1998) also develop a general equilibrium search model where workers' skill depreciation generates high persistence of unemployment in response to transitory shocks. However, their model does not generate hysteresis effects, that is, unemployment eventually converges—albeit slowly—to the unique steady state once the shock wears off.

[6]However, these authors conclude that the loss of skill mechanism is more likely to capture persistent effects, rather than permanent hysteresis effects, due to the entry of skilled workers and the exit of the unskilled long-term unemployed from the labour force (see also Yagan, 2019).

Asymmetric fluctuations. The two most closely related papers that study asymmetric fluctuations due to downward wage rigidity, or stickiness, are those of Abbritti and Fahr (2013) and Dupraz, Nakamura, and Steinsson (2025). Abbritti and Fahr (2013) consider a New Keynesian model with search and matching frictions and asymmetric wage adjustment costs. These costs affect the wage through the bargaining power of workers, which is assumed to be countercyclical (i.e. negatively related to the wage). The asymmetric adjustment cost (paid by the firm) implies that workers capture a larger fraction of the surplus during recessions relative to expansions, leading to downward wage stickiness and asymmetric effects of shocks on unemployment. Dupraz et al. (2025) consider a modern incarnation of Friedman's "plucking model" of the business cycle, by analysing a search and matching framework with endogenous separations and an exogenous constraint on downward nominal wage adjustments. As a consequence, negative shocks result in higher unemployment while positive shocks result in wage increases: unemployment sometimes rises far above its steady state, but never falls below. The present paper contributes to this literature by studying asymmetric fluctuations that are due to endogenous, path-dependent wage rigidity. This model not only generates richer and more realistic wage dynamics—capturing both stickiness and rigidity depending on the size of shocks—but also enables a more transparent quantitative assessment of their implications, thanks to its evidence-based theoretical foundation.[7]

2 Baseline Model

This section develops a model of path-dependent wage rigidity under incomplete employment contracts embedded into an otherwise standard general equilibrium model of the economy with search and matching frictions and without capital. There is a representative large household consisting of a continuum of members with risk-neutral preferences over consumption and labour supply; and a representative large, price-taking firm employing household members to produce output according to a linear production function with labour as the only input. The model departs from an otherwise standard framework as follows: i) the firm's output is determined by its employed workers' effort and the employment contract is incomplete: workers' effort is discretionary and not contractible; ii) employed workers evaluate wage contracts with respect to a reference "fair" wage, which varies based on their employment status, and they are particularly averse to unfair wages; iii) wages are unilaterally set by firms anticipating the effect that these can have on effort and therefore on output.

The only stochastic element is a shock to the firm's revenue productivity. Inflation is normalised to zero, meaning that real and nominal variables are the same in the baseline model (Section 6.1 develops an extension of the model considering inflation and the role of

[7]Section 5.2 compares the wage dynamics implied by the present framework with other forms of downward wage rigidity and stickiness considered in the business cycle literature (e.g. Benigno and Ricci, 2011; Blanchard and Galí, 2010; Schmitt-Grohé and Uribe, 2016; Michaillat, 2012; Martins et al., 2010).

monetary policy). Following a standard approach, general equilibrium results from the firm being owned by the household and by the government collecting taxes from the household to finance unemployment income.

2.1 Labour market environment

The labour force is constant and normalized to unity. In each period, there is a measure n_t of employed workers and a measure $u_t = 1 - n_t$ of unemployed workers. At the beginning of each period, unemployed workers carried over from the previous period u_{t-1} search for jobs, while a fraction of the existing employed workers carried over from the previous period n_{t-1} are separated with constant probability $\rho \in (0, 1)$. Separated workers cannot be immediately re-hired. Denote *newly hired* workers by n_{nt}, that is, the fraction of unemployed workers that are hired in each period, and by $n_{it} \equiv [1 - \rho]n_{t-1}$ the fraction of *incumbent* workers, so that in each period $n_t = n_{it} + n_{nt}$. Henceforth, the subscript n stands for newly hired, while the subscript i stands for incumbent. Newly hired workers can immediately produce output for the firm. At the beginning of each period, a worker can be in one of three states: unemployed and searching for a job; employed as a new hire; or employed as an incumbent if they were employed in the previous period and were not laid off. This implies a worker is a new hire only in their first employment period.

Search is random. The number of job matches taking place in each period is $m_t \equiv m(u_{t-1}, v_t) = \overline{m}u_{t-1}^{\alpha}v_t^{1-\alpha}$ where $v_t \in \mathbb{R}_+$ is the number of vacancies, $\overline{m} \in \mathbb{R}_{++}$ is the efficiency of the matching process, and $\alpha \in (0, 1)$ is the elasticity of the matching function with respect to unemployment. Let $\theta_t \equiv v_t/u_{t-1}$ define labour market tightness, then, the probability that a vacant job is matched with a worker is $h_t \equiv h(\theta_t) = \frac{m_t}{v_t} = \overline{m}\theta_t^{-\alpha}$, while the probability that an unemployed worker makes contact with a vacancy is $f_t \equiv f(\theta_t) = \frac{m_t}{u_{t-1}} = \overline{m}\theta_t^{1-\alpha}$. These assumptions imply that $m(u_{t-1}, v_t) = h(\theta_t)v_t = f(\theta_t)u_{t-1}$. The parameters of the model are such that every worker-vacancy match is mutually advantageous for the household and the firm (the precise conditions for this will be spelled out later). Hence, $f(\theta_t)$ captures the job-finding rate in the model, implying that for given $\{n_0, u_0\}$ employment and unemployment evolve according to

$$n_t = [1 - \rho]n_{t-1} + f(\theta_t)u_{t-1} \tag{1}$$

$$u_t = [1 - f(\theta_t)]u_{t-1} + \rho n_{t-1} \tag{2}$$

and that newly hired workers are given by $n_{nt} \equiv h(\theta_t)v_t = f(\theta_t)u_{t-1}$.

2.2 Household

There is a representative large household consisting of a continuum of infinitely lived members of measure 1.

Household member preferences. Each household member $j \in [0,1]$ has time-separable preferences over consumption $c_t^j \in \mathbb{R}_+$, labour supply $n_t^j \in \mathbb{R}_+$ (which is fixed and normalised to 1) and effort $e_t^j \in \mathbb{R}_+$, given by $U(c_t^j, e_t^j) = c_t^j + n_t^j \nu(e_t^j)$, where $n_t^j \in \{n_{nt}^j, n_{it}^j\}$ depending on whether the household member is employed as new hire or as an incumbent, and $n_t^j = 0$ if unemployed. The function $\nu(e_t^j) \in \{\nu(e_{nt}^j), \nu(e_{nt}^j)\}$ captures the household member's *net payoff from employment*. This crucially depends on their choice of effort and employment status. The specific form of $\nu(e_t^j)$ is described next.

Net payoff from employment. Employed household members evaluate wage contracts with respect to a reference "fair" wage $r_t^j \in \mathbb{R}_+$ which depends on their employment status. New hires find it fair to be paid the same wage that is currently paid to incumbent workers, while incumbent workers reference wage is given by their most recent wage contract. More formally:

$$r_t^j = \begin{cases} r_{nt}^j = w_{it}^j & \text{if } n_{nt}^j = 1 \\ r_{it}^j = w_{t-1}^j & \text{if } n_{it}^j = 1 \end{cases} \qquad r_{i0} \text{ given.} \tag{3}$$

Since employment contracts are incomplete, workers have discretion over the effort they exert when employed, which is chosen to maximise their net payoff from employment, considering the wage they are paid in relation to their reference wage:

$$\nu(e_t^j) \equiv \nu(e_t^j, w_t^j, r_t^j) = -\frac{e_t^{j2}}{2} + \bar{e}e_t^j + e_t^j \mu\left(q(w_t^j) - q(r_t^j)\right). \tag{4}$$

The first two terms represent the worker's intrinsic psychological net cost of effort, where the parameter $\bar{e} \in \mathbb{R}_{++}$ captures the marginal benefit of effort when workers are paid their reference wage.[8] The third term is referred to as the *morale function*, that depends on the worker's evaluation of the wage in relation to the reference wage: μ is a piece-wise linear gain loss function such that $\mu(x) = \eta x$ if $x > 0$ and $\mu(x) = \lambda\eta x$ if $x < 0$, where $\eta \in \mathbb{R}_+$ and $\lambda \geq 1$. Finally, $q(x) = x^{1-\gamma}/[1-\gamma]$ if $\gamma \in (0,1)$, and $q(x) = \log x$ if $\gamma = 1$. The specific form of ν is in the spirit of existing models of reference-dependent workers' preferences in the labour market.[9] Hence, the parameter η captures the sensitivity of relative wage considerations on workers' morale, while λ captures the degree of workers' aversion to "unfair" wages below their reference wage. In combination with the reference wage formation in (3), the net payoff in (4) implies that, if $\lambda > 1$, new hires will be averse to disadvantageous pay inequality, while incumbents will be averse to wage cuts. For brevity, these asymmetries are referred

[8] The inclusion of a psychological marginal benefit of exerting effort reflects the notion that workers have intrinsic motivation to work, i.e. they are willing to exert a strictly positive amount of effort that is independent of monetary incentives (see Cassar and Meier (2018) and the discussion in Altmann, Falk, Grunewald, and Huffman (2014, Appendix)). As shown later, the implication of this assumption is also consistent with the finding that it is not wage *levels* but wage *changes* that affect workers' effort (see, e.g., Bewley, 2007).

[9] See, for instance, the worker's payoff modeled in Breza, Kaur, and Shamdasani (2016), Sliwka and Werner (2017), Kaur (2019), as well as the one used by Danthine and Kurmann (2007, 2010). The payoff in (4) is very close in spirit to the worker payoff in the general model of asymmetric reference-dependent reciprocity developed by Dickson and Fongoni (2019).

to as workers' *concerns for relative wages.* The morale function captures the psychological cost/benefit of effort associated with the worker's perception of fairness. If the wage exceeds the reference wage (it is perceived as a gift) the worker gains some additional benefit of effort and an increase in effort (a gift to the firm) will increase utility. If the wage falls short of the reference wage (it is perceived as unfair) there is a psychological cost of effort and a reduction in effort (an unkind action towards the firm) increases utility. As such, the morale function implies the worker's net payoff from employment exhibits reciprocity.[10]

For the remainder of this section, denote the optimal choice of effort of a household member with $\tilde{e}_t^j \equiv \tilde{e}(w_t^j, r_t^j, \lambda)$, and the resulting optimised net payoff from employment by $\tilde{\nu}_t^j \equiv \tilde{\nu}(w_t^j, r_t^j)$. Henceforth, the tilde over a function denotes the outcome of an optimal choice.

Household optimisation problem. Since household members are otherwise identical, it turns out that all new hires will be paid the same wage, implying $\int w_{nt}^j n_{nt}^j dj = w_{nt} n_{nt}$, and all incumbents will be paid the same wage, implying $\int w_{it}^j n_{it}^j dj = w_{it} n_{it}$. Hence, the resulting optimal effort choices will also be the same among workers of the same employment status. Taking these and the resulting optimised net payoffs from employment $\tilde{\nu}_{nt}$ and $\tilde{\nu}_{it}$ as given, the household pools the income of all its members together before deciding on consumption. With additively separable utility between consumption and employment, this implies the household equalises utility across members and acts as if it has utility function $U(c_t, n_t) = c_t + \tilde{\nu}_{nt} n_{nt} + \tilde{\nu}_{it} n_{it}$, where the dependency on the respective wages and reference wages of new hires and incumbents has been suppressed to ease exposition.

The household allocates total consumption in each period to maximise the expected discounted sum of utility, taking as given the employment and unemployment flows, (1) and (2), and being subject to the budget constraint

$$c_t = w_{nt} n_{nt} + w_{it} n_{it} + b u_t + d_t - \tau_t, \tag{5}$$

where $b \in \mathbb{R}_+$ is the per-period constant value of unemployment income, d_t is the firm's per-period profit and $\tau_t \in \mathbb{R}_+$ is a lumpsum tax from the government. Future payoffs are discounted by a factor $\beta \in (0, 1)$. The value function of the household is denoted by $\mathcal{V}(n_t, u_t)$.

Employed members' participation constraints. Denote the household value of having one more member employed as a new hire by $\mathcal{E}_n(w_{nt}, r_{nt})$; the value of having one more member employed as an incumbent by $\mathcal{E}_i(w_{it}, r_{it})$; and the value of having one more member unemployed by $\mathcal{U}_t$ (all in terms of consumption units). These value functions are derived in detail in Appendix B. Then, denote by $\mathcal{S}_n(w_{nt}, r_{nt}) \equiv \mathcal{E}_n(w_{nt}, r_{nt}) - \mathcal{U}_t$ the marginal utility for a household of having a member employed as a new hire rather than unemployed, and by $\mathcal{S}_i(w_{it}, r_{it}) \equiv \mathcal{E}_i(w_{it}, r_{it}) - \mathcal{U}_t$ the marginal utility for a household of having a member

[10] See, for instance, Rabin (1993) and Dufwenberg and Kirchsteiger (2000, 2004) for more general treatments of "intention-based" reciprocity.

employed as an incumbent rather than unemployed. In each period, the household will prefer to have members employed as new hires rather than unemployed if $\mathcal{S}_n(w_{nt}, r_{nt}) \geq 0$, and will prefer to have members employed as incumbents rather than unemployed if $\mathcal{S}_i(w_{it}, r_{it}) \geq 0$, which together characterise new hires' and incumbents' participation constraints.

2.3 Firm

There is a representative large firm operating with a linear production technology using labour as the only input and employing a measure of n_t workers in each period.

Technology and profit. At the beginning of each period the firm employs a measure n_{nt} of new hires, obtained by posting v_t vacancies at a constant cost $\kappa \in \mathbb{R}_{++}$, and a measure $n_{it} = [1-\rho]n_{t-1}$ of incumbents inherited from the previous period. Incomplete employment contracts imply that all workers have discretion over effort. The firm's output is therefore given by

$$y_t \equiv z_t[e_{nt}n_{nt} + e_{it}n_{it}], \tag{6}$$

where z_t is a real shock affecting the revenue product of the firm, and follows an AR(1) process (to be specified below). The firm's per-period profit is $d_t \equiv y_t - w_{nt}n_{nt} - w_{it}n_{it} - \kappa v_t$. Note, since workers' effort is endogenous to both the wage and the reference wage, and new hires and incumbents are characterised by different reference wages, they are not *de facto* perfect substitutes in production.[11]

Firm's optimisation problem. The firm chooses the number of vacancies to post and the wages to pay its newly hired and incumbent workers in each period to maximise the expected discounted sum of profits, taking as given the real shock z_t, labour market tightness θ_t, its employed workers' reference wages $\{r_{nt}, r_{it}\}$ and endogenous effort choices $\{e_{nt}, e_{it}\}$, and the number of incumbent workers $n_{it} = [1-\rho]n_{t-1}$ carried over from the previous period.[12] Denote the value function of the firm by $\mathcal{J}(r_{nt}, r_{is}, z_s, n_{s-1})$. While the firm is owned by the household, it is assumed that it discounts future profit with a factor $\hat{\beta} \equiv \delta\beta$, where $\delta \in (0,1]$ captures an additional source of discounting. This assumption is innocuous for the theory of hysteresis developed in this paper but, as will become clearer in the calibration of Section 4, it helps to match empirically plausible wage elasticities while maintaining a conventional value

[11] A tractable form of imperfect substitutability would assume that new hires are strictly less productive than incumbents (e.g. due to skill loss during unemployment). For a sufficiently large discrepancy in relative productivity, however, the present model would imply a non-degenerate distribution of incumbent workers' wages. This poses two challenges. First, it would require changing the assumption on new hires' reference wage formation. Second, it would require to keep track of the distribution of the wage of new hires in each period—becoming incumbents in the next period—along the equilibrium path. Pursuing this route will substantially increase the complexity of the model at a loss of greater tractability and transparency, without effectively improving our understanding of the key mechanisms behind the paper's central results. For these reasons, this extension is beyond the scope of the paper.

[12] That the representative large firm does not internalise the congestion externality it creates when posting vacancies is a standard simplification adopted in the literature to avoid the need to model a continuum of otherwise identical firms (see, e.g., Shimer, 2010; Petrosky-Nadeau and Wasmer, 2017).

of the household discount factor.[13]

Timing of decisions. The timing of decisions is important. At the beginning of each period, the firm initially sets the wage of its existing workers. Then, it simultaneously decides on the number of vacancies to post and on the wage to be paid to each potential new hire. To ease exposition, we consider the firm setting the hiring wage first, taking vacancies as given, and then deciding on the optimal number of vacancies, knowing how much each additional worker will cost.[14]

Choice of wages. The firm sets the wages of its newly hired and incumbent workers anticipating the effect that these can have on their effort $\{\tilde{e}_{nt}, \tilde{e}_{it}\}$ due to concerns for relative wages. The linearity of $\mathcal{J}$ with respect to both n_{nt} and n_{it} (inherited by the linear production function), implies the firm's optimal wage setting problem can be formalised as if the firm is setting the wage of each individual worker separately. As such, in each employment period and for each worker, wage setting can be thought of as the outcome of a two-stage game of complete and perfect information in which the firm makes take–it–or–leave–it wage offers to a worker, who subsequently chooses how much discretionary effort to exert. Since choices are made sequentially and the firm is motivated only by profit, the game can be solved by backward induction. Under the assumption on the timing of decisions, the firm will choose the wage of all its incumbent workers first, subject to $r_{it} = w_{t-1}$ for all t and the participation constraint $\mathcal{S}_i(w_{it}, r_{it}) \geq 0$, and then the wage of each new hire, subject to $r_{nt} = \tilde{w}_{it}$ in the first period, $r_{it+1} = w_t$ from the second period onwards, and the participation constraint $\mathcal{S}_n(w_{nt}, r_{nt}) \geq 0$. Denote the resulting optimal wages by $\tilde{w}_{nt} \equiv \tilde{w}(r_{nt}, z_t)$ and $\tilde{w}_{it} \equiv \tilde{w}(r_{it}, z_t)$ respectively for new hires and incumbents. The properties of these will be established in the next section.

Choice of vacancies. For given optimal wages and resulting effort choice of workers, the firm's optimisation problem over vacancies can be expressed recursively as

$$\mathcal{J}(z_t, n_{t-1}) = \max_{v_t} \Big\{ z_t[\tilde{e}_{nt} n_{nt} + \tilde{e}_{it} n_{it}] - \tilde{w}_{nt} n_{nt} - \tilde{w}_{it} n_{it} - \kappa v_t + \hat{\beta} \mathbb{E}_t \mathcal{J}(z_{t+1}, n_t) \Big\} \tag{7}$$

subject to $v_t \geq 0$, $n_{nt} = h_t v_t$, and $n_{it} = [1-\rho] n_{t-1}$. The dependency of $\mathcal{J}$ on the workers' reference wages has been suppressed to ease the exposition. The vacancy posting decision of the firm will therefore result in an optimal vacancy-unemployment ratio in each period,

[13] As discussed below, δ effectively captures the extent to which the firm internalises the effect of the updating of the reference point by its employed workers, on top of a more standard discount factor. Nevertheless, the parameter δ can be thought as capturing, in reduced form, additional sources of discounting such as firm-specific survival risk, agency frictions, or exogenous financial or regulatory wedges.

[14] If the firm were to choose the wage of all its workers (new hires and incumbents) simultaneously, it would optimally internalise that a higher wage of its existing workers implies a greater cost of hiring new workers. This, in turn, would result in the firm offering relatively lower wages to all its workers. Putting aside the empirical validity of this additional form of wage compression, it will unnecessarily complicate the model without fundamentally affecting the paper's central results. Further note it is equally plausible that there are delays between the firm's decision to set the wage of its existing workforce and that of any potential new hire when vacancy posting takes place, which lends support to the timing of decisions assumed here.

which is denoted by $\tilde{\theta}_t \equiv \tilde{\theta}(r_{nt}, z_t)$.

2.4 Government

It is assumed that the value bu_t in the budget constraint is financed by transfer payments. Hence, the government runs a balanced budget in each period: $\tau_t = bu_t$.

3 Equilibrium Analysis

This section begins by characterising the out–of–steady-state optimal choices of the household and the firm under the assumption of perfect foresight. This approach enables an analytical characterisation of the model equations, and in particular of the optimal wage setting policy of the firm, while also preserving their key dynamic properties. This solution is also the one implemented later to simulate the model impulse responses in Section 4. Next, the existence and (non-)uniqueness of the model steady state will be characterised. Finally, the analysis will formally establish the main mechanisms for the presence of hysteresis effects and asymmetric fluctuations.

Assume the real shock z_t is characterised by the following AR(1) process:

$$z_{t+1} = \bar{z}^{1-\rho_z} z_t^{\rho_z} \exp(\varepsilon_{t+1}), \quad z_0 \text{ given}, \tag{8}$$

where $\rho_z \in (0, 1)$ and ε_{t+1} is i.i.d. and distributed according to $\mathcal{N}(0, \sigma_z^2)$, and $\bar{z} \in \mathbb{R}_{++}$ denotes the steady state. The model equilibrium from any initial condition in period $t = s$ can be defined as follows.

Definition 1. *Given initial z_s, u_s and r_{is}, an equilibrium consists of paths for wages $\{w_{nt}, w_{it}\}$, reference wages $\{r_{nt}, r_{it}\}$, effort levels $\{e_{nt}, e_{it}\}$, market tightness θ_t, unemployment u_t and employment n_t rates, consumption c_t, output y_t, and transfers τ_t, such that z_t evolves according to (8); employed households choose effort that maximise the net payoff from employment (4), for given wage and reference wage; the household chooses consumption that maximise the expected discounted sum of utility subject to the budget constraint (5) and the dynamics of employment and unemployment, given its members' wages and reference wages; the firm chooses wages and vacancies to maximise its expected discounted sum of profits subject to the workers' participation constraints; employed workers' reference wages are given by (3); the employment and unemployment rate are given by (1) and (2); the government budget is balanced; and the goods market clears.*

The optimal decisions of the firm—namely, wage setting and vacancy posting—can be solved independently, as explained in the previous section.[15] Once the optimal wages paid

[15]Other prominent search and matching models in the literature that share this property are those, for instance,

to new hires and incumbent workers are characterised as functions of the state variables $\{r_{nt}, r_{it}\}$ and z_t, it is possible to solve for the optimal choice of vacancies, which in turn determines employment, unemployment and, together with employed workers' choice of effort, the output produced by the firm and the consumption of the household.

3.1 Optimal choices under perfect foresight

Consider an economy away from the steady state (due to some exogenous shock, or an initial condition).

Definition 2. *An equilibrium path under perfect foresight is such that, starting from any period $t = s$, where $z_s \neq \overline{z}$ and $\varepsilon_{t+1} = 0$ for all $t \geq s$, then $\mathbb{E}_t z_{t+1} = \overline{z}^{1-\rho_z} z_t^{\rho_z}$ for all $t \geq s$. Hence, if $z_s < \overline{z}$ then $\{z_t\}$ is an increasing sequence bounded above by $\overline{z}$, while if $z_s > \overline{z}$ then $\{z_t\}$ is a decreasing sequence bounded below by $\overline{z}$.*

Henceforth, this section considers the case of $\gamma = 1$, such that $q(x) = \log x$, and imposes the normalisation $\overline{z} = 1$.

Optimal discretionary effort. Each worker chooses effort that maximises their net payoff from employment, taking as given the wage they are paid w_t^j in relation to their reference wage r_t^j.

Proposition 1. *The optimal effort of an employed member is explicitly given by*

$$\tilde{e}_t^j \equiv \tilde{e}(w_t^j, r_t^j, \lambda) = \begin{cases} \bar{e} + \eta \left[\log w_t^j - \log r_t^j\right] & \text{if } w_t^j > r_t^j \\ \bar{e} & \text{if } w_t^j = r_t^j \\ \bar{e} + \lambda\eta \left[\log w_t^j - \log r_t^j\right] & \text{if } w_t^j < r_t^j \end{cases}$$

where $r_t^j = w_{it}^j$ if a new hire, and $r_t^j = w_{t-1}^j$ if an incumbent.

Optimal discretionary effort is a continuous and strictly increasing function of the wage in relation to the reference wage and displays a kink at $w_t^j = r_t^j$ if $\lambda > 1$. When $w_t^j = r_t^j$, workers exert $\tilde{e}(w_t^j, r_t^j, \lambda) = \bar{e}$, which is constant and referred to as *normal effort*. When $w_t^j > r_t^j$, workers will exert *supra-normal* effort $\tilde{e}(w_t^j, r_t^j, \lambda) > \bar{e}$, and when $w_t^j < r_t^j$, workers will exert *sub-normal* effort $\tilde{e}(w_t^j, r_t^j, \lambda) < \bar{e}$. Since household members are otherwise identical, workers' heterogeneity is captured by their reference wage, which reflects their employment status. All new hires, for which $r_{nt} = w_{it}$, will be paid the same wage and exert optimal effort $\tilde{e}(w_{nt}, w_{it}, \lambda)$, and all incumbents, for which $r_{it} = w_{t-1}$, will be paid the same wage and exert optimal effort $\tilde{e}(w_{it}, w_{t-1}, \lambda)$. Hence, optimal effort reflects how workers respond to

of Eliaz and Spiegler (2014), who assume that workers have no bargaining power, and Hall and Milgrom (2008) who assume that the workers' threats of quitting to unemployment are not credible if wage bargaining does not break down exogenously.

relative wage considerations: if $\lambda > 1$, new hires respond more strongly to disadvantageous pay inequality, while incumbents respond more strongly to wage cuts.

Optimal wage setting. It is possible to generalise the firm's optimal wage setting problem with an employed worker $\omega \in \{n, i\}$, with given reference wage $r_{\omega t}$, and optimal effort $\tilde{e}(r_{\omega t}, w_{\omega t}, \lambda)$ as follows:

$$\mathcal{J}_\omega(r_{\omega t}, z_t) = \max_{w_{\omega t}} \left\{ z_t \tilde{e}(w_{\omega t}, r_{\omega t}, \lambda) - w_{\omega t} + \hat{\beta}[1-\rho]\mathbb{E}_t \mathcal{J}_i(w_{\omega t}, z_{t+1}) \right\} \tag{9}$$

subject to $r_{it+1} = w_{\omega t}$ and $\mathcal{S}_\omega(w_{\omega t}, r_{\omega t}) \geq 0$. In this expression, $\mathcal{J}_\omega(r_{\omega t}, z_t)$ denotes the marginal value to the firm of employing an additional worker (new hire if $\omega = n$ or incumbent if $\omega = i$) in period t, and $\mathbb{E}_t \mathcal{J}_i(w_{\omega t}, z_{t+1})$ is the expected continuation value of an employment relationship with that worker being incumbent in the next period (recall that a worker is a new hire only in their first employment period). To ease the exposition, the analysis considers the case in which the workers' participation constraint is not binding, which is equivalent to assuming that unemployment income b is sufficiently low.[16] The more general treatment accounting for such possibility, as well as details on the derivations of the firm marginal values $\mathcal{J}_n$ and $\mathcal{J}_i$, is contained in Appendix B.

The first-order condition characterising the optimal wage paid to an employed worker is given by

$$z_t \frac{\partial \tilde{e}(w_{\omega t}, r_{\omega t}, \lambda)}{\partial w_{\omega t}} - 1 + \hat{\beta}[1-\rho]\mathbb{E}_t \frac{\partial \mathcal{J}_i(w_{\omega t}, z_{t+1})}{\partial r_{t+1}} = 0 \quad \forall w_{\omega t} \neq r_{\omega t} \tag{10}$$

where $\frac{\partial \tilde{e}(w_{\omega t}, r_{\omega t}, \lambda)}{\partial w_{\omega t}} = \mu' w_{\omega t}^{-1}$ is the marginal effect of a wage change on the worker's effort in period t and

$$\mathbb{E}_t \frac{\partial \mathcal{J}_i(w_{\omega t}, z_{t+1})}{\partial r_{t+1}} = \mathbb{E}_t z_{t+1} \frac{\partial \tilde{e}(w_{t+1}, w_{\omega t}, \lambda)}{\partial r_{t+1}} = -z_t^{\rho_z} \mu' w_{\omega t}^{-1} \tag{11}$$

is the marginal effect of a wage change in t on the expected continuation value of the employment relationship in $t+1$ (obtained by iterating forward the envelope condition). Note that in any period t, $\mu' = \eta$ if $w_{\omega t} > r_{\omega t}$ and $\mu' = \lambda\eta$ if $w_{\omega t} < r_{\omega t}$. Since (11) is strictly negative, the first-order condition (10) implies that the optimal wage equalises the current marginal benefit, in terms of effort, to the marginal cost of paying a higher wage in the current period, net of the additional expected marginal cost of employing a worker with a higher reference wage in the future. This additional expected marginal cost generates "wage compression", a distinctive feature of this class of models (see, e.g., Elsby, 2009). These considerations suggest that the firm discounting wedge δ, which affects $\hat{\beta}$, effectively captures the extent to which the firm internalises the effect of the updating of the reference point by its employed workers, on top of a more standard discount factor.

