Wage and Unemployment Dynamics under Incomplete Employment Contracts

Marco Fongoni and Alexandros Loukas

Aix-Marseille Université, CNRS, AMSE

March 2026

WORK IN PROGRESS

PLEASE DO NOT CITE OR CIRCULATE WITHOUT PERMISSION FROM THE AUTHORS

Abstract

We develop an analytic theory of downward nominal wage rigidity under incomplete employment contracts—worker effort is discretionary and not contractible. Our theory is grounded on well-known facts about firms' wage-setting behavior and can quantitatively replicate several empirical regularities of wage dynamics. Employed workers' aversion to nominal wage cuts induces firms to optimally freeze wages following moderately negative shocks, rendering nominal increases partially irreversible. We derive closed-form solutions to the value function and the resulting "Ss" wage policy of the firm, and formally establish both optimality and uniqueness. Our framework provides novel predictions about path dependence and hysteresis of wage adjustments, it can be quantitatively evaluated to perform counterfactual analyses, and it is portable to be integrated into existing workhorse macroeconomic models with or without heterogeneous agents. In ongoing work, we are exploring the general-equilibrium implications of the inherent asymmetry and irreversibility generated by our model for the dynamics of wages and unemployment.

Keywords: Downward nominal wage rigidity, state dependence, path dependence, irreversibility, incomplete contracts, work effort.

JEL codes: E24, E32, J23, J31, J41, J64.

I. Introduction

Downward nominal wage rigidity is considered a fundamental friction in macroeconomics, explaining cyclical fluctuations in unemployment (e.g. Schmitt-Grohé and Uribe, 2016; Dupraz, Nakamura, and Steinsson, 2025) and justifying the role of stabilization policy (e.g. Erceg, Henderson, and Levin, 2000; Christiano, Eichenbaum, and Evans, 2005; Benigno and Ricci, 2011). Accordingly, understanding its *sources* and *implications* has been a research priority for decades. Empirical explanations of downward nominal wage rigidity emphasize managers' concerns that wage cuts will damage morale and productivity; see, e.g., the seminal work of Bewley (1999) and the survey by Howitt (2002). As long as workers are able to respond to pay cuts by withholding effort (i.e. under incomplete employment contracts), downward adjustments become effectively costly.[1] While the existing macroeconomic literature frequently cites this evidence as motivation, the resulting models typically rely on complete contracts with exogenous adjustment frictions—either time-dependent or state-dependent—to generate wage rigidity and to study its aggregate implications.[2] As much as these approaches provide many valuable insights, a more complete and nuanced account of aggregate wage dynamics calls for a theory of wage-setting behavior that more closely captures the incomplete and relational nature of labor markets. In this paper, we fill this gap by developing a parsimonious analytical model of downward nominal wage rigidity under incomplete employment contracts that yields closed-form solutions and thus is readily portable to the analysis of aggregate fluctuations.

In our framework employment contracts are incomplete, meaning that workers' productivity (effort) is discretionary and not contractible. Firms have monopsony power over wage setting, but recognize that nominal wage cuts can be costly to implement as they trigger a disproportionate reduction in worker effort.[3] This asymmetric effort response is the fundamental friction in our model, inducing firms to optimally freeze wages following moderately negative shocks and to dampen wage cuts following large negative shocks, rendering nominal increases

[1] These explanations are based on robust findings from a large number of surveys to compensation managers (see, e.g., the pioneering work of Blinder and Choi (1990); Campbell and Kamlani (1997); Bewley (1999) and the more recent evidence in Du Caju et al. (2015); Bertheau et al. (2025)), and align well with recent advances in behavioral and experimental economics on fairness and reciprocity in labor relations (see the surveys by Bewley (2007) and Fehr, Goette, and Zehnder (2009)). Employees perceive wage cuts as particularly unfair, and the negative reciprocity in response to unfair treatment is typically more pronounced than the positive reciprocity response to a wage "gift". Employers that are aware of these patterns hesitate to cut their existing workers' wages (see Dickson and Fongoni (2019) for a micro-foundation of workers' asymmetric reciprocity based on reference-dependent preferences and loss aversion). In addition, Fongoni, Schaefer, and Singleton (2025) provide quantitative observational evidence that compensation managers' beliefs about contract incompleteness reduces the likelihood of nominal wage cuts to job stayers.

[2] For instance, in Benigno and Ricci (2011) and Dupraz, Nakamura, and Steinsson (2025) wage cuts are assumed to be infeasible; in the recent papers by Elsby et al. (2025) and Blanco and Drenik (2023), wage rigidity arises due to exogenous renegotiation costs; while in Gertler and Trigari (2009) and Blanco et al. (2025), wages are rigid due to staggered contracts à la Calvo (1983). Notable exceptions are Elsby (2009) and Fongoni (2024), in which downward wage rigidity arises endogenously as an optimal "inaction" policy of firms, anticipating their workers' disproportionate drop in effort. As we explain in detail below, we substantially extend this latter class of models and embed them into a general equilibrium framework.

[3] In the microeconomic literature, workers' aversion to nominal wage cuts has been linked to loss aversion, where current wages serve as the reference "fair" wage for evaluating new contracts (see, e.g., Dickson and Fongoni, 2019; Sliwka and Werner, 2017, and the references therein.).

partially irreversible. Within this framework, which builds on the work of Elsby (2009) and Fongoni (2024), we make three key contributions. First, we recast this class of models in continuous time and derive fully analytical (closed-form) solutions to the wage policy and value function of the firm, formally establishing both optimality and uniqueness of the resulting "Ss" policy. Second, we show how the model can be calibrated to replicate key empirical regularities of wage dynamics, while also providing novel predictions—such as path dependence and hysteresis in wage adjustments—thereby serving as a rigorous framework for quantitative counterfactual analyses.[4] Finally (work in progress), we exploit the portability of our framework to embed it into a benchmark general-equilibrium model with heterogeneous agents, in order to study the aggregate and distributional implications of its inherent asymmetry and irreversibility for wage and unemployment dynamics.

Our model is parsimonious in capturing the main rationale behind employers' tendency to avoid nominal wage cuts, enabling a transparent economic interpretation of the fundamental frictions that generate downward nominal wage rigidity and related implications. At the same time, it is also sufficiently rich to quantitatively replicate several empirical regularities of wage cyclicality and the distribution of wage changes in the data.[5] In fact, the workers' effort function introduces three important parameters, which underlie the net costs of wage adjustments to firms: the sensitivity of effort to wage changes, the relative sensitivity to wage cuts, and a curvature parameter. We show that these can be jointly calibrated to achieve the probability of no annual wage change, the average rent-sharing elasticity, and the average labor share of income as observed in the data. As such, the model generates an empirically plausible spike at zero together with a missing mass of positive wage changes (see, e.g., Grigsby, Hurst, and Yildirmaz, 2021; Stüber and Beissinger, 2012), as well as more dampened wage decreases (see, e.g., Holden and Wulfsberg, 2009, 2014). A key advantage of our micro-founded approach—relative to existing, reduced-form models of nominal wage adjustments—is that it provides a clear and meaningful mapping between the magnitude of frictions required to replicate empirical wage dynamics and other, directly observable labor market characteristics, such as rent sharing and the labor share. Moreover, we are able to precisely disentangle the contribution

[4]The notions of path dependence and hysteresis are inherently tied in the structure proposed in our study. Path dependence means that the current value of the exogenous state is not enough to characterize the position of the system in a given point in time. Put simply, whether a lower realization of the state results in a wage cut or a wage freeze will depend on its value relative to the state in the previous period via its relation to the current wage level. This notion is also in line with how path dependence is defined in the literature on investment dynamics under capital irreversibility; see e.g. Baley and Blanco (2026). Specifically, wage increases beget future wage increases and wage cuts beget future wage cuts. As will become clearer below, our framework generates hysteresis in wage dynamics: transitory shocks can leave wages permanently higher; see Quach (2025) for recent evidence of wage hysteresis.

[5]The empirical literature on downward nominal wage rigidity has mainly focused on job stayers. For up-to-date surveys of this extensive literature see Elsby and Solon (2019) and ?Appendix A.1]Jo2025. To avoid concerns about measurement error in studies using household survey data (see, e.g., Barattieri, Basu, and Gottschalk, 2014, and references therein), we take as the main reference the empirical regularities reported in Grigsby, Hurst, and Yildirmaz (2021), who use comprehensive administrative payroll data covering millions of US workers. Grigsby, Hurst, and Yildirmaz (2021), Gertler, Huckfeldt, and Trigari (2020) and Hazell and Taska (2025) show that these regularities, namely the asymmetry in wage adjustments and their relatively low cyclicality, are also features of the wage paid to new hires—i.e. the hiring wage.

of incomplete employment contracts to the probability of no annual wage change, relative to a standard reservation wage constraint—a decomposition that has eluded both the empirical and theoretical literature.

An important innovation of our model is the recasting, in continuous time, of the discrete-time models of Elsby (2009) and Fongoni (2024). This approach enables us to employ recent advances in mathematical methods of stochastic impulse control problems involving time-varying free boundaries to: i) provide fully analytical (closed form) solutions to both the wage policy and value function of the firm; and ii) formally establish both optimality and uniqueness of the firm's Ss policy. This formulation expresses the firm's problem as a choice of a sequence of intervention times—*when* to adjust the wage—and a sequence of optimal impulse controls—by *how much* to adjust the wage. The time-varying free boundaries capture the shock thresholds governing the firm's inaction region (range of rigidity) over time. These results place our contribution beyond the literature on wage dynamics. In fact, this class of wage-setting models—which share notable similarities with the literature on "kinked adjustment costs" and optimal Ss policies (e.g., Bertola and Caballero, 1990)[6]—present technical challenges that have perhaps reduced their tractability and application to more standard macroeconomic frameworks, despite their solid microfoundations and empirical appeal.[7] In contrast, our approach yields a model that is highly tractable and portable to other macroeconomic environments, while also offering substantial gains in computational efficiency.

In ongoing work, we are embedding our framework into an otherwise standard general equilibrium model of the economy with search frictions and capital accumulation (as in, for instance, Gertler and Trigari, 2009). Our framework will feature: random search, heterogeneous job matches, endogenous job creation and destruction by firms, endogenous acceptance and quitting of jobs by workers, and investment in capital. By performing state-of-the-art quantitative analyses, our approach allows for the exploration of counterfactual scenarios by varying the strength of workers' discretionary effort responses. This enables us to disentangle the effects of contract incompleteness and costly wage cuts on hiring, layoffs, and wage and unemployment dynamics. The resulting analysis will provide a substantial contribution to the literature on labor market dynamics, by exploring both the micro and macroeconomic implications of path dependent downward nominal wage rigidity under incomplete employment contracts. As such, our approach provides a complementary view to existing theories of labor market dynamics that are grounded on complete employment contracts—i.e. where workers' productivity is both observable and contractible and wages are determined by bargaining (for recent discussions of this literature, see Elsby et al. (2025) and Blanco et al. (2025)).

[6]See also Caballero and Engel (1999) and Elsby, Michaels, and Ratner (2019) for applications on the dynamics of investment and employment respectively. We discuss the relation of our model to this literature more in detail below.

[7]Namely, these models require shocks to follow a geometric random walk to obtain closed-form solutions for the optimal policy, and numerical methods to solve for the threshold functions governing the relevant range of inaction, i.e. the range of wage rigidity (see Elsby, 2009; Fongoni, 2024). Moreover, optimality of the resulting Ss policy is typically hard to establish, while a formal proof of its uniqueness has to date been elusive (see, for instance Caballero and Engel, 1999, footnote 14, p. 791).

The remainder of the paper is organized as follows. After a discussion of the related literature below, Section II presents our baseline model of optimal wage setting under incomplete employment contracts, while III explores its qualitative and quantitative properties for wage dynamics. Section IV (work in progress) embeds this framework into an otherwise standard general equilibrium model. Section V provides some preliminary concluding remarks. All proofs and detailed derivations are contained in Appendix A.

Related literature. This paper relates and contributes to three main strands of literature. Our model is most closely related to the discrete-time wage-setting model developed by Elsby (2009) and further extended by Fongoni (2024). Elsby (2009) studies a partial-equilibrium model of firm wage-setting behavior where worker effort is endogenous to the wage such that wage cuts are particularly costly to implement. The model generates downward nominal wage rigidity and is used to show that a forward-looking firm will compress both wage cuts and wage increases, implying that downward nominal wage rigidity has no effect on aggregate wage growth in the long run. Fongoni (2024) builds on the asymmetric reference-dependent reciprocity model developed by Dickson and Fongoni (2019), and extends this framework by considering the role of both new hires and incumbent workers in an otherwise standard search and matching framework à la Pissarides (1985). The model aligns with evidence on hiring and incumbents' wage cyclicality, and formally demonstrates that when effort is endogenous, hiring wage cyclicality has no first-order effect on job creation and is therefore irrelevant for unemployment volatility.[8] Both papers study representative one-worker-one-firm employment relationships which are assumed to be mutually advantageous, therefore abstracting from endogenous layoff decisions by firms or participation constraints by workers. As explained above, the approach we take in this paper substantially extends this class of models by formally establishing optimality and uniqueness of the optimal wage policy of the firm, as well as deriving fully analytical (closed-form) solutions to both the wage policy and value function of the firm. These results pave the way to build a richer model including heterogeneous job-match productivity, and workers' and firms' participation decisions, in a general equilibrium framework.

Our paper is also complementary to the recent theoretical literature considering the role of frictions impeding wage adjustments—under *complete* employment contracts—on the dynamics of labor markets. Elsby et al. (2025) develop a labor market model where wage adjustments occur sporadically due to the presence of renegotiation costs and show that this is consistent with job separations being bilaterally efficient.[9] Blanco et al. (2025) study an equilibrium labor market model with staggered wage contracts à la Calvo (1983) and show that the presence of wage rigidity can lead to endogenous, inefficient layoffs and quits. These papers provide

[8] In a recent paper, Fongoni (2025) further explores the implications of this framework when new hires are concerned about pay inequality within the firm, showing that the downward rigidity in the wage of job stayers spills over to the hiring wage with fundamental implications for the cyclical behavior of unemployment. However, for analytical tractability, the model is solved under perfect foresight where the only sources of disturbance are unanticipated, temporary, and persistent shocks to the firm's value of output.

[9] See also Blanco and Drenik (2023) for a model of wage rigidity under Nash bargaining due to renegotiation costs.

many valuable insights on the *consequences* of frictions to wage adjustments for the dynamics of labor markets, but remain agnostic about the *sources* of these frictions. Our contribution to this literature is to i) develop a general equilibrium model of the labor market under incomplete employment contracts; and ii) explore the implications of one of the main sources of downward nominal wage rigidity as established in the empirical literature, that is, the role of workers' discretionary effort and aversion to nominal wage cuts.

Finally, conceptually and structurally, our analysis shares a theoretical lineage with the literature on (partially) irreversible investment under uncertainty.[10] This body of work typically attributes infrequent factor adjustments to two sets of assumptions: (i) The presence of non-convex or "kinked" adjustment cost technologies (e.g., Bertola and Caballero, 1990; Bentolila and Bertola, 1990; Caballero, Engel, and Haltiwanger, 1997); (ii) The presence of a wedge between the acquisition and resale price of inputs (e.g., Abel and Eberly, 1994, 1996; Abel et al., 1996; Baley and Blanco, 2026).[11] Our framework can be viewed as an incomplete-contract labor market isomorphism of these classic investment mechanisms. In a nutshell, the interaction between workers' reference dependence, aversion to wage cuts, and stochastic productivity generates sluggish wage dynamics—regions of inaction followed by asymmetric responses—effectively treating wage adjustments as partially irreversible investment. In our setting the firm inherits a "stock" of wage entitlements embodied in the current reference wage and chooses wage adjustments that determine the value of that stock tomorrow. Hence the reference wage functions as a stock variable—analogous to physical capital—while the wage deviation from that reference point functions as a flow variable—analogous to positive or negative investment. The asymmetry in workers' reference-dependent effort is the "friction" that induces partial irreversibility and forces the firm to treat wage setting as a long-term commitment rather than a spot adjustment to productivity. Absent this friction, we obtain the case of full reversibility where the firm adjusts wages continuously in line with productivity shocks. Equivalently, this asymmetry creates a discrepancy between the cost of creating worker effort (wage hikes) and the value recovered from dismantling it (wage cuts). Just as a wedge between the purchase and resale price of capital generates inertia in investment, the wedge between worker gratitude and resentment generates inertia in wages. Therefore the firm is compelled to buffer productivity fluctuations in both directions: it avoids wage cuts to escape the immediate penalty of loss aversion (downward rigidity), while simultaneously delaying wage hikes to minimize the risk of future lock-in (upward rigidity).

[10] See Dixit and Pindyck (1994) and Stokey (2009) for systematic treatments of using option pricing and stochastic control methods to characterize optimal behavior in these settings.

[11] Many studies additionally assume fixed or sunk costs which, notably, we abstain from in order to keep things as transparent as possible. Incorporating them is straightforward and would pose no complications in our setting.

II. A Model of Wage Setting under Incomplete Contracts

We develop a parsimonious partial-equilibrium model of wage setting within the context of a firm–worker employment relationship. The employment contract is incomplete, i.e., workers' effort is discretionary and non-contractible. The output/revenue of firms is determined by workers' endogenous effort, and firms set nominal wages unilaterally in every period. Workers are reference dependent, i.e., their effort depends on the evaluation of their current wage contract relative to an endogenous reference point that updates only when the firms resets pay. In addition, each firm-worker match productivity evolves stochastically over time, which generates a non-trivial intertemporal tradeoff between current profits and the future cost (benefit) of raising (lowering) the worker's reference wage. This structure turns the firm's problem into an analytic stochastic impulse-control problem with free boundaries, which we fully characterize and solve in closed form for the value and policy functions.

II.1. Environment

The model is set in continuous time $t \geq 0$ under an infinite horizon and common knowledge of all primitives. We assume a constant inflation rate $\pi \in \mathbb{R}$; hence the price level at time t is governed by $dP_t = \pi P_t dt$ for some initial condition $P_0 > 0$. We consider an employment relationship between a price-taking firm and a worker. Nominal revenue is generated by a linear technology $Y_t = Z_t f_t$, where $Z_t \in \mathbb{R}_{++}$ is an idiosyncratic "match productivity" and $f_t = f(W_t, R_t) \in \mathbb{R}_+$ denotes the effort exerted by the worker. Effort depends on the nominal wage $W_t \in \mathcal{W} \subset \mathbb{R}_+$ in relation to a reference wage $R_t \in \mathcal{R} \subset \mathbb{R}_+$, which is taken as given by both the worker and the firm. The properties of f_t, as well as the evolution of R_t, are crucial features of the model and are spelled out in detail in the next section. The employment contract is incomplete, meaning that effort is discretionary and not contractible, and the wage is set unilaterally by the firm. The state vector (R_t, Z_t) takes values in $\mathcal{X} = \mathcal{R} \times \mathcal{Z}$.

We work on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions of right-continuity and completeness. The firm's control variables are the nominal wage W_t, chosen from the space of square-integrable $\mathcal{F}_t$-adapted processes, as well as a sequence of intervention times $\tau_0 < \tau_1 < \ldots \widetilde{\tau}$, together with a layoff/termination time $\widetilde{\tau}$.

Given an initial $z_0 \triangleq Z_0/P_0$, real idiosyncratic match productivity is subject to exogenous shocks and evolves according to a geometric Brownian motion,

$$dz_t = \mu z_t dt + \sigma z_t dB_t, \tag{1}$$

with drift $\mu \in \mathbb{R}$, volatility $\sigma \in \mathbb{R}_+$, and B_t denoting a standard Wiener process (Brownian motion). The stochastic nature of z_t is the sole source of uncertainty for the firm. Accordingly, revenue per unit of effort Z_t follows a geometric Brownian motion with the same diffusion and an inflation-augmented drift,

$$dZ_t = (\mu + \pi) Z_t dt + \sigma Z_t dB_t. \tag{2}$$

The process $\{Z_t\}$ encapsulates all exogenous determinants of the match's nominal revenue capacity per unit of effort, conflating idiosyncratic real factors with aggregate nominal trends. The real drift parameter μ captures the expected "life-cycle" profile of the match's productivity, which may be interpreted as the rate of job-specific human capital accumulation or returns to tenure as the worker gains experience and organizational knowledge. The volatility parameter σ captures idiosyncratic shocks to the firm's specific demand or production technology. The composite drift $(\mu + \pi)$ thus governs the expected growth of the firm's nominal ability to pay, driven by the internal accumulation of real skills and the external rise in the general price level.

II.2. Worker

The behavior of the worker is characterized by three elements: a law of motion for the evolution of their reference wage; an effort function capturing the worker's discretionary effort choice in reduced form; and a reservation wage governing their decision to quit the employment relationship.

Reference wage formation. A central premise of our theory is that workers are reference dependent. Here we assume that a worker's reference wage is anchored to the wage level of their last employment contract, reflecting a form of reference wage adaption.[12] Specifically, let $\{\tau_i\}_{i=1}^{\infty}$ be a strictly increasing sequence of *intervention times* at which the firm sets a wage W_{τ_i} *distinct* from the prevailing reference wage, and let $\{W_{\tau_i}\}_{i=1}^{\infty}$ be the corresponding sequence of chosen wages. Given an initial reference point R_s, the reference wage R_t evolves according to:

$$R_t = \begin{cases} R_s, & \text{for } t \in [s, \tau_1] \\ W_{\tau_i}, & \text{for } t \in (\tau_i, \tau_{i+1}],\ \forall i = 1, 2, \dots \end{cases} \tag{3}$$

The law of motion (3) implies the reference wage may evolve endogenously as a *discrete jump process* rather than a continuous one. This has important implications for our subsequent solution method and for optimal wage setting. Note that at each intervention time τ_i, by definition, $W_{\tau_i} \neq R_{\tau_i}$, since R_t is fixed at the moment of adjustment and updates only an instant dt later. Between intervention times the reference wage remains constant and equal to the wage set at the last adjustment. Moreover, the integral representation of (3) clarifies that R_t is essentially a (cumulative) stock variable, evolving through discrete "investments" equal to the difference between the prevailing wage and the existing reference point.[13]

[12] This assumption is consistent with a substantial body of evidence documented in the literature on adaptation in social psychology (see, e.g., Baucells and Sarin, 2010; Baucells, Weber, and Welfens, 2011; Baucells and Sarin, 2013), reference wage formation in behavioral economics (see, e.g., Bewley, 2007; Kahneman, Knetsch, and Thaler, 1986; Kahneman and Thaler, 1991; Sliwka and Werner, 2017), and contracts as reference points in the literature on incomplete contracts (see, e.g., Bartling and Schmidt, 2015; Fehr, Hart, and Zehnder, 2011; Herz and Taubinsky, 2013).

[13] In particular, letting dN_t denote the differential of the counting process associated with intervention times,

Discretionary effort. The worker's effort is characterized by an *asymmetric reference-dependent effort function* $f : \mathcal{W} \times \mathcal{R} \to \mathbb{R}$. We assume the following functional form:

$$f(W_t, R_t; \lambda) = \bar{e} + \frac{\eta}{1-\gamma} \begin{cases} (W_t^{1-\gamma} - R_t^{1-\gamma}), & \text{if } W_t \geq R_t, \\ \lambda(W_t^{1-\gamma} - R_t^{1-\gamma}), & \text{if } W_t < R_t, \end{cases} \tag{4}$$

where $\bar{e} > 0$ denotes "normal effort"; $\eta > 0$ parametrizes the sensitivity of effort to wage changes relative to the reference wage; $\gamma \in (0, 1)$ governs the concavity of the effort function with respect to the wage; and $\lambda \geq 1$ determines the relative strength of effort to wage changes below the reference wage.[14] Combined with our assumption on the law of motion of the reference wage, the effort function (4) postulates a positive wage-effort relationship such that worker effort only responds to wage changes—otherwise constant and equal to $\bar{e}$—and that if $\lambda > 1$, these responses are asymmetric. The decrease in effort below $\bar{e}$ resulting from a wage cut will be greater than its increase above $\bar{e}$ resulting from an equivalent-sized wage increase. These properties reflect the well-established facts that, under incomplete employment contracts, workers' aversion to wage cuts lead to disproportionally low morale and labor productivity, making wage cuts costly to implement; see, for instance, Bewley (1999) and the survey in Bewley (2007).

Reservation wage. Denote by $\mathcal{E}(R_t, W_t)$ the value of the employment relationship to a worker with reference wage R_t who is paid the wage W_t, and by $\mathcal{U}$ their reservation value, i.e., the value of unemployment or of being employed elsewhere. The worker will prefer to remain employed as long as $\mathcal{E}(R_t, W_t) \geq \mathcal{U}$. As such, we define the worker's reservation wage as

$$\underline{W}(R_t) \triangleq \{W_t : \mathcal{E}(R_t, W_t) = \mathcal{U}\} \tag{5}$$

II.3. Firm

The firm is risk neutral and discounts the future at a constant rate $\rho > 0$. In addition to endogenous separation—layoff or quit—the employment relationship can also terminate exogenously at the Poisson rate $\delta > 0$. With constant inflation, the effective discount rate for valuing nominal flows is thus $\rho + \delta + \pi$.

R_t has the equivalent integral form:

$$R_t = R_s + \sum_{s<\tau_i<t} (W_{\tau_i} - R_{\tau_i}) = R_s + \int_s^{t^-} (W_u - R_u) dN_u.$$

[14] The functional form (4) is from Fongoni (2024). Other labor market models that study the implications of a kinked, reduced-form effort function in the spirit of the one we consider here are those of Elsby (2009), Eliaz and Spiegler (2014), and Kaur (2019). Dickson and Fongoni (2019) show how an asymmetric reference-dependent effort function such as the one in (4) can be derived more generally from a worker's optimal choice; see also Sliwka and Werner (2017) for an alternative approach.