[16] The numerical simulations of the model performed in Section 4 confirm that the worker participation constraint is never binding. This is also due to the fact that, in contrast with a canonical model, employed workers gain a strictly positive net payoff from employment $\tilde{\nu}_t$ on top of the wage.

The discontinuity of the first-order condition at $w_{\omega t} = r_{\omega t}$ when $\lambda > 1$ implies the optimal wage policy follows an "Ss" policy characterised by two thresholds. As it is formally shown below, these thresholds can be expressed in terms of either one of the two state variables $r_{\omega t}$ and z_t. Following the timing of decisions, we first characterise the optimal wage paid to incumbents, and then that of new hires.

Incumbents. It is convenient to express the thresholds in terms of the real shock z_t as functions of the reference wage $r_{it} = w_{t-1}$. The lower threshold $z^l(w_{t-1})$ is such that if $z_t < z^l(w_{t-1})$, then the first-order condition is satisfied at a wage strictly below the reference wage, and the upper threshold $z^u(w_{t-1})$ is such that if $z_t > z^u(w_{t-1})$, then the first-order condition is satisfied at a wage exceeding the reference wage. If $z_t \in [z^l(w_{t-1}), z^u(w_{t-1})]$, the sum of profits is maximized at the kink, where $w_{it} = w_{t-1}$.

Proposition 2. *For all $t \geq s$ where $z_s \neq \overline{z}$, $\varepsilon_{t+1} = 0$, and $\mathbb{E}_t z_{t+1} = z_t^{\rho_z}$ (perfect foresight), the optimal wage policy of the firm employing an incumbent worker with reference wage $r_{it} = w_{t-1}$ is given by*

$$\tilde{w}_{it} \equiv \tilde{w}(w_{t-1}, z_t) = \begin{cases} \eta\{z_t - \hat{\beta}[1-\rho]z_t^{\rho_z}\} & \text{if } z_t > z^u(w_{t-1}) \\ w_{t-1} & \text{if } z_t \in [z^l(w_{t-1}), z^u(w_{t-1})] \\ \lambda\eta\{z_t - \hat{\beta}[1-\rho]z_t^{\rho_z}\} & \text{if } z_t < z^l(w_{t-1}) \end{cases} \tag{12}$$

where the thresholds are uniquely characterised by

$$z^u(w_{t-1}) - \hat{\beta}[1-\rho]z^u(w_{t-1})^{\rho_z} - \frac{w_{t-1}}{\eta} = 0 \tag{13}$$

$$z^l(w_{t-1}) - \hat{\beta}[1-\rho]z^l(w_{t-1})^{\rho_z} - \frac{w_{t-1}}{\lambda\eta} = 0 \tag{14}$$

and where $z^u(w_{t-1}) > z^l(w_{t-1})$ if $\lambda > 1$.

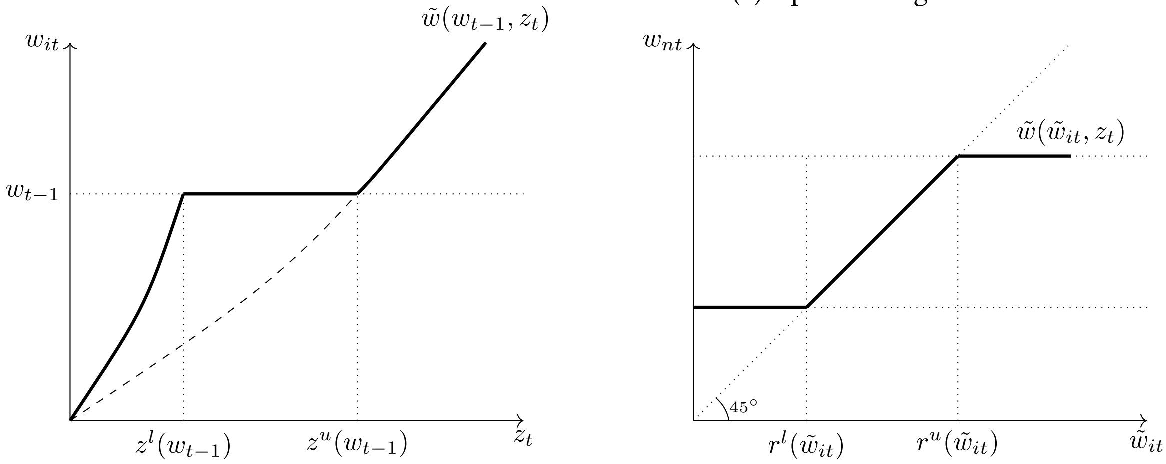
Proposition 2 establishes that in each period the optimal wage is nondecreasing in z_t and that there is a range of z_t within which the wage is not adjusted. This is referred to as the *range of rigidity*, which is nonempty if incumbent workers' are averse to wage cuts ($\lambda > 1$): in this region, the benefit of reducing the wage will be offset by the cost generated from the worker's response in the form of sub-normal effort. As it will be established later throughout the paper, this property can generate both downward *and* upward wage rigidity depending on the *size* and *history* of shocks. As such, the optimal wage setting policy established in Proposition 2 is referred to as *path-dependent wage rigidity.*

Another important property of the optimal wage policy is that for large enough shocks, the firm will optimally cut the wage of incumbent workers. However, the cut will be more *muted* relative to the case in which workers' effort response was not asymmetric (i.e. relative to the case in which $\lambda = 1$). If $\lambda > 1$ wage cuts are particularly costly due to incumbent workers' disproportionate drop in effort. As such, the firm optimally dampens wage decreases to attenuate this cost.[17] The presence of muted wage cuts is illustrated by the solid line being strictly above the dashed line when $z_t < z^l(w_{t-1})$ in Figure 1a below, which illustrates the optimal wage policy for incumbent workers as a function of the real shock z_t.

Figure 1: Optimal Wage Policies: Incumbents and New Hires

(a) Optimal Wage: Incumbents

(b) Optimal Wage: New Hires



New hires. In this case, it turns out to be more convenient to express the thresholds in terms of the reference wage r_{nt} as functions of the real shock z_t (recall that new hires' reference wage is $r_{\omega t} = r_{nt} = \tilde{w}_{it}$ in their first employment period, and then $r_{\omega t+1} = r_{it+1} = w_t$ in every subsequent period). Hence, the lower threshold $r^l(z_t) \equiv z^{u-1}(z_t)$ is such that if $r_{nt} < r^l(z_t)$, then the optimal wage is $w_{nt} > r_{nt}$, and the upper threshold $r^u(z_t) \equiv z^{l-1}(z_t)$ is such that if $r_{nt} > r^u(z_t)$, then the optimal wage is $w_{nt} < r_{nt}$. If $r_{it} \in [r^l(z_t), r^u(z_t)]$, then the sum of profits is maximized by $w_{nt} = r_{nt}$.

Proposition 3. *For all $t \geq s$ where $z_s \neq \overline{z}$, $\varepsilon_{t+1} = 0$, and $\mathbb{E}_t z_{t+1} = z_t^{\rho_z}$ (perfect foresight), the optimal wage policy of the firm employing a newly hired worker with reference wage $r_{nt} = \tilde{w}_{it}$ in period $t = s$ and $r_{it+1} = w_t$ for all $t > s$, is such that*

$$\tilde{w}_{nt} = r_{nt} = \tilde{w}_{it} \tag{15}$$

for any z_t.

Proposition 3 establishes that the firm will always pay new hires the same wage as incumbents. While this may not seem surprising when newly hired and incumbents seem *a priori* perfect substitutes in production, it is in fact a consequence of relative wage considerations and new hires' aversion to disadvantageous pay inequality.[18] This result also bears

[17] This implication finds empirical support in the work of Holden and Wulfsberg (2009, 2014) who show that, while the presence of downward wage rigidity prevents some small wage cuts, it also results in larger wage cuts being reduced to a smaller size.

[18] For instance, if new hires did not care about the wage paid to existing workers and were arriving at firms

important implications for the dynamics of vacancy creation and the relative (endogenous to the wage) productivity of new hires and incumbent workers. In fact, while the output produced by incumbent workers will endogenously respond to wage changes—wage increases triggering supra-normal effort, and wage cuts triggering sub-normal effort—new hires will always perceive the wage they are paid as "fair" and will therefore always exert normal effort $\tilde{e}_{nt} = \bar{e}$ in their first employment period. The optimal wage policy for newly hired workers, as a function of the optimal wage paid to incumbent workers, is illustrated in Figure 1b above.

Optimal vacancy posting. For given optimal wages $\tilde{w}_{nt}$ and $\tilde{w}_{it}$ and respective optimal effort choices $\tilde{e}_{nt}$ and $\tilde{e}_{it}$ of newly hired and incumbent workers, the optimal choice of vacancies yields the following job creation condition (see Appendix B)

$$\tilde{\theta}_t \equiv \tilde{\theta}(r_{nt}, z_t) = \left[\frac{\bar{m}}{\kappa}\mathcal{J}_n(r_{nt}, z_t)\right]^{\frac{1}{\alpha}}, \tag{16}$$

where the results established thus far imply that

$$\mathcal{J}_n(r_{nt}, z_t) = z_t\bar{e} - \tilde{w}_{nt} + \hat{\beta}[1-\rho]\mathbb{E}_t\mathcal{J}_i(\tilde{w}_{nt}, z_{t+1}). \tag{17}$$

$$\mathcal{J}_i(w_{t-1}, z_t) = z_t\tilde{e}_{it} - \tilde{w}_{it} + \hat{\beta}[1-\rho]\mathbb{E}_t\mathcal{J}_i(\tilde{w}_{it}, z_{t+1}) \tag{18}$$

Job creation crucially depends on the reference wage of the (potential) new hire. This, in turn, is determined by the optimal wage paid to incumbent workers, meaning that any friction in the adjustment of the wage of incumbent workers will "spill over" to the hiring wage and crucially affect job creation. Moreover, since new hires and incumbents are equally paid $w_{nt} = w_{it}$, and this is perceived as fair by new hires, the effort the firm obtains from a new hire is constant $\tilde{e}_{nt} = \bar{e}$, specifically, it is independent of the wage level, while that of incumbents is endogenous to wage changes $\tilde{e}_{it} \equiv \tilde{e}(w_{it}, w_{t-1})$.

3.2 Range of steady states

This section characterises the model steady state. For simplicity, denote $\varphi \equiv 1 - \hat{\beta}[1-\rho]$.

Definition 3. *A steady-state equilibrium is an equilibrium path in which $z_{t+1} = z_t = \overline{z}$, $r_{it+1} = r_{it} = \overline{r}$, and $u_{t+1} = u_t = \overline{u}$, for all t.*

The next proposition follows from the logic of the previous section.

Proposition 4 (Range of Steady States)**.** *For a given initial r_{is}, if $\lambda > 1$, there exists a range of steady-state wages paid to both new hires and incumbents $\overline{w}_n = \overline{w}_i = \overline{w} \in [\overline{w}^+, \overline{w}(\lambda)^-]$ such*

with a sufficiently low reference wage, such that $r_{nt} < r^l(z_t)$ in some period, they might have been paid less than incumbent workers in response to negative shocks. While if they had a sufficiently high reference wage, such that $r_{nt} > r^u(z_t)$ in some period, they might have been paid more than incumbent workers in response to positive shocks. Hence, it is new hires' concern for relative wages, and their aversion to disadvantageous inequality, that drives the result of Proposition 3.

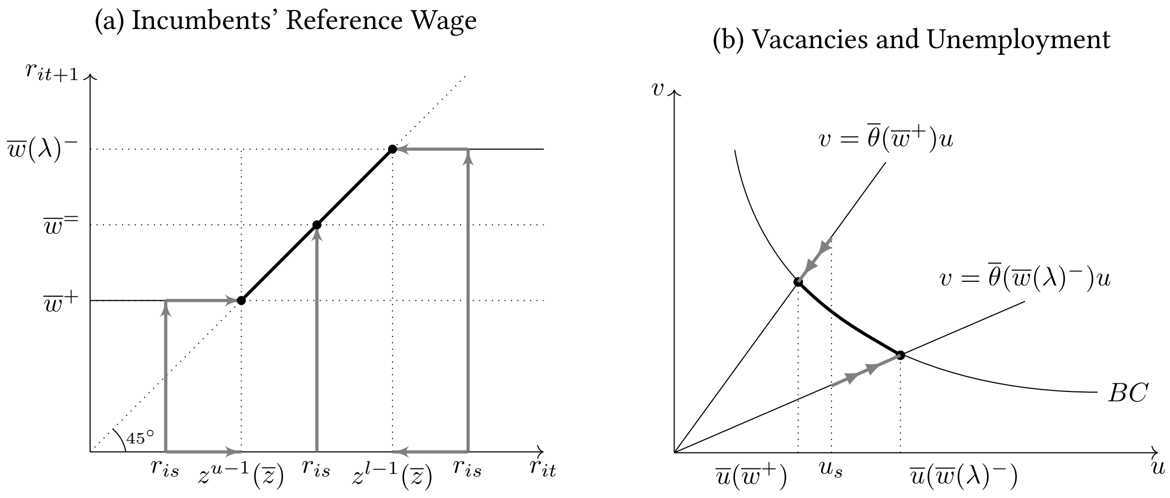
that if $r_{is} \le z^{u-1}(\bar{z})$, then $\overline{w} = \overline{w}^+ \equiv \eta\bar{z}\varphi$; if $r_{is} \ge z^{l-1}(\bar{z})$, then $\overline{w} = \overline{w}(\lambda)^- \equiv \lambda\eta\varphi\bar{z}$; and if $r_{is} \in (z^{u-1}(\bar{z}), z^{l-1}(\bar{z}))$, then $\overline{w} = r_{is}$. Employed workers are paid their reference wage, $\bar{r}_n = \overline{w}_i = \overline{w}$ and $\bar{r}_i = \overline{w}_i = \overline{w}$, and exert normal effort $\bar{e}_n = \bar{e}_i = \bar{e}$. Hence, there exists a range of steady-state labour market tightness $\overline{\theta} \in [\overline{\theta}(\overline{w}(\lambda)^-), \overline{\theta}(\overline{w}^+)]$, unemployment rates $\overline{u} \in [\overline{u}(\overline{w}^+), \overline{u}(\overline{w}(\lambda)^-)]$, and output levels $\overline{y} \in [\overline{y}(\overline{w}^+), \overline{y}(\overline{w}(\lambda)^-)]$, such that $\overline{\theta} = \{[\bar{m}/\kappa][\bar{z}\bar{e} - \overline{w}]/\varphi\}^{1/\alpha}$, $\overline{u} = \rho/[\rho + \bar{m}\overline{\theta}^{1-\alpha}]$, and $\overline{y} = \bar{z}\bar{e}[1 - \overline{u}]$. If $\lambda = 1$, there exists a unique steady-state wage $\overline{w} = \overline{w}^+ = \eta\varphi\bar{z}$, such that $\bar{r}_n = \bar{r}_i = \overline{w}^+$ and $\bar{e}_n = \bar{e}_i = \bar{e}$, and a unique steady-state labour market tightness $\overline{\theta} = \overline{\theta}(\overline{w}^+)$, unemployment rate $\overline{u} = \overline{u}(\overline{w}^+)$, and output $\overline{y} = \overline{y}(\overline{w}^+)$.

The economic intuition behind the range of steady states is as follows. When incumbent workers are characterised by a relatively high reference wage (i.e. $r_{is} \ge z^{u-1}(\bar{z})$), as long as this is not *too high*, i.e. $r_{is} \le z^{l-1}(\bar{z})$, the firm finds it optimal to pay them their reference wage rather than pay them less and endure the cost of substantially lower sub-normal effort. Since λ governs the extent of this cost, a greater λ will increase the range of r_{is} within which the wage $\overline{w} = r_{is}$ is optimal. On the other hand, if $\lambda = 1$, this range is empty and the optimal steady-state wage is unique for all r_{is}. Moreover, since $\overline{w}(\lambda)^- > \overline{w}^+$ when $\lambda > 1$, the steady-state equilibria within the range can be ranked. Henceforth, the steady state in which workers are paid the relatively lower wage $\overline{w}^+$ and unemployment is at its lowest $\overline{u}(\overline{w}^+)$ will be referred to as the *low unemployment* equilibrium. Analogously, the one in which the wage is high $\overline{w} = \overline{w}(\lambda)^-$ and unemployment is at its highest $\overline{u}(\overline{w}(\lambda)^-) > \overline{u}(\overline{w}^+)$ as the *high unemployment* equilibrium. If $\lambda = 1$ there exists a unique steady-state equilibrium which coincides with the low unemployment steady state.

The transitional dynamics of the model are simple, absent shocks to z_t. Figure 2 illustrates this for the two main state variables of the model: the reference wage of incumbent workers (Figure 2a) and the unemployment rate (Figure 2b). The existence of a range of steady-state equilibria established by Proposition 4 is not to be confused with indeterminacy. In fact, there exists a one-to-one mapping between the initial condition r_{is}, the wage $\overline{w}$, and the resulting unemployment $\overline{u}$ in the steady state. The multiplicity of equilibria is due to the non-linearity of this mapping, and the fact that depending on the given value of r_{is} in relation to the range $[z^{u-1}(\bar{z}), z^{l-1}(\bar{z})]$, there exists a unique, yet distinctively different, steady state. As it is formally shown next, what is crucial is that the reference wage of incumbent workers r_{is} itself can be a function of past shocks in any given initial period $t = s$.[19]

[19] An alternative version of this model could assume that the initial reference wage of incumbents is given by their rational expectation of the wage they will get in equilibrium. In this case, the model will generate a range of perfect-foresight equilibria as in Bhaskar (1990), with the possibility of endogenous perfect-foresight fluctuations within the range, or sunspot-induced cycles, if incumbent workers were to condition their expectations on a stochastic signal.

Figure 2: Transitional Dynamics



Note: The label 'BC' in Figure 2b stands for 'Beveridge Curve', and corresponds to the locus $u_{t+1} - u_t = 0$. The arrows describe the implied transitional dynamics. Paths without intervening arrows connote jump dynamics; paths with intervening arrows connote persistent dynamics.

3.3 Business cycles and unemployment hysteresis

This section analyses the *qualitative* properties of the model by studying its response to unanticipated, symmetric economic cycles of different magnitudes. This approach enables us to clearly identify under which conditions the model of this paper generates hysteresis effects.[20] The section begins with a general definition of hysteresis, and of the type of economic cycles we consider. Next, it provides a detailed illustration of the case in which an expansion followed by a recession can result in wage and unemployment hysteresis. The section concludes with the two main theoretical results of the paper, spelling out precisely under which conditions costly wage cuts and relative wage comparisons can generate hysteresis and asymmetric fluctuations.

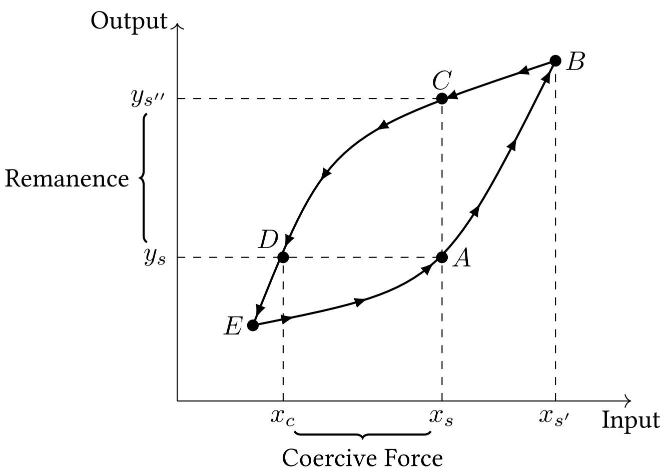
Hysteresis. The first paragraph of the introduction stated that a dynamic system is hysteretic whenever temporary disturbances have permanent effects. Following Cross, Grinfeld, and Lamba (2009), the following statement provides a more formal definition based on the general theory of systems with hysteresis.

Definition 4. *Consider an input-output system with input x_t output y_t and some initial state $\{x_s, y_s\}$. Then suppose that the input x_t changes from x_s to some value $x_{s'}$ and then reverts back to $x_{s''} = x_s$. The system is hysteretic if for each x_s there are values $x_{s'}$ such that, after the excursion, the output y_t does not return to y_s, but to some different value $y_{s''} \neq y_s$. This phenomenon is known as "remanence". To return the output to its original value y_s, the input needs to be changed by an additional amount called "coercive force".*

[20]This qualitative approach is in the spirit of Pissarides (1985), who evaluates the cyclical properties of the search and matching model by analysing the effects of a series of unanticipated permanent shocks.

These concepts are illustrated in Figure 3 below. While general, this definition can imme-

Figure 3: A Hysteresis Loop



Note: This figure is adapted from Cross et al. (2009).

diately be mapped into a dynamic macroeconomic model that starting from an initial steady state is hit by a transitory shock. In the present model, the "input" is the real shock z_t and the "output" can be the unemployment rate u_t. Macroeconomic models based on the "natural rate" hypothesis (virtually every workhorse model of the business cycle) are not hysteretic: as the shock wears off, unemployment will return to the original steady state (the natural rate). Models of unemployment persistence via unit-root (zero-root) processes are also not hysteretic. Hence, to detect hysteresis it is necessary to consider shocks that are transitory.[21] In the remainder of the section this is modelled with two unanticipated shocks of opposite sign but equal magnitude.

Economic Cycles. We proceed by providing a clear definition of a symmetric economic cycle.

Definition 5. *A symmetric boom-bust cycle, between $t = s$ and $t = s''$, is a series of two unanticipated permanent shocks to z_t such that $\overline{z}_{s'} > \overline{z}_s$ and $\overline{z}_{s''} = \overline{z}_s$. Conversely, a symmetric bust-boom cycle, between $t = s$ and $t = s''$, is a series of two unanticipated permanent shocks to z_t such that $\overline{z}_{s'} < \overline{z}_s$ and $\overline{z}_{s''} = \overline{z}_s$*

Further, informed by the optimal wage setting policy of the firm, the analysis differentiates between economic cycles that generate wage rigidity (upward, downward, or both) from

[21]Equivalently, note that in macroeconomic systems that are not hysteretic there exists a one-to-one mapping between the state vector and the control vector, implying that knowledge of the state is sufficient to identify the position of the system—that is, the system is history independent, or memoryless. If a system is hysteretic, the one-to-one mapping between state and control breaks down, implying that its position is now path-dependent. However, hysteresis does not necessarily imply strong history dependence, but can rather be described by the system having a "selective memory" (see, for instance, the discussion on the economics of discontinuous adjustment in Cross (1994)).

cycles that are large enough to generate optimal wage changes, even if these are costly to the firm.

Definition 6. *Let* $\Delta\bar{z}' \equiv [\bar{z}_{s'} - \bar{z}_s]/\bar{z}_s$*. A moderate boom-bust cycle is such that* $\Delta\bar{z}' \in (0, \lambda - 1)$*, while a large boom-bust cycle is such that* $\Delta\bar{z}' > \lambda - 1$*. A moderate bust-boom cycle is such that* $\Delta\bar{z}' \in ([1-\lambda]/\lambda, 0)$*, while a large bust-boom cycle is such that* $\Delta\bar{z}' < [1-\lambda]/\lambda$*.*

In what follows, the analysis begins by illustrating the model response to boom-bust cycles of different magnitudes when the economy is initially at the low unemployment steady state. Then, it provides a complete characterisation of the model dynamics and hysteretic properties in response to all the cycles just defined.

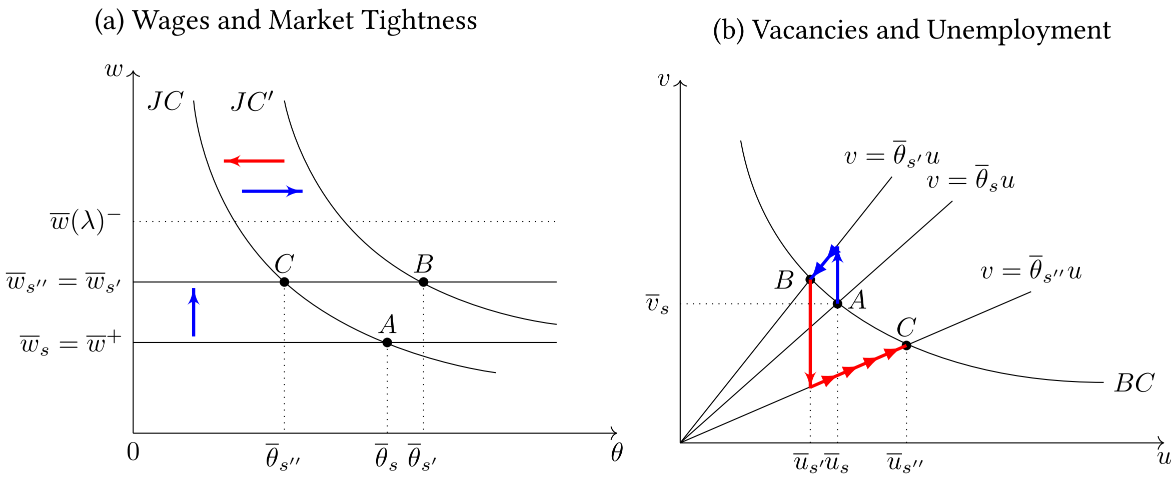
Illustration: Boom-bust cycles and hysteresis. Consider an economy in the low unemployment steady state in period $t = s$, in which workers are paid the lower wage $\bar{w}_s = \bar{w}_s^+ = \eta\varphi\bar{z}_s$, unemployment is at the lower bound of the range of steady states $\bar{u} = \bar{u}(\bar{w}_s^+)$, and per-worker output is $\bar{z}_s\bar{e}$. This situation is illustrated by point A in Figures 4 and 5 below. We start with the analysis of a moderate cycle, and subsequently that of a large cycle.

Moderate cycle and downward wage rigidity. Consider a moderate boom-bust cycle such that the economy experiences a positive shock in some period $t = s' > s$, such that $\bar{z}_{s'} > \bar{z}_s$. Since $\bar{z}_{s'} > z^u(r_{is'})$, the firm optimally increases the wage of both new hires and incumbents, where $\bar{w}_{s'} = \eta\varphi\bar{z}_{s'}$. The positive shock results in greater value of per-worker output, which stimulates hiring: market tightness jumps to a higher steady-state value $\bar{\theta}_{s'} > \bar{\theta}_s$ and unemployment slowly decreases to a lower steady state $\bar{u}_{s'} < \bar{u}_s$. These effects are shown by the blue arrows in Figure 4: in the left panel, the wage curve shifts up, while the job creation curve shifts to the right; in the right panel, the job creation curve rotates counter-clockwise. The economy is at the new steady-state equilibrium, point B.

Next, the economy experiences a recession in some period $t = s'' > s'$, such that the shock reverts back to its initial state $\bar{z}_{s''} = \bar{z}_s$. In a moderate cycle: $\bar{z}_{s''} > \bar{z}_{s'}/\lambda$. It follows that $\bar{z}_{s''} \in [z^l(\bar{r}_{is''}), z^u(\bar{r}_{is''})]$, resulting in the firm optimally freezing the wage of its existing workers: $\bar{w}_{s''} = \bar{w}_{s'}$. Further, relative wage comparisons by new hires imply that such downward wage rigidity "spills over" to the hiring wage. Hence, while per-worker output is back at its initial level $\bar{z}_s\bar{e}$, $\bar{w}_{s''} = \eta\varphi\bar{z}_{s'} > \bar{w}_s = \eta\varphi\bar{z}_s$. The value of hiring an additional worker is lower than at the start of the cycle. Market tightness jumps to a lower steady state $\bar{\theta}_{s''} < \bar{\theta}_s$ and unemployment slowly increases to a higher rate $\bar{u}_{s''} > \bar{u}_s$. These effects are shown by the red arrows in Figure 4: in the left panel, the wage curve does not shift downward (i.e. downward wage rigidity), while the job creation curve shifts back to its initial position; in the right panel, the job creation curve rotates clockwise to a new *lower* position. While the shock is back at its initial steady state, the economy is at a new steady-state equilibrium, point C, where both wages and unemployment are permanently higher.