Wage setting problem. The nominal flow profit of the firm at time t is given by

$$\Pi(W, R, Z) = Z_t f(W_t, R_t; \lambda) - W_t. \tag{6}$$

At each given time s, the firm seeks to maximize the expected present-discounted value of its flow profits by choosing an admissible sequence of wages:

$$\mathcal{J}(R, Z) = \sup_{W \in \mathcal{W}} \mathbb{E}\left[\int_s^\infty e^{-(\rho+\delta+\pi)(t-s)} \Pi(W, R, Z) dt \,\middle|\, \mathcal{F}_s \right], \tag{7}$$

where the expectation is taken conditional on the information set at time s, given some initial state vector (R_s, Z_s). To ensure that the firm's optimization problem is well-posed and that its expected present value remains finite as time goes to infinity, we henceforth assume that the following parameter condition holds:

$$\rho + \delta + \pi > \frac{\mu + \pi}{\gamma} + \frac{1}{2} \frac{\sigma^2(1-\gamma)}{\gamma^2}. \tag{8}$$

This is a typical "no-bubble" condition requiring the firm's effective discount rate to be sufficiently high to dominate the amplified growth rate of its own value, which is driven by both stochastic productivity trends and the option value of volatility. As we will see later on, this is indeed the appropriate condition to ensure that the transversality condition is satisfied under any realized productivity path. We also need to consider appropriate boundary conditions to render the problem well defined, which we do below.

Reservation productivity. In addition to the exogenous separation rate, a match can terminate whenever the worker decides to quit, or the firm decides to lay off the worker. From the firm side, this occurs exactly when the value of the firm is equal to its outside option, which is normalized to zero. We define the productivity cutoff below which the firm will optimally decide to lay off the worker as

$$\underline{Z}^{layoff}(r_t) \triangleq \{Z_t : \mathcal{J}(R_t, Z_t) = 0\}. \tag{9}$$

At the same time, the firm knows that any wage below the reservation wage $\underline{W}(R_t)$ will result in the worker quitting. Hence for any given wage $\widehat{W}(W_t, Z_t)$ set by the firm, we can define the productivity cutoff below which the worker will quit as

$$\underline{Z}^{quit}(r_t) \triangleq \{Z_t : \widehat{W}(R_t, Z_t) = \underline{W}(R_t)\}. \tag{10}$$

The productivity threshold at which the firm-worker match is *de facto* terminated is then

$$\underline{Z}(r_t) \triangleq \max\{\underline{Z}^{quit}(R_t), \underline{Z}^{layoff}(R_t)\}. \tag{11}$$

II.4. Recursive formulation

One may be tempted to proceed by setting up a Hamilton-Jacobi-Bellman (HJB) equation over the entire state space, effectively treating the above as a typical stochastic optimal control problem. But in fact, it is not. The reason is that the law of motion for the reference wage is not a continuous process governed by a continuous control, but rather a jump process driven by discrete impulses at optimally chosen times. To show this more formally, let us first define the *impulse payoff* function $G : \mathcal{W} \times \mathcal{X} \to \mathbb{R}$,

$$G(W, R, Z) \triangleq \Pi(W, R, Z) - \Pi(R, R, Z) = Z\left(f(W, R; \lambda) - \bar{e}\right) - (W - R). \tag{12}$$

The function $G(.)$ quantifies the net payoff of an intervention, i.e., the instantaneous change in firm's flow profit elicited by setting a wage W_t distinct from the reference R_t. Given a realization of Z_t, this payoff amounts to the revenue gain/loss driven by worker's reciprocity $(Z_t\left(f(W_t, R_t; \lambda) - \bar{e}\right)$, minus the corresponding increase/decrease in the wage bill $(W_t - R_t)$. We may then rewrite the value function (7) as

$$\mathcal{J}(R, Z) = \sup_{\{\tau_i, W_{\tau_i}\}_{i\geq 1}} \mathbb{E}\left[\int_s^\infty e^{-(\rho+\delta+\pi)(t-s)}\Pi(R_t, R_t, Z_t)dt + \sum_{i=1}^{\infty} e^{-(\rho+\delta+\pi)(\tau_i - s)}G(W_{\tau_i}, R_{\tau_i}, Z_{\tau_i}) \,\middle|\, \mathcal{F}_s\right].$$

Now the supremum is taken over all *intervention times* $\tau_i \in [0, \infty]$ with respect to the filtration $\mathcal{F}_t$, as well as over admissible wage policies chosen at τ_i. This expression shows how the value function is properly defined by optimizing over the timing and magnitude of wage adjustments, incorporating the constant part of the flow profit between intervention times and the instantaneous profit difference $G(.)$ after each intervention.

The next step is crucial. It is useful to introduce the non-local *intervention operator* $\mathcal{M}$:

$$\mathcal{M}\mathcal{J}(R, Z) \triangleq \sup_{W\in\mathcal{W}} \{G(W, R, Z) + \mathcal{J}(W, Z)\}. \tag{13}$$

Intuitively, $\mathcal{M}\mathcal{J}(R, Z)$ represents the maximum value the firm can attain if it chooses to intervene at the current state (R, Z). This value consists of two components: the immediate impulse payoff $G(W, R, Z)$ from a wage adjustment, and the continuation value $J(W, Z)$ encompassing the firm's future prospects starting from the new reference state $R = W$. By maximizing over W, the operator characterizes the optimal wage policy that balances the immediate cost/benefit of an adjustment against the future impact of a lower/higher reference wage.

Equipped with this formulation along with the recursive structure of the problem, we can invoke the dynamic programming principle to express the firm's value function as

$$\mathcal{J}(R, Z) = \sup_{\tau\geq 0} \mathbb{E}\left[\int_s^\tau e^{-(\rho+\delta+\pi)(t-s)}\Pi(R, R, Z)dt + e^{-(\rho+\delta+\pi)(\tau-s)}\mathcal{M}\mathcal{J}(R, Z) \,\middle|\, \mathcal{F}_s\right]. \tag{14}$$

It is now even more clear that the firm will attain its maximum value by recursively choosing

a sequence of $\mathcal{F}_t$-adapted optimal intervention times—*when* to adjust the wage—as well as a sequence of optimal impulse controls—by *how much* to adjust the wage. This is why the above formulation is instructive to illustrate the nature of the problem, namely, a *stochastic impulse control problem* involving *free boundaries*.

To be more precise, the value function $\mathcal{J}$ must satisfy the following conditions for all (R, Z): (i) The value of being in any state is at least as high as the value attained by optimally intervening; i.e., $\mathcal{J} \geq \mathcal{M}\mathcal{J}$. (ii) $\mathcal{J}$ satisfies a HJB inequality over the entire state space; i.e., $(\rho + \delta + \pi)\mathcal{J}(R, Z) - (\mu + \pi)Z\mathcal{J}_Z(R, Z) - \frac{1}{2}\sigma^2 Z^2 \mathcal{J}_{ZZ}(R, Z) - \Pi(R, R, Z) \geq 0$. (iii) Importantly, optimality requires that the two previous conditions must hold with complementarity, hence at least one of the inequalities in (i) or (ii) must hold with equality.

It turns out that equation (14) is the integral form of a *quasi-variational inequality* that the value function has to satisfy for all t. The classical reference for both theory and a range of applications of quasi-variational inequalities is the foundational work of Bensoussan and Lions (1984); see also Bardi and Capuzzo-Dolcetta (1997), Pham (2009), and Øksendal and Sulem (2019) for more recent textbook treatments of related problems.

Denote the infinitesimal generator of the nominal productivity process by $\mathcal{L}\mathcal{J} \triangleq (\mu + \pi)Z\mathcal{J}_Z + \frac{1}{2}\sigma^2 Z^2 \mathcal{J}_{ZZ}$. The firm's value function must satisfy the following *Hamilton-Jacobi-Bellman quasi-variational inequality (HJBQVI)* along with two boundary conditions:

$$\min\Big\{(\rho + \delta + \pi)\mathcal{J}(R, Z) - \mathcal{L}J(R, Z) - \Pi(R, R, Z),\ \mathcal{J}(R, Z) - \mathcal{M}\mathcal{J}(R, Z)\Big\} = 0, \quad (15)$$

$$\lim_{T\to\infty} \mathbb{E}\left[e^{-(\rho+\delta+\pi)(T-s)}\mathcal{J}(R_{\tau\wedge T}, Z_{\tau\wedge T}) \,|\, \mathcal{F}_s\right] = 0, \quad (16)$$

$$\text{for all } R_t < \underline{W}(R_t),\ \ \mathcal{J}(R, Z) = 0. \quad (17)$$

The first term of the HJBQVI (15) can be interpreted as the "cost of inaction" at state (R, Z), and the second term represents the "cost of intervention" at the same state (R, Z). The HJBQVI succintly formalizes the fact that, at the optimum, both costs are non-negative and at least one must be zero. In the inaction region the cost of inaction is zero while the cost of adjustment is positive, whereas in the intervention region the cost of intervention is zero while the cost of inaction is positive.

Equation (16) is a standard transversality condition (TVC): a minimally sufficient boundary condition that the value function has to satisfy at the optimum, as well as a tool to identify it among potential solutions. The TVC requires that there is no residual discounted value at infinity along non-terminated paths. Technically, the TVC is required in our verification theorem in order to remove the terminal term after applying Itô's lemma and optional sampling to the process $e^{-(\rho+\delta+\pi)(t-s)}\mathcal{J}(R, Z)$ on $[s, \tau \wedge T]$. The TVC holds trivially on paths where $\tau < \infty$ since $\mathcal{J}(R, Z) = 0$ at and after termination. When $\Pr(\tau = \infty) > 0$, it amounts to the mild requirement that the discounted $\mathcal{J}(R_T, Z_T)$ vanishes in expectation as $T \to \infty$, which in our case is ensured under growth conditions granted by the parameter condition (8).

Finally, equation (17) is a simple boundary condition stating that if the firm offers a wage

strictly lower than the worker's reservation wage $\underline{W}(R_t)$, the worker quits and the match is dissolved (its value is normalized to zero).

II.5. Solution

Inaction region and wage rigidity. Building on the structure of the HJBQVI, we now dissect the optimal intervention induced by the $\mathcal{M}(.)$ operator. To proceed formally we introduce the set $\mathcal{I}$, which we call the *inaction region*, $\mathcal{I} = \{(R, Z) \in \mathcal{X} : \mathcal{J}(R, Z) > \mathcal{M}\mathcal{J}(R, Z)\}$, as well as its complement set $\mathcal{A}$, which we call the *action region*, $\mathcal{A} = \{(R, Z) \in \mathcal{X} : \mathcal{J}(R, Z) = \mathcal{M}\mathcal{J}(R, Z)\}$.[15] The sets above partition the state space accordingly: in the inaction region it is optimal for the firm to maintain the current wage contract, while in the action region it is optimal to adjust the wage immediately. This partition highlights the possibility of a non-trivial inaction region, but under what circumstances is it optimal for the firm to pursue inaction? The following proposition provides the condition under which wage rigidity arises as optimal policy.

PROPOSITION 1: *There exists a non-trivial inaction region, i.e., $\mathcal{I}$ is a non-empty open set with positive Lebesgue measure, if and only if $\lambda > 1$.*

Proposition 1 establishes that asymmetric reciprocity is necessary and sufficient for the existence of a non-trivial inaction region. As this is the condition that generates the empirically relevant case of nominal downward wage rigidity along with interesting wage dynamics, we henceforth retain the assumption $\lambda > 1$.

Optimal wage policy. The optimal policy is determined by maximizing the value of the intervention operator defined in (13). Conditional on intervening, the firm chooses a new wage to optimize the sum of the immediate impulse payoff and the continuation value. The necessary first-order condition (FOC) that characterizes the optimal wage policy in the action region $(\widetilde{W}_t \neq R_t)$ thus corresponds to

$$\underbrace{Z_t f_W(W, R; \lambda)}_{\substack{\text{current} \\ \text{marginal benefit}}} - \underbrace{1}_{\substack{\text{current} \\ \text{marginal cost}}} + \underbrace{\mathcal{J}_R(W, Z)}_{\substack{\text{future} \\ \text{marginal cost}}} = 0, \text{ for } W \neq R, \tag{18}$$

where $f_W(W, R; \lambda) \triangleq \frac{\partial f(W,R;\lambda)}{\partial W}$, and $\mathcal{J}_R(W, Z) \triangleq \frac{\partial J(R,Z)}{\partial R}\big|_{W=R}$ represents the shadow price of the current reference wage R evaluated at the new reference wage W, thereby capturing the intertemporal marginal cost of the firm's action.[16] The first-order condition (18) for an interior

[15] It is perhaps more common to encounter the equivalent terms "continuation" region and "intervention" region in the mathematics or applied probability literature. Nevertheless, we choose the terms "inaction" and "action" as they are more frequently used in economics, see e.g., Stokey (2009).