The key mechanism generating hysteresis is the *partial irreversibility* of wage increases. The relatively lower wages at the start of the cycle warrant fully procyclical wage increases

Figure 4: Moderate Cycle: Downward Wage Rigidity



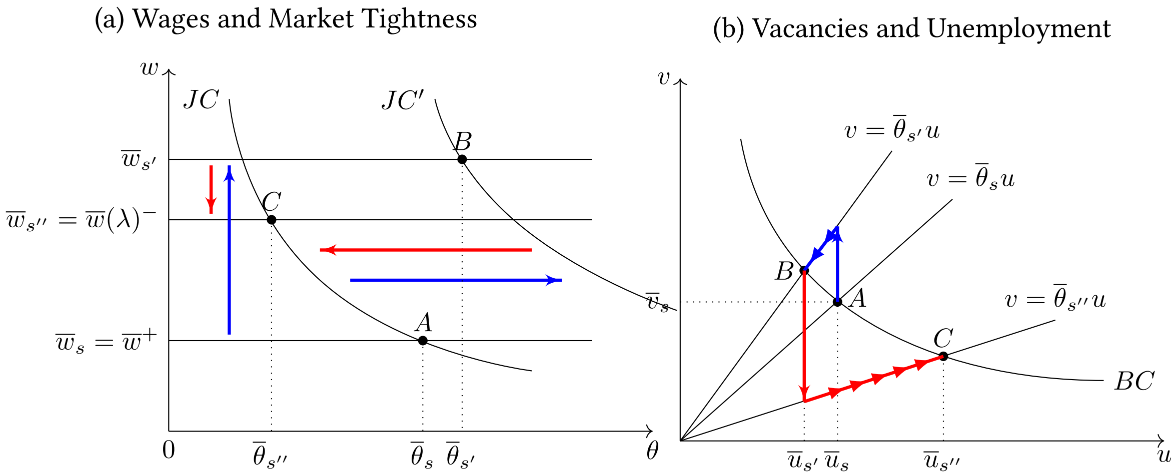
Note: The label 'JC' in Figure 4a stands for 'Job creation Curve', and corresponds to the locus $\tilde{\theta}_{t+1} - \tilde{\theta}_t = 0$. The label 'BC' in Figure 4b stands for 'Beveridge Curve', and corresponds to the locus $u_{t+1} - u_t = 0$. Blue arrows indicate the effect of the expansion phase, while red arrows indicate the effect of the recession phase. In Figure 4b the arrows describe the implied transitional dynamics. Paths without intervening arrows connote jump dynamics; paths with intervening arrows connote persistent dynamics.

during the initial expansion phase. Then, concerns for relative wages by workers—incumbents' aversion to wage cuts and new hires' aversion to disadvantageous pay inequality—imply that during the subsequent recession phase labour is relatively more costly than at the start of the cycle.

Large cycle and muted wage cuts. Next, consider a large boom-bust cycle, such that the economy experiences a large positive shock in $t = s' > s$. As before, the firm optimally increases its workers' wages, market tightness jumps to a higher steady-state value, and unemployment decreases to a lower steady state (point B in Figure 5). Next, the economy experiences a large recession phase such that, even if the shock reverts back to its initial state $\overline{z}_{s''} = \overline{z}_s$, by definition of a large cycle $\overline{z}_{s''} < z^l(\overline{r}_{is''}) = \overline{z}_{s'}/\lambda$. The negative shock is large enough to warrant optimal wage cuts to incumbent workers, where $\overline{w}_{s''} = \lambda\eta\varphi\overline{z}_{s''}$. However, as discussed in the previous section, wage cuts are muted: the firm optimally dampens them to attenuate the sub-normal effort response of its existing workers, resulting in the wage being strictly above the prevailing wage at the start of the cycle $\overline{w}_{s''} > \overline{w}_s = \eta\varphi\overline{z}$. This is shown by the shorter downward shift of the wage curve in the left panel of Figure 5. New hires concerns' for relative wages again imply an optimal hiring wage that is relatively higher than at the start of the cycle. As before, per-worker output is back to its initial level but the job value of a new hire is lower, resulting in lower vacancies $\overline{\theta}_{s''} < \overline{\theta}_s$ and a higher unemployment rate $\overline{u}_{s''} > \overline{u}_s$ in the steady state (point C). This case shows that unemployment hysteresis can result even if wages are downward flexible. The key mechanism is once again the partial irreversibility of wage increases.

Central results: Hysteresis and asymmetric fluctuations. The two scenarios anal-

Figure 5: Large Cycle: Muted Wage Cuts



Note: The label 'JC' in Figure 5a stands for 'Job creation Curve', and corresponds to the locus $\tilde{\theta}_{t+1} - \tilde{\theta}_t = 0$. The label 'BC' in Figure 5b stands for 'Beveridge Curve', and corresponds to the locus $u_{t+1} - u_t = 0$. Blue arrows indicate the effect of the expansion phase, while red arrows indicate the effect of the recession phase. In Figure 5b the arrows describe the implied transitional dynamics. Paths without intervening arrows connote jump dynamics; paths with intervening arrows connote persistent dynamics.

ysed in the previous section have shown that expansion phases featuring wage increases can "plant the seeds" for subsequent recession-induced unemployment hysteresis, due to the partial irreversibility of wage increases implied by workers' concerns about relative wages (when $\lambda > 1$). Next, it will be formally established that the necessary (and also sufficient) condition for such irreversibility is that wages are *fully procyclical* during the expansion phase, i.e. they respond to the shock with an elasticity of (at least) one.[22]

Since the optimal wage implies a jump in dw/dz depending on the size of the shock, define its elasticity as $\varepsilon_{w,z} \equiv \Delta\overline{w}/\Delta\overline{z}$, where $\overline{w} \in [\overline{w}^{+}, \overline{w}(\lambda)^{-}]$ and Δ denotes a percentage change between two steady states. The following proposition establishes the conditions under which the hiring wage $\overline{w}$ can be fully procyclical (i.e. $\epsilon_{w,z} = 1$), sticky (i.e. $\epsilon_{w,z} \in (0,1)$), or fully rigid (i.e. $\epsilon_{w,z} = 0$) depending on the size and direction of shocks.[23]

Proposition 5 (Wage Dynamics)**.** *Consider the case of $\lambda > 1$.*

i) If the economy is at the low unemployment steady state, the elasticity of the wage is $\epsilon_{w,z} = 1$ if $\Delta\overline{z} > 0$ (positive shock); $\epsilon_{w,z} = 0$ if $\Delta\overline{z} \in \left(\frac{1-\lambda}{\lambda}, 0\right)$ (moderate negative shock); and $\epsilon_{w,z} = \frac{\lambda\overline{z}' - \overline{z}}{\overline{z}' - \overline{z}} < 1$ if $\Delta\overline{z} < \frac{1-\lambda}{\lambda}$ (large negative shock).

[22] In a version of the model in which $\gamma \in (0,1)$, or under the assumption of transitory and persistent shocks with perfect foresight, the elasticity of the wage with respect to z can be even greater than one (see the calibration section below). In that model, the condition would be on the elasticity being *equal or greater* than 1.

[23] The qualitative properties established in this section for the boundaries of the range of steady states will be preserved for any of the steady states within the range, depending on the relationship between the initial reference wage r_{is} and the parameter governing the size of the cycles, that is, λ. Since there is a continuum of possible initial conditions within the range $[\overline{w}^{+}, \overline{w}(\lambda)^{-}]$, a complete analysis will be extremely tedious without adding much in terms of economic intuition. Nevertheless, numerical simulations of the model, such as those performed in Section 4, can be used to analyse specific cases of interest.

ii) If the economy is at the high unemployment steady state, the elasticity of the wage is $\epsilon_{w,z} = \frac{\overline{z}'-\lambda\overline{z}}{\lambda[\overline{z}'-z]} < 1$ if $\Delta\overline{z} > \lambda - 1$ (large positive shock); $\epsilon_{w,z} = 0$ if $\Delta\overline{z} \in (0, \lambda - 1)$ (moderate positive shock); and $\epsilon_{w,z} = 1$ if $\Delta\overline{z} < 0$ (negative shock).

If $\lambda = 1$, the elasticity of the wage is $\epsilon_{w,z} = 1$ for any $\Delta\overline{z} \in \mathbb{R}$.

Starting from a situation in which workers are paid relatively low wages $\overline{w} = \overline{w}^{+}$, in response to positive shocks the hiring wage is fully procyclical with an elasticity of one. In response to negative shocks, the firm will either optimally freeze the wage of both incumbents and new hires (downward wage rigidity) or, if the shock is large enough to warrant (muted) wage cuts, the hiring wage is procyclical, but downward sticky, with an elasticity of less than one. Analogous implications apply, but in the opposite direction, when workers are paid relatively higher wages $\overline{w} = \overline{w}(\lambda)^{-}$. In response to negative shocks the firm optimally cuts the wage of its existing workers, implying the hiring wage is procyclical with an elasticity of one. In response to positive shocks, the firm will either keep the wage fixed (upward wage rigidity) or, if the shock is large enough, it will increase both existing and new hires' wages, but less than proportionally to the shock.

Unemployment Hysteresis. Building on the results of Proposition 5, it is then possible to formally characterise the hysteretic properties of the model. In particular, the following proposition will establish that the necessary *and* sufficient condition for the existence of unemployment hysteresis—when workers are concerned about relative wages—is the full procyclicality of wages during a given phase of an economic cycle.

Proposition 6 (Hysteresis)**.** *Consider the case of $\lambda > 1$.*

i) Boom-bust cycles generate hysteresis with permanently higher wages and unemployment if and only if $\epsilon_{w,z} = 1$ during the initial expansion phase in which $\Delta\overline{z} > 0$.

ii) Bust-boom cycles generate hysteresis with permanently lower wages and unemployment if and only if $\epsilon_{w,z} = 1$ during the initial recession phase in which $\Delta\overline{z} < 0$.

If $\lambda = 1$, economic cycles do not generate hysteresis.

Proposition 6 establishes the first of the two main theoretical results of the paper when $\lambda > 1$. Indeed, when $\lambda = 1$ wage changes are fully reversible, the steady state is unique, and there are no hysteresis effects. Case (i) of Proposition 6 was covered in detail by the illustration above: fully procyclical wage increases during the expansion phase imply that, during the subsequent recession, the firm will either optimally freeze the wage (if the cycle is moderate) or implement wage cuts that are muted (if the cycle is large). In both cases, hiring wages will be higher than at the start of the cycle, resulting in permanently higher unemployment. Case (ii), instead, establishes the possibility that a recession followed by an expansion can lead to permanently lower wages and unemployment. In a moderate bust-boom cycle, if the firm fully cuts the wage of its existing workers during the initial recession, it will then find it optimal to keep them constant during the subsequent (moderate) expansion (upward wage rigidity). The dynamics of a large bust-boom cycle is similar, with the only difference that, after an initial wage cut, the (large) expansion phase now warrants wage increases that are more "compressed". In both cases, wages will be lower than at the start of the cycle, implying workers are relatively cheaper to hire. The intuition, which stems from Proposition 5, is that at the start of the cycle wages were already relatively high (e.g. due to hysteresis effects in the past), warranting fully procyclical wage cuts following a negative shock. In summary, in case (ii), it is initial wage cuts that are partially irreversible, enabling the firm to hire workers at a relatively lower cost during the subsequent expansion phase.

These results provide a nuanced perspective on the sources and consequences of costly wage cuts and relative wage comparisons for unemployment hysteresis. In particular, they reveal one important condition, on the elasticity of the hiring wage, which determines the presence or absence of hysteresis effects. Hysteresis only arises when the initial wage change (be it a raise or a cut) is one-to-one proportional to the shock, that is, only as a consequence of wages being procyclical with an elasticity of (at least) one. In fact, this is what makes wage changes *partially irreversible*, planting the seeds for either wage "stickiness" (muted wage cuts, compressed wage increases) or wage "rigidity" (downward, upward acyclicality), when the shock reverses, preventing the labor market from fully adjusting.

Asymmetric fluctuations. These results also bear implications for the volatility of unemployment. Since the influential work of Shimer (2005), the literature on labor market fluctuations has placed considerable emphasis on the extent of the cyclicality of the hiring wage as a source of amplification. The asymmetries in the cyclicality of the hiring wage just discussed suggest that the present model has the prerequisites to generate asymmetries in the amplitude of vacancies and unemployment fluctuations.

To see this, consider the steady-state elasticity of labor market tightness θ with respect to z as a function of the elasticity of the hiring wage in the model of this paper:[24]

$$\epsilon_{\theta,z} = \frac{1}{\alpha} \frac{\overline{z}\overline{e} - \epsilon_{w,z}\overline{w}}{\overline{z}\overline{e} - \overline{w}} \tag{19}$$

where $\overline{w} \in [\overline{w}^{+}, \overline{w}(\lambda)^{-}]$, and $\epsilon_{w,z} \equiv \Delta\overline{w}/\Delta\overline{z}$ as above. Equation (19) reflects one familiar expression for this elasticity in the literature (see Pissarides (2009); Elsby et al. (2015); Ljungqvist and Sargent (2017)), which emphasises the role of the elasticity of the hiring wage as one key determinant of unemployment volatility. In fact, $\epsilon_{\theta,z}$ achieves its minimum when $\epsilon_{w,z}$ approaches one (full wage procyclicality), and its maximum when $\epsilon_{w,z}$ approaches zero (full wage rigidity)—for given $\overline{y} = \overline{z}\overline{e}$ and $\overline{w}$ (determining the "fundamental surplus" in the

[24] The elasticity of market tightness in the steady state is regarded by the existing literature as a good approximation of the volatility of vacancies and unemployment when the labor market is hit by exogenous shocks (Mortensen and Nagypál, 2007; Elsby, Michaels, and Ratner, 2015; Ljungqvist and Sargent, 2017).

present model (Ljungqvist and Sargent, 2017)).

Using the results established in Proposition 5, it is then relatively straightforward to characterise the presence of asymmetric fluctuations.

Corollary 1 (Asymmetric Fluctuations)**.** *Consider the case of $\lambda > 1$.*

i) If the economy is at the low unemployment steady state, then the elasticity of the wage is $\epsilon_{w,z} = 1$ if $\Delta\overline{z} > 0$; and $\epsilon_{w,z} \in [0, 1)$ if $\Delta\overline{z} < 0$; implying the elasticity of market tightness $\epsilon_{\theta,z}$ is larger in response to negative shocks than to positive shocks.

ii) If the economy is at the high unemployment steady state, then the elasticity of the wage is $\epsilon_{w,z} \in [0, 1)$ if $\Delta\overline{z} > 0$; and $\epsilon_{w,z} = 1$ if $\Delta\overline{z} < 0$; implying the elasticity of market tightness $\epsilon_{\theta,z}$ is larger in response to positive shocks than to negative shocks.

If $\lambda = 1$, the elasticity of market tightness is at its lowest $\epsilon_{\theta,z} = 1/\alpha$ and fluctuations are symmetric.

Starting from a situation in which workers are paid relatively low wages $\overline{w} = \overline{w}^{+}$, vacancies respond more strongly to negative shocks than positive shocks, with an elasticity of market tightness reaching its maximum in response to moderate negative shocks (due to downward wage rigidity, $\epsilon_{w,z} = 0$). In response to negative shocks that are large enough to warrant (muted) wage cuts, the hiring wage is procyclical, but downward sticky ($\epsilon_{w,z} \in (0, 1)$); while in response to positive shocks the hiring wage fully adjusts ($\epsilon_{w,z} = 1$), implying a lower elasticity of market tightness. Analogous implications apply, but in the case of positive shocks, when workers are paid relatively higher wages $\overline{w} = \overline{w}(\lambda)^{-}$.

Comparative statics. The results established in this section are driven by the optimal wage setting decision of the firm anticipating the effort response of its employed workers. First, existing workers' aversion to wage cuts, governed by λ, is the primary source of irreversibility and asymmetry in the model. A greater λ not only implies a larger range of rigidity, but also a lower elasticity of the wage whenever wages are sticky. These considerations highlight that a greater λ increases: i) the range of steady-state wages, and therefore of unemployment rates; and ii) the amplitude of the asymmetric response of vacancies to shocks. Second, for a given λ, the effect of these considerations on the firm optimal wage setting decision crucially depends on the sensitivity of workers' net payoff from employment—and therefore of their effort—to relative wage comparisons, governed by η. As in every other efficiency wage model, the stronger is the effort response to the wage, the greater is the incentive for the firm to set a relatively higher wage. In our model, the incentive is to exploit the benefit of supra-normal effort when $w > r$, and to dampen the cost of sub-normal effort when $w < r$. However, recall that, independently of the hiring wage $\overline{w}$, the firm always obtains the same quantity of output $\overline{z}\overline{e}$ from a new hire (since new hires are always paid their reference wage). Hence, the greater η the higher the hiring wage relative to the output from a new match, resulting in: i) a lower profit flow from a new match; and ii) a greater elasticity of market tightness whenever the hiring wage is either sticky or fully rigid.

Corollary 2. *The range of steady-state unemployment rates* $[\overline{u}(\overline{w}^{+}), \overline{u}(\overline{w}(\lambda)^{-})]$ *is strictly increasing in* λ *and* η. *Moreover, whenever the hiring wage is sticky* $\epsilon_{w,z} \in (0, 1)$ *or rigid,* $\epsilon_{w,z} = 0$, *the elasticity of market tightness* $\epsilon_{\theta,z}$ *is strictly increasing in* λ *and* η.

The discussion above highlights two important amplification mechanisms that align with well-established insights on the amplification properties of models with search and matching frictions. Specifically, a greater λ works on reducing the elasticity of the wage, an amplification channel originally identified by Shimer (2005). While a greater η increases the hiring wage relative to the output of a new match. This resonates with more recent appraisals emphasizing the importance of a wage-output ratio w/y close to one (Elsby et al., 2015) or of a small "fundamental surplus fraction" $[y - w]/y$ (Ljungqvist and Sargent, 2017) to generate a large elasticity of market tightness.

The next section provides a quantitative evaluation the model distinctive mechanisms for hysteresis and asymmetric fluctuations. Afterwards, in Section 5, the paper will discuss the evidence in support of both the assumptions, and resulting wage dynamics, of the wage setting model proposed.

4 Quantitative Exploration

This section quantitatively evaluates the extent of hysteresis effects due to costly wage cuts and relative wage comparisons. Henceforth, the baseline model *with* costly wage cuts (i.e. in which $\lambda > 1$) is referred to as the *rigid-wage model* and the baseline model *without* costly wage cuts (i.e. in which $\lambda = 1$) as the *flexible-wage model.*

If not otherwise stated, the results presented below are impulse responses to unanticipated, transitory, but persistent shocks under the assumption of perfect foresight once the shock has realised. This approach is pursued for two main reasons. First, to perform a more standard stochastic simulation, and therefore to solve for the optimal wage policy under uncertainty, one would require that z_t follows a geometric random walk (see Elsby (2009),Fongoni (2024a), and Appendix C.1). This approach, however, will prevent a study of the model response to transitory but persistent shocks, which is the appropriate method to detect, and quantitatively assess, hysteresis effects. Second, this approach substantially simplifies the numerical computation of the firm's job creation condition and enables to simulate the exact solution of the model, preserving all its non-linear properties (for details on the computation approach, see Appendix Section C.2).

4.1 Benchmark calibration

Since the rigid-wage model features a range of steady states, the lower boundary of which is equivalent to the unique steady state of the flexible-wage model, the calibration of certain parameters is pinned down from the latter. This naturally sets the flexible-wage model as a benchmark, where deviations from it can be attributed to the two identified channels for hysteresis: costly wage cuts and relative wage comparisons.

The time period is a month. The parameter space consists of a set of conventional parameters $\{\rho, \alpha, \bar{m}, \kappa, b, \rho_z, \sigma_z\}$ and a set of *behavioural* parameters $\{\beta, \delta, \bar{e}, \gamma, \lambda, \eta\}$ which characterise the household and firm preferences—and in particular, the functional form of the net payoff from employment and resulting optimal effort function.

Conventional block. The mean and steady-state value of z_t are normalised to $\overline{z} = 1$, while its persistence and conditional standard deviation are set to $\rho_z = 0.949$ and $\sigma_z = 0.0065$ (yielding an unconditional standard deviation of 0.02). These values reflect standard calibrations of the exogenous stochastic component of output in the literature (see, e.g., Hagedorn and Manovskii, 2008). The calibration of the remaining conventional parameters seeks consistency with the literature. The job destruction rate is set to $\rho = 0.034$, the elasticity of the matching function is set to $\alpha = 0.5$, and the flow value of unemployment is set to $b = 0.4$.[25] The efficiency of matching $\bar{m}$ and the cost of posting a vacancy κ are calibrated within the model to match a steady-state job finding rate of 0.45 in the flexible-wage model, after normalising the steady-state v/u ratio to 1.[26]

Behavioural block. Normal effort $\bar{e}$ is normalised to 1, implying that per-worker output is also equal to 1 in the steady state (or, in the absence of relative wage considerations). The parameter γ is set to 1, implying $q(x) = \log x$.[27] The parameters η and λ are jointly calibrated as follows. The sensitivity of workers' morale to relative wages η is set to achieve an elasticity of employed workers' effort with respect to the wage (of both new hires and incumbents) of 0.7 in the baseline flexible-wage model. Under the current set of parameter values, it can be shown that $\epsilon_{e,w} = \eta$. This figure reflects the average across the empirical estimates for the elasticity of *discretionary* effort with respect to the wage (see Section III.B. in Fongoni (2024a) for a survey of the available evidence). Informed by the results established in Corollary 2, we then perform sensitivity analysis for values of η yielding an elasticity of effort from 0.1 to 1.1 (the latter corresponding to the highest elasticity of effort admitted by the model under perfect foresight).[28] The relative sensitivity of workers' morale and effort to unfair wages λ,

[25]The numerical value of b is not important for the quantitative results of the paper, since the worker participation constraint is never binding for values of $b \in [0, 1]$ under the calibrations of the model adopted in the following analyses.

[26]More precisely, $\bar{m}$ is set such that $f(\overline{\theta}(\overline{w}^+)) = \bar{m}\overline{\theta}(\overline{w}^+)^{1-\alpha} = 0.45$ when $\overline{\theta}(\overline{w}^+) = 1$. Then κ is calibrated within the model to achieve $\overline{\theta}(\overline{w}^+) = 1$ using the steady-state job creation condition when $\overline{w} = \overline{w}^+$. A similar strategy is used by Kudlyak (2014).

[27]Values of $\gamma \in (0, 1)$ will result in the wage being unrealistically elastic with respect to shocks at impact (i.e with an elasticity greater than 1) in the impulse response analysis. See also the discussion below.

[28]For $\eta > 1.1$, the optimal wage would violate the firm's participation constraint: job values with both new

which crucially determines the cost of wage cuts, is calibrated to achieve an annual frequency of (hiring and existing workers') wage freezes of 28%, corresponding to the average of the estimates in the literature.[29] This is achieved within a stochastic simulation of the wage setting model, yielding $\lambda = 1.014$ (this procedure is explained in Appendix Section C.1). Sensitivity analysis is then performed for values of λ generating a frequency of wage freezes from 0% to 50%.

The final parameters that remain to be set are the household discount factor β which, in general equilibrium, also affect the discount factor used by the firm, via the assumed relationship $\hat{\beta} = \delta\beta$. The wage setting model adopted in this paper, together with the assumption of stationary shocks and perfect foresight, requires a special treatment for the firm discount factor. In this class of models, the firm fully anticipates the reversal and persistence of the shock as well as the effect of a higher wage in the present on the worker's reference wage and effort on the whole subsequent transition path. Under conventional assumptions on time discounting, the resulting "wage compression" incentive will be extremely strong (resulting in extremely low wages) and the implied elasticity of the wage implausibly large. In fact, the elasticity of the wage with respect to shocks in the baseline flexible–wage model under perfect foresight (with $\gamma \in (0, 1]$) is

$$\epsilon_{w,z} = \frac{1}{\gamma}\frac{1-\hat{\beta}[1-\rho]\rho_z}{1-\hat{\beta}[1-\rho]}. \tag{20}$$

Under a conventional calibration of $\beta = 0.996$ (with the remaining parameters set as above), if the firm discounts the future at the same rate as the household (i.e. if $\delta = 1$), the implied elasticity of the wage would be $\epsilon_{w,z} = 2.3$, which is implausibly large.[30] To account for this, while β is calibrated to a standard value (to match a monthly real interest rate of 0.4%) $\hat{\beta}$ is calibrated to match an elasticity of wages with respect to shocks equal to its lower bound of one in the flexible-wage model. This strategy yields $\beta = 0.996$ and $\delta = 0.156$. This implies the firm discounts the future substantially more than would be suggested by observed interest rates. Nevertheless, this calibration is consistent with the estimate of an analogous discount factor in Ehrlich and Montes (2024, p. 189), who consider a reduced-form model wage setting with costly wage cuts and active wage compression, in the spirit of the class of models like the one considered in this paper.[31]

hires and incumbents become negative, implying the firm would shut down entirely. As previously stated, the formal analysis of this case is beyond the scope of the present paper and left to further research.

[29]See for instance the evidence surveyed in Elsby and Solon (2019) and Grigsby et al. (2021).

[30]From equation (20) it is also possible to see why $\gamma < 1$ will further increase this elasticity. That this feature is specific of the assumption of perfect foresight when z_t follows a stationary AR(1) process, can be seen by noticing that, if z_t were to follow a geometric random walk, $\rho_z = 1$, then $\epsilon_{w,z} = 1/\gamma$, which is independent of the discount factor $\hat{\beta}$. Even in this case a plausible calibration would suggest γ to be close to one. Further note that such wage compression incentive is present even in the absence of costly wage cuts. See Dickson and Fongoni (2019) for a thorough analysis of this point in a two-period model.

[31]Ehrlich and Montes (2024) estimate the parameters of their model via an indirect inference approach to match a set of moments in the data on wage and hiring elasticities at the establishment level, including measures of wage rigidity and wage elasticities. While their model differs from the one used in this paper, it nevertheless

The benchmark calibration is summarized in Table 1. The remainder of the section provides

Table 1: Monthly calibration

Parameter	Value	Explanation	Source / Reason
ρ_z	0.949	Persistence of z	Conventional
σ_z	0.0065	Standard deviation of ε	Conventional
$\overline{z}$	1.0	Steady state level for z	Normalisaton
$\bar{e}$	1.0	Steady state level for e	Normalisaton
$\overline{\theta}^{+}$	1.0	Steady state level for θ	Normalisaton
ρ	0.034	Separation rate	Shimer (2005)
α	0.5	Elasticity of matching	Literature
$\bar{m}$	0.45	Efficiency of matching	Job-finding rate = 0.45
κ	0.214	Cost of posting a vacancy	Job-finding rate = 0.45
b	0.4	Flow value of unemployment	Shimer (2005)
η	0.7	Sensitivity of effort to relative wages	Elasticity of effort = 0.7
λ	1.014	Relative sensitivity of effort to unfair wages	Frequency of wage freezes = 28%
β	0.996	Household discount factor	interest rate = 0.4%
δ	0.156	Firm discounting wedge	Elasticity of wages = 1

a quantitative evaluation of the model hysteresis effects, and subsequently assess the extent of the model asymmetric response to positive or negative shocks. Throughout, all simulations will include a sensitivity analysis along the two key parameters λ and η, governing the cost of wage cuts and sensitivity to relative wage comparisons.