[16] Note that $J_R(W, Z)$ does not hinge on a separate envelope condition, but is simply the result of applying the chain rule: $\frac{d}{dW}\mathcal{J}(W, R) = \frac{\partial \mathcal{J}(R,Z)}{\partial R}\big|_{W=R}\frac{dW}{dW} + \frac{\partial \mathcal{J}(R,Z)}{\partial Z}\big|_{W=R}\frac{dZ}{dW} = \mathcal{J}_R(W, Z)$, since $\frac{dZ}{dW} = 0$.

optimum thus translates to

$$\widetilde{W}(R,Z) = \left(\frac{Z\eta}{1-\mathcal{J}_R(\widetilde{W},Z)}\right)^{1/\gamma}, \quad \text{for } \widetilde{W} > R, \tag{19}$$

$$\widetilde{W}(R,Z) = \left(\frac{Z\lambda\eta}{1-\mathcal{J}_R(\widetilde{W},Z)}\right)^{1/\gamma}, \quad \text{for } \widetilde{W} < R. \tag{20}$$

These equations characterize the optimal wage as an implicit function of the state, and they are economically meaningful only if they define a wage policy within the set of admissible actions. Clearly, the term $1 - J_R(\widetilde{W}, Z)$ must be strictly positive to ensure the optimal wage is both real and strictly positive. The following lemma proves that this is indeed the case and that the firm's value function is strictly decreasing in the worker's current reference wage.

LEMMA 1: *The firm's value function in the inaction region is strictly decreasing in the reference wage, i.e.,* $\mathcal{J}_R(W,Z) < 0, \forall (R,Z) \in \mathcal{X}$. *Therefore, any interior optimal impulse control results in a strictly positive wage schedule,* $\widetilde{W}(R,Z) > 0$.

Within the inaction region the firm keeps the nominal wage at the prevailing reference wage and the value function solves the corresponding HJB equation, which is a linear second-order PDE in (R, Z) and thus remains highly tractable. The next proposition provides the general solution in closed form. This result is critical because it serves as the core building block for the candidate value function used in the verification theorem, which is essential for obtaining the optimal value and policy functions.

PROPOSITION 2: *The value function in the inaction region* $(\widetilde{W}_t = R_t)$ *admits the general solution*

$$\mathcal{J}(R,Z) = \frac{Z\bar{e}}{\rho+\delta-\mu} - \frac{R}{\rho+\delta+\pi} + \Omega_1 Z^{\beta_1} R^{1-\gamma\beta_1} + \Omega_2 Z^{\beta_2} R^{1-\gamma\beta_2}, \tag{21}$$

where Ω_1, Ω_2 *are real constants to be determined, and* $\beta_1 > 1, \beta_2 < 0$ *are the roots of the characteristic quadratic equation* $\frac{1}{2}\sigma^2\beta(\beta-1) + (\mu+\pi)\beta - (\rho+\delta+\pi) = 0$ *given by*

$$\beta_{1,2} = \frac{-(\mu+\pi-\frac{1}{2}\sigma^2) \pm \sqrt{(\mu+\pi-\frac{1}{2}\sigma^2)^2 + 2\sigma^2(\rho+\delta+\pi)}}{\sigma^2} \tag{22}$$

Equation (21) reveals that the inaction-region value consists of two fundamental parts. The first two terms (the particular solution) capture the present value of flow profits when the wage is held fixed at R_t. The two remaining terms (the homogeneous solution) capture the option value created by diffusive productivity and the possibility of future wage adjustments. As we show right below, the constants Ω_1 and Ω_2 are pinned down in closed form by the FOCs for the optimal wage, and their values are such that the smooth-pasting conditions at the free

boundaries are always satisfied for any R_t.

Now that $\mathcal{J}(R,Z)$ in the inaction region is known, we can calculate $\mathcal{J}_R(\widetilde{W},Z)$ showing up in the denominator of (19)–(20) and characterize the wage setting decision of the firm. In fact, we derive the optimal wage policy in closed form together with the productivity thresholds (free boundaries) that separate the upward/downward action regions from the inaction region. This is achieved analytically because the representation of $\mathcal{J}(R,Z)$ embeds a useful scaling invariance: the marginal value $\mathcal{J}_R(\widetilde{W},Z)$ is homogeneous of degree zero in $(\widetilde{W}^\gamma,Z)$; see proof of Proposition 2 for details. We summarize these findings below.

PROPOSITION 3: *If $Z_t > \underline{Z}(R_t)$, the interior (non-layoff) optimal wage policy for every $t \geq s$ must take the form*

$$\widetilde{W}(R_t,Z_t) = \max\left\{\widehat{W}(R_t,Z_t), \underline{W}(R_t)\right\}, \tag{23}$$

$$\widehat{W}(R_t,Z_t) = \begin{cases} (Z_t/U)^{1/\gamma} > R_t, & \text{if } Z_t > Z_U(R_t), \\ R_t, & \text{if } Z_t \in [Z_L(R_t), Z_U(R_t)], \\ (Z_t/L)^{1/\gamma} < R_t, & \text{if } Z_t < Z_L(R_t), \end{cases} \tag{24}$$

where U and L are real constants to be determined. The productivity thresholds (free boundaries) are determined by the value-matching conditions for all R_t, and are given by

$$Z_U(R_t) = UR_t^\gamma \quad \text{and} \quad Z_L(R_t) = LR_t^\gamma. \tag{25}$$

The constants Ω_1, Ω_2 in the value function (21) are determined by the first-order conditions for the optimal wage, ensuring that the smooth-pasting conditions at the free boundaries are satisfied for all R_t, and are given by

$$\Omega_1 = \frac{U^{\beta_2}\left(\frac{1+\rho+\delta+\pi}{\rho+\delta+\pi} - \eta\lambda L\right) - L^{\beta_2}\left(\frac{1+\rho+\delta+\pi}{\rho+\delta+\pi} - \eta U\right)}{(1-\gamma\beta_1)\left(L^{\beta_1}U^{\beta_2} - U^{\beta_1}L^{\beta_2}\right)}, \tag{26}$$

$$\Omega_2 = \frac{L^{\beta_1}\left(\frac{1+\rho+\delta+\pi}{\rho+\delta+\pi} - \eta U\right) - U^{\beta_1}\left(\frac{1+\rho+\delta+\pi}{\rho+\delta+\pi} - \eta\lambda L\right)}{(1-\gamma\beta_2)\left(L^{\beta_1}U^{\beta_2} - U^{\beta_1}L^{\beta_2}\right)}. \tag{27}$$

If $Z_t \leq \underline{Z}(R_t) = \underline{Z}^{layoff}(R_t)$, the firm will lay off the worker, while if $Z_t \leq \underline{Z}(R_t) = \underline{Z}^{quit}(R_t)$ the worker will quit.

Proposition 3 establishes that the firm's interior (non-layoff) optimal wage-setting policy is a state-dependent trigger policy characterized by two free boundaries, i.e., an "S,s" rule with two time-varying thresholds. The optimal wage exhibits a three-region structure: when productivity is sufficiently high ($Z_t > Z_U(R_t)$), the firm raises the wage to $(Z_t/U)^{1/\gamma}$; when productivity is sufficiently low ($Z_t < Z_L(R_t)$), the firm cuts the wage to $(Z_t/L)^{1/\gamma}$; and when productivity lies in the intermediate "range of rigidity" $[Z_L(R_t), Z_U(R_t)]$, the firm optimally

freezes the wage at the reference level R_t. The latter is exactly the inaction region delineated in Proposition 1, which is nonempty if and only if $\lambda > 1$ reflecting that the benefit of wage cuts is offset by the cost of asymmetric reciprocity.

It is critical to note that the free boundaries $Z_L(R)$ and $Z_U(R)$ scale with the reference wage non-linearly (modulated by the curvature parameter γ), implying that the width of the inaction region grows or shrinks with the wage level. Accordingly, downward wage rigidity is far from a static phenomenon but rather evolves dynamically throughout the employment relationship. Since past wage increases become embedded in the worker's current reference wage through adaptation, the firm's reluctance to cut wages applies at any wage level the firm-worker match reaches, not merely at its initial level. Consequently, moderate productivity declines will likely trigger wage freezes rather than cuts, as the firm optimally tolerates compressed profit margins to avoid the disproportionate effort loss from negative reciprocity.

Another important feature of the wage policy is the role of the worker's reservation wage $\underline{W}(R_t)$ as a binding constraint in relatively low productivity states. Whenever the interior solution $\widehat{W}(R_t, Z_t)$ falls below $\underline{W}(R_t)$, the firm sets the wage exactly at the reservation wage instead of allowing the worker to quit, provided the match continues to generate positive surplus (if $\mathcal{J}(R_t, Z_t) > 0$). In this region, the wage is determined by the worker's outside option and not by the first-order condition, but the firm's willingness to pay it reflects genuine economic surplus from the match. Only when productivity falls below the threshold $\underline{Z}(R_t)$ (where $\mathcal{J}(R_t, Z_t) \leq 0$) does separation occur through layoff or quit. In that sense, the optimal wage-setting behavior of the firm insulates the model from the critique put forth by Barro (1977).

Let the candidate optimal policy $\widetilde{W}(R, Z)$ be given by Proposition 3. Henceforth we consider the following *candidate value function*:

$$\widetilde{\mathcal{J}}(R,Z) = G(\widetilde{W},R,Z) + \overline{\mathcal{J}}(\widetilde{W},Z) = \begin{cases} \mathcal{J}^U(R,Z;\widetilde{W}), & \text{if } Z_t > Z_U(R) \\ \overline{\mathcal{J}}(R,Z), & \text{if } Z_t \in [Z_L(R), Z_U(R)]\,, \\ \mathcal{J}^L(R,Z;\widetilde{W}), & \text{if } Z_t \in (\,\underline{Z}(R), Z_L(R)\,)\,, \\ 0 \text{ (layoff)} & \text{if } Z_t \leq \underline{Z}(R), \end{cases} \tag{28}$$

$$\mathcal{J}^U(R,Z;\widetilde{W}) \triangleq \frac{\eta Z}{1-\gamma}\Big(\widetilde{W}^{1-\gamma} - R^{1-\gamma}\Big) - (\widetilde{W} - R) + \frac{Z\bar{e}}{\rho+\delta-\mu} - \frac{\widetilde{W}}{\rho+\delta+\pi} + \Omega_1 Z^{\beta_1}\widetilde{W}^{1-\gamma\beta_1} + \Omega_2 Z^{\beta_2}\widetilde{W}^{1-\gamma\beta_2}, \tag{29}$$

$$\overline{\mathcal{J}}(R,Z) \triangleq \frac{Z\bar{e}}{\rho+\delta-\mu} - \frac{R}{\rho+\delta+\pi} + \Omega_1 Z^{\beta_1}R^{1-\gamma\beta_1} + \Omega_2 Z^{\beta_2}R^{1-\gamma\beta_2} \tag{30}$$

$$\mathcal{J}^L(R,Z;\widetilde{W}) \triangleq \frac{\lambda\eta Z}{1-\gamma}\Big(\widetilde{W}^{1-\gamma} - R^{1-\gamma}\Big) - (\widetilde{W} - R) + \frac{Z\bar{e}}{\rho+\delta-\mu} - \frac{\widetilde{W}}{\rho+\delta+\pi} + \Omega_1 Z^{\beta_1}\widetilde{W}^{1-\gamma\beta_1} + \Omega_2 Z^{\beta_2}\widetilde{W}^{1-\gamma\beta_2}. \tag{31}$$

The next step is one of the most crucial in establishing that our candidate value function is indeed the true value function of the problem. Below we prove that (28) is the unique viscosity solution of the HJBQVI, as well as a classical solution of the HJBQVI almost everywhere on the state space, and identify the necessary and sufficient conditions that pin down the constants U and L. The notion of viscosity solutions is essential and far from a mere technicality: it is the appropriate weak solution concept for degenerate elliptic partial differential equations in this setting given that classical (twice) differentiability fails at the free boundaries $Z_U(R)$ and $Z_L(R)$, as well as at the layoff boundary $\underline{Z}(R)$; see e.g. Bensoussan and Lions (1984), Bardi and Capuzzo-Dolcetta (1997), Crandall, Ishii, and Lions (1992), and Fleming and Soner (2006).

PROPOSITION 4: *The candidate value function (28) is:*

1) *A classical solution of the HJBQVI (15) almost everywhere on $\mathcal{X}$;*

2) *A viscosity solution of the HJBQVI (15) everywhere on $\mathcal{X}$;*

3) *The unique viscosity solution of the HJBQVI (15) everywhere on $\mathcal{X}$;*

if and only if $U \geq \dfrac{1+\rho+\delta+\pi}{(\rho+\delta-\mu)\eta}$ *and* $L \leq \dfrac{1+\rho+\delta+\pi}{(\rho+\delta-\mu)\eta\lambda}$.

Proposition 4 reveals that our candidate value function is indeed the unique viscosity solution to the HJBQVI under the simple necessary and sufficient conditions that U and L are greater/less than a deterministic constant. We remind the reader that we have not yet derived the *optimal* values of U and L, but we will do so shortly and show that they always satisfy the above conditions. Armed with this uniqueness result we can now proceed with a formal verification procedure. In what follows we apply a step-by-step verification-theorem argument to prove that the viscosity solution identified right above coincides with the firm's true value function, and that the associated policy is indeed optimal and unique.

PROPOSITION 5: *For every $t \geq s$, the value function $\mathcal{J}(R,Z)$ of the firm employing a worker with reference wage R and nominal idiosyncratic match productivity Z is given by (28), and the unique optimal wage policy is given by (23), if and only if*

$$U = \frac{\frac{1+\rho+\delta+\pi}{\rho+\delta+\pi}\Big[\beta_1(\beta_2-1)x^{\beta_1} - \beta_2(\beta_1-1)x^{\beta_2} + (\beta_1-\beta_2)x^{\beta_1+\beta_2}\Big]}{\eta\,(\beta_1-1)(\beta_2-1)\big(x^{\beta_1}-x^{\beta_2}\big)} > U^*, \tag{32}$$

$$L = \frac{\frac{1+\rho+\delta+\pi}{\rho+\delta+\pi}\Big[\beta_1(\beta_2-1)x^{\beta_1} - \beta_2(\beta_1-1)x^{\beta_2} + (\beta_1-\beta_2)x^{\beta_1+\beta_2}\Big]}{x\,\eta(\beta_1-1)(\beta_2-1)\big(x^{\beta_1}-x^{\beta_2}\big)} < \frac{U^*}{\lambda}, \tag{33}$$

where $x \triangleq U/L$ solves equation (A115) in the Appendix and admits the analytic representation

$$x = 1 + \sum_{n=1}^{\infty} \alpha_n (\lambda-1)^{n/3}, \quad \alpha_n = \frac{1}{n!}\left[\frac{d^{n-1}}{d\varepsilon^{n-1}}\left[\left(\frac{F(1+\varepsilon)-1}{\varepsilon^3}\right)^{-n/3}\right]\right|_{\varepsilon=0}, \tag{34}$$

and where $U = L = U^* = \dfrac{1+\rho+\delta+\pi}{(\rho+\delta-\mu)\eta}$ *when* $\lambda = 1$.

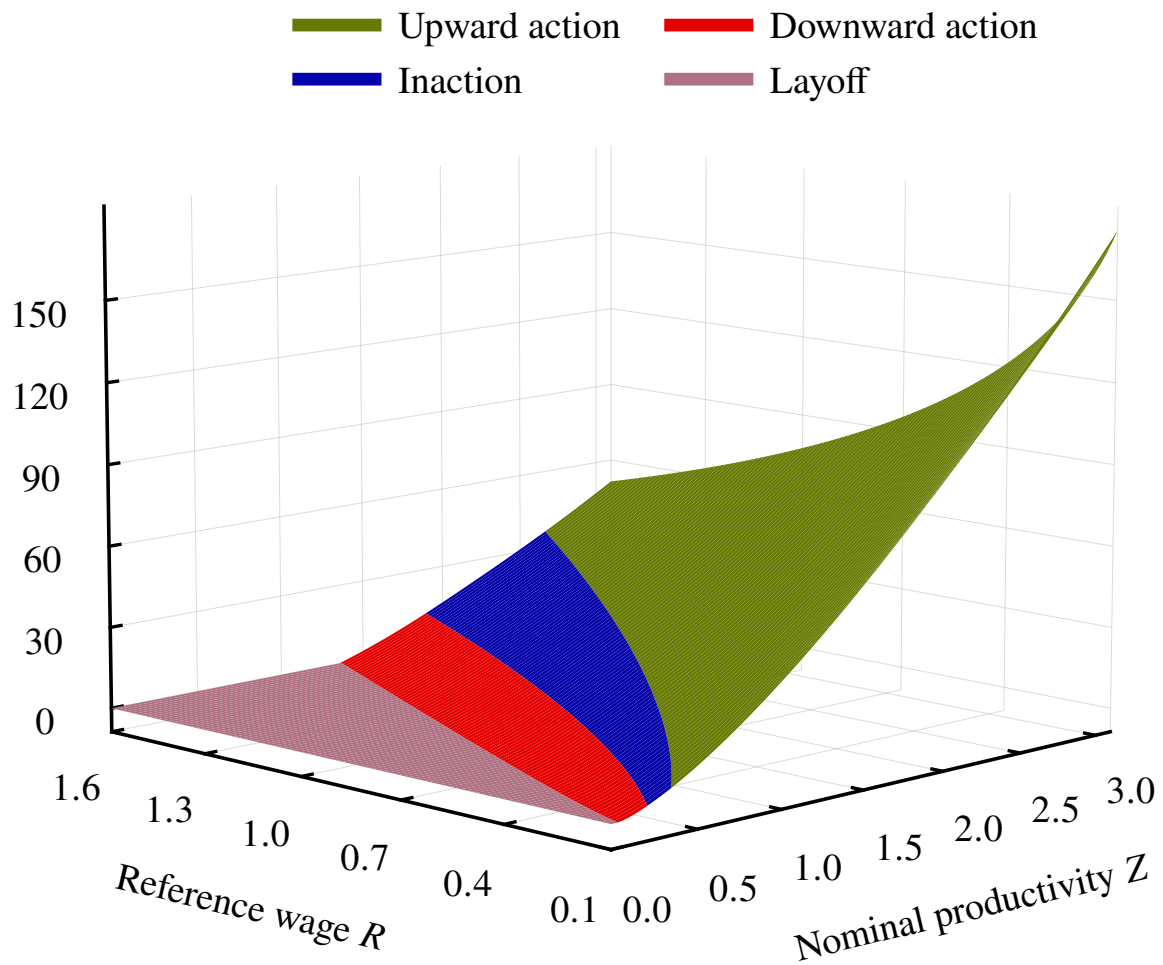
Proposition 5 concludes the full characterization of the firm's problem. The solution delivers closed-form expressions for both the value function and the wage policy across the entire state space, conditional on the worker's reservation wage $\underline{W}(R_t)$ and the separation threshold $\underline{Z}(R_t)$. The verification procedure confirms that the candidate viscosity solution to the HJBQVI satisfies the transversality condition and coincides with the firm's true value function, thereby establishing that the solution characterized in Proposition 4 is indeed optimal and that the associated policy $\widetilde{W}(R, Z)$ maximizes the firm's expected discounted profits. Importantly, the constants U and L are uniquely and unconditionally determined by the model's primitives. The asymmetry in worker reciprocity manifests directly in the ratio $x = U/L > 1$, capturing why downward rigidity arises: the firm requires a proportionally larger productivity decline to justify a wage cut than the increase needed for a wage raise. It is evident from the analytic representation that higher values of λ lead to the optimal U being larger and the optimal L being smaller, which means that the range of inaction expands for any given reference wage R, making the occurrence of wage rigidity more likely in expectation. Of course, when $\lambda = 1$, the inaction region disappears ($U = L = U^*$) and trivially collapses to a singleton.

Figure 1 illustrates the value function of the firm $\mathcal{J}(R, Z)$ and its partition into four distinct regions across the state space of reference wage R and nominal productivity Z. The layoff region occupies the low-productivity/high-reference-wage part where the match generates negative surplus and separation is optimal. The downward action and upward action regions surround the inaction region with the time-varying boundaries $Z_L(R)$ and $Z_U(R)$ clearly visible as curves that scale with R^γ. The value function is smooth within and across (up to the first derivative) each region, consistent with the smooth-pasting conditions that pin down the constants Ω_1 and Ω_2. Notably, $\mathcal{J}(R, Z)$ is strictly decreasing in R for any given Z—the same realization of Z may result in upward action, inaction, downward action, or layoff depending on how high the current R is—and strictly increasing in Z for any given R, reflecting both the contemporaneous and option value of higher productivity.

III. Wage Dynamics

This section explores the dynamic implications of the firm's optimal wage-setting policy in partial equilibrium. We begin by deriving a series of results that delineate some qualitative properties of wage dynamics, providing analytical insights into how the model's structural parameters shape wage rigidity, adjustment patterns, and rent-sharing elasticities. We then turn to our framework's quantitative performance, calibrating the model using a simulated panel of heterogeneous agents to match key empirical moments from the distribution of nominal wage changes and evaluating its ability to replicate additional untargeted statistics.

FIGURE 1: THE FIRM'S VALUE FUNCTION



Notes: The figure depicts the value function of a firm $\mathcal{J}(R, Z)$ employing a worker with reference wage R and nominal match productivity Z. The parametrization follows the baseline calibration reported in Table 1.

III.1. Qualitative implications

We begin by characterizing how the extent of nominal wage rigidity—captured by the width of the inaction region—responds to the model's structural parameters. This comparative-statics exercise lays out the primitives that determine the range of wage rigidity and provides testable predictions for how this range varies across different economic environments.

COROLLARY 1: *Let* $x \triangleq Z_U(R_t)/Z_L(R_t) = U/L$ *denote the width of the inaction region capturing the range of nominal wage rigidity. We have that:*

— x *is strictly increasing in the reciprocity asymmetry parameter* λ*;*

— x *is strictly decreasing in the discount rate* ρ*;*

— x *is strictly decreasing in the separation rate* δ*;*

— x *is invariant to the effort sensitivity parameter* η*;*

— x *is invariant to the curvature parameter* γ*.*

Corollary 1 reveals how the range of nominal wage rigidity changes according to the model's state and primitives. The reciprocity asymmetry λ is the main parameter that amplifies the range of rigidity by lowering the downward threshold L disproportionately more relative to the upward threshold U, making wage cuts more costly and thus widening the inaction region. Conversely, a higher discount rate ρ or separation rate δ reduces the present value of future reference wage costs, which naturally tilts the firm's trade-off toward more frequent adjustment. In contrast, the invariance of x to the effort sensitivity parameter η and the curvature parameter γ arises because these parameters affect the level mapping between productivity and wages, but not the relative spacing between the two adjustment thresholds: they shift the two margins proportionately, so their effects cancel in the ratio U/L.

While Corollary 1 characterized the width of the inaction region and its comparative statics, the following result characterizes the *dynamics* of wages when the firm does choose to adjust. The width of the inaction region determines the frequency of wage adjustments, while the stochastic process governing wages in the action region determines the magnitude and volatility of those adjustments. We continue by showing that the optimal wage dynamics in the action region inherit the stochastic properties of Z_t, but with modified drift and volatility due to the non-linear transformation induced by the optimality conditions.

COROLLARY 2: *The firm's optimal wage policy in the action region* ($Z_t \notin [Z_L(R), Z_U(R)]$ *and* $\widetilde{W}_t > \underline{W}(R)$) *follows a geometric Brownian motion with modified drift and volatility:*

$$\frac{d\widetilde{W}_t}{\widetilde{W}_t} = \left(\frac{\mu+\pi}{\gamma} + \frac{1}{2}\frac{\sigma^2(1-\gamma)}{\gamma^2}\right)dt + \frac{\sigma}{\gamma}dB_t, \tag{35}$$

with the corresponding integral representation:

$$\widetilde{W}_t = \widetilde{W}_{\tau_i}\exp\left\{\frac{1}{\gamma}\left(\mu+\pi-\frac{\sigma^2}{2}\right)(t-\tau_i) + \frac{\sigma}{\gamma}(B_t - B_{\tau_i})\right\}, \quad \text{for } t \in (\tau_i, \tau_{i+1}], \tag{36}$$

where τ_i *denotes the most recent intervention time, and* $\widetilde{W}_{\tau_i} = (Z_{\tau_i}/U)^{1/\gamma}$ *if* $Z_{\tau_i} > Z_U(R_{\tau_i})$ *or* $\widetilde{W}_{\tau_i} = (Z_{\tau_i}/L)^{1/\gamma}$ *if* $Z_{\tau_i} < Z_L(R_{\tau_i})$*. Under full reversibility* ($\lambda = 1$)*, the inaction region vanishes* ($U = L$) *and the wage follows* (35)–(36) *continuously for all* $t \geq 0$ *and* $Z_t > \underline{Z}(R_t)$*.*

There are several useful insights to be drawn from Corollary 2. First, the effort sensitivity parameter η affects only the *level* of optimal wages but not their growth rates or volatility. This

is evident from η entering the solution (36) only through the initial condition $\widetilde{W}_{\tau_0}$, but does not appear in the differential form (35). Higher effort sensitivity leads to proportionally higher wages at all times, but leaves percentage wage changes unaffected. At the same time, η remains important as it determines the level of the value function and thus affects both the endogenous layoff threshold $\underline{Z}^{layoff}(R_t)$ and as the state-space where the reservation wage binds.