4.2 Hysteresis effects

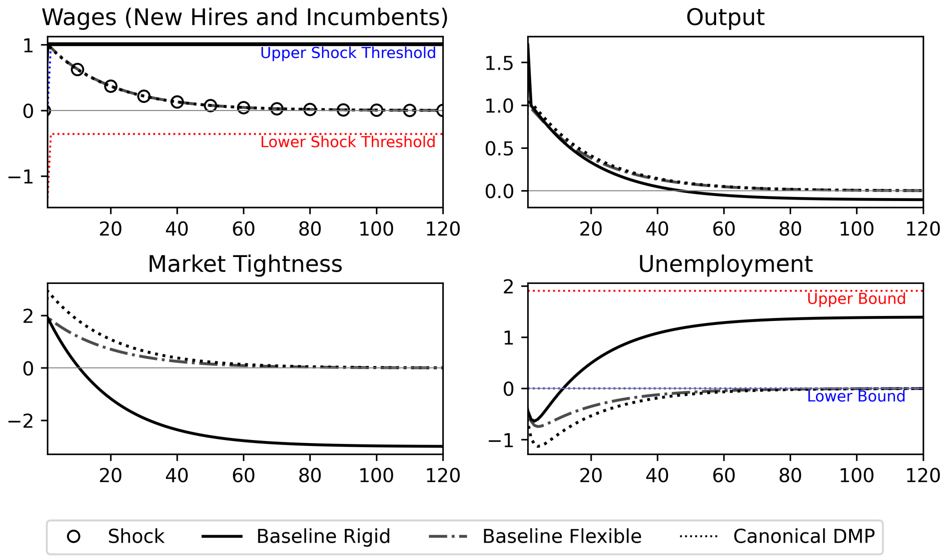
The analysis of Section 3 has formally shown that hysteresis effects only arise subsequent wage changes that are at least one-to-one proportional to shocks. The results of Proposition 6 established that this happens as a consequence of positive transitory shocks when the economy is at the low unemployment steady state, and negative transitory shocks when the economy is at the high unemployment steady state. The impulse responses of a selection of key macroeconomic variables (wages, output, market tightness, and unemployment) to transitory shocks of 1% are presented in Figure 6 and Figure 7 for both cases respectively.[32] The figures also display analogous results in the flexible-wage model and in a canonical search and matching model ("canonical DMP") where workers' bargaining power is calibrated to achieve a unit elasticity of the wage with respect to shocks (the correct baseline of comparison).

Benchmark results. In response to a positive transitory shock (Figure 6), wages increase at impact across all models, with a common elasticity of one (by design). However, as the

features an analogous wage compression incentive due to a reduced-form cost of wage cuts, which the firm rationally anticipates in expectations. Within the context of their model and estimation procedure, they estimate an establishment (annual) discount factor of 0.53 (corresponding to an annual real interest rate of 85.19%).

[32]Impulse responses for all variables of the model are presented in Figures C.1-C.4 in Appendix Section C.2.

Figure 6: Hysteresis effects: Low steady state

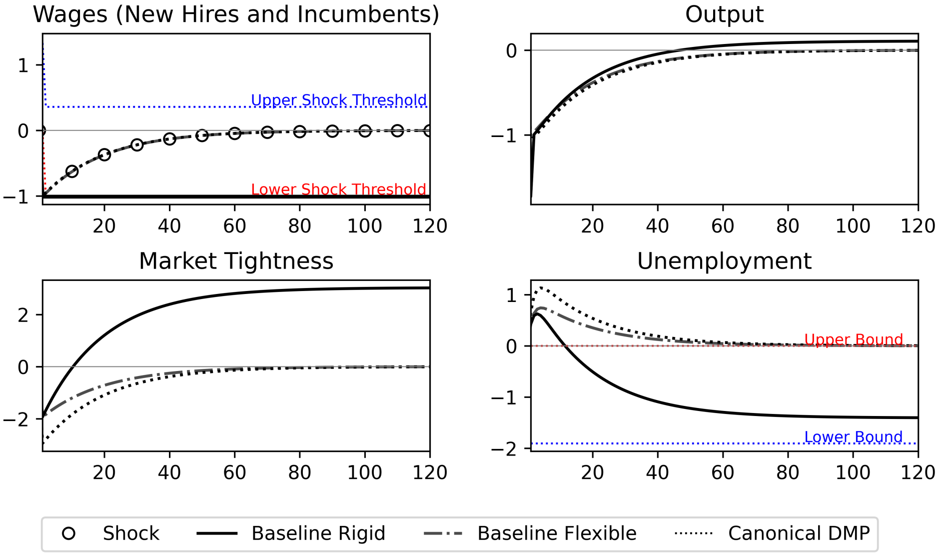


Note: Response to a 1% positive, transitory shock starting from the low unemployment steady state. The x-axis is months; the y-axis is percent deviations from the initial steady state. In the upper-left panel the blue and red dotted lines respectively indicate the thresholds z_t^u and z_t^l, which characterise the optimal wage setting policy of incumbent workers. In the lower-right panel the blue and red dotted lines respectively indicate the lower and upper bound of the range of steady states for the unemployment rate.

shock reverses, the rigid-wage model features *downward* wage rigidity: the initial wage increase is irreversible, and both the hiring and incumbent workers' wages remain permanently higher. Market tightness jumps upwards at impact across all models (with the canonical DMP model displaying a slightly higher elasticity), and as the shock reverses, it decreases faster in the rigid-wage model, it undershoots the initial steady state after about 10 months, and then it remains more than 2% lower than at the start of the shock. This dynamics imply that while unemployment initially decreases at impact across all models, in the rigid-wage model it overshoots the initial steady state and subsequently converges to a permanently higher rate, generating a *remanence* of about 1.4% (in percentage deviation from the initial steady state). Hysteresis effects are indeed transmitted to output as well, but are less pronounced, generating a remanence of about -0.11%. Results for the response to a negative transitory shock (Figure 7) are the mirror image of those just discussed. The obvious qualitative difference being that it is the initial wage cut that is partially irreversible, resulting in *upward* wage rigidity when the shock reverses, and a permanently *lower* unemployment rate and *higher* output.

These hysteresis effects are small. For example, a remanence of 1.4% means that if unemployment was initially at 7%, after a boom-bust cycle with hysteresis, it will increase to 7.1%, which is a negligible change. The theoretical analysis of Section 3 has emphasised the importance of the relative cost of wage cuts (governed by $\lambda > 1$) and the sensitivity of workers' morale and effort to relative wage comparisons (governed by η) for the size of hysteresis effects. The remainder of this section provides a comprehensive quantitative assessment of

Figure 7: Hysteresis effects: High steady state



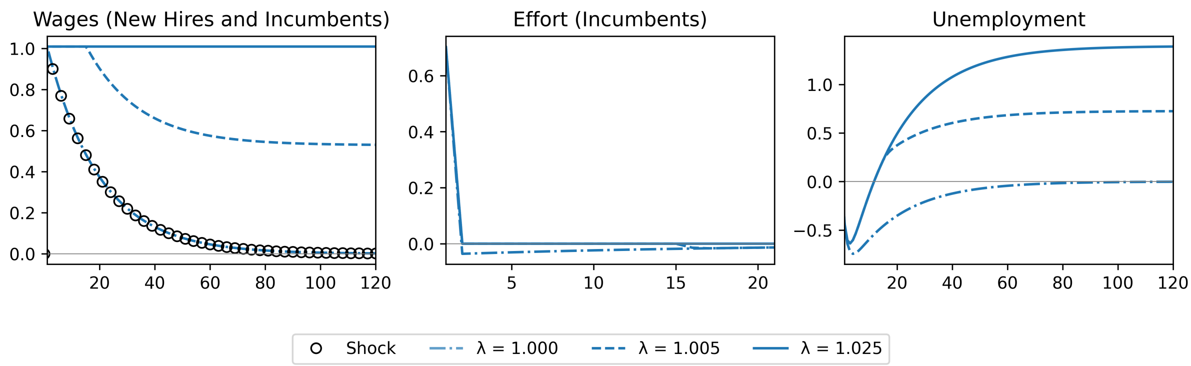
Note: Response to a 1% negative, transitory shock starting from the high unemployment steady state. The x-axis is months; the y-axis is percent deviations from the initial steady state. In the upper-left panel the blue and red dotted lines respectively indicate the thresholds z_t^u and z_t^l, which characterise the optimal wage setting policy of incumbent workers. In the lower-right panel the blue and red dotted lines respectively indicate the lower and upper bound of the range of steady states for the unemployment rate.

their implications for unemployment hysteresis. Since the hysteresis effects stemming from a negative shock are the mirror image of those stemming from a positive one, in the interest of brevity the following analysis is focused on the latter case only.[33]

Cost of wage cuts. The following analysis investigates how different costs of wage cuts affect the wage and unemployment dynamics in the rigid-wage model. For this exercise λ is calibrated to achieve an annual frequency of (incumbents and new hires) wage freezes of 0% ($\lambda = 1$, like in the flexible-wage model), 10% ($\lambda = 1.005$) and 40% ($\lambda = 1.025$)—approximately the lowest and highest estimates in the empirical literature. The impulse responses for a selection of variables (wages, effort of incumbents, and unemployment) are presented in Figure 8. First, as established in Section 3, a greater cost of wage cuts increases the extent of hysteresis effects, with the remanence ranging from 0%, to 0.7% and 1.4%. Interestingly, the latter is equivalent to the one resulting from the benchmark calibration when $\lambda = 1.014$ (annual frequency of wage freezes of 28%). The reason is that, in both cases, a 1% positive transitory shock corresponds to a *moderate cycle*. As such, after the initial wage increase (which is independent of λ), wages are downward rigid and therefore equivalent across calibrations, resulting in equivalent dynamics of vacancies and unemployment. On the other hand, when $\lambda = 1.005$, a positive transitory shock of 1% now corresponds to a *large cycle*, inducing the firm to begin cutting wages after about 15 months. This is accompanied by a drop in incumbents' effort at the same time, which is visible in the middle panel after period 15. Nevertheless, muted

[33]The vacancy posting cost κ is re-calibrated in each simulation to maintain a steady-state job finding rate of 0.45 across models throughout the whole sensitivity analysis.

Figure 8: Hysteresis: Cost of wage cuts

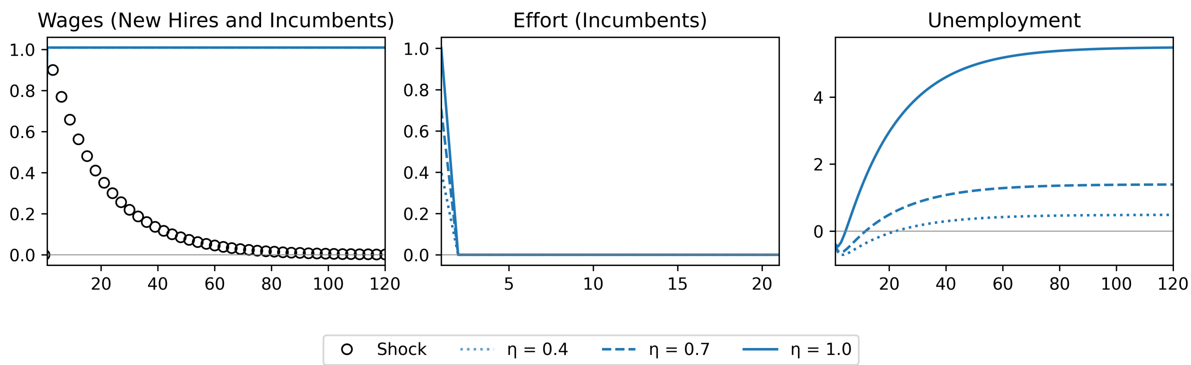


Note: Response to a 1% positive, transitory shock starting from the 'low' unemployment steady state. The x-axis is months; the y-axis is percent deviations from the initial steady state.

wage cuts imply that incumbents' and hiring wages do not return to the initial steady state and unemployment remains permanently higher. These results are not surprising in light of the analysis of Section 3: it is indeed the size of shocks, relative to the cost of wage cuts, that matters for the size of hysteresis effects. This point is taken up again at the end of the section.

Relative wage comparisons. What follows investigates the implications of different sensitivities of workers' morale and effort to relative wages. For this exercise η is calibrated to achieve three values for the elasticity of effort with respect to the wage: 0.4, 0.7 (benchmark), and 1. The impulse responses are displayed in Figure 9. The wage response is equiva-

Figure 9: Hysteresis: Relative wage considerations



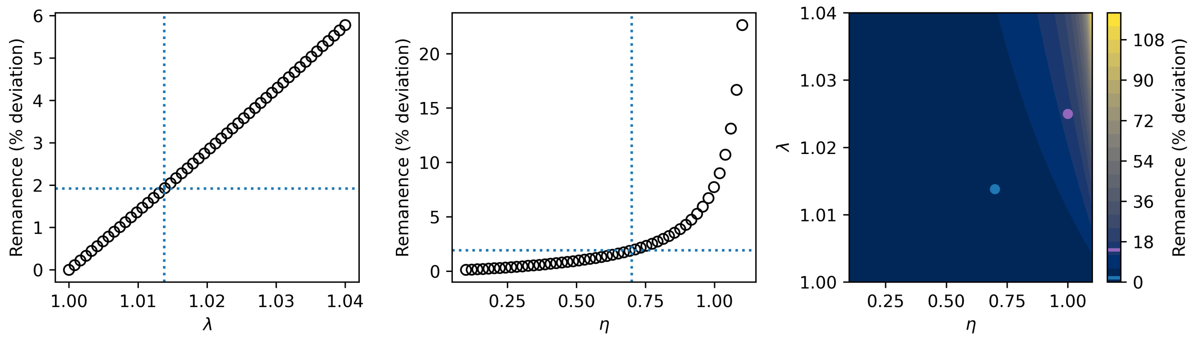
Note: Response to a 1% positive, transitory shock starting from the low unemployment. The x-axis is months; the y-axis is percent deviations from the initial steady state.

lent across parameterisations, displaying a unit elasticity at impact and full downward wage rigidity thereafter. However, the greater responsiveness of incumbents' effort to wage changes across models (evident in the middle panel) induces relatively higher wages (in absolute value), and generates greater hysteresis effects. For values of the elasticity of effort that range between 0.4 to 1, the model generates an unemployment remanence (in percentage deviations) ranging from 0.5% to 5.6%. The latter means that if unemployment was initially at 7%, after the boom-bust cycle it will increase to 7.4%. This magnitude is quantitatively significant, especially because it is the result of a positive, and fully transitory shock of 1%.

Maximum remanence. The previous analyses have focused on transitory shocks of 1%. In most of the simulations, this corresponded to a moderate boom-bust cycle. In the following analysis we provide a more comprehensive assessment of hysteresis effects also accounting for large cycles. A large boom-bust cycle is such that, for given values of η and λ, the reversal of the shock is large enough to warrant wage cuts that are dampened, which in turn imply unemployment moving from the lower bound $\overline{u}(\overline{w}^{+})$ to the upper bound of the range of steady states $\overline{u}(\overline{w}(\lambda)^{-})$. Hence, the greatest extent of hysteresis effects can be calculated by directly computing the model *maximum remanence* $[\overline{u}(\overline{w}(\lambda)^{-}) - \overline{u}(\overline{w}^{+})]/\overline{u}(\overline{w}^{+})$, for given combinations of λ and η. [34]

Results are reported in Figure 10, where maximum remanence is computed for a range of λ from 1 to 1.04 (the latter corresponding to an annual frequency of wage freezes as high as 50%) and for a range of η from 0.1 to 1.1. The left panel shows the relationship between

Figure 10: Hysteresis: Maximum remanence



Note: Blue dotted lines correspond to the values of η, λ and maximum remanence under the benchmark calibration. In the right panel, the blue dot is the point $(\eta, \lambda) = (0.7, 1.014)$ corresponding to the benchmark calibration, while the purple dot is the point $(\eta, \lambda) = (1, 1.025)$ corresponding to the highest values of η and λ in the previous sensitivity analyses.

the cost of wage cuts and the maximum remanence keeping the sensitivity to relative wages fixed at the benchmark value of $\eta = 0.7$. The plot uncovers an increasing linear relationship, showing that for costs of wage cuts resulting in an annual frequency of wage freezes as high as 50%, the benchmark calibration can generate a maximum remanence of 5.8%. In the middle panel, the cost of wage cuts is fixed at the benchmark value of $\lambda = 1.014$, while workers' sensitivity to relative wages varies. The plot uncovers an increasing and convex relationship (mainly due to the properties of the matching function) such that for high values of η, the model generates a maximum remanence of 22.6%. The right panel provides a heat-map where

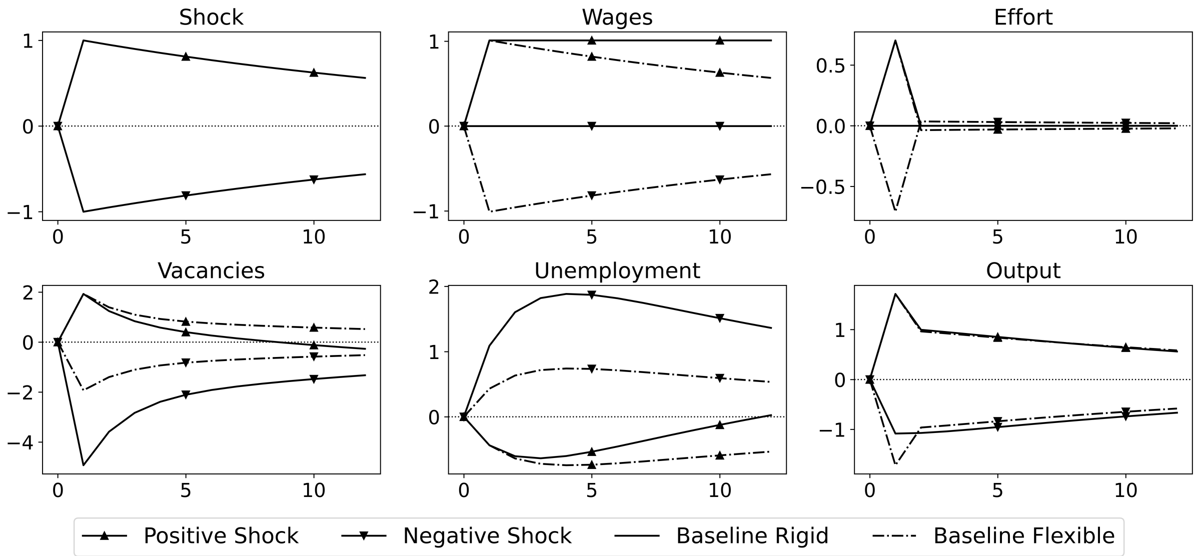
[34]Indeed this figure will be equivalent to the absolute value of the maximum unemployment remanence resulting from a large bust-boom cycle when the initial steady stage is the high unemployment.

both key parameters are allowed to vary. The blue dot corresponds to a maximum remanence of 1.9% (when η and λ are as in the benchmark calibration), while the purple dot corresponds to a maximum remanence of 15.2% (when $\eta = 1$ and $\lambda = 1.025$, the highest values used in the previous sensitivity analyses). To give a concrete example, a remanence of 15.2% would imply that if unemployment was initially at 7%, after a large boom-bust cycle unemployment would permanently increase to 8.1%, i.e. more than 1 percentage point. The heat-map also shows that for economies featuring high costs of wage cuts and high workers' sensitivities to relative wages, large boom-bust cycles might result in large hysteresis effects on unemployment. For instance, a remanence as high as 115% means a permanent increase in the rate of unemployment from 7% to 15% after a positive, fully transitory shock. This figure is perhaps unrealistically high, but it provides a perspective of the quantitative potential of the mechanisms behind the hysteresis effects considered in this paper.

4.3 Asymmetric fluctuations

This section studies impulse responses of a selection of variables to both positive and negative transitory shocks of different magnitudes, starting from the low unemployment steady state under the benchmark calibration.[35] The responses to moderate shocks (i.e. of 1%) are displayed in Figure 11, while those to large shocks (i.e. of 2%, corresponding to a one standard deviation shock) are displayed in Figure 12, for both the rigid- and flexible-wage models.

Figure 11: Asymmetric fluctuations: Moderate shocks

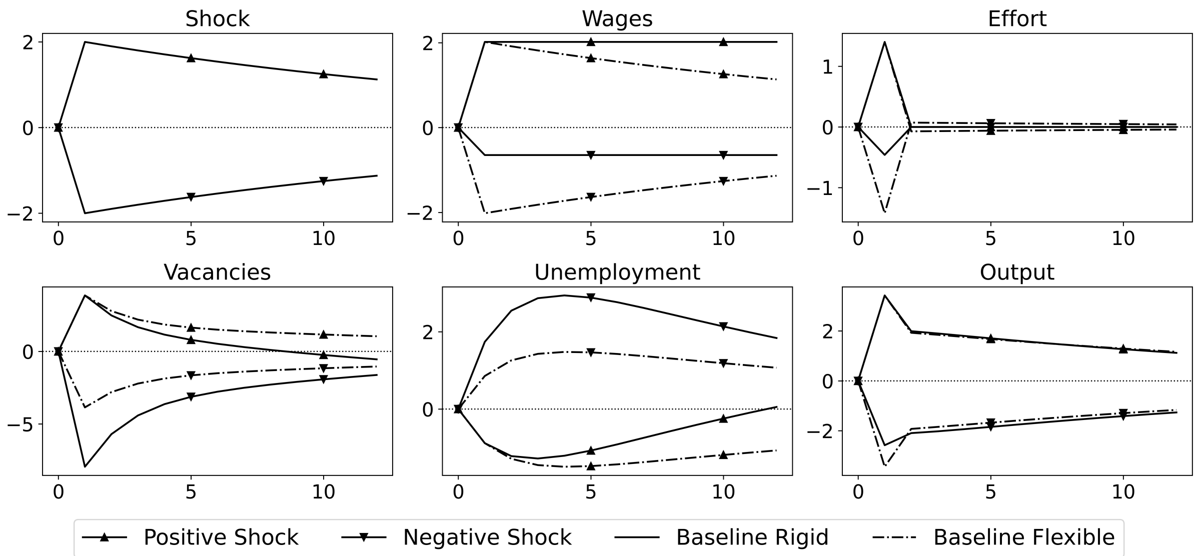


Note: Responses to positive and negative, transitory shocks of 1% starting from the low unemployment; The x-axis is months; the y-axis is percent deviations from the initial steady state.

[35]Once again, the dynamics of the model starting from the high unemployment steady state will be the mirror image of the one analysed here, but with upward wage rigidity or stickiness. See Proposition 1 for details.

Benchmark results. These results highlight strong asymmetries implied by the rigid-wage model. In response to positive shocks (moderate or large), both wages and effort jump upwards with elasticities of 1 and 0.7. At impact, these responses, as well as those of all other variables, are equivalent between the rigid-wage and flexible-wage models. On the other hand, in response to negative shocks, the rigid-wage model exhibits much larger responses. A *moderate* negative shock (Figure 11) results in downward wage rigidity. This substantially depresses hiring, causing a much larger increase in the unemployment rate. Notably, the response of vacancies is more than twice as large. However, due to incumbent workers' effort remaining constant, the response of output is rather dampened. A *large* negative shock (Figure 12) generates quantitatively similar effects on vacancies and unemployment, with the only difference that wage cuts are muted at impact. The response of output is also dampened, but less than that displayed in Figure 11, due to incumbent workers' slightly decreasing effort at impact due to the wage cut. This analysis suggests that the inherent asymmetries generated by

Figure 12: Asymmetric fluctuations: Large shocks



Note: Responses to positive and negative, transitory shocks of 2% starting from the low unemployment; The x-axis is months; the y-axis is percent deviations from the initial steady state.

the path-dependent rigid-wage model are quantitatively important, especially for vacancies and unemployment.

Sensitivity analysis. To conclude, the following exercise compares the model response to both positive and negative shocks of one unconditional standard deviation (i.e. 2%), under different calibrations of the cost of wage cuts and sensitivity to relative wage comparisons. Table 2 summarises the results. The first column reports model elasticities *at impact* for three calibrations of λ, keeping η fixed at its benchmark value of 0.7 respectively. Analogously, the second column reports elasticities for three calibrations of η keeping λ fixed at its benchmark value of 1.014 (thus, the middle row shows results for the benchmark calibration, reporting the exact figures behind Figure 12 above). Looking at the first column, as wage cuts become

Table 2: Asymmetries at impact

	$+\sigma$	$-\sigma$	ratio	$+\sigma$	$-\sigma$	ratio
	$\lambda = 1.005$			$\eta = 0.4$		
$\epsilon_{w,z}$	1.01	0.75	0.75	1.01	0.34	0.34
$\epsilon_{e,w}$	0.67	0.68	1.02	0.38	0.39	1.02
$\epsilon_{v,z}$	1.93	2.69	1.39	1.96	2.66	1.35
$\epsilon_{u,z}$	-0.44	-0.59	1.35	-0.45	-0.59	1.31
$\epsilon_{y,z}$	1.71	1.56	0.91	1.42	1.18	0.83
	$\lambda = 1.014$			$\eta = 0.7$		
$\epsilon_{w,z}$	1.01	0.34	0.34	1.01	0.34	0.34
$\epsilon_{e,w}$	0.67	0.69	1.03	0.67	0.69	1.03
$\epsilon_{v,z}$	1.93	3.91	2.03	1.93	3.91	2.03
$\epsilon_{u,z}$	-0.44	-0.86	1.94	-0.44	-0.86	1.94
$\epsilon_{y,z}$	1.71	1.30	0.76	1.71	1.30	0.76
	$\lambda = 1.025$			$\eta = 1$		
$\epsilon_{w,z}$	1.01	0.00	n.a.	1.01	0.34	0.34
$\epsilon_{e,w}$	0.67	0.00	n.a.	0.96	0.99	1.03
$\epsilon_{v,z}$	1.93	4.97	2.58	1.78	9.66	5.43
$\epsilon_{u,z}$	-0.44	-1.08	2.45	-0.41	-2.03	4.99
$\epsilon_{y,z}$	1.71	1.08	0.63	1.99	1.50	0.75

Note: Responses of the model to a positive and negative transitory shock of one unconditional standard deviation σ (i.e. 2%), starting from the low unemployment steady state, under different calibrations of η and λ. All other parameters are as in the benchmark calibration if not stated otherwise. The elasticities of effort and wages are based on the average effort and average wage (between new hires and incumbents) generated by model.

more costly to implement, the firm tends to dampen them, generating wage stickiness in response to negative shocks. For $\lambda = 1.025$ (corresponding to a 40% annual frequency of wage freezes) wages are instead downward rigid and the asymmetries are even more pronounced with vacancies declining by more than twice in response to negative shocks, and effort of both new hires and incumbents remaining constant, implying a *dampened* output response about 60%. Throughout the second column, the cost of wage cuts is such that wage decreases are muted across all calibrations of η. Nevertheless, the results show that for an elasticity of workers' effort approaching one, the response of vacancies and unemployment to negative shocks can be about 5 times larger (see also the discussion around Corollary 2), while the output response remains dampened by just 75%.

5 Empirical plausibility of the wage mechanism

The theory of unemployment hysteresis advanced in this paper is based on a wage setting model implying that wage changes are partially irreversible: any rigidity or stickiness in the hiring wage is path-dependent. The next section discusses the empirical plausibility of the *assumptions* leading to such dynamics, while Section 5.2 discusses the empirical plausibility of the resulting *implications* for wage dynamics.

5.1 Discussion of the main assumptions

The wage setting model of this paper is based on the following assumptions: i) the employment contract is incomplete, and the wage is set unilaterally by the firm; ii) the firm anticipates that workers care about fairness and that "unfair" wages can be particularly damaging for morale and effort; iii) existing workers evaluate wages with respect to their most recent wage, while new hires find it fair to be paid the same wage of their incumbent co-workers.