Second, while η acts as a scaling parameter, the curvature parameter γ plays a fundamentally different role as it shapes the transmission of productivity shocks into wage dynamics. Specifically, it governs the elasticity of the wage with respect to productivity, $1/\gamma$, and thereby determines both the drift and volatility of wage changes, since γ captures the concavity of the effort function—how rapidly the marginal effort response to wage increases diminishes. When γ is small (high curvature), effort exhibits sharply diminishing returns to wages, so the firm must adjust wages more aggressively in response to productivity shocks to maintain optimal effort incentives. Conversely, when $\gamma \to 1$, effort responds nearly linearly to wages and the optimal wage inherits the exact stochastic properties of productivity.

Finally, notice that the drift term in (35) is exactly the right-hand side of the parameter condition (8) which ensures the TVC is satisfied. This coincidence is not accidental: along non-terminated paths, the firm's value is asymptotically dominated by its fastest-growing component (upward adjustment region), which is proportional to $\widetilde{W}_t$. The transversality condition therefore requires the firm's effective discount rate $(\rho+\delta+\pi)$ to exceed this *maximal* expected growth rate $\mathbb{E}[d\widetilde{W}_t/\widetilde{W}_t] = \frac{\mu+\pi}{\gamma} + \frac{1}{2}\frac{\sigma^2(1-\gamma)}{\gamma^2}$, so that the discounted continuation value vanishes in expectation as $t \to \infty$. See Step 2 in the proof of Proposition 5 for more details.

The next result characterizes the firm's wage-setting behavior over finite time intervals of length $\Delta > 0$. Here we work with the firm's fundamental state variable, $X_t \triangleq Z_t/R_t^\gamma$, which evolves as a geometric Brownian motion over each time interval since the reference wage remains constant until the next (potential) adjustment. This yields analytical and intuitive expressions for state-dependent adjustment probabilities that depend on the current state X_t, the boundaries of the inaction region, and the primitives of the productivity process.

PROPOSITION 6: *Define the firm's state variable $X_t \triangleq Z_t/R_t^\gamma$ for any $t \geq 0$. Over a time interval of length $\Delta > 0$, the optimal (non-layoff) wage evolves according to:*

$$\widehat{W}_{t+\Delta}(X_t) = \begin{cases} W_U > \widehat{W}_t & \textit{with probability } \theta_U(X_t, \Delta), \\ \widehat{W}_t & \textit{with probability } 1 - \theta_U(X_t, \Delta) - \theta_L(X_t, \Delta), \\ W_L < \widehat{W}_t & \textit{with probability } \theta_L(X_t, \Delta), \end{cases} \tag{37}$$

where the state-dependent adjustment probabilities are given by

$$\theta_U(x,\Delta) = \Phi\left(\frac{\ln(x/U) + (\mu + \pi - \sigma^2/2)\Delta}{\sigma\sqrt{\Delta}}\right), \tag{38}$$

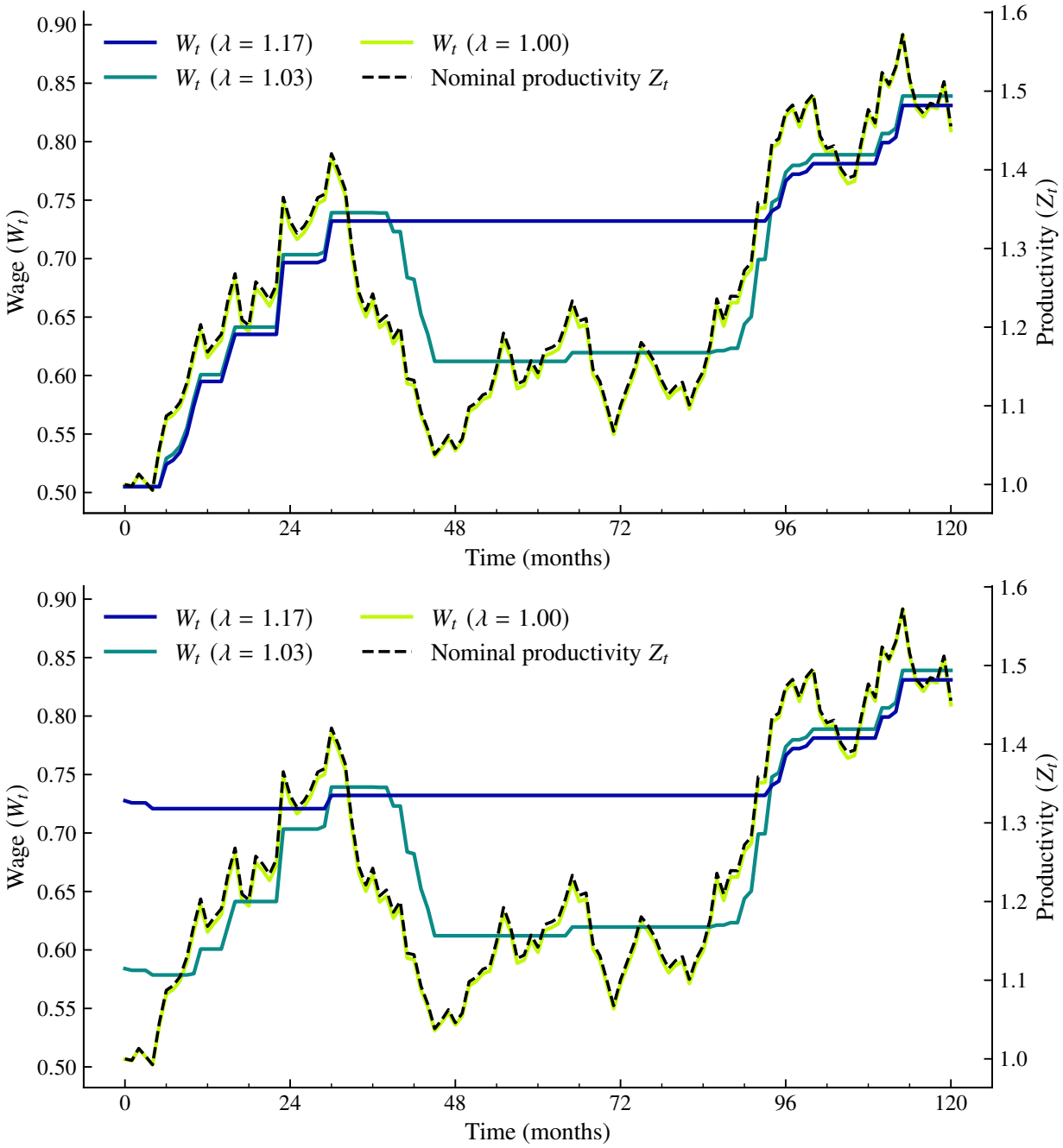
$$\theta_L(x,\Delta) = \Phi\left(\frac{\ln(L/x) - (\mu + \pi - \sigma^2/2)\Delta}{\sigma\sqrt{\Delta}}\right), \tag{39}$$

the reset wages are $W_U = \left(\frac{Z_{t+\Delta}}{U}\right)^{1/\gamma}$ *and* $W_L = \left(\frac{Z_{t+\Delta}}{L}\right)^{1/\gamma}$, *and* $\Phi(\cdot)$ *denotes the standard Gaussian cumulative distribution function.*

Proposition 6 reveals a rich state-dependent structure in the evolution of wage adjustments and further demonstrates the presence of path dependence and hysteresis in wage dynamics. In our setting, the implied adjustment probabilities are endogenous functions of economic fundamentals and thus inherently time-varying. Firms that are currently raising wages because of high productivity *relative to reference wages* ($X_t > U$) are much more likely to raise wages again in the next period; those with low relative productivity ($X_t < L$) are much more likely to cut wages again next period; and those currently in the inaction region ($L < X_t < U$) should be expecting to keep wages frozen in the near future. As a result, two workers with identical productivity trajectories but sufficiently different initial reference wages may exhibit wildly dissimilar wage levels and adjustment patterns over time. These features contrast sharply with time-dependent (Calvo-type) frameworks used in many macroeconomic models in which the implied hazard functions are constant.

Figure 2 plots the firm's wage policy for three values for the parameter $\lambda = \{1.17, 1.03, 1.00\}$ under the *same* simulated path for nominal productivity Z_t. The top panel illustrates how the degree of asymmetric reciprocity shapes wage dynamics along identical productivity paths; the underlying $R_0 = 0.5$ is such that in all three cases the initial optimal wages are numerically identical. The benchmark case of full reversibility ($\lambda = 1.0$) exhibits continuous wage adjustments that exactly track nominal productivity fluctuations. In stark contrast, the case of strong asymmetry ($\lambda = 1.17$) generates pronounced downward nominal wage rigidity. Despite a sustained productivity decline lasting nearly 60 months, the wage remains completely frozen throughout as the threshold $Z_L(R_t)$ is sufficiently low that even large productivity declines fail to breach it. Wages only adjust upward once productivity recovers strongly enough to exceed $Z_U(R_t)$. Crucially, this upward adjustment is itself compressed relative to the frictionless case, illustrating the model's property of *upward wage compression.* The intermediate case ($\lambda = 1.03$) exhibits moderate downward nominal wage rigidity. The wage then remains frozen until productivity falls sufficiently far below the respective threshold $Z_L(R_t)$ to trigger a downward adjustment, but the resulting cut is substantially smaller in magnitude than the concurrent decline in Z_t and attenuated relative to the full-reversibility benchmark. This illustrates the model's property of *downward wage compression.* Notably, all three policies converge to similar wage levels by the end of the simulation as productivity trends upward, but the paths differ dramatically—this is precisely the asymmetry that variations in λ are designed to capture.

Figure 2: Wage Dynamics Under Different Values for λ



Notes: The figure depicts the firm's wage policy under three different values for the parameter governing workers' negative reciprocity to nominal wage cuts, $\lambda = \{1.45, 1.2, 1.0\}$. All wage policies use the same simulated path for nominal productivity (dashed line, $Z_0 = 1.0$), as well as the same initial reference wage: $R_0 = 0.5$ in the top panel, $R_0 = 0.8$ in the bottom panel. Apart from the different values of λ, the rest of the parametrization follows the baseline calibration reported in Table 1.

The bottom panel of Figure 2 reveals an additional dimension of the model's implications for path dependence. Here we consider a higher initial reference wage, $R_0 = 0.8$, so that the three optimal policies now start from different initial wage levels. Given the state-dependent nature of wage setting, R_0 is a state variable that influences not only the initial wage level but also the entire subsequent wage trajectory. Under strong asymmetric reciprocity ($\lambda = 1.17$) the wage remains almost completely rigid for approximately 100 months, absorbing nearly all productivity fluctuations within the inaction region until a sufficiently large positive shock triggers an upward adjustment. Under the intermediate case ($\lambda = 1.03$) the initial wage is higher, as the elevated R_0 shifts the entire inaction region upward. The wage implements modest adjustments in both directions as productivity fluctuates, but each adjustment is again attenuated relative to the frictionless benchmark. Taken together, the two panels illustrate that R_0 is not merely an initial condition but a persistent determinant of wage dynamics. Workers facing identical productivity trajectories but different initial reference wages will exhibit persistently different wage levels and dynamics. This feature captures the empirical regularity that wages display substantial and persistent dispersion even among observably similar workers, and that the frequency and magnitude of wage adjustments vary considerably in ways that are difficult to attribute to contemporaneous productivity differences alone.

We now turn to yet another element our model can speak to: the elasticities of wages with respect to key firm-level outcomes. These quantities provide a direct link between our theoretical framework and the empirical literature on "rent-sharing elasticities", which typically estimates reduced-form relationships between wages and measures of firm performance such as productivity, output, or profits; see the comprehensive survey of Card et al. (2018).

COROLLARY 3: *For any given (R_t, Z_t) such that Z_t is not in the inaction region $[Z_L(R), Z_U(R)]$, for a given firm-specific labor (wage) share of output $s_L = \frac{W_t}{Y_t}$, and worker's elasticity of effort with respect to the wage given by*

$$\varepsilon_{f,W} = \begin{cases} \frac{\eta W_t^{1-\gamma}}{f(W_t,R_t)}, & \text{if } W_t \geq R_t \\ \frac{\lambda\eta W_t^{1-\gamma}}{f(W_t,R_t)}, & \text{if } W_t < R_t, \end{cases} \tag{40}$$

the instantaneous wage elasticities with respect to productivity $(\varepsilon_{W,Z})$, firm output $(\varepsilon_{W,Y})$, and firm profits $(\varepsilon_{W,\Pi})$ are given by:

$$\varepsilon_{W,Z} = \frac{d\log W_t}{d\log Z_t} = \frac{1}{\gamma} \tag{41}$$

$$\varepsilon_{W,Y} = \frac{d\log W_t}{d\log Y_t} = \frac{1}{\gamma + \varepsilon_{f,W}} \tag{42}$$

$$\varepsilon_{W,\Pi} = \frac{d\log W_t}{d\log \Pi_t} = \frac{1 - s_L}{\gamma + \varepsilon_{f,W} - s_L} \tag{43}$$

Therefore we have that $\varepsilon_{W,\Pi} < \varepsilon_{W,Y} < \varepsilon_{W,Z}$ for all t when $\varepsilon_{f,W} > 0$.

Corollary 3 clarifies that there is a certain hierarchy in the magnitude of rent-sharing elasticities embedded in our framework. The wage elasticity with respect to productivity $\varepsilon_{W,Z} = 1/\gamma$ is indeed the most fundamental elasticity implied by our model. Nevertheless, we should consider alternative notions of rent-sharing elasticities from an empirical standpoint, as the true productivity process of a firm-worker match can rarely be observed in actual data. When we calculate the elasticity with respect to output $\varepsilon_{W,Y}$, the denominator increases by $\varepsilon_{f,W}$ because output depends not only on Z_t but also on endogenous effort, which itself responds to wage changes. This effectively creates a feedback loop: productivity increases raise the optimal wage, which in turn boosts effort and amplifies the output response, thereby attenuating the wage-output elasticity. Accordingly, the elasticity with respect to profits $\varepsilon_{W,\Pi}$ is even lower because profits are a residual after wage payments—determined by the current burden of the wage bill relative to output—making them more sensitive to wage changes than output. The ranking $\varepsilon_{W,\Pi} < \varepsilon_{W,Y} < \varepsilon_{W,Z}$ thus reflects the progressive compounding of endogenous responses as we move from the exogenous productivity shocks to the firm's bottom line.

These results speak directly to the extensive empirical literature on rent-sharing elasticities, which documents substantial heterogeneity in estimated elasticities depending on the measure of "quasi-rent" per worker used. Most empirical studies regress observed worker wages either on value-added per worker or on profits per worker (plus a number of controls together with worker and time fixed effects). Our theoretical prediction is consistent with the typical finding that output-based elasticities exceed (often by a lot) profit-based elasticities; see Card et al. (2018) and the numerous references therein. Our framework also provides a structural interpretation of this inherent heterogeneity. Specifically, the different magnitudes of estimated elasticities are not merely alternative empirical specifications but rather reflect distinct economic objects that should theoretically differ due to the endogeneity of effort.

III.2. Quantitative implications

We now assess the model's quantitative performance by calibrating it to match key empirical moments of the wage change distribution along with additional relevant moments. The calibration exercise serves two purposes: first, to demonstrate that our parsimonious theoretical framework can quantitatively account for salient features of observed wage dynamics; and second, to explore the model's predictions for additional statistics not directly targeted in the calibration, thereby providing a test of its internal consistency and empirical relevance.

The model economy is calibrated to US data. The length of a period is taken to be one month and the results are based on a simulated panel of $100,000$ firm-worker matches per period over a time span of 120 months (10 years). The calibration strategy lies on the side of parsimony: we reduce the number of degrees of freedom by preassigning values to three conventional and well-accepted parameters in the literature, and normalize the values of two more. The model is then disciplined via the joint calibration of five parameters targeting five relevant moments that are sufficiently informative about the calibrated parameters.

The set of parameters that are assigned standard values is the following. The annual time discount rate is chosen to be $\rho = 0.04$ and the annual inflation rate is $\pi = 0.02$.. The exogenous annual separation rate is set to $\delta = 0.40$ consistent with observed annual job separation rates (Shimer, 2005). We also normalize $\bar{e} = 1.0$ and $\underline{W} = 0.00$.

Table 1 summarizes the calibration output and resulting parameter values. The five targeted moments (at annual frequency) are drawn from and consistent with empirical evidence on job-stayers' wage dynamics. The first three targeted moments—the probability of no wage change (33.2%), the median unconditional wage change (2.4%), and the standard deviation of unconditional wage changes (6.3%)—are taken directly from Grigsby, Hurst, and Yildirmaz (2021), who document these statistics using comprehensive administrative payroll data from ADP covering millions of US workers. These moments capture the essential features of wage rigidity: a substantial mass of workers experiencing no nominal wage adjustment over a 12-month period, modest median wage growth consistent with low inflation and moderate productivity growth, and considerable dispersion in wage changes reflecting heterogeneity in firm-level shocks. The fourth targeted moment, an average rent-sharing elasticity 8% (with respect to firm profits as a measure of quasi-rent per worker), is consistent with the empirical literature surveyed by Card et al. (2018), which typically estimates rent-sharing elasticities in the range of $5-15\%$ depending on the measure of quasi-rents per worker and estimation strategy. Moreover, this choice falls between the estimates of around 4% by Juhn et al. (2018) and the IV estimate of around 16% by Barth et al. (2016), both for the US economy. Finally, we target an average labor share of 60%, a standard value in macroeconomic calibrations.

To be consistent with the empirical rent-sharing literature, we estimate the elasticity of wages with respect to firm profit using the two-way fixed effects estimator:

$$\log W_{it} = \alpha_i + \alpha_t + \beta \log \Pi_{it} + \varepsilon_{it}, \tag{44}$$

where $\Pi_{it} \equiv Z_{it} f(W_{it}, R_{it}) - W_{it}$ is per-worker profit, α_i and α_t are worker and time fixed effects respectively, and the coefficient β identifies the rent-sharing elasticity.

The calibration jointly determines five structural parameters: the reciprocity asymmetry λ, the productivity drift μ and volatility σ, the effort curvature γ, and the effort sensitivity η. The reciprocity asymmetry $\lambda = 1.17$ is the key parameter governing downward wage rigidity and is identified primarily by the spike at zero in the wage change distribution. The productivity parameters $\mu = 0.02$ and $\sigma = 0.09$ (both annualized) are highly informative of the central tendency and dispersion of wage changes. The curvature parameter $\gamma = 0.80$ and the effort sensitivity $\eta = 17.3$ primarily determine the average rent-sharing elasticity and pin down the average labor share, consistent with the findings of Corollary 3.

Table 2 reports additional statistics generated by the calibrated model at monthly frequency, providing a more granular view of wage dynamics and a test of the model's ability to plausibly capture untargeted moments. The selection of these moments is inspired by Table 4 in Grigsby, Hurst, and Yildirmaz (2021). The model generates a median unconditional monthly

wage change of zero, consistent with the high degree of wage rigidity observed in the data, while the mean unconditional change of 0.4% reflects the positive drift in productivity and inflation. The standard deviation of 1.3% at monthly frequency is consistent with the annualized value of 6.3% when accounting for the time aggregation of idiosyncratic shocks. Notably, the model produces substantial positive skewness (3.5) and excess kurtosis (17.9) in the unconditional wage change distribution, capturing the empirical regularity that wage changes are characterized by a large mass at zero combined with fat tails particularly on the positive side. Conditional on any wage change occurring, the model generates strongly positive mean and median changes, along with lower skewness and kurtosis, reflecting a less asymmetric distribution once the mass point at zero is removed. The model also successfully captures the inherent asymmetry between positive and negative wage changes: conditional on a positive change, the mean and median are 2.5% and 2.1%, while conditional on a negative change they are -2.1% and -1.7%, with the magnitudes being roughly comparable whereas wage cuts occur far less frequently than wage increases due to the asymmetric reciprocity parameter $\lambda > 1$.

Figure 3 illustrates the model's predicted distribution of wage changes under the baseline calibration ($\lambda = 1.17$) and the full-reversibility counterfactual ($\lambda = 1.00$), for both annual (top panel) and monthly (bottom panel) time periods. Under full reversibility, both distributions are Gaussian centered at the drift of the wage percentage growth process shown in (35), with the monthly distribution being considerably narrower due to the shorter aggregation window. Under the baseline calibration, both distributions instead exhibit a pronounced spike at zero and are strongly right-skewed with an extensive missing mass of downward adjustments. The two distributions differ in one additional respect. At the monthly horizon, the calibrated distribution exhibits strictly lower density than the flexible benchmark at *every* non-zero bin: mass that would otherwise be spread across small monthly adjustments is absorbed entirely into the zero spike, with no compensating redistribution elsewhere. At the annual horizon, by contrast, the calibrated case displays higher density than the flexible benchmark for small non-zero changes and lower density for larger ones. This reflects a compounding effect: in the fully flexible case small monthly movements accumulate into moderate annual changes, spreading mass away from zero; under DNWR, extended spells of inaction followed by compressed upward adjustments produce annual changes that are either zero or relatively small, thereby concentrating mass near zero at the expense of the moderate- or large-change region. This comparison highlights the central role of asymmetric reciprocity in generating the distinctive features of observed wage change distributions. The model's ability to replicate these features with a parsimonious set of structural parameters—and to generate realistic predictions for untargeted moments—provides further support for the theoretical mechanism at the heart of our analysis.

IV. Macroeconomic Implications

WORK IN PROGRESS

V. Conclusion

We have developed an analytically tractable theory of downward nominal wage rigidity grounded in the incomplete and relational nature of employment contracts. Our framework yields closed-form solutions to the firm's value function and associated "Ss" wage policy, with formal proofs of both optimality and uniqueness rigorously established—results that have eluded previous work in this class of models. Moreover, the model's parsimony does not come at the expense of empirical realism. Calibrated to match key moments from the distribution of wage changes, our simulations replicate the pronounced spike at zero, the missing mass of small positive adjustments, and the dampened response to negative shocks observed in data.

Beyond matching established facts, our framework generates novel predictions about the dynamic consequences of asymmetric reciprocity. Downward rigidity during negative shocks necessarily induces upward rigidity during subsequent recoveries, as wages remain elevated relative to productivity. This creates path dependence and hysteresis: the current state alone does not determine wage adjustment; rather, the history of shocks matters through its effect on the gap between current wages and worker reference points. These predictions align with recent empirical evidence on wage hysteresis and suggest that transitory shocks can have persistent—or even permanent—effects on wage levels.

The continuous-time formulation we employ offers substantial gains in tractability relative to discrete-time predecessors, and almost eliminates the need for numerical solutions, which can be challenging and often imprecise. By recasting the firm's problem as a stochastic impulse control problem with time-varying free boundaries, we leverage mathematical advances to derive fully analytical solutions. This portability is crucial: our framework can be readily embedded into existing workhorse macroeconomic models, whether representative-agent or heterogeneous-agent, to explore the aggregate implications of contract incompleteness for wage and unemployment dynamics.

Our ongoing work aims to integrate this wage-setting framework into a general-equilibrium model with heterogeneous agents, search frictions, endogenous job creation and destruction, and capital accumulation. This extension will allow us to quantify how the strength of workers' discretionary effort responses shapes hiring, layoffs, and the cyclical behavior of wages and unemployment. By providing a structural lens through which to interpret observed wage rigidity, our approach offers a complementary perspective to theories based on complete contracts and bargaining. Incorporating these microfoundations into quantitative macroeconomic models represents, in our opinion, an important step toward a more complete and nuanced account of wage and unemployment dynamics in general equilibrium.

Table 1: Partial-Equilibrium Model Calibration Summary

Targeted Moments (annual frequency)	Model	Data	Data Source
Probability of no wage change (%)	33.2	33.2	Grigsby et al. (2021)
Median unconditional wage change (%)	2.4	2.4	Grigsby et al. (2021)
Std. dev. of unconditional wage change (%)	6.3	6.3	Grigsby et al. (2021)
Average rent-sharing elasticity (%)	8.0	8.0	See evidence discussed
Average labor share of income (%)	60	60	Standard
Calibrated Parameters	**Value**		**Comment**
Relative sensitivity of effort to wage cuts λ	1.17		See discussion in text
Real productivity drift μ (annual)	0.02		See discussion in text
Real productivity volatility σ (annual)	0.09		See discussion in text
Effort function curvature parameter γ	0.80		See discussion in text
Sensitivity of effort to wage changes η	17.3		See discussion in text
Assigned Parameters	**Value**		**Comment**
Time discount rate ρ (annual)	0.04		Standard
Exogenous job separation rate δ (annual)	0.40		40% annual separation rate
Inflation rate π (annual)	0.02		2% annual inflation rate
Workers' normal effort $\bar{e}$	1.00		Normalization
Workers' reservation wage $\underline{W}$	0.00		Normalization

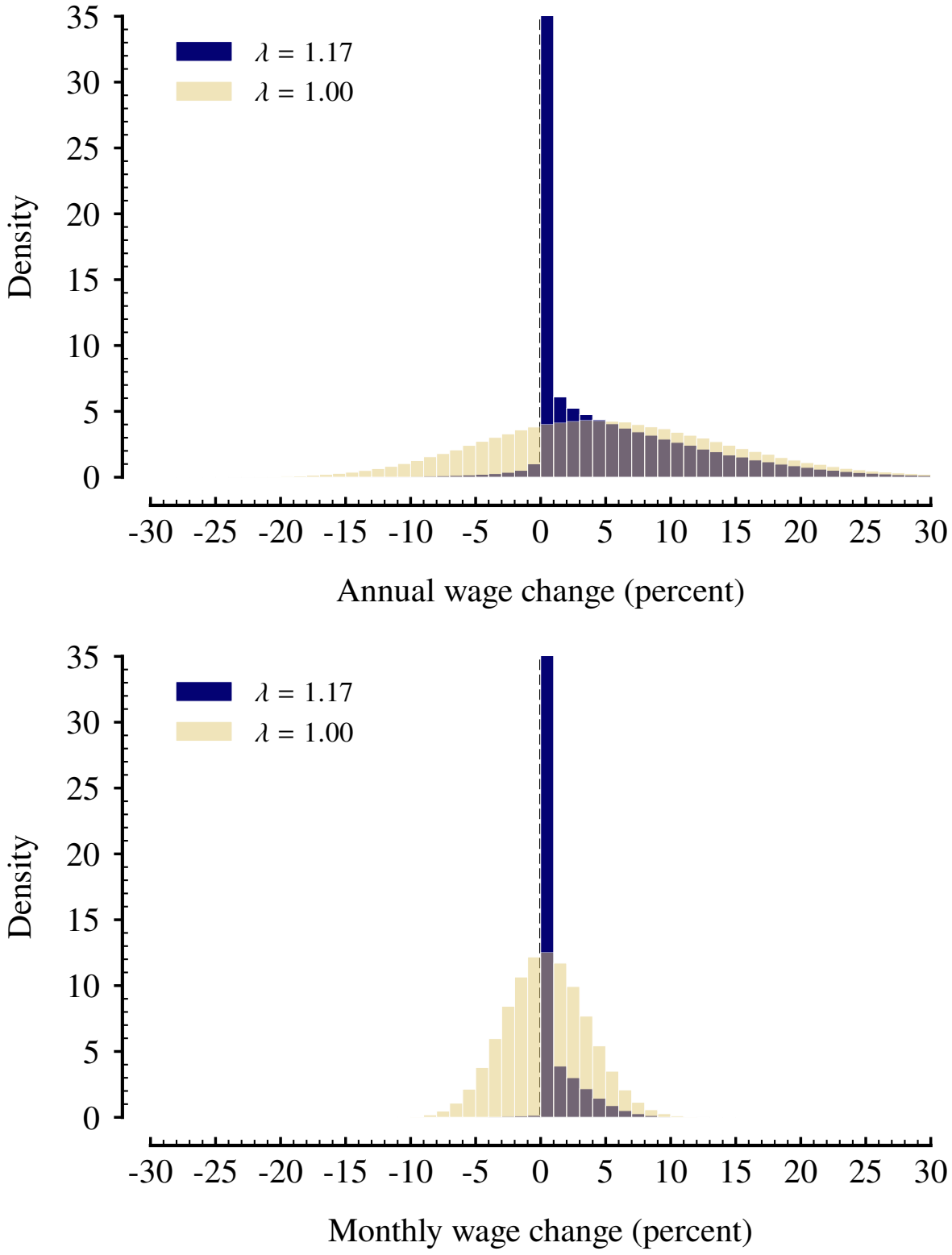
Notes: The results are based on a simulated panel of $100,000$ firm-worker matches per period over a time span of 120 months (10 years), after discarding a burn-in panel of the same size.

Table 2: Partial-Equilibrium Model Calibration; Additional Statistics

Moments (monthly frequency)	Model
Mean unconditional wage change (%)	0.4
Median unconditional wage change (%)	0.0
Std. dev. of unconditional wage change (%)	1.3
Skewness of unconditional wage change	3.5
Kurtosis of unconditional wage change	17.9
Mean conditional wage change (%)	2.4
Median conditional wage change (%)	2.0
Std. dev. of conditional wage change (%)	2.1
Skewness of conditional wage change	0.7
Kurtosis of conditional wage change	4.7
Mean wage change conditional on positive (%)	2.5
Median wage change conditional on positive (%)	2.1
Std. dev. of wage change conditional on positive (%)	1.9
Skewness of wage change conditional on positive	1.1
Kurtosis of wage change conditional on positive	4.3
Mean wage change conditional on negative (%)	−2.1
Median wage change conditional on negative (%)	−1.7
Std. dev. of wage change conditional on negative (%)	1.6
Skewness of wage change conditional on negative	−1.1
Kurtosis of wage change conditional on negative	4.2

Notes: The results are based on a simulated panel of $100,000$ firm-worker matches per period over a time span of 120 months (10 years), after discarding a burn-in panel of the same size.

Figure 3: The Model's Distribution of Annual Wage Changes



Notes: The figure depicts the simulated distribution of annual wage changes (top panel) and monthly wage changes (bottom panel) under the baseline model calibration reported in Table 1 ($\lambda = 1.17$), as well as for the counterfactual simulation under $\lambda = 1.00$. The results are based on a simulated panel of $100,000$ firm-worker matches per period over a time span of 120 months (10 years), after discarding a burn-in panel of the same size.

REFERENCES

Abel, Andrew B., and Janice C. Eberly. 1994. "A Unified Model of Investment Under Uncertainty." *American Economic Review*, 84(5): 1369–1384.

Abel, Andrew B., and Janice C. Eberly. 1996. "Optimal Investment with Costly Reversibility." *Review of Economic Studies*, 63(4): 581–593.

Abel, Andrew B., Avinash K. Dixit, Janice C. Eberly, and Robert S. Pindyck. 1996. "Options, the Value of Capital, and Investment." *Quarterly Journal of Economics*, 111(3): 753–777.

Baley, Isaac, and Andrés Blanco. 2026. "The Macroeconomics of Irreversibility." *The Review of Economic Studies*, rdag001.

Barattieri, Alessandro, Susanto Basu, and Peter Gottschalk. 2014. "Some Evidence on the Importance of Sticky Wages." *American Economic Journal: Macroeconomics*, 6(1): 70–101.

Bardi, Martino, and Italo Capuzzo-Dolcetta. 1997. *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations.* Birkhäuser Boston, MA.

Barro, Robert J. 1977. "Long-term Contracting, Sticky Prices, and Monetary Policy." *Journal of Monetary Economics*, 3(3): 305–316.

Barth, Erling, Alex Bryson, James C. Davis, and Richard Freeman. 2016. "It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States." *Journal of Labor Economics*, 34(S2): S67–S97.

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

Baucells, Manel, and Rakesh K. Sarin. 2010. "Predicting Utility Under Satiation and Habit Formation." *Management Science*, 56(2): 286–301.

Baucells, Manel, and Rakesh K Sarin. 2013. "Determinants of experienced utility: Laws and implications." *Decision Analysis*, 10(2): 135–151.

Baucells, Manel, Martin Weber, and Frank Welfens. 2011. "Reference Point Formation and Updating." *Management Science*, 57(3): 506–519.

Benigno, Pierpaolo, and Luca Antonio Ricci. 2011. "The Inflation-Output Trade-Off with Downward Wage Rigidities." *American Economic Review*, 101(June): 1436–1466.

Bensoussan, Alain, and Jacques-Louis Lions. 1984. *Impulse Control and Quasi-Variational Inequalities.* μ Gauthier-Villars, Montrouge. English translation of "Contrôle impulsionnel et inéquations quasi-variationnelles" by J.M. Cole.

Bentolila, Samuel, and Giuseppe Bertola. 1990. "Firing Costs and Labour Demand: How Bad is Eurosclerosis?" *Review of Economic Studies*, 57(3): 381–402.

Bertheau, Antoine, Marianna Kudlyak, Birthe Larsen, and Morten Bennedsen. 2025. "Why Firms Lay Off Workers Instead of Cutting Wages: Evidence from linked survey-administrative data." IZA Discussion Papers.

Bertola, Giuseppe, and Ricardo J. Caballero. 1990. "Kinked Adjustment Costs and Aggregate Dynamics." In *NBER Macroeconomics Annual 1990, Volume 5*. 237–296. MIT Press.

Bewley, Truman F. 1999. *Why Wages Don't Fall during a Recession.* Cambridge, MA and London, England: Harvard University Press.

Bewley, Truman F. 2007. "Fairness, Reciprocity, and Wage Rigidity." In *Behavioral Economics and Its Applications.* , ed. Peter Diamond and Hannu Vartiainen. Princeton University Press.

Blanco, Andrés, and A Drenik. 2023. "How does Inflation 'Grease the Wheels' in a Frictional Labor Market." Working Paper.

Blanco, Julio, Andrés Drenik, Christian Moser, and Emilio Zaratiegui. 2025. "The Macroeconomics of Wage Rigidity and Job Separations." *Available at SSRN 4124811.*

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

Caballero, Ricardo J., and Eduardo M.R.A. Engel. 1999. "Explaining Investment Dynamics in U.S. Manufacturing: A Generalized (S,s) Approach." *Econometrica*, 67(4): 783–826.

Caballero, Ricardo J., Eduardo M. R. A. Engel, and John Haltiwanger. 1997. "Aggregate Employment Dynamics: Building from Microeconomic Evidence." *The American Economic Review*, 87(1): 115–137.

Calvo, Guillermo A. 1983. "Staggered Prices in a Utility-Maximizing Framework." *Journal of monetary Economics*, 12(3): 383–398.

Campbell, Carl M, and Kunal S Kamlani. 1997. "The Reasons for Wage Rigidity: Evidence From Survey of Firms." *Quarterly Journal of Economics*, 112: 759–789.

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. "Firms and Labor Market Inequality: Evidence and Some Theory." *Journal of Labor Economics*, 36(S1): S13–S70.

Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy*, 113(1): 1–45.

Crandall, Michael G., Hitoshi Ishii, and Pierre-Louis Lions. 1992. "User's Guide to Viscosity Solutions of Second Order Partial Differential Equations." *Bulletin of the American Mathematical Society*, 27(1): 1–67.

Dickson, Alex, and Marco Fongoni. 2019. "Asymmetric Reference-Dependent Reciprocity, Downward Wage Rigidity, and the Employment Contract." *Journal of Economic Behavior & Organization*, 163(C): 409–429.

Dixit, Avinash K., and Robert S. Pindyck. 1994. *Investment Under Uncertainty.* Princeton University Press.

Du Caju, Philip, Theodora Kosma, Martina Lawless, Julián Messina, and Tairi Rõõm. 2015. "Why Firms Avoid Cutting Wages: Survey Evidence from European Firms." *ILR Review*, 68(4): 862–888.

Dupraz, Stéphane, Emi Nakamura, and Jón Steinsson. 2025. "A Plucking Model of Business Cycles." *Journal of Monetary Economics*, 103766.

Eliaz, Kfir, and Ran Spiegler. 2014. "Reference Dependence and Labor Market Fluctuations." *NBER Macroeconomics Annual*, 28(1): 159–200.

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

Elsby, Michael W. L., and Gary Solon. 2019. "How Prevalent Is Downward Rigidity in Nominal Wages? International Evidence from Payroll Records and Pay Slips." *Journal of Economic Perspectives*, 33(3): 185–201.

Elsby, Michael W. L., Axel Gottfries, Pawel Krolikowski, and Gary Solon. 2025. "Wage Adjustment in Efficient Long-Term Employment Relationships."

Elsby, Michael W, Ryan Michaels, and David Ratner. 2019. "The Aggregate Effects of Labor Market Frictions." *Quantitative Economics*, 10(3): 853–891.

Erceg, Christopher J, Dale W Henderson, and Andrew T Levin. 2000. "Optimal Monetary Policy with Staggered Wage and Price Contracts." *Journal of Monetary Economics*, 46(2): 281–313.

Fehr, Ernst, Lorenz Goette, and Christian Zehnder. 2009. "A Behavioral Account of the Labor Market: The Role of Fairness Concerns." *Annual Review of Economics*, 1(Volume 1, 2009): 355–384.

Fehr, Ernst, Oliver Hart, and Christian Zehnder. 2011. "Contracts as Reference Points—Experimental Evidence." *American Economic Review*, 101(April): 493–525.

Fleming, Wendell H., and H. Mete Soner. 2006. *Controlled Markov Processes and Viscosity Solutions.* Vol. 25 of *Stochastic Modelling and Applied Probability*. 2nd ed., New York:Springer-Verlag.

Fongoni, Marco. 2024. "Asymmetric Reciprocity and the Cyclical Behavior of Wages, Effort, and Job Creation." *American Economic Journal: Macroeconomics*, 16(3): 52–89.

Fongoni, Marco. 2025. "Costly Wage Cuts, Relative Wage Comparisons, and Unemployment Hysteresis." *AMSE Working Paper Nr. 40.*

Fongoni, Marco, Daniel Schaefer, and Carl Singleton. 2025. "Why Wages Don't Fall in Jobs with Incomplete Contracts." *Management Science*, 71(8): 6319–6339.

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

Gertler, Mark, Christopher Huckfeldt, and Antonella Trigari. 2020. "Unemployment fluctuations, match quality, and the wage cyclicality of new hires." *The Review of Economic Studies*, 87(4): 1876–1914.

Grigsby, John, Erik Hurst, and Ahu Yildirmaz. 2021. "Aggregate Nominal Wage Adjustments: New Evidence from Administrative Payroll Data." *American Economic Review*, 111(2).

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

Herz, Holger, and Dmitry Taubinsky. 2013. "Market Experience is a Reference Point in Judgments of Fairness