Incompleteness of the employment contract. A basic premise of the present model is that the firm cannot precisely specify, or enforce, the productivity of its employees in every possible contingency. The output of the firm depends on its workers' effort which, importantly, is discretionary and not contractible. Economists refer to this as the *incompleteness* of the employment contract, and consider it to be an inherent feature of labour markets. For instance, contracts based on time-rate payments (e.g. hourly wage contracts) may specify the amount of hours of work, but rarely the *pace* at which each hour must be worked. Similarly, piece-rate contracts seldom specify the *rate* at which output must be produced. This characteristic of labour markets—unlike, for example, goods markets—has long been recognised. Simon (1951) and later Williamson (1985) were among the first to acknowledge that the employment contract is an "incomplete agreement", while Okun (1981) referred to the labour market as being governed by an "invisible handshake", rather than by an invisible hand.[36] Indeed, contract incompleteness is the main premise that provides a rationale for efficiency wages (Malcomson, 1981).[37] Moreover, it has been shown to prevent firms from offering low (market-clearing) wages in experimental labour markets (Fehr and Falk, 1999) and from cutting wages to job stayers (Fongoni, Schaefer, and Singleton, 2025).[38]

Fairness, morale, and reciprocity. If workers have some discretion on the pace, quality, and amount of work, they could use this margin to retaliate against (or positively reciprocate) anything that they perceive as unfair (or as kind) treatment by their employer. Hence, if wages are unilaterally set by firms (Card, 2022; Manning, 2021), it is in their profit-maximising interest to internalise the effects that wage changes may have on employees' perceptions of fairness, morale, and therefore effort. These considerations reflect a well-established consensus on the wage setting behaviour of firms, grounded in surveys and interviews with compensation

[36] Quoting Herbert Simon: "in an employment contract certain aspects of the worker's behavior are stipulated in the contract terms, certain other aspects are placed within the authority of the employer, and still other aspects are left to the worker's choice." (Simon, 1951, p. 305).

[37] One might view the wage setting model of this paper as offering a more nuanced, dynamic incarnation of the "morale model" of Akerlof (1982) and Akerlof and Yellen (1990).

[38] In the introduction to their paper, Fehr and Falk (1999) provide a thorough discussion of contract incompleteness in labour markets, quoting from Oliver Williamson, Paul Milgrom, Roberts John, and Truman Bewley.

managers (pioneered by Bewley (1999), Blinder and Choi (1990), and Campbell and Kamlani (1997)) and reinforced by more recent advances in behavioural and experimental economics on fairness and reciprocity in labour relations (Bewley, 2007; Fehr et al., 2009). Among the key findings of this extensive empirical literature are that employees perceive both wage cuts and disadvantageous pay inequality as particularly unfair; that negative reciprocity in response to unfair treatment is typically stronger than positive reciprocity; and that employers, aware of these patterns, hesitate to cut incumbent workers' wages (see, e.g., Bertheau, Kudlyak, Larsen, and Bennedsen, 2025) and to pay new hires differently from incumbents, even when the labour market is slack (see, e.g., Galuscak, Keeney, Nicolitsas, Smets, Strzelecki, and Vodopivec, 2012).

Relative wage comparisons. That workers care about relative wages by evaluating offers or contracts in relation to a reference "fair" wage has long been understood by economists.[39] What is perhaps less understood is the determination of such a reference point. The model of this paper assumes that incumbent workers' reference wage is their most recent wage contract. This assumption is consistent with a large body of empirical evidence that past contracts serve as reference points, both in the context of labour markets (see, e.g., Kahneman, Knetsch, and Thaler, 1986; Sliwka and Werner, 2017) and in the context of incomplete contracts (see, e.g., Bartling and Schmidt, 2015; Herz and Taubinsky, 2017). The key implication is that incumbent workers are particularly averse to wage cuts, which is the source of the partial irreversibility of wage increases. Newly hired workers, coming from the unemployment pool, are rather assumed to use the wage paid to the existing workers as a reference. This assumption is also consistent with evidence documenting that workers' perceptions of fairness and effort are influenced by their co-workers' wage (see, e.g., Gächter, Nosenzo, and Sefton, 2012; Cohn, Fehr, Herrmann, and Schneider, 2014; Flynn, 2022) as well as with "internal pay equity" concerns expressed by compensation managers (see, e.g., Blinder and Choi (1990) and the survey evidence cited above). As such, the established irreversibility in the wage of existing workers might spill over to that of the hiring wage.[40]

5.2 Wage dynamics

The degree of rigidity (acyclicality) and stickiness (less-than-proportional cyclicality) of the hiring wage carry important implications for the dynamics of wages and unemployment in the present model.[41] The following exercise aims to evaluate the empirical plausibility of the dynamic and cyclical properties of the optimal wage policy considered here, by also comparing

[39]For instance, a notion of the "fair wage" is already present in Alfred Marshall's Principles (Marshall, 1890).

[40]The strands of literature mentioned in this section are vast. The reader is referred to the surveys in Bewley (2007), Fehr et al. (2009), Esteves-Sorenson (2018) and Fongoni (2024b) for more thorough reviews of the literature on morale, fairness, reciprocity, and wage setting in labour markets.

[41]The terms "rigidity" and "stickiness" have been used somewhat interchangeably by the existing literature. For instance, in the well known paper by Hall (2005) wages are fully rigid (both upward and downward) while the author refers to this as "wage stickiness". On the other hand, in the models analysed by Shimer (2010), Michaillat (2012), and Abbritti and Fahr (2013), wages are in fact sticky, responding less-than-proportinally to shocks, while the authors refer to this as "wage rigidity".

it with that of other wage setting mechanisms commonly assumed, or derived, in the literature of labour market fluctuations.

Consider four wage setting models: a frictionless wage model, a sticky wage model, a downwardly rigid-wage model, and the path-dependent one derived in this paper. These are compared in the context of a stochastic simulation where the shock to the firm's revenue productivity follows a geometric random walk, denoted by $\hat{z}_t$ (see Appendix Section C.1 for details). Under the assumption of zero inflation, these policies capture both real and nominal wages, unless otherwise stated. Define the *frictionless wage* as a fully procyclical wage with an elasticity of one, $\hat{w}_t^* = \hat{z}_t$, and the *sticky wage* as a procyclical wage with an elasticity less than one, $\hat{w}_t^{Sticky} = \hat{z}_t^{\phi}$, with $\phi \in (0,1)$. This wage equation has been adopted, for instance, by Blanchard and Galí (2010) and Michaillat (2012) and generally captures—in reduced form—other forms of wage stickiness (such as the one of a canonical search and matching model with acyclical unemployment income).[42] If $\phi = 0$ the wage is fully rigid as in Hall (2005). Next, define the *downward rigid wage (DWR)* as a wage rule that constraints wage cuts: $\hat{w}_t^{DWR} = \max\{\hat{w}_t^*, \varphi \hat{w}_{t-1}^{DWR}\}$. With zero inflation, there are two interpretations. If $\varphi = 1$, it captures absolute downward *real* wage rigidity (a case considered, for instance, by Martins et al. (2010)). If $\varphi < 1$, it captures absolute downward *nominal* wage rigidity: the real wage $\hat{w}_t^{DWR}$ is allowed to decrease by a factor φ (which is typically the inverse of the gross inflation rate). This wage rule has been adopted by, among others, Benigno and Ricci (2011) and Dupraz et al. (2025).[43] The parameters ϕ and φ are set to match the wage elasticity of the path-dependent rigid-wage model under the benchmark calibration of Section 4 (see also Appendix C.1).

Panels (a-c) of Figure 13 show the distribution of annual log-wage changes implied by these models. A simple visual inspection reveals key distinctive features of each wage setting model relative to the frictionless wage. The sticky wage model (panel (a)) results in a missing mass of large wage changes (both cuts and raises) relative to the frictionless case, while it also features a *larger* proportion of wage changes of smaller magnitudes. Under downward real wage rigidity (panel (b), red), wage cuts are absent, while wage increases are extremely rare, resulting in a frequency of wage freezes beyond 90%. Under downward nominal wage rigidity (panel (b), green), the real wage is allowed to fall, but wage cuts larger than $\ln(\varphi)$ are prevented.[44] The path-dependent rigid-wage model (panel (c)) results in a missing mass over both wage cuts and raises relative to the frictionless case, and presents a spike at zero.[45] Figure

[42] An alternative sticky-wage rule could be: $\hat{w}_t^{Sticky} = \phi \hat{w}_t^* + [1-\phi]\hat{w}_{t-1}^{Sticky}$, as in Shimer (2010).

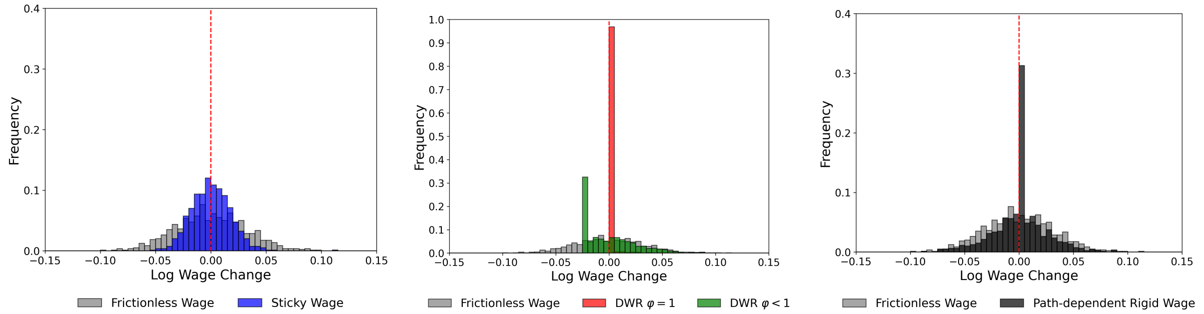
[43] The papers of Schmitt-Grohé and Uribe (2016) and Acharya et al. (2022) consider a hybrid version of DWR where $\hat{w}_t^{DWR}$ is the nominal wage and φ can take values between 0 and 1 included. See also Abbritti and Fahr (2013) for a model of asymmetric wage stickiness, where the wage is less responsive to negative shocks than positive shocks.

[44] As previously mentioned, under downward nominal wage rigidity $\varphi \equiv 1/\Pi$ where Π is the gross inflation rate, meaning that the real wage change distribution will have a spike at $\ln(1/\Pi) = -\ln(\Pi)$, i.e. at minus the inflation rate. Under the present calibration, $\varphi = 0.9980$ corresponds to a 0.002 (0.024) monthly (annual) inflation rate, generating a spike at -0.024 in the *annual* log-wage change distribution.

[45] In panel (c) the frequency of wage freezes is of about 30%. This is because, to ease the illustration, in Figure 13 the histogram bins are of length 0.005, rather than 0.001, which is the actual length used to calibrate the model

Figure 13: Annual log-wage change distributions across wage models

(a) Wage stickiness (b) Downward wage rigidity (c) Path-dependent wage rigidity



Note: The distributions of annual log-wage changes displayed in Panels (a-c) are computed via stochastic simulations of the four wage models considered as explained in detail in Appendix Section C.1. The bins are of length 0.005 (which explain the visual discrepancy with the figures in Table 3 below). The parameters $\phi = 0.54$ and $\varphi = 0.99805$ (corresponding to a 2.4% annual inflation rate) are set to match the same elasticity of the wage generated by the path-dependent rigid-wage model.

13 is complemented by Table 3 which reports the implied frequency of wage freezes (defined as absolute log-wage changes of less than 0.001) and wage elasticities across the models considered (also differentiating between positive (+) and negative (-) shocks). With the exception of the real DWR policy, all wage models present striking differences in the frequency of wage freezes and responsiveness of wages to positive and negative shocks (despite them sharing the same average wage elasticity by design).

Table 3: Frequency of wage freezes and wage cyclicality

Wage Policy	Wage Freezes (%)	Elasticity	Elasticity $(+)$	Elasticity $(-)$
Frictionless Wage	2.13	1.00	1.00	1.00
Path-dependent Rigid Wage	28.04	0.54	0.59	0.57
Sticky Wage	4.63	0.54	0.54	0.54
DWR $\varphi < 1$	1.63	0.54	0.67	0.34
DWR $\varphi = 1$	96.37	0.02	0.02	0.01

Note: The figures stem from a stochastic simulations of the four wage setting policies considered as explained in detail in Appendix Section C.1. The parameters $\phi = 0.54$ and $\varphi = 0.99805$ (corresponding to a 2.4% annual inflation rate) are set to match the same elasticity of the wage generated by the dynamic rigid-wage model. Wage freezes also include absolute log wage changes of less than 0.1%.

This exercise suggests that evaluating the empirical validity of wage setting models on the basis of their implied average elasticity can hide other important dynamic and cyclical properties of wages, which the present paper has shown can be crucial for unemployment dynamics. To this end, the path-dependent rigid-wage model has a greater degree of flexibility, since λ and γ can be jointly calibrated to achieve a given target frequency of wage freezes (λ) and a given average elasticity of the wage (γ). Moreover, this model appears to be

to a frequency of 28%.

consistent with several empirical facts on the dynamics of wages. In terms of wage change distributions, the model generates an empirically plausible spike at zero (see, e.g., Elsby and Solon, 2019), a missing mass of positive log-wage changes (Elsby, 2009; Stüber and Beissinger, 2012), and more dampened wage decreases (Holden and Wulfsberg, 2009, 2014; Kurmann and McEntarfer, 2024). In addition, it generates state-dependent wage rigidity (Grigsby et al., 2021; Schaefer and Singleton, 2023; Kurmann and McEntarfer, 2024), and empirically plausible average wage cyclicality (see, e.g., Haefke, Sonntag, and Rens, 2013), and due to relative wage comparisons, similar asymmetries between the wage of hiring and existing workers (Gertler, Huckfeldt, and Trigari, 2020; Grigsby et al., 2021; Hazell and Taska, 2025).

Since the wage dynamics implied by the path-dependent rigid-wage model is *the* mechanism at the heart of hysteresis effects, these considerations, together with the evidence discussed in the previous section, provide strong support for the theory of unemployment hysteresis advanced in this paper.

6 Discussion and Conclusions

This section develops an extension of the baseline model that generates endogenous path-dependent nominal wage rigidity to discuss the role of inflation and monetary policy in response to hysteresis and asymmetric fluctuations. Next, the section concludes with a summary of the main results and a discussion of future research avenues.

6.1 Inflation, money illusion, and monetary policy

This section presents an extension of the baseline model that includes inflation and a role for monetary policy (for details of the model derivations, see Appendix Section B.5).

Suppose the household can now borrow and save via a nominal (government) bond $A_t \equiv P_t a_t$ that pays a nominal interest rate i_t in each period, where P_t is the price level. Risk neutrality implies the Euler equation for bonds is an inverse relationship between the price and the nominal interest rate: $\mathbb{E}_t P_{t+1}/P_t = \beta[1+i_t]$. Next, suppose that employed household members evaluate wage contracts relative to a *nominal* reference wage $R_t^j \equiv P_t r_t^j$. This assumption implies that household members are subject to "money illusion" (Shafir et al., 1997),[46] and that incumbent workers are averse to nominal wage cuts. Then, optimal wage setting by the firm results in path-dependent *nominal wage rigidity* for existing workers which, due to relative wage comparisons of new hires', spills over to the hiring wage. The key implication is that, whenever the nominal wage is optimally freezed $\widetilde{W}_t = W_{t-1}$, the *real* wage decreases by a factor equal to the gross inflation rate: $\widetilde{W}_t/P_t = w_{t-1}/\Pi_t$, where $\Pi_t \equiv P_t/P_{t-1}$. Otherwise, whenever the nominal wage is adjusted, the optimal wage is equivalent to the one resulting

[46] See also Akerlof, Dickens, Perry, Bewley, and Blinder (2000).

from a model with *either* zero inflation *or* no money illusion. Hence, money illusion is the additional necessary assumption to generate downward nominal wage rigidity in the model of this paper. Importantly, despite inflation "greasing the wheels" when wages are not adjusted, muted wage cuts imply the real wage still fails to fully adjust as much as in the flexible-wage model.

In this version of the model a central bank with direct control over the nominal interest rate can affect prices and therefore real wages.[47] In this economy, the central bank might want to offset hysteresis effects or counteract asymmetric fluctuations. Since both situations are caused by frictions in the adjustment of the real wage in response to negative shocks, a monetary policy objective would be to target the price level $P_t^\star$ such that the real wage in the path-dependent rigid-wage model (i.e. for $\lambda > 1$) is equal to the frictionless real wage (i.e. for $\lambda = 1$):[48]

$$\tilde{w}(W_{t-1}/P_t^\star, z_t) = \tilde{w}^\star(z_t). \tag{21}$$

Such a policy, if *feasible*, would offset both hysteresis effects and asymmetric fluctuations in response to negative shocks.[49] The following proposition summarises the qualitative behaviour of a central bank setting the nominal interest rate to ensure that (21) holds.

Proposition 7. *Consider an economy where nominal wage cuts are costly $\lambda > 1$ and such that in some period $t = s$, $z_s < z_{s-1}$ and $\tilde{w}(W_{s-1}/P_s, z_s) > \tilde{w}^\star(z_s)$ due to downward nominal rigidity or muted wage cuts. The central bank achieves $\tilde{w}(W_{t-1}/P_t, z_t) = \tilde{w}^\star(z_t)$ for all $t \geq s$ by targeting of the price level*

$$P_t^\star \equiv P^\star(z_t) = \frac{P_{t-1}\tilde{w}_{t-1}}{\tilde{w}^\star(z_t)} \quad \forall t \geq s, \tag{22}$$

therefore setting the nominal interest rate

$$i_t^\star \equiv i^\star(z_t) = \frac{\mathbb{E}_t P_{t+1}}{P^\star(z_t)} \frac{1}{\beta} - 1 \quad \forall t \geq s, \tag{23}$$

where $P_t^\star$ is strictly decreasing in z_t and $i_t^\star$ is strictly increasing in z_t.

[47] The baseline model of Section 2 can be thought as that of an economy with a "passive" central bank setting the nominal interest rate $\bar{i} \equiv 1/\beta - 1$ to keep inflation constant at zero (i.e. $\Pi_t = \overline{\Pi} = 1$ for all t).

[48] The frictionless wage, i.e. the wage stemming from the flexible-wage model, is such that the unique steady-state equilibrium coincides with the low unemployment steady state. That is, $\tilde{w}_t^\star = \tilde{w}_t^+$ using the notation of Section 3. The steady state prevailing absent nominal rigidities has been referred as the "natural rate" by some authors (see, e.g., Blanchard, 2018).

[49] The analysis of the present paper abstracts from the zero lower bound (ZLB) to keep the exposition concise. The qualitative consequences of such a constraint are straightforward to understand—see the discussion below—while its quantitative relevance is beyond the scope of the present paper, but worth of further research. The policy considered here has been shown by Acharya et al. (2022) to be optimal under discretion for a planner who minimises deviations of unemployment and wage inflation from their frictionless benchmarks (both zero in their model), regardless of the relative weights put by the central bank on these deviations. For a more thorough treatment of monetary policy subject to the ZLB constraint and downward nominal wage rigidity, the reader is referred to Acharya et al. (2022) and references therein.

The monetary policy strategy resulting from Proposition 7 is intuitive and in line with how advanced economies central banks are described to operate: in response to negative shocks, due to downward nominal rigidity or muted nominal wage cuts, the central bank cuts the interest rate to increase the price level, generating price inflation $\Pi_s > 1$ and reducing the real wage to be equal to the frictionless wage. In so doing, the central bank is minimising deviations of the unemployment rate from the low unemployment steady state, while keeping wage inflation at zero.

What happens to the nominal interest rate and price inflation in subsequent periods crucially depends on the history of shocks up to $t = s$ (as per the analysis of Section 3). If the negative shock is the initial phase of a bust-boom cycle (negative transitory shock, that would otherwise generate larger unemployment fluctuations), after an initial interest rate cut, the central bank will reverse course by slowly increasing the nominal interest rate to reduce inflation, until the economy is back at the low unemployment steady state. Here, monetary policy fully offsets asymmetric unemployment fluctuations. If the negative shock is the reversal of a boom-bust cycle (positive transitory shock, that would otherwise generate hysteresis effects), instead, the central bank will keep decreasing the nominal interest rate, to increase the price level, along the equilibrium path, until the initial positive shock has fully reversed. However, while following this strategy enables the central bank to offset the hysteresis effects on unemployment, these are eventually absorbed by the price level, implying that the nominal interest rate remains permanently lower (hence, in light of Figure 3, one can interpret the resulting deviation from the initial interest rate as *coercive force*).[50]

The presence of a zero lower bound might therefore limit the central bank's ability to offset hysteresis effects or asymmetric fluctuations. The discussion above also suggests that the hysteresis effects generated by the path-dependent rigid-wage model may increase the likelihood of hitting the zero lower bound after a boom-bust cycle that leaves the nominal interest rate permanently lower. A proper analysis would require the quantitative assessment of a richer macroeconomic framework, possibly including the role of physical and human capital accumulation, and a more standard treatment of optimal monetary policy. While this is beyond the scope of the present paper, the preliminary analysis presented in this section suggests this as a promising direction for future research.

6.2 Summary and future directions

This paper advances a theory of unemployment hysteresis based on a model of endogenous, path-dependent rigidity in the hiring wage under incomplete employment contracts. Due to incumbents' aversion to wage cuts and the pay equality concerns of new hires, fully procyclical wage increases are partially irreversible, limiting the firm's ability to adjust hiring wages

[50]Further note that, in both cases, if the cycles are large—as previously defined—monetary policy effectively induces downward nominal wage rigidity even if the negative shock would have warranted (muted) nominal wage cuts.

downward in response to negative shocks. These properties—resulting in an optimal Ss wage policy—constitute the mechanism at the heart of the hysteresis effects and asymmetric fluctuations studied in this paper. This wage setting model is based on well-documented features of employment relationships and generates wage dynamics that are consistent with several empirical facts about wages over the business cycle. These considerations lend strong support to the path-dependent rigid-wage model and to the resulting theory of unemployment hysteresis.

An important feature of the theory is that periods of economic expansion can plant the seeds for hysteresis during subsequent recessions. The necessary and sufficient condition for this is the full procyclicality of wages during the expansion phase. Analogously, the model establishes that periods of economic contractions, where wages adjust fully downwards, can also generate hysteresis effects during subsequent expansions, resulting in permanently lower wages and unemployment. These theoretical considerations are novel, and have potential to guide future empirical work aimed at identifying hysteresis effects in macroeconomic time series. For instance, hysteresis effects are more likely to arise in labour markets, industries, or sectors characterised by incomplete employment contracts, and after periods of highly procyclical wage increases (or decreases).

The paper also provides a quantitative assessment of this mechanism for the magnitude of hysteresis effects and the resulting asymmetries in unemployment fluctuations. The analysis suggests that hysteresis effects can be quantitatively significant, generating unemployment remanence as high as 15% after a 1% transitory shock. Moreover, the model generates vacancy and unemployment fluctuations that are more than twice as large in response to negative shocks. This assessment is, nevertheless, limited. In fact, the simple general equilibrium framework adopted by this paper abstracts from important propagation mechanisms, such as physical and human capital accumulation. Hence, a promising avenue for future research would be to enrich the present framework by incorporating these elements, enabling the analysis of how skill loss due to long-term unemployment interacts with the hysteresis mechanism identified here, and study the resulting implications for capital accumulation and output dynamics. As discussed in the final section of the paper, future research could also consider the role of monetary policy in counteracting hysteresis effects. In this context, a necessary condition for the presence of path-dependent nominal wage rigidities—and therefore for the effectiveness of monetary policy—is that workers are subject to money illusion. While this assumption can be justified when inflation is relatively low, it is not obvious that it holds when inflation is high. Future research could investigate the consequences of the interactions between inflation, money illusion, hysteresis effects, and the zero lower bound on nominal interest rates for wage and unemployment dynamics.

The theory of unemployment hysteresis advanced in this paper ultimately relies on the asymmetry and partial irreversibility of wage adjustments. The paper aims to provide solid, evidence-based foundation for these dynamics into a general equilibrium framework, but has only began to explore their implications for the theory of equilibrium unemployment and the business cycle. This paper can therefore be seen as a starting point for a broader research programme that incorporates asymmetry, irreversibility, and hysteresis into business cycle models of unemployment.

References

Abbritti, M. and S. Fahr (2013). Downward wage rigidity and business cycle asymmetries. *Journal of Monetary Economics 60*(7), 871 – 886.

Acharya, S., J. Bengui, K. Dogra, and S. L. Wee (2022). Slow recoveries and unemployment traps: Monetary policy in a time of hysteresis. *The Economic Journal 132*(646), 2007–2047.

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

Akerlof, G. A., W. T. Dickens, G. L. Perry, T. F. Bewley, and A. S. Blinder (2000). Near-rational wage and price setting and the long-run phillips curve. *Brookings papers on economic activity 2000*(1), 1–60.

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

Altmann, S., A. Falk, A. Grunewald, and D. Huffman (2014). Contractual Incompleteness, Unemployment, and Labour Market Segmentation. *Review of Economic Studies 81*(October 2013), 30–56.

Amable, B., J. Henry, F. Lordon, and R. Topol (1994). Strong hysteresis versus zero-root dynamics. *Economics Letters 44*(1-2), 43–47.

Amable, B., J. Henry, F. Lordon, and R. Topol (1995). Hysteresis revisited: a methodological approach. In R. Cross (Ed.), *The natural rate of unemployment: reflections on 25 years of the hypothesis.* Cambridge University Press.

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

Benigno, P. and L. A. Ricci (2011). The Inflation-Output Trade-Off with Downward Wage Rigidities. *American Economic Review 101*(June), 1436–1466.

Bertheau, A., M. Kudlyak, B. Larsen, and M. Bennedsen (2025). Why firms lay off workers instead of cutting wages: Evidence from linked survey-administrative data. Technical report, IZA Discussion Papers.

Bertola, G. and R. J. Caballero (1990). Kinked adjustment costs and aggregate dynamics. *NBER Macroeconomics Annual 5*, 237–288.

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

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

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

Blanchard, O. (2018). Should we reject the natural rate hypothesis? *Journal of Economic Perspectives 32*(1), 97–120.

Blanchard, O. and J. Galí (2010). Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment. *American Economic Journal Macroeconomics*(2), 1–30.

Blanchard, O. J. and L. H. Summers (1986). Hysteresis and the european unemployment problem. *NBER Macroeconomics Annual 1*, 15–78.

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

Breza, E., S. Kaur, and Y. Shamdasani (2016). The Morale Effects of Pay Inequality. *NBER Working Paper No. 22491*.

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

Caplin, A. and J. Leahy (1991). State-dependent pricing and the dynamics of money and output. *The Quarterly Journal of Economics 106*(3), 683–708.

Card, D. (2022). Who Set Your Wage? *American Economic Review 112*(4), 1075–1090.

Cassar, L. and S. Meier (2018). Nonmonetary incentives and the implications of work as a source of meaning. *Journal of Economic Perspectives 32*(3), 215–238.

Cerra, V., A. Fatás, and S. C. Saxena (2023). Hysteresis and business cycles. *Journal of Economic Literature 61*(1), 181–225.

Cohn, A., E. Fehr, B. Herrmann, and F. Schneider (2014). Social Comparison and Effort Provision: Evidence from a Field Experiment. *Journal of the European Economic Association 12*, 877–898.

Cross, R. (1993). On the foundations of hysteresis in economic systems. *Economics & Philosophy 9*(1), 53–74.

Cross, R. (1994). The macroeconomic consequences of discontinuous adjustment. *Scottish Journal of Political Economy 41*(2).

Cross, R. (2014). Unemployment: Natural Rate Epicycles or Hysteresis? *European Journal of Economics and Economic Policies 11*(2).

Cross, R. and A. Allan (1988). On the hystory of hysteresis. In R. Cross (Ed.), *Unemployment, Hysteresis & the Natural Rate Hypothesis*. Basil Blackwell Ltd.

Cross, R., M. Grinfeld, and H. Lamba (2009). Hysteresis and economics. *IEEE Control Systems Magazine 29*(1), 30–43.

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

Danthine, J.-P. and A. Kurmann (2010, oct). The Business Cycle Implications of Reciprocity in Labor Relations. *Journal of Monetary Economics 57*(7), 837–850.

Dickson, A. and M. Fongoni (2019). Asymmetric reference-dependent reciprocity, downward wage rigidity, and the employment contract. *Journal of Economic Behavior & Organization 163*, 409 – 429.

Dixit, A. (1992). Investment and hysteresis. *Journal of Economic Perspectives 6*(1), 107–132.

Dufwenberg, M. and G. Kirchsteiger (2000). Reciprocity and wage undercutting. *European Economic Review 44*(4-6), 1069–1078.

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

Dupraz, S., E. Nakamura, and J. Steinsson (2025). A plucking model of business cycles. *Journal of Monetary Economics*, 103766.

Ehrlich, G. and J. Montes (2024). Wage rigidity and employment outcomes: Evidence from administrative data. *American Economic Journal: Macroeconomics 16*(1), 147–206.

Eliaz, K. and R. Spiegler (2014). Reference Dependence and Labor-Market Fluctuations. *NBER Macroeconomics Annual 28*, 159–200.

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

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

Elsby, M. W. L. and G. Solon (2019, August). How prevalent is downward rigidity in nominal wages? international evidence from payroll records and pay slips. *Journal of Economic Perspectives 33*(3), 185–201.

Esteves-Sorenson, C. (2018). Gift exchange in the workplace: Addressing the conflicting evidence with a careful test. *Management Science 64*(9), 4365–4388.

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

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

Flynn, J. (2022). Salary disclosure and individual effort: Evidence from the national hockey league. *Journal of Economic Behavior & Organization 202*, 471–497.

Fongoni, M. (2024a). Asymmetric reciprocity and the cyclical behavior of wages, effort, and job creation. *American Economic Journal: Macroeconomics 16*(3), 52–89.

Fongoni, M. (2024b). Does pay inequality affect worker effort? an assessment of experimental designs and evidence. *Journal of Economic Behavior & Organization 220*, 697–716.

Fongoni, M., D. Schaefer, and C. Singleton (2025). Why wages don't fall in jobs with incomplete contracts. *Management Science 71*(8), 6319–6339.

Fukui, M. (2020). A theory of wage rigidity and unemployment fluctuations with on-the-job search. *Working Paper*.

Furlanetto, F., A. Lepetit, Ø. Robstad, J. Rubio-Ramírez, and P. Ulvedal (2025). Estimating hysteresis effects. *American Economic Journal: Macroeconomics 17*(1), 35–70.

Gächter, S., D. Nosenzo, and M. Sefton (2012). The impact of social comparisons on reciprocity. *The Scandinavian Journal of Economics 114*(4), 1346–1367.

Galí, J. (2022). Insider–outsider labor markets, hysteresis, and monetary policy. *Journal of Money, Credit and Banking 54*(S1), 53–88.

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

Gertler, M., C. Huckfeldt, and A. Trigari (2020, 02). Unemployment Fluctuations, Match Quality, and the Wage Cyclicality of New Hires. *The Review of Economic Studies 87*(4), 1876–1914.

Gertler, M. and A. Trigari (2009). Unemployment Fluctuations with Staggered Nash Wage Bargaining Nash Wage Bargaining. *Journal of Political Economy 117*(1), 38–86.

Göcke, M. (2002). Various Concepts of Hysteresis Applied in Economics. *Journal of Economic Surveys 16*(2), 167–188.

Grigsby, J., E. Hurst, and A. Yildirmaz (2021). Aggregate Nominal Wage Adjustments: New Evidence from Administrative Payroll Data. *American Economic Review 111*(2), 428–471.

Haefke, C., M. Sonntag, and T. V. Rens (2013). Wage Rigidity and Job Creation. *Journal of Monetary Economics 60*(8), 887–899.

Hagedorn, M. and I. Manovskii (2008, September). The cyclical behavior of equilibrium unemployment and vacancies revisited. *American Economic Review 98*(4), 1692–1706.

Hall, R. E. (2005). Employment Fluctuations with Equilibrium Wage Stickiness. *The American Economic Review 95*(1), 50–65.

Hall, R. E. and P. R. Milgrom (2008). The limited influence of unemployment on the wage bargain. *American Economic Review 98*(4), 1653–74.

Hazell, J. and B. Taska (2025). Downward rigidity in the wage for new hires. *American Economic Review 115*(12), 4183–4217.

Herz, H. and D. Taubinsky (2017). What Makes a Price Fair? An Experimental Study of Transaction Experience and Endogenous Fairness Views. *Journal of the European Economic Association 16*(2), 316–352.

Holden, S. and F. Wulfsberg (2009). How Strong is the Macroeconomic Case for Downward Real Wage Rigidity? *Journal of Monetary Economics 56*, 605–615.

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

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

Kaur, S. (2019). Nominal wage rigidity in village labor markets. *American Economic Review 109*(10), 3585–3616.

Koenig, F., A. Manning, and B. Petrongolo (2024). Reservation wages and the wage flexibility puzzle. *Review of Economics and Statistics*, 1–32.

Kudlyak, M. (2014). The Cyclicality of the User Cost of Labor. *Journal of Monetary Economics 68*, 53–67.

Kurmann, A. and E. McEntarfer (2024). Downward Wage Rigidity in the United States: New Evidence from Administrative Data. *Working Paper*.

Lindbeck, A. and D. J. Snower (1987). Union Activity, Unemployment Persistence and Wage-employment Ratchets. *European Economic Review 31*(1), 157 – 167.

Ljungqvist, L. and T. J. Sargent (1998). The european unemployment dilemma. *Journal of political Economy 106*(3), 514–550.

Ljungqvist, L. and T. J. Sargent (2017). The Fundamental Surplus. *American Economic Review 107*(9), 2630–65.

Malcomson, J. M. (1981). Unemployment and the efficiency wage hypothesis. *The Economic Journal 91*(364), 848–866.

Manning, A. (2021). Monopsony in Labor Markets: A Review. *ILR Review 74*(1), 3–26.

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

Martins, P. S., A. Snell, and J. P. Thomas (2010). Downward Wage Rigidity in A Model of Equal Treatment Contracting. *Scandinavian Journal of Economics 112*(4), 841–863.

Michaillat, P. (2012). Do Matching Frictions Explain Unemployment? Not in Bad Times. *American Economic Review 102*(4), 1721–1750.

Mortensen, D. T. and E. Nagypál (2007). More on Unemployment and Vacancy Fluctuations. *Review of Economic Dynamics 10*(3), 327–347.

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

Petrosky-Nadeau, N. and E. Wasmer (2017). *Labor, Credit, and Goods Markets: The macroeconomics of search and unemployment.* MIT press.

Phelps, E. S. (1972). *Inflation Policy and Unemployment Theory: The Cost-Benefit Approach to Monetary Planning.* Macmillan, London.

Pissarides, C. A. (1985). Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages. *American Economic Review 75*(4), 676–690.

Pissarides, C. A. (1992). Loss of skill during unemployment and the persistence of employment shocks. *The Quarterly Journal of Economics 107*(4), 1371–1391.

Pissarides, C. A. (2009). The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer? *Econometrica 77*(5), 1339–1369.

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

Schaefer, D. and C. Singleton (2023). The extent of downward nominal wage rigidity: New evidence from payroll data. *Review of Economic Dynamics 51*, 60–76.

Schmitt-Grohé, S. and M. Uribe (2016). Downward nominal wage rigidity, currency pegs, and involuntary unemployment. *Journal of Political Economy 124*(5), 1466–1514.

Shafir, E., P. Diamond, and A. Tversky (1997). Money illusion. *The Quarterly Journal of Economics 112*(2), 341–374.

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

Shimer, R. (2010). *Labor Markets and Business Cycles.* Princeton University Press.

Simon, H. A. (1951). A formal theory of the employment relationship. *Econometrica: Journal of the Econometric Society*, 293–305.

Sliwka, D. and P. Werner (2017). Wage Increases and the Dynamics of Reciprocity. *Journal of Labor Economics 35*(2), 299–344.

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

Stüber, H. and T. Beissinger (2012, may). Does downward nominal wage rigidity dampen wage increases? *European Economic Review 56*(4), 870–887.

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

Yagan, D. (2019). Employment hysteresis from the great recession. *Journal of Political Economy 127*(5), 2505–2558.

Yellen, J. L. (2016). Macroeconomic research after the crisis. *Speech at the Conference in Honor of Thomas J. Sargent, Federal Reserve Bank of Boston, Boston, Massachusetts.*

Appendix

Marco Fongoni
Aix-Marseille Univ, CNRS, AMSE

A Mathematical Proofs

Proof of Proposition 2. Denote the combined discount factor of the firm by

$$\psi \equiv \hat{\beta}[1-\rho].$$

The optimal wage policy for an incumbent worker is characterised by the following Euler equation

$$z_t \mu'_t w_{it}^{-1} - 1 - \psi z_t^{\rho_z} \mu'_{t+1} w_{it}^{-1} = 0 \quad \forall w_{it} \neq r_{it} \quad \forall w_{it+1} \neq w_{it} \tag{A.1}$$

where

$$\mu'_t \equiv \begin{cases} \eta & \text{if } w_t > w_{t-1} \\ \lambda\eta & \text{if } w_t < w_{t-1} \end{cases} \qquad \mu'_{t+1} \equiv \begin{cases} \eta & \text{if } w_{t+1} > w_t \\ \lambda\eta & \text{if } w_{t+1} < w_t. \end{cases}$$

Hence it is possible to solve for the optimal wage explicitly: $\tilde{w}_{it} = \mu'_t z_t - \psi \mu'_{t+1} z_t^{\rho_z}$. The optimal wage is increasing in z_t if

$$\frac{\partial w_{it}}{\partial z_t} = \mu'_t - \psi \mu'_{t+1} \rho_z z_t^{\rho_z - 1} > 0 \quad \Longleftrightarrow \quad z_t > \left[\psi \rho_z \frac{\mu'_{t+1}}{\mu'_t}\right]^{\frac{1}{1-\rho_z}} \equiv \underline{z} \tag{A.2}$$

The proof proceeds as follows. First we conjecture that the condition on $z_t > \underline{z}$ holds, and we characterise the optimal wage policy. Then we verify that this condition holds under a wide range of plausible parameterisations of the shock.

Increasing sequences. If w_{it} is increasing in z_t, and z_t is an increasing sequence, then it must be that w_{it} is also an increasing sequence, and therefore that $\mu'_t = \mu'_{t+1} = \eta$. This implies the threshold $z^u(w_{t-1})$ is such that $w_{it} > w_{t-1}$ for all $z_t > z^u(w_{t-1})$, and $w_{it} = w_{t-1}$ for $z_t = z^u(w_{t-1})$ is given by the following condition

$$g^u(w_{t-1}, z_t) \equiv [z_t - \psi z_t^{\rho_z}] - \frac{w_{t-1}}{\eta} = 0 \tag{A.3}$$

since $r_{it} = w_{t-1}$ for incumbent workers. Then $g^u(w_{t-1}, 0) = -\frac{w_{t-1}^{\gamma}}{\eta} < 0, \lim_{z\to\infty} g^u(w_{t-1}, z_t) = \infty$ and $\frac{\partial g^u(w_{t-1}, z_t)}{\partial z_t} = 1 - \psi \rho_z z_t^{\rho_z - 1}$ is negative for all $z_t < \underline{z}$ and positive for all $z_t > \underline{z}$ where $\underline{z} \equiv [\psi\rho_z]^{\frac{1}{1-\rho_z}}$ was defined in (A.2) and corresponding to $\frac{\partial g^u(w_{t-1}, \underline{z})}{\partial z_t} = 0$ when $\mu'_t = \mu'_{t+1} = \eta$. Hence there exists a unique z_t at which $g^u(w_{t-1}, z_t) = 0$, which corresponds to $z_t^u(w_{t-1})$. We can also show that

$$\frac{\partial z_t^u(w_{t-1})}{\partial r_t} = -\frac{-\gamma w_{t-1}^{\gamma-1}\eta^{-1}}{\frac{\partial g^u(w_{t-1},z_t)}{\partial z_t}} > 0 \quad \Longleftrightarrow \quad z_t > \underline{z}$$

Hence, the optimal wage when $w_{it} > w_{t-1}$ is explicitly given by $w_{it} = \eta\{z_t - \psi z_t^{\rho_z}\}$. If the firm is behaving optimally in every period, then the next period optimal wage will be $w_{it+1} = \eta\{z_{t+1} - \psi z_{t+1}^{\rho_z}\}$ which will be greater than $\tilde{w}_{it}$ if $z_{t+1} > z^u(w_{it})$, which is implicitly determined by the value of z_{t+1} such that

$$[z_{t+1} - \psi z_{t+1}^{\rho_z}] - [z_t - \psi z_t^{\rho_z}] = 0$$

after having substituted in for the reference wage $r_{t+1} = \tilde{w}_{it}$ where $\tilde{w}_{it}$ is given above. From this, it is straightforward to verify that the only value of z_{t+1} satisfying the condition is $z_{t+1} = z_t$ implying that $z^u(w_{it}) = z_t$. Hence, if z_t is an increasing sequence, it will be the case that in every period $z_t > z^u(w_{it-1})$ and therefore that the optimal wage is indeed $w_{it} = \eta\{z_t - \psi z_t^{\rho_z}\}$. We now check the validity of the condition $z_t > [\psi\rho_z]^{\frac{1}{1-\rho_z}}$ for all t. Since z_t is increasing for all $t \geq s$, it is sufficient that the initial deviation from the steady state in period $t = s$ is $z_s > \underline{z}$, which is confirmed under plausible parameterisations of negative deviations from the steady state (note that even in the extreme case of $\psi \to 1$, and $\rho_z = 0.5$, $[\psi\rho_z]^{\frac{1}{1-\rho_z}} \approx 0.25$. Hence for z_s to be smaller than 0.25 would require a negative shock larger than 75%, which is implausible).

Decreasing sequences. If w_{it} is increasing in z_t, and z_t is a decreasing sequence, then it must be that w_{it} is also a decreasing sequence, and that $\mu'_t = \mu'_{t+1} = \lambda\eta$. This implies the threshold $z^l(w_{t-1})$ is such that $w_{it} < w_{t-1}$ for all $z_t < z^l(w_{t-1})$, and $w_{it} = w_{t-1}$ for $z_t = z^l(w_{t-1})$ is given by

$$g^l(w_{t-1}, z_t) \equiv [z_t - \psi z_t^{\rho_z}] - \frac{w_{t-1}}{\lambda\eta} = 0. \tag{A.4}$$

By analogy with the proof of the uniqueness of $z^u(w_{t-1})$, we can show that $z^l(w_{t-1})$ is unique and increasing in $r_{it} = w_{t-1}$ if $z_t > \underline{z}$. Hence, the optimal wage when $w_{it} < w_{t-1}$ is $w_{it} = \lambda\eta\{z_t - \psi z_t^{\rho_z}\}$. If the firm is behaving optimally in every period, then the next period optimal wage will be $w_{it+1} = \lambda\eta\{z_{t+1} - \psi z_{t+1}^{\rho_z}\}$ which will be smaller than w_{it} if $z_{t+1} < z^l(w_{it})$, which is implicitly determined by the value of z_{t+1} such that

$$[z_{t+1} - \psi z_{t+1}^{\rho_z}] - [z_t - \psi z_t^{\rho_z}] = 0$$

From this, it is straightforward to verify that the only value of z_{t+1} consistent with the equation is $z_{t+1} = z_t$ implying that $z^l(w_{it}) = z_t$. Hence, if z_t is a decreasing sequence, it will be the case that in every period $z_t < z^l(w_{it-1})$ and therefore that the optimal wage is indeed $w_{it} = \lambda\eta\{z_t - \psi z_t^{\rho_z}\}$. We know check the validity of the condition that $z_t > \underline{z}$ for all t. Since z_t is decreasing for all $t \geq s$ and bounded below by $\overline{z}$, it is sufficient that $\overline{z} > [\psi\rho_z]^{\frac{1}{1-\rho_z}}$, which our normalisation $\overline{z} = 1$ ensures (see also the numerical example discussed above). ■

Proof of Proposition 3. The proof follows the same logic as the proof of Proposition 2 with the only difference that the functions g^u and g^l, defined in (A.3) and (A.3) respectively, are now functions of $r_{nt} = \tilde{w}_{it}$ rather than w_{t-1}. This in turn implies we can express the thresholds (explicitly) in terms of the reference wage as a function of the shock:

$$r^u(z_t) = \lambda\eta\{z_t - \hat{\beta}[1-\rho]z_t^{\rho_z}\} \tag{A.5}$$
$$r^l(z_t) = \eta\{z_t - \hat{\beta}[1-\rho]z_t^{\rho_z}\} \tag{A.6}$$

where $r^u(z_t) > r^l(z_t)$ if $\lambda > 1$. The resulting optimal wage policy is then given by ($\psi \equiv \hat{\beta}[1-\rho]$):

$$\tilde{w}_{nt} \equiv \tilde{w}(r_{nt}, z_t) = \begin{cases} \eta\{z_t - \psi z_t^{\rho_z}\} & \text{if } r_{nt} < r^l(z_t) \\ r_{nt} & \text{if } r_{nt} \in [r^l(z_t), r^u(z_t)] \\ \lambda\eta\{z_t - \psi z_t^{\rho_z}\} & \text{if } r_{nt} > r^u(z_t). \end{cases} \tag{A.7}$$

Next, since $r_{nt} = \tilde{w}_{it}$, from (12) it is straightforward to verify that $r_{nt} = \tilde{w}_{it} \in [r^l(z_t), r^u(z_t)]$ for any z_t, which implies the firm will always pay new hires the same wage as incumbents: $\tilde{w}_{nt} = r_{nt} = \tilde{w}_{it}$. ■

Proof of Proposition 4. To solve for the steady state we first characterise the optimal wage policy for incumbent workers. For a given initial r_{is}, from (12) it follows that

$$\overline{w}_i = \begin{cases} \eta\varphi\overline{z} & \text{if } \overline{z} > z^u(r_{is}) = \dfrac{r_{is}}{\eta\varphi} \\ r_{is} & \text{if } \overline{z} \in [z^l(r_{is}), z^u(r_{is})] \\ \lambda\eta\varphi\overline{z} & \text{if } \overline{z} < z^l(r_{is}) = \dfrac{r_{is}}{\lambda\eta\varphi} \end{cases} \tag{A.8}$$

where $\varphi \equiv 1 - \hat{\beta}[1-\rho]$. Hence, depending on the relationship between r_{is} and $\overline{z}$, the optimal steady-state wage paid to incumbent workers is a continuous function of r_{is} and can take three distinct values: $\overline{w}_i^+ \equiv \eta\varphi\overline{z} > r_{is}$, $\overline{w}_i = r_{is}$, or $\overline{w}_i(\lambda)^- \equiv \lambda\eta\varphi\overline{z} < r_{is}$. Using this result, and the fact that $\overline{w}_n = \overline{w}_i = \overline{w}$, it is then relatively straightforward to characterise the model steady state, since vacancies and unemployment are crucially pinned down by the steady-state wage $\overline{w} \in [\overline{w}^+, \overline{w}(\lambda)^-]$. From $s+1$ onwards, incumbents' reference wage is in the steady state $r_{is+1} = \overline{r}_i = \overline{w}_i$, implying that incumbents exert normal effort $\overline{e}_i = \overline{e}$. And since new hires are paid the same $\overline{w}_n = \overline{w}_i$, it follows that new hires and incumbents are identical in the steady state, as it is their marginal value to the firm: $\bar{\mathcal{J}} = [\overline{ze} - \overline{w}]/\varphi$. ■

Proof of Proposition 5. Denote percentage changes between two steady states as $\Delta\overline{x} =$

$[\overline{x}' - \overline{x}]/\overline{x}$. When the economy is at the low unemployment steady state, then $\overline{w} = \overline{w}^+$. From the wage policy in (A.8) it is the case that $z^u(\overline{r}') = z^u(\overline{w}^+) = \overline{z}$ and $z^l(\overline{r}') = z^l(\overline{w}^+) = \overline{z}/\lambda$. It follows that $\Delta\overline{z} > 0$

$$\epsilon_{w,z} = \frac{\Delta\overline{w}}{\Delta\overline{z}} = \frac{\frac{\overline{w}(\overline{z}')^+ - \overline{w}(\overline{z})^+}{\overline{w}(\overline{z})^+}}{\frac{\overline{z}'-\overline{z}}{\overline{z}}} = \frac{\frac{\eta\varphi\overline{z}' - \eta\varphi\overline{z}}{\eta\varphi\overline{z}}}{\frac{\overline{z}'-\overline{z}}{\overline{z}}} = 1$$

for $\Delta\overline{z} \in [\frac{1-\lambda}{\lambda}, 0]$ the wage is unchanged, while for $\Delta\overline{z} < \frac{1-\lambda}{\lambda}$ then

$$\epsilon_{w,z} = \frac{\Delta\overline{w}}{\Delta\overline{z}} = \frac{\frac{\overline{w}(\overline{z}',\lambda)^- - \overline{w}(\overline{z})^+}{\overline{w}(\overline{z})^+}}{\frac{\overline{z}'-\overline{z}}{\overline{z}}} = \frac{\frac{\lambda\eta\varphi\overline{z}' - \eta\varphi\overline{z}}{\eta\varphi\overline{z}}}{\frac{\overline{z}'-\overline{z}}{\overline{z}}} = \frac{\lambda\overline{z}' - \overline{z}}{\overline{z}' - \overline{z}} < 1$$

since $\overline{z}' < \overline{z}$ and if $\lambda > 1$. Analogously, when the economy is at the high unemployment steady state, then $\overline{w} = \overline{w}(\lambda)^-$, and from the wage policy in (A.8) it is the case that $z^u(\overline{r}') = z^u(\overline{w}(\lambda)^-) = \lambda\overline{z}$ and $z^l(\overline{r}') = z^l(\overline{w}(\lambda)^-) = \overline{z}$. It follows that if $\Delta\overline{z} < 0$ then

$$\epsilon_{w,z} = \frac{\Delta\overline{w}}{\Delta\overline{z}} = \frac{\frac{\overline{w}(\overline{z}',\lambda)^- - \overline{w}(\overline{z},\lambda)^-}{\overline{w}(\overline{z},\lambda)^-}}{\frac{\overline{z}'-\overline{z}}{\overline{z}}} = \frac{\frac{\lambda\eta\varphi\overline{z}' - \lambda\eta\varphi\overline{z}}{\lambda\eta\varphi\overline{z}}}{\frac{\overline{z}'-\overline{z}}{\overline{z}}} = 1$$

when $\Delta z \in [0, \lambda - 1]$ then $\epsilon_{w,z} = 0$, while when $\Delta z > \lambda - 1$ then

$$\epsilon_{w,z} = \frac{\Delta w}{\Delta z} = \frac{\frac{\overline{w}(z')^+ - \overline{w}(z,\lambda)^-}{\overline{w}(z,\lambda)^-}}{\frac{z'-z}{z}} = \frac{\frac{\eta\varphi z' - \lambda\eta\varphi z}{\lambda\eta\varphi z}}{\frac{z'-z}{z}} = \frac{z' - \lambda z}{\lambda z' - \lambda z} < 1$$

since $\overline{z}' > \ sz$ and $\lambda > 1$. Finally, it is straightforward to show that $\epsilon_{w,z} = 1$ for any Δz if $\lambda = 1$. ■

Proof of Proposition 6. We start by considering the case of $\lambda > 1$.

Part i). Consider the case of the initial steady state being characterised by the low unemployment steady state: $\overline{w}_s = \overline{w}_s^+ = \eta\varphi\overline{z}_s$. From the optimal wage policy in the steady state (A.8) we know that in period $t = s'$ the shock thresholds are $z^u(\overline{r}_{is'}) = z^u(\overline{w}_s^+) = \overline{z}_s$ and $z^l(\overline{r}_{is'}) = z^l(\overline{w}_s^+) = \overline{z}_s/\lambda$. In a boom-bust cycle (moderate or large) since $\overline{z}_{s'} > \overline{z}_s$, during the expansion phase the optimal wage in period $t = s'$ is $\overline{w}_{s'}^+ = \eta\varphi\overline{z}_{s'}$, with an elasticity of one. This in turn implies that in period $t = s''$ the shock thresholds are $z^u(\overline{r}_{is''}) = z^u(\overline{w}_{s'}) = \overline{z}_{s'}$ and $z^l(\overline{r}_{is''}) = z^u(\overline{w}_{s'}) = \overline{z}_{s'}/\lambda$. Hence, in a moderate boom-bust cycle, in the recession phase in period $t = s''$, by its definition $\overline{z}_{s''} = \overline{z} > \overline{z}_{s'}/\lambda$ implying that the optimal wage is a wage freeze $\overline{w}_{s''} = \overline{w}_{s'} > \overline{w}_s$, with an elasticity of zero. While in a large boom-bust cycle, by its definition $\overline{z}_{s''} = \overline{z} < \overline{z}_{s'}/\lambda$ implying that the optimal wage is a wage cut $\overline{w}_{s''} = \lambda\eta\varphi\overline{z}_{s''} = \lambda\eta\varphi\overline{z}_s = \overline{w}(\lambda)^-$, but that is such that $\overline{w}_{s'} > \overline{w}_{s''} > \overline{w}_s$, with an elasticity of less than one. Next consider the case of the initial steady state being characterised by the high unemployment steady state: $\overline{w}_s = \overline{w}_s(\lambda)^- = \lambda\eta\varphi\overline{z}_s$. From the optimal wage policy in the steady state (A.8) we know that in period $t = s'$ the shock thresholds are $z^u(\overline{r}_{is'}) = z^u(\overline{w}_s(\lambda)^-) = \lambda \overline{z}_s$ and $z^l(\overline{r}_{is'}) = z^l(\overline{w}_s(\lambda)^-) = \overline{z}_s$. In a boom-bust cycle the response of the wage during the expansion phase depends on its magnitude. In a moderate boom-bust cycle, even if $\overline{z}_{s'} > \overline{z}_s$ by definition $\overline{z}_{s'} < \lambda \overline{z}_s$ which implies that during the expansion phase in period $t = s'$ the optimal wage is a wage freeze $\overline{w}_{s'} = \overline{w}_s$, with an elasticity of zero. This in turn implies that in period $t = s''$ the shock thresholds are unchanged $z^u(\overline{r}_{is''}) = \lambda \overline{z}_{s'}$ and $z^l(\overline{r}_{is''}) = \overline{z}_{s'}$. Hence, during the recession phase in period $t = s''$, as the shock reverses to $\overline{z}_{s''} = \overline{z}_s$, it implies that $\overline{z}_{s''} = z^l(\overline{r}_{is''})$ and the optimal wage is again a wage freeze $\overline{w}_{s''} = \overline{w}_{s'}$. Since the wage has not changed, as the shock reverses, unemployment will be back at its initial high steady state. In a large boom-bust cycle, by its definition $\overline{z}_{s'} > \lambda \overline{z}_s$ which implies that during the expansion phase in period $t = s'$ the optimal wage is a wage increase to $\overline{w}_{s'} = \eta\varphi \overline{z}_{s'} > \overline{w}_s$, with an elasticity of less than one. This in turn implies that in period $t = s''$ the shock thresholds are $z^u(\overline{r}_{is''}) = z^u(\overline{w}_{s'}) = \overline{z}_{s'}$ and $z^l(\overline{r}_{is''}) = z^u(\overline{w}_{s'}) = \overline{z}_{s'}/\lambda$. Hence, in the recession phase in period $t = s''$, the shock reverses to $\overline{z}_{s''} = \overline{z}_s$ which by definition $\overline{z}_s < \overline{z}_{s'}/\lambda$ implying that the optimal wage is a wage cut to $\overline{w}_{s''} = \lambda\eta\varphi \overline{z}_{s''} = \overline{w}_s$ since $\overline{z}_{s''} = \overline{z}_s$. Since the wage at the end of the cycle is back to its initial steady state, as the shock reverses unemployment will be back at its initial high steady state. Hence, boom-bust cycles results in permanently higher wages and unemployment only when, in the boom phase, wage increases are fully procyclical with an elasticity of one.

Part ii). Consider the case of the initial steady state being characterised by the low unemployment steady state: $\overline{w}_s = \overline{w}_s^+ = \eta\varphi \overline{z}_s$ In a bust-boom cycle the response of the wage during the recession phase depends on its magnitude. In a moderate bust-boom cycle, by its definition $\overline{z}_{s'} > z_s/\lambda$, which implies the optimal wage in period $t = s'$ is a wage freeze $\overline{w}_{s'} = \overline{w}_s$, with an elasticity of zero. This in turn implies that in period $t = s''$ the shock thresholds are unchanged $z^u(\overline{r}_{is''}) = \overline{z}_{s'}$ and $z^l(\overline{r}_{is''}) = \overline{z}_{s'}/\lambda$. Hence, in the expansion phase in period $t = s''$, the shock reverses to $\overline{z}_{s''} = \overline{z}_s = z^u(\overline{r}_{is''})$ which implies the optimal wage is again a wage freeze $\overline{w}_{s''} = \overline{w}_{s'}$. Since the wage has not changed, as the shock reverses unemployment will be back at its initial low steady state. In a large bust-boom cycle, by its definition $\overline{z}_{s'} < \overline{z}_s/\lambda$, which implies the optimal wage in period $t = s'$ is a muted wage cut $\overline{w}_{s'} = \lambda\eta\varphi \overline{z}_{s'} < \overline{w}_s$, with an elasticity of less than one. This in turn implies that in period $t = s''$ the shock thresholds are $z^u(\overline{r}_{is''}) = z^u(\overline{w}_{s'}) = \lambda \overline{z}_{s'}$ and $z^l(\overline{r}_{is''}) = z^u(\overline{w}_{s'}) = \overline{z}_{s'}$. Hence, in the expansion phase in period $t = s''$, the shock reverses to $\overline{z}_{s''} = \overline{z}_s$ which by definition $\overline{z}_s > \lambda \overline{z}_{s'}$ implying that the optimal wage is a wage increase to $\overline{w}_{s''} = \eta\varphi \overline{z}_{s''} = \overline{w}_s$ since $\overline{z}_{s''} = \overline{z}_s$. Since the wage at the end of the cycle is back to its initial steady state, as the shock reverses unemployment will be back at its initial low steady state. Next consider the case of the initial steady state being characterised by the high unemployment steady state: $\overline{w}_s = \overline{w}_s(\lambda)^- = \lambda\eta\varphi \overline{z}_s$. In a bust-boom cycle (moderate or large) since $\overline{z}_{s'} < \overline{z}_s$ during the recession phase in period $t = s'$ the optimal wage decreases to $\overline{w}_{s'}(\lambda)^- = \lambda\eta\varphi \overline{z}_{s'} < \overline{w}_s(\lambda)^-$, with an elasticity of one. This in turn implies that in period $t = s''$ the shock thresholds are

$z^u(\overline{r}_{is''}) = z^u(\overline{w}_{s'}) = \lambda\overline{z}_{s'}$ and $z^l(\overline{r}_{is''}) = z^u(\overline{w}_{s'}) = \overline{z}_{s'}$. Hence, in a moderate bust-boom cycle, in the expansion phase in period $t = s''$, by its definition $\overline{z}_{s''} = \overline{z} < \lambda\overline{z}_{s'}$ which implies the optimal wage is a wage freeze $\overline{w}_{s''} = \overline{w}_{s'} < \overline{w}_s$, with an elasticity of zero. While in a large bust-boom cycle, by its definition $\overline{z}_{s''} = \overline{z} > \lambda z_{s'}$ implying that the optimal wage is a wage increase $\overline{w}_{s''} = \eta\varphi\overline{z}_{s''} < \lambda\eta\varphi\overline{z}_s = \overline{w}_s(\lambda)^-$, since $\overline{z}_{s''} = \overline{z}_s$, with an elasticity of less than one. Hence, bust-boom cycles results in permanently lower wages and unemployment only when, in the recession phase, wage decreases are fully procyclical with an elasticity of one.

The case of $\lambda = 1$ is trivial, since there is a unique steady state to which the economy will always return after a symmetric cycle of any type and magnitude. ■

Proof of Proposition 7. Consider an initial situation in $t = s - 1$ such that

$$\widetilde{W}_{s-1} = \widetilde{W}^\star_{s-1} = \eta P_{s-1}\{z_{s-1} - \hat{\beta}[1-\rho]z_{s-1}^{\rho_z}\}$$

This could capture an economy that is initially at the low unemployment steady state, in which case $\widetilde{W}_{s-1} = \overline{W} = P_{s-1}\eta\{1 - \hat{\beta}[1-\rho]\}$, or an economy during the expansion phase of a positive, temporary shock in period $t = s - 1$ in which wages are fully procyclical. Next, consider a negative shock such that $z_s < z_{s-1}$. In the rigid-wage model the wage either remains constant, or it is cut, but not fully:

$$\widetilde{W}_s = \begin{cases} \widetilde{W}_{s-1} = \eta P_{s-1}\{z_{s-1} - \hat{\beta}[1-\rho]z_{s-1}^{\rho_z}\} & \text{if } z_s \in [z^l(W_{s-1}/P_s), z^u(W_{s-1}/P_s)] \\ \lambda\eta P_s\{z_s - \hat{\beta}[1-\rho]z_s^{\rho_z}\} & \text{if } z_s < z^l(W_{s-1}/P_s), \end{cases}$$

while in the frictionless model the wage is fully cut:

$$\widetilde{W}^\star_s = \eta P_s\{z_s - \hat{\beta}[1-\rho]z_s^{\rho_z}\} < \widetilde{W}^\star_{s-1} \quad \Leftrightarrow \quad P_s = P_{s-1}, z_s < z_{s-1}.$$

If $P_s = P_{s-1}$ then it follows that $\widetilde{W}_s > \widetilde{W}^\star_s$ for all $z_s < z_{s-1}$.

Downward nominal rigidity. In the case of downward nominal wage rigidity, the price level directly affects the real wage, and it will therefore be set such that:

$$\widetilde{W}_{s-1} = \widetilde{W}^\star_s \;\Leftrightarrow\; P_{s-1}w_{s-1} = P_s\tilde{w}^\star_s \;\Leftrightarrow\; P_s = \frac{P_{s-1}w_{s-1}}{\tilde{w}^\star_s} \equiv P^\star_s$$

Muted wage cuts. In the case of muted wage cuts, the central bank can increase P_s such that the upper threshold of the optimal wage policy decreases enough to warrant wage cuts equal to the flexible-wage model, that is

$$P_s z_s \geq P_s z^u(W_{s-1}/P_s) \;\Leftrightarrow\; z_s \geq z^u(W_{s-1}/P_s)$$

Denote the price at which the equality above is strict by $P^\star_s$. Any price $P_s \geq P^\star_s$ will be such that $\widetilde{W}_s = \widetilde{W}_s^\star$. However, the price $P_s > P_s^\star$ will generate wage inflation, since we know that at $P_s = P_s^\star$ then $z_s = z^u(W_{s-1}/P_s)$ and $\widetilde{W}_s = \widetilde{W}_s$ by definition, while at any $P_s > P_s^\star$ it follows that $\widetilde{W}_s > \widetilde{W}_s$. Hence we assume the central bank will prefer to keep wage-inflation at zero, by targeting $P_s = P_s^\star$. Indeed, it can be shown that $P_s^\star$ is equivalent to the one obtained in the downward nominal rigidity case, since at this price

$$\widetilde{W}_s = \widetilde{W}_{s-1} = \widetilde{W}_s^\star \Leftrightarrow P_{s-1}w_{s-1} = P_s\tilde{w}_s^\star \Leftrightarrow P_s = \frac{P_{s-1}w_{s-1}}{\tilde{w}_s^\star} \equiv P_s^\star$$

as above.

Both cases imply that $P_s^\star > P_{s-1}$, since $w_s^\star < w_{s-1}$. From the Euler equation, $i_t^\star$ is decreasing in $P_t^\star$ and therefore the central bank will cut the interest rate $i_s^\star < i_{s-1}^\star$ generating inflation $\Pi_s > 1$ in period $t = s$. The subsequent path of prices and interest rates depends on the subsequent path of z_t and of w_t relative to w_{t-1}. Since $\tilde{w}^\star = \tilde{w}^\star(z_t)$ is increasing in z_t, it follows that $P_t^\star$ is decreasing in z_t, and $i_t^\star$ increasing in z_t. ■

B Model Derivations

B.1 Household values of employment over unemployment

The household's optimisation problem can be expressed recursively:

$$\mathcal{V}(n_t, u_t) = \max_{c_t} \{c_t + \tilde{\nu}_{nt} n_{nt} + \tilde{\nu}_{it} n_{it} + \beta \mathbb{E}_t \mathcal{V}(n_{t+1}, u_{t+1})\} \tag{B.9}$$

subject to (1), (2), and (5). Since the household is risk neutral, and $n_t = n_{nt} + n_{it}$, the household value of having one more member employed as a new hire (in terms of consumption units) is given by $\mathcal{E}_n(w_{nt}, r_{nt}) \equiv \frac{\partial V(n_t,u_t)}{\partial n_{nt}}$ the value of having one more member employed as an incumbent by $\mathcal{E}_i(w_{it}, r_{it}) \equiv \frac{\partial V(n_t,u_t)}{\partial n_{it}}$ and the value of having one more member unemployed by $\mathcal{U}_t \equiv \frac{\partial V(n_t,u_t)}{\partial u_t}$. These are therefore respectively given by

$$\begin{aligned}
\frac{\partial V(n_t, u_t)}{\partial n_{nt}} &= w_{nt} + \tilde{\upsilon}_{nt} + \beta \mathbb{E}_t \frac{\partial V(n_{t+1}, u_{t+1})}{\partial n_{it+1}} \frac{\partial n_{it+1}}{\partial n_{nt}} + \beta \frac{\partial V(n_{t+1}, u_{t+1})}{\partial u_{t+1}} \frac{\partial u_{t+1}}{\partial n_{nt}} \\
\frac{\partial V(n_t, u_t)}{\partial n_{it}} &= w_{it} + \tilde{\upsilon}_{it} + \beta \mathbb{E}_t \frac{\partial V(n_{t+1}, u_{t+1})}{\partial n_{it+1}} \frac{\partial n_{it+1}}{\partial n_{it}} + \beta \frac{\partial V(n_{t+1}, u_{t+1})}{\partial u_{t+1}} \frac{\partial u_{t+1}}{\partial n_{it}} \\
\frac{\partial V(n_t, u_t)}{\partial u_t} &= b + \beta \mathbb{E}_t \frac{\partial V(n_{t+1}, u_{t+1})}{\partial n_{nt+1}} \frac{\partial n_{nt+1}}{\partial u_t} + \beta \frac{\partial V(n_{t+1}, u_{t+1})}{\partial u_{t+1}} \frac{\partial u_{t+1}}{\partial u_t}
\end{aligned}$$

where $\partial n_{it+1}/\partial n_{nt} = 1 - \rho, \partial u_{t+1}/\partial n_{nt} = \rho, \partial u_{t+1}/\partial n_{it} = \rho, \partial n_{nt+1}/\partial u_t = f_{t+1}$, and $\partial u_{t+1}/\partial u_t = 1 - f_{t+1}$. Hence, using the definitions above, these can be expressed recursively as

$$\mathcal{E}_n(w_{nt}, r_{nt}) = w_{nt} + \tilde{\upsilon}_n(w_{nt}, r_{nt}) + \beta[1-\rho]\mathbb{E}_t \mathcal{E}_i(w_{it+1}, r_{it+1}) + \beta\rho \mathbb{E}_t \mathcal{U}_{t+1} \tag{B.10}$$

$$\mathcal{E}_i(w_{it}, r_{it}) = w_{it} + \tilde{\upsilon}_i(w_{it}, r_{it}) + \beta[1-\rho]\mathbb{E}_t \mathcal{E}_i(w_{it+1}, r_{it+1}) + \beta\rho \mathbb{E}_t \mathcal{U}_{t+1} \tag{B.11}$$

$$\mathcal{U}_t = b + \beta \mathbb{E}_t f_{t+1} \mathbb{E}_t \mathcal{E}_n(w_{nt+1}, r_{nt+1}) + \beta \mathbb{E}_t [1 - f_{t+1}] \mathbb{E}_t \mathcal{U}_{t+1} \tag{B.12}$$

which, after some algebra, together imply that the household members' surpluses from employment are given by

$$\mathcal{S}_n(w_{nt}, r_{nt}) = w_{nt} + \tilde{\nu}(w_{nt}, r_{nt}) - b + \beta \left\{[1-\rho]\mathbb{E}_t \mathcal{S}_i(w_{it+1}, r_{it+1}) - \mathbb{E}_t f_{t+1} \mathbb{E}_t \mathcal{S}_n(w_{nt+1}, r_{nt+1})\right\} \tag{B.13}$$

$$\mathcal{S}_i(w_{it}, r_{it}) = w_{it} + \tilde{\nu}(w_{it}, r_{it}) - b + \beta \left\{[1-\rho]\mathbb{E}_t \mathcal{S}_i(w_{it+1}, r_{it+1}) - \mathbb{E}_t f_{t+1} \mathbb{E}_t \mathcal{S}_n(w_{nt+1}, r_{nt+1})\right\}. \tag{B.14}$$

for new hires and incumbents respectively.

B.2 Firm job values with new hires and incumbents

The firm's optimisation problem can be expressed more generally as

$$\mathcal{J}(r_{nt}, r_{is}, z_s, n_{s-1}) = \max_{\{w_{it}, w_{nt}, v_t\}} \mathbb{E}_s \sum_{t=s}^{\infty} \hat{\beta}^{t-s} \left\{ z_t[e_{nt}n_{nt} + e_{it}n_{it}] - w_{nt}n_{nt} - w_{it}n_{it} - \kappa v_t \right\}, \tag{B.15}$$

From the assumption on the timing of decisions, suppose that the firm has already optimally set its employed workers' wages. Then, for given number of vacancies v_t, optimal wages $\{\tilde{w}_{nt}, \tilde{w}_{it}\}$, and optimal effort choices $\{\tilde{e}_{nt}, \tilde{e}_{it}\}$ we can express the expected discounted sum of profits for the firm as

$$\mathcal{J}(r_{nt}, r_{it}, z_t, n_{t-1}) = \max_{v_t} \left\{ z_t[\tilde{e}_{nt}n_{nt} + \tilde{e}_{it}n_{it}] - \tilde{w}_{nt}n_{nt} - \tilde{w}_{it}n_{it} - \kappa v_t + \hat{\beta}\mathbb{E}_t \mathcal{J}(r_{nt+1}, r_{it+1}, z_{t+1}, n_t) \right\}.$$

where $n_t = n_{it} + n_{nt}$, in which $n_{it} = [1-\rho]n_{t-1}$ and $n_{nt} = h_t v_t$.

The marginal value of employing an additional incumbent worker is

$$\frac{\partial \mathcal{J}(r_{nt}, r_{it}, z_t, n_{t-1})}{\partial n_{it}} = z_t \tilde{e}_{it} - w_{it} + \hat{\beta}\mathbb{E}_t \frac{\partial \mathcal{J}(r_{nt+1}, r_{it+1}, z_{t+1}, n_t)}{\partial n_t} \frac{\partial n_t}{\partial n_{it}} \tag{B.16}$$

where $\partial n_t / \partial n_{it} = 1$ in any t, and where

$$\begin{aligned} \frac{\partial \mathcal{J}(r_{nt+1}, r_{it+1}, z_{t+1}, n_t)}{\partial n_t} &= z_{t+1}\tilde{e}_{it+1}\frac{\partial n_{it+1}}{\partial n_t} - \tilde{w}_{it+1}\frac{\partial n_{it+1}}{\partial n_t} + \hat{\beta}\mathbb{E}_{t+1}\frac{\partial \mathcal{J}(r_{nt+2}, r_{it+2}, z_{t+2}, n_{t+1})}{\partial n_{t+1}}\frac{\partial n_{t+1}}{\partial n_{it+1}}\frac{\partial n_{it+1}}{\partial n_t} \\ &= z_{t+1}\tilde{e}_{it+1}[1-\rho] - \tilde{w}_{it+1}[1-\rho] + \hat{\beta}\mathbb{E}_{t+1}\frac{\partial \mathcal{J}(r_{nt+2}, r_{it+2}, z_{t+2}, n_{t+1})}{\partial n_{t+1}}[1-\rho] \\ &= [1-\rho]\left[z_t\tilde{e}_{it+1} - \tilde{w}_{it+1} + \hat{\beta}\mathbb{E}_{t+1}\frac{\partial \mathcal{J}(r_{nt+2}, r_{it+2}, z_{t+2}, n_{t+1})}{\partial n_{t+1}} \right] \\ &= [1-\rho]\frac{\partial \mathcal{J}(r_{nt+1}, r_{it+1}, z_{t+1}, n_t)}{\partial n_{it+1}} \end{aligned}$$

since $\partial n_{it+1}/\partial n_t = 1-\rho$ and where the last equality follows from (B.16) iterated forward one period. Denoting $\mathcal{J}_i(r_{it}, z_t) \equiv \frac{\partial \mathcal{J}(r_{nt}, r_{it}, z_t, n_{t-1})}{\partial n_{it}}$ as in the main body of the paper, equation (B.16) can be expressed as

$$\mathcal{J}_i(r_{it}, z_t) = z_t\tilde{e}_{it} - \tilde{w}_{it} + \hat{\beta}[1-\rho]\mathbb{E}_t\mathcal{J}_i(r_{it+1}, z_{t+1}) \tag{B.17}$$

which is equivalent to (9) in which $\omega = i$ and $r_{it+1} = \tilde{w}_{it}$, if $\tilde{w}_{it}$ is the optimal wage.

Next, note that $\mathbb{E}_t\mathcal{J}_i(r_{it+1}, z_{t+1})$ in which $r_{it+1} = \tilde{w}_{it}$ is the expected continuation value of employing an incumbent worker that is paid the wage w_{it} in period t. Hence, since every newly hired worker becomes incumbent after one period of employment, $\mathbb{E}_t\mathcal{J}_i(r_{it+1}, z_{t+1})$ after substituting for $r_{it+1} = \tilde{w}_{nt}$, is the continuation value of hiring a newly hired worker that is paid the wage $\tilde{w}_{nt}$ in period t.

Then, the marginal value of employing an additional new hire is

$$\frac{\partial \mathcal{J}(r_{nt}, r_{it}, z_t, n_{t-1})}{\partial n_{nt}} = z_t \tilde{e}_{nt} - w_{nt} + \hat{\beta}\mathbb{E}_t \frac{\partial \mathcal{J}(r_{nt+1}, r_{it+1}, z_{t+1}, n_t)}{\partial n_t}\frac{\partial n_t}{\partial n_{nt}} \tag{B.18}$$

where $\partial n_t/\partial n_{nt} = 1$ in any t as before, and $\mathbb{E}_t \frac{\partial \mathcal{J}(r_{nt+1}, r_{it+1}, z_{t+1}, n_t)}{\partial n_t} = [1-\rho]\mathbb{E}_t \mathcal{J}_i(r_{it+1}, z_{t+1})$ as just derived, and in which $r_{it+1} = \tilde{w}_{nt}$. Denoting $\mathcal{J}_n(r_{nt}, z_t) \equiv \frac{\partial \mathcal{J}(r_{nt}, r_{it}, z_t, n_{t-1})}{\partial n_{nt}}$ as in the main body of the paper, equation (B.18) can be expressed as

$$\mathcal{J}_n(r_{nt}, z_t) = z_t \tilde{e}_{nt} - w_{nt} + \hat{\beta}[1-\rho]\mathbb{E}_t \mathcal{J}_i(r_{it+1}, z_{t+1}) \tag{B.19}$$

which is equivalent to (9) in which $\omega = n$ and $r_{it+1} = \tilde{w}_{nt}$, if $\tilde{w}_{nt}$ is the optimal wage.

B.3 Optimal wage policy with worker participation constraint

This section characterises the first order conditions for the firm's optimisation problem considering the worker's participation constraint, either by directly employing the participation constraint, or by using an expression for the reservation wage.

Preliminaries. Before characterising the optimal wage policy it is necessary to study whether the household's surplus from having a member employed (as a new hire or incumbent $\omega \in \{n, i\}$) rather than unemployed is increasing in the wage (which would ensure the existence of a reservation wage). In contrast to a canonical model, in the model of this paper this is not straightforward to establish. In fact, it might depend on whether the worker internalises the adaptation of their reference wage from the second employment period onwards ($r_{\omega t+1} = w_t$). In such a case, the worker anticipates that a higher wage in period t, which increases utility (in terms of consumption and morale) for any given effort level e_t (and wage w_t), will result in them having a higher reference wage in $t+1$, which decreases utility (in terms of morale) for any given effort level e_{t+1} (and wage w_{t+1}). To see this, express the marginal effect of a higher wage on the household's member value $\mathcal{S}_\omega$ as

$$\frac{\partial \mathcal{S}_\omega(w_t, r_{\omega t})}{\partial w_t} = \frac{\partial \mathcal{E}_\omega(w_t, r_{\omega t})}{\partial w_t} = 1 + \frac{\partial \tilde{v}_\omega(w_t, r_{\omega t})}{\partial w_t} + \beta[1-\rho]\mathbb{E}_t \frac{\partial \mathcal{E}_i(w_{t+1}, w_{\omega t})}{\partial r_{t+1}} \tag{B.20}$$

where by the envelope theorem $\partial \tilde{v}_{\omega t}/\partial e_t = 0$, which implies that

$$\frac{\partial \tilde{v}_\omega(w_t, r_{\omega t})}{\partial w_t} = \tilde{e}_t \mu' w_t^{-1}$$

and

$$\mathbb{E}_t \frac{\partial \mathcal{E}_i(w_{t+1}, w_{\omega t})}{\partial r_{t+1}} = \mathbb{E}_t \frac{\partial \tilde{v}_i(w_{t+1}, w_t)}{\partial r_{t+1}} = -\mathbb{E}_t \tilde{e}_{t+1}\mu'_{t+1} w_t^{-1}$$

Hence, if the worker internalises the reference wage adaptation rule, the last term in (B.20) is non-zero and captures an additional marginal cost for the worker of starting the next employment period with a higher reference wage (which decreases utility for any given effort level and wage). The resulting expression for $\frac{\partial \mathcal{S}_\omega(w_t, r_{\omega_t})}{\partial w_t}$ is then

$$\frac{\partial \mathcal{S}_\omega(w_t, r_{\omega_t})}{\partial w_t} = 1 + w_t^{-1}\{\tilde{e}_t \mu_t' - \beta[1-\rho]\mathbb{E}_t \tilde{e}_{t+1}\mu_{t+1}'\}. \tag{B.21}$$

From this it is possible to deduce that if the wage is a decreasing sequence, effort is also a decreasing sequence, and the second term will be positive (i.e. an additional marginal benefit of a higher wage). While if the wage is an increasing sequence, effort will be an increasing sequence, and the second term could possibly be negative (i.e. an additional marginal cost of a higher wage). Nevertheless, it is possible to show that a sufficient condition for $\frac{\partial \mathcal{S}_\omega(w_t, r_{\omega_t})}{\partial w_t} > 0$ is

$$\frac{\mathbb{E}_t \tilde{e}_{t+1} - \tilde{e}_t}{\tilde{e}_t} \leq \frac{\mu_t' - \beta[1-\rho]\mathbb{E}_t \mu_{t+1}'}{\beta[1-\rho]\mathbb{E}_t \mu_{t+1}'}$$

That is, if the growth of effort is not too high relative to the inverse of the worker discount factor, a higher wage will increase the value of being employed to the worker.

On the other hand, if the worker does not internalise the adaptation of the reference wage, then the last term in (B.20) is zero, implying that

$$\frac{\partial \mathcal{S}_\omega(w_t, r_{\omega_t})}{\partial w_t} = 1 + w_t^{-1}\tilde{e}_t \mu_t' > 0 \tag{B.22}$$

Whether it is plausible to assume that the worker internalises the adaptation of their reference wage, therefore anticipating the effects of a higher wage today on their work morale in the future, would require further discussion and analysis which is beyond the scope of the present paper. In what follows, results are established for both cases.

One example of a model dealing with a similar issue is that of Koenig, Manning, and Petrongolo (2024). They assume that workers are reference dependent when searching for a job and that their reference wage is a weighted average of the past wage and the steady-state wage. However, they also assume that when considering the option value of unemployment, the worker does not update their reference wage to the most recent wage contract, but rather keep using their past wage as reference. In the context of our discussion above, they assume that workers do not internalise the adaptation of their reference wage in the event they become unemployed in the future.

Participation constraint. Denote with Λ_t the lagrange multiplier associated with the worker $\omega \in \{n, i\}$ participation constraint $S_\omega(w_t, r_{\omega_t}) \geq 0$. The first-order condition characterising the optimal wage is given by

$$z_t \frac{\partial \tilde{e}(w_t, r_{\omega t}, \lambda)}{\partial w_t} - 1 + \hat{\beta}[1-\rho]\mathbb{E}_t \frac{\partial \mathcal{J}_i(w_t, z_{t+1})}{\partial r_{t+1}} + \Lambda_t \frac{\partial \mathcal{S}_\omega(w_t, r_{\omega_t})}{\partial w_t} = 0 \quad \forall w_t \neq r_{\omega t} \tag{B.23}$$

and the Kuhn-Tucker conditions are

$$\mathcal{S}_\omega(w_t, r_{\omega_t}) \geq 0, \quad \Lambda_t \geq 0, \quad \text{and} \quad \Lambda_t \mathcal{S}_\omega(w_t, r_{\omega_t}) = 0.$$

As in the main body of the paper

$$\frac{\partial \tilde{e}(w_t, r_{\omega t}, \lambda)}{\partial w_t} = \mu' w_t^{-1}$$

$$\mathbb{E}_t \frac{\partial \mathcal{J}_i(w_t, z_{t+1})}{\partial r_{t+1}} = \mathbb{E}_t z_{t+1} \frac{\partial \tilde{e}(w_{t+1}, w_t, \lambda)}{\partial r_{t+1}} = -z_t^{\rho_z} \mu' w_t^{-1}$$

while from the results above:

$$\frac{\partial \mathcal{S}_\omega(w_t, r_{\omega_t})}{\partial w_t} = \begin{cases} 1 + w_t^{-1}\{\tilde{e}_t \mu'_t - \beta[1-\rho]\mathbb{E}_t \tilde{e}_{t+1} \mu'_{t+1}\} & \text{with internalisation} \\ 1 + w_t^{-1} \tilde{e}_t \mu'_t & \text{without internalisation.} \end{cases}$$

Hence, if the participation constraint is binding in some period, then $\Lambda_t > 0$. Using the results established in the proof of Proposition 2, the first-order condition for the wage is

$$\mu' w_{\omega t}^{-1}\{z_t - \hat{\beta}[1-\rho] z_t^{\rho_z}\} - 1 + \Lambda_t \langle 1 + \mu' w_{\omega t}^{-1}\{\tilde{e}_t - \beta[1-\rho]\mathbb{E}_t \tilde{e}_{t+1}\}\rangle = 0 \tag{B.24}$$

with related Kuhn-Tucker conditions as above.

Reservation wage. Once established that $\lim_{w \to 0} \mathcal{S}_\omega(w_t, r_{\omega_t}) < 0$ and $\frac{\partial \mathcal{S}_\omega(w_t, r_{\omega_t})}{\partial w_t} > 0$ then it is possible to establish that there exists a reservation wage, denoted by $\underline{w}(r_{\omega t}) \geq 0$ such that $\mathcal{S}_\omega(\underline{w}(r_{\omega t}), r_{\omega_t}) = 0$ and the relevant constraint on the firm's wage problem becomes

$$w_t - \underline{w}(r_{\omega t}) \geq 0 \tag{B.25}$$

Denoting with Ω_t the lagrange multiplier associated with the worker $\omega \in \{n, i\}$ reservation wage constraint, the first-order condition characterising the optimal wage is then given by

$$z_t \frac{\partial \tilde{e}(w_t, r_{\omega t}, \lambda)}{\partial w_t} - 1 + \hat{\beta}[1-\rho]\mathbb{E}_t \frac{\partial \mathcal{J}_i(w_t, z_{t+1})}{\partial r_{t+1}} + \Omega_t = 0 \quad \forall w_t \neq r_{\omega t} \tag{B.26}$$

and the Kuhn-Tucker conditions are

$$w_t - \underline{w}(r_{\omega t}) \geq 0, \quad \Omega_t \geq 0, \quad \text{and} \quad \Omega_t[w_t - \underline{w}(r_{\omega t})] = 0.$$

Hence, if the participation constraint is binding in some period, then $\Omega_t > 0$. Using the results established above and in the proof of Proposition 2, the first-order condition for the wage can be written as

$$\mu' w_{\omega t}^{-1}\{z_t - \hat{\beta}[1-\rho] z_t^{\rho_z}\} - 1 + \Omega_t = 0 \tag{B.27}$$

with related Kuhn-Tucker conditions as above.

B.4 Optimal vacancy posting decision

For given optimal wages $\tilde{w}_{nt}$ and $\tilde{w}_{it}$ and respective optimal effort choices $\tilde{e}_{nt}$ and $\tilde{e}_{it}$ of newly hired and incumbent workers, the first-order condition characterising the optimal vacancy posting decision of the firm is

$$z\tilde{e}_{nt}h_t - \tilde{w}_{nt}h_t - \kappa + \hat{\beta}\mathbb{E}_t \frac{\partial \mathcal{J}(z_{t+1}, n_t)}{\partial n_t} \frac{\partial n_t}{\partial n_{nt}} h_t = 0 \tag{B.28}$$

where

$$\mathbb{E}_t \frac{\partial \mathcal{J}(z_{t+1}, n_t)}{\partial n_t} = [1-\rho]\mathbb{E}_t \frac{\partial \mathcal{J}(z_{t+1}, n_t)}{\partial n_{it+1}} = [1-\rho]\mathbb{E}_t \mathcal{J}_i(r_{it+1}, z_{t+1}) \tag{B.29}$$

is the marginal effect of hiring an additional worker in period t on the expected continuation value of the firm, and can be derived from iterating forward one period the envelope condition for vacancies. In fact, (B.29) corresponds to the expected marginal value of a newly hired worker in period t that is not separated at the end of the period, with probability $1-\rho$, and becomes incumbent in the next period with expected marginal value $\mathbb{E}_t \mathcal{J}_i(r_{it+1}, z_{t+1})$. By combining the first-order (B.28) and envelope (B.29) conditions and solving for θ using $h_t \equiv h(\theta_t) = \frac{m_t}{v_t} = \overline{m}\theta_t^{-\alpha}$ and the definition of the marginal value of employing an additional new hire in (B.18) yields the job creation condition (16) in the text.

B.5 Extension: Inflation and money illusion

Household. Suppose the household can borrow or save via a nominal bond A_t from the government, which pays the nominal interest rate i_t in each period. The household budget constraint can be expressed as

$$c_t + a_t = w_{nt}n_{nt} + w_{it}n_{it} + bu_t + d_t\tau_t + [1 + i_{t-1}]\frac{P_{t-1}}{P_t}a_{t-1} \tag{B.30}$$

where P_t is the price level in period t, and P_0 is given. The household optimisation problem can now be expressed as the choice of real assets a_t

$$\mathcal{V}(n_t, u_t, a_{t-1}) = \max_{a_t} \{c_t + \tilde{\nu}_{nt}n_{nt} + \tilde{\nu}_{it}n_{it} + \beta\mathbb{E}_t\mathcal{V}(n_{t+1}, u_{t+1}, a_t)\} \tag{B.31}$$

subject to the real budget constraint (B.30). By combining the first-order and envelope condition we obtain the familiar Euler equation under linear utility

$$\frac{\mathbb{E}_t P_{t+1}}{P_t} = \beta[1 + i_t] \tag{B.32}$$

which defines an inverse relationship between the price P_t and the nominal interest rate i_t. Denote by $\Pi_t \equiv P_t/P_{t-1}$ the gross inflation rate.

Next, it is assumed that household members are subject to "money illusion", that is, they evaluate wages with respect to a *nominal* reference wage $R^j \equiv P_t r_t^j$. Combined with our assumption on reference wage formation of newly hired and incumbent workers implies that

$$R_t^j = \begin{cases} R_{nt}^j = W_{it}^j \equiv P_t w_{it}^j & \text{if } n_{nt}^j = 1 \\ R_{it}^j = W_{t-1}^j \equiv P_{t-1} w_{t-1}^j & \text{if } n_{it}^j = 1 \end{cases} \quad R_{i0} \text{ given} \tag{B.33}$$

Firm. As in the baseline model without inflation, it is possible to generalise the firm optimal wage setting problem with an employed worker $\omega \in \{n, i\}$, but in nominal terms, where the firm takes the entire price path as given. Hence, for a worker with reference wage $R_{\omega t}$, we denote the nominal value of a job in period t in terms of period t prices as $\mathcal{P}_\omega(R_{\omega t}, P_t z_t) \equiv P_t \mathcal{J}_\omega(r_{\omega t}, z_t)$ The wage setting problem can now be written as follows:

$$\mathcal{P}_\omega(R_{jt}, P_t z_t) = \max_{W_t} \left\{ P_t z_t \tilde{e}(W_t, R_t, \lambda) - W_t + \hat{\beta}[1-\rho]\mathbb{E}_t \Pi_{t+1}^{-1} \mathcal{P}_\omega\left(W_t, P_{t+1} z_{t+1}\right) \right\} \tag{B.34}$$

$$\text{s.t.} \quad R_{t+1} = W_t, \; R_{\omega t} \text{ given} \tag{B.35}$$

$$\mathcal{E}_\omega(W_t, R_{\omega t}) - \mathcal{U}_t \geq 0 \tag{B.36}$$

where $R_{\omega t} = \{W_{t-1}, W_{\omega t}\}$, and nominal shocks evolve according to

$$Z_{t+1} = P_{t+1} \bar{z}^{1-\rho_z} [Z_t/P_t]^{\rho_z} \exp(\varepsilon_{t+1}) \tag{B.37}$$

where, under perfect foresight, $\mathbb{E}_t Z_{t+1} = \mathbb{E}_t P_{t+1}[Z_t/P_t]^{\rho_z} = \mathbb{E}_t P_{t+1} z_t^{\rho_z}$ if $\varepsilon_{t+1} = 0$ for all t.

Using analogous arguments to those used in the proof of Propositions 2 and 3 it is possible to show that the resulting optimal wage policy for incumbent workers can be expressed in nominal terms as

$$\widetilde{W}_{it} = \widetilde{W}(W_{t-1}, P_t z_t) = \begin{cases} \eta P_t \{z_t - \hat{\beta}[1-\rho] z_t^{\rho_z}\} & \text{if } z_t > z^u(W_{t-1}/P_t) \\ W_{t-1} & \text{if } z_t \in [z^l(W_{t-1}/P_t), z^u(W_{t-1}/P_t)] \\ \lambda \eta P_t \{z_t - \hat{\beta}[1-\rho] z_t^{\rho_z}\} & \text{if } z_t < z^l(W_{t-1}/P_t) \end{cases} \tag{B.38}$$

where the thresholds are the uniquely characterised by

$$P_t z^u(W_{t-1}/P_t) - \hat{\beta}[1-\rho] P_t z^u(W_{t-1}/P_t)^{\rho_z} - \frac{W_{t-1}}{\eta} = 0 \tag{B.39}$$

$$P_t z^l(W_{t-1}/P_t) - \hat{\beta}[1-\rho] P_t z^l(W_{t-1}/P_t)^{\rho_z} - \frac{W_{t-1}}{\lambda \eta} = 0 \tag{B.40}$$

and that new hires are paid the same wage as incumbents

$$\widetilde{W}_{nt} = R_{nt} = \widetilde{W}_{it} \tag{B.41}$$

These results imply that whenever the nominal wage is adjusted, the optimal wage policy is equivalent to that resulting from a model with zero inflation or no money illusion. The only difference is that now the model generates *nominal* wage rigidity for a range of shocks, such that whenever the nominal wage is optimally freezed by the firm, the *real* wage decreases by a factor equal to the gross inflation rate: $\widetilde{W}_t/P_t = W_{t-1}/P_t = w_{t-1}/\Pi_t$ whenever $z_t \in [z^l\left(w_{t-1}/\Pi_t\right), z^u\left(w_{t-1}/\Pi_t\right)]$.

Government. The government budget is balanced, such that

$$P_t b u_t + [1 + i_{t-1}]A_{t-1} = P_t \tau_t + A_t \tag{B.42}$$

Central bank. The nominal interest rate i_t is under the control of the central bank. The baseline model in the main body of the paper can be thought as an economy in which a "passive" central bank would set the nominal interest rate to keep inflation constant at zero (i.e. $\Pi_t = 1$ for all t): $\tilde{i}_t = \bar{i} \equiv [1-\beta]/\beta$ in every period. In this extension we consider the role of an active central bank that can chose the nominal interest rate to affect the price level, with the objective to keep the path-dependent rigid wage ($\lambda > 1$) equal to the frictionless wage ($\lambda = 1$).

C Computational Approach

C.1 Stochastic model

This section explains the details behind the stochastic simulations of the optimal wage setting policy performed in Section 5.2. The section largely draws on the analytical results and numerical methods developed by Elsby (2009), which have been more recently adopted and extended by Fongoni (2024a). The reader is referred to these papers for a more formal treatment and discussion.

Optimal wage policy under uncertainty. Denote by $\hat{z}_t$ the real shock to be used in the stochastic simulation of the model (i.e. with uncertainty). The stochastic process for $\hat{z}_t$ is assumed to be given by the geometric random walk

$$\ln \hat{z}_{t+1} = \ln \hat{z}_t + \varepsilon_{t+1} \tag{C.43}$$

where $\hat{z}_0$ is given and $\varepsilon_{t+1} \sim \mathcal{N}(0, \sigma_{\hat{z}}^2)$ are i.i.d. shocks. Hence, $\hat{z}_{t+1}$ follows a log-normal distribution where $\ln \hat{z}_{t+1} \sim \mathcal{N}(\ln \hat{z}_t, \sigma_{\hat{z}}^2)$. Under this assumption, it can be shown that the optimal wage policy for a worker $\omega \in \{n, i\}$ takes the following "Ss" form:

$$\hat{w}_{\omega t} \equiv \hat{w}(r_{\omega t}, \hat{z}_t) = \begin{cases} [\hat{z}_t/\psi^U]^{\frac{1}{\gamma}} > r_{\omega t} & \text{if } \hat{z}_t > \hat{z}^u(r_{\omega t}) \\ r_{\omega t} & \text{if } \hat{z}_t \in [\hat{z}^l(r_{\omega t}), \hat{z}^u(r_{\omega t})] \\ [\hat{z}_t/\psi^L]^{\frac{1}{\gamma}} < r_{\omega t} & \text{if } \hat{z}_t < \hat{z}^l(r_{\omega t}); \end{cases} \tag{C.44}$$

where $\psi^U > \psi^L$ if $\lambda > 1$;

$$\hat{z}^u(r_{\omega t}) = \psi^U r_{\omega t}^{\gamma}; \quad \text{and} \quad \hat{z}^l(r_{\omega t}) = \psi^L r_{\omega t}^{\gamma}. \tag{C.45}$$

The coefficients ψ^U and ψ^L are functions of the model parameters $\{\beta, \sigma_{\hat{z}}, \eta, \lambda\}$ and can be numerically computed.

Stochastic simulations. For the wage policy simulations performed in Section 5.2, the stochastic process for $\hat{z}_t$ is calibrated such that its mean and steady-state value are normalised to 1 and $\sigma_{\hat{z}} = 0.009$ is chosen to match a standard deviation of 0.02 This is consistent with the calibration of the AR(1) process for z_t in the main body of the paper. To achieve this, a series of n stochastic processes $\hat{z}_n^T = (\hat{z}_1, \ldots, \hat{z}_T)$ of $T = 10636$ periods are simulated, and the first 1000 processes with mean 1 and average quarterly standard deviation of 0.02 in the last 636 periods are collected and can be used to simulate the model. For the wage setting policy in (C.44), the firm discounting wedge is set to $\delta = 1$, implying that $\hat{\beta} = \beta = 0.996$ (since here there is no issue with active wage compression under perfect foresight as discussed in the main body of the paper, and the elasticity of the wage with respect to shocks, whenever the wage is adjusted, is equal to $1/\gamma$). The remaining parameters are set to $\gamma = 1$, $\lambda = 0.014$, and $\eta = 0.7$, consistent with the benchmark calibration. This yields a frequency of wage freezes of 28% and an average elasticity of the wage of 0.54. The remaining parameters featuring the sticky wage policy and the DWR wage policy with $\varphi \in (0, 1)$ are set to $\phi = 0.54$ and $\varphi = 0.99805$ (corresponding to a 2.4% annual inflation rate) in order to match the same average elasticity of the wage of 0.54 generated by the path-dependent rigid-wage model.

Calibration of λ. To calibrate λ to achieve a desired frequency of wage freezes, the exercise adopts the following procedure. Starting from a value of λ close to one the wage setting policy in (C.44) is simulated over a series of shocks following (C.43) and calibrated as above, for 9600 months (corresponding to 800 years). The resulting time series is aggregated to a yearly frequency, then logged and hp-filtered. From this, the frequency of log-wage changes that are less than 0.001 is computed—this corresponds to the definition of wage freezes in the text (note it is also possible to compute the frequency of log-wage changes of exactly zero). If this is less than the desired target frequency, the procedure above is repeated as many times as necessary by iteratively augmenting λ until the desired frequency of wage freezes is achieved.

C.2 Impulse response analysis

Backward recursion algorithm. Since the firm's optimal wage setting policy can be solved independently of its vacancy posting decision, the approach to solve the model's impulse response to unanticipated transitory shocks in a given period $t = s$ under perfect foresight is the following.

First, for any given sequence of $\{z_t\}$ and $\{r_{it}\}$ with z_s and r_{is} given, we can solve for the optimal wage policy sequence $\{\tilde{w}_{it}\}$ for the wage paid to incumbent workers, as well as their optimal effort sequence $\{\tilde{e}_{it}\}$ from the period of the shock s until convergence to steady state. Denote the period in which the system converges to the steady state by T. Next, it is possible to compute the optimal wage sequence paid to new hires $\{\tilde{w}_{nt}\}$, since this is equal to the wage paid to incumbents. Also, this implies they new hires exert normal effort along the equilibrium path $e_{nt} = \bar{e}$ for all $t = s, \ldots, T$. With this information it is then possible to compute the firm's per-period, per-worker profit for new hires and incumbents respectively as $\tilde{d}_{nt} = z_t\bar{e} - \tilde{w}_{nt}$ and $\tilde{d}_{it} = z_t\tilde{e}_{it} - \tilde{w}_{it}$ for all $t = s, \ldots, T$.

Next, to compute the job values $\mathcal{J}_{it}$ and $\mathcal{J}_{nt}$ it is possible to employ a backward recursion algorithm that leverages the recursive structure of the value functions (B.17) and (B.19) after having substituted in for the optimal wage policy. That is, computing the following (where $\psi \equiv \hat{\beta}[1 - \rho]$):

$$\mathcal{J}_{it} = \tilde{d}_{it} + \psi\mathcal{J}_{it+1} \quad \forall t = s, \ldots, \infty$$

$$\mathcal{J}_{nt} = \tilde{d}_{nt} + \psi\mathcal{J}_{it+1} \quad \forall t = s, \ldots, \infty$$

with initial condition given by $\mathcal{J}_{is-1} = \mathcal{J}_{ns-1} = \frac{\overline{z}\overline{e} - \overline{w}_{s-1}}{1-\psi}$ and terminal condition $\mathcal{J}_{iT} = \mathcal{J}_{nT} = \frac{\overline{z}\overline{e} - \overline{w}_T}{1-\psi}$, and where $\overline{w}_{s-1}$ and $\overline{w}_T$ can be any of the steady-state wages within the range established by Proposition 4. Hence, since the paths of $\tilde{d}_{nt}$ and $\tilde{d}_{it}$ are known for all $t = s, \ldots, T$, the algorithm iteratively applies the backward recursion

$$\mathcal{J}_{it} = \tilde{d}_{it} + \psi \mathcal{J}_{it+1} \quad \forall t = T-1, T-2 \ldots, s$$

which can then be used to calculate $\mathcal{J}_{nt} = \tilde{d}_{nt} + \psi \mathcal{J}_{it+1}$ in each period $t = s \ldots, T$ as required. This, in turn, can be used to calculate the optimal sequence of market tightness $\{\tilde{\theta}_t\}$.

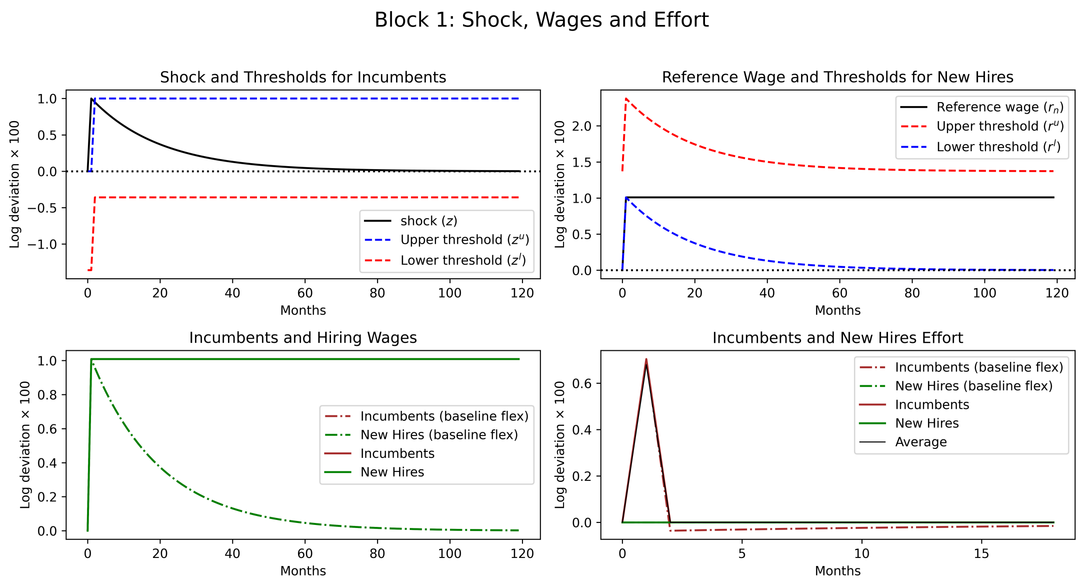
Finally, the same technique can be applied to compute the household's marginal utilities, or worker surpluses, $\mathcal{S}_{nt}$ and $\mathcal{S}_{it}$ as expressed in (B.13) and (B.14). This can be done after having computed the sequences for the probability of matching with a vacancy $\{\bar{m}\tilde{\theta}_t^{1-\alpha}\}$ and the optimised net payoffs from employment $\tilde{v}_{nt}$ and $\tilde{v}_{it}$ using the information on the wage and effort paths computed as above. The backward recursion will then simultaneously solve for

$$\begin{aligned}
\mathcal{S}_{nt} &= \tilde{w}_{nt} + \tilde{v}_{nt} - b + \beta[1-\rho]\mathcal{S}_{it+1} - \beta f(\tilde{\theta}_{t+1})\mathcal{S}_{nt+1} \quad t = T-1, T-2, \ldots, s \\
\mathcal{S}_{it} &= \tilde{w}_{it} + \tilde{v}_{it} - b + \beta[1-\rho]\mathcal{S}_{it+1} - \beta f(\tilde{\theta}_{t+1})\mathcal{S}_{nt+1} \quad t = T-1, T-2, \ldots, s
\end{aligned}$$

with initial and terminal conditions given by $\mathcal{S}_{ns-1} = \mathcal{S}_{is-1} = \frac{\overline{w}_{s-1} + \bar{e}^2/2 - b}{1-\beta[1-\rho-f(\overline{\theta}_{s-1})]}$ and $\mathcal{S}_{nT} = \mathcal{S}_{iT} = \frac{\overline{w}_T + \bar{e}^2/2 - b}{1-\beta[1-\rho-f(\overline{\theta}_T)]}$ where $\overline{w}_{s-1}$ and $\overline{w}_T$ can be any of the steady-state wages within the range established by Proposition 4, and $\overline{\theta}$ can be computed once they are known.

Full benchmark results. The figures below report results of the benchmark simulations for all the variables of the model. More precisely, Figures C.1 and C.2 complement Figure 6 in the main body of the paper, while Figures C.3 and C.4 complement Figure 7.

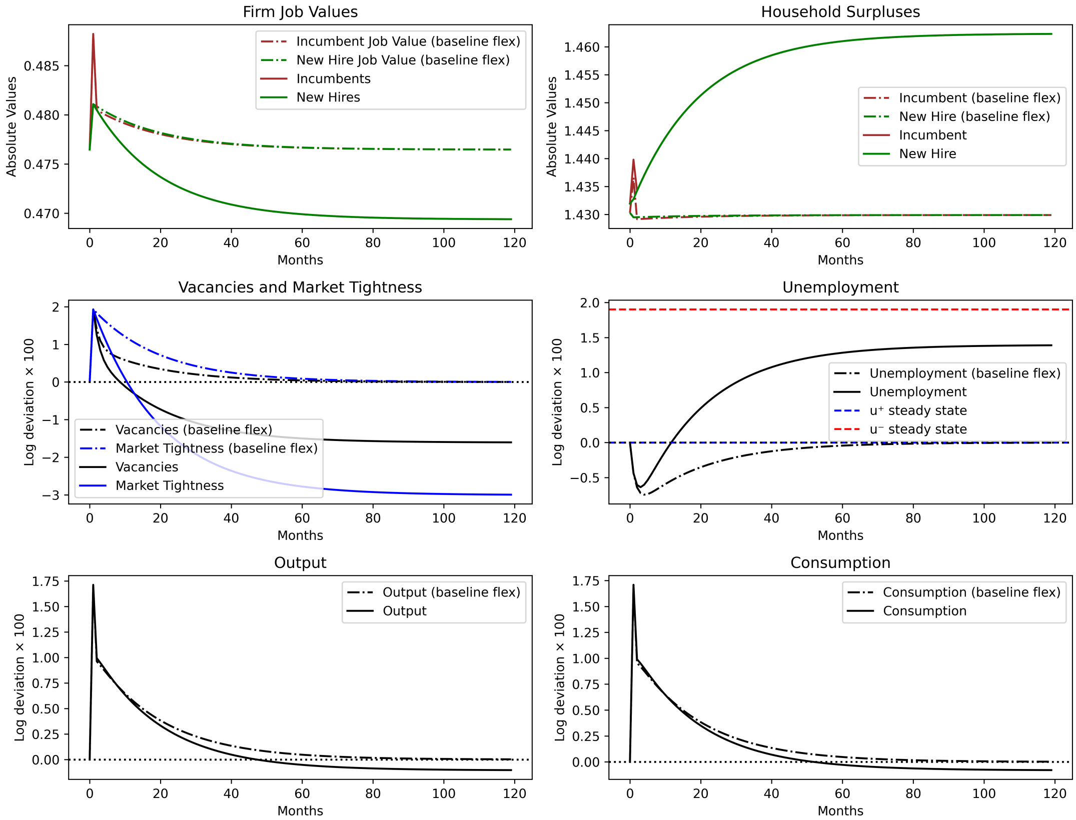
Figure C.1: Hysteresis effects: Low steady state



Note: Response to a 1% positive, transitory shock starting from the low unemployment steady state.

Figure C.2: Hysteresis effects: Low steady state

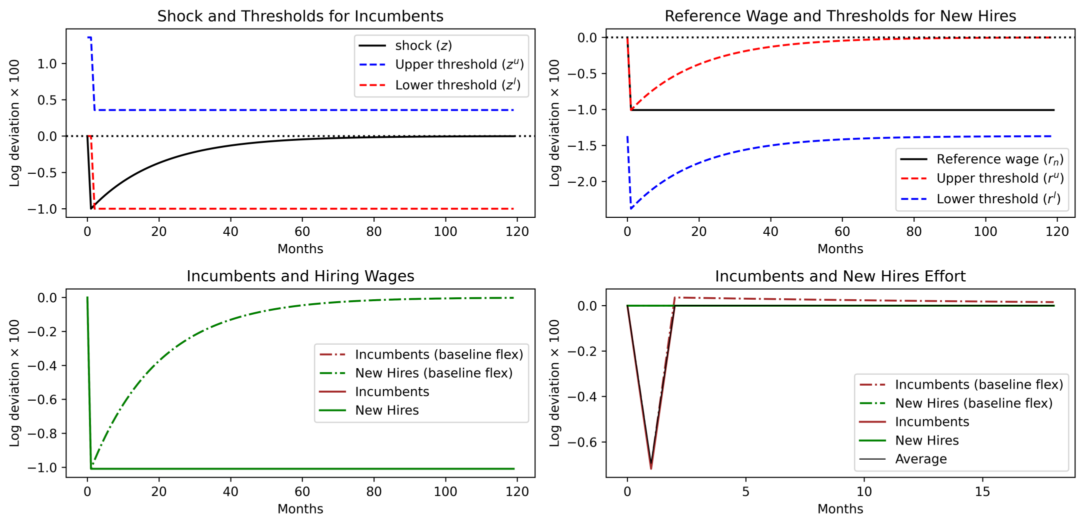
Block 2: Job Values, Labor Market and Output



Note: Response to a 1% positive, transitory shock starting from the low unemployment steady state.

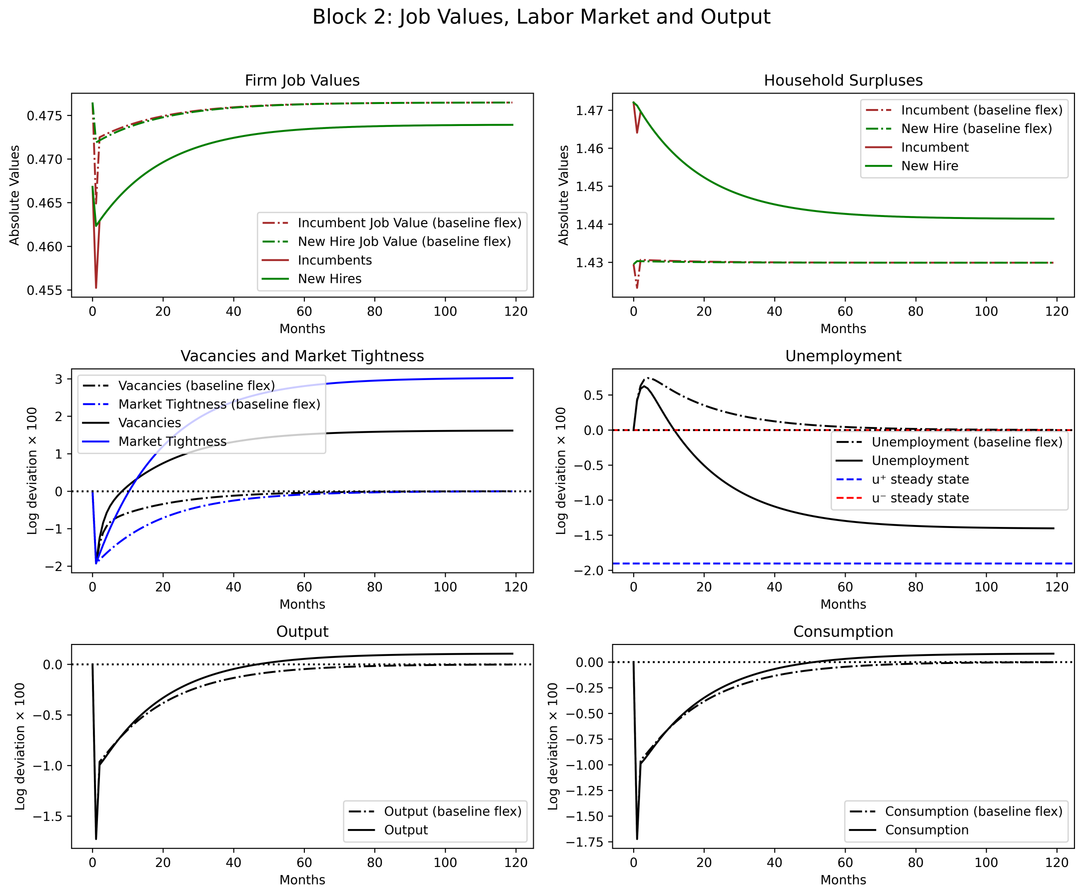
Figure C.3: Hysteresis effects: High steady state

Block 1: Shock, Wages and Effort



Note: Response to a 1% negative, transitory shock starting from the high unemployment steady state.

Figure C.4: Hysteresis effects: High steady state



Note: Response to a 1% negative, transitory shock starting from the high unemployment steady state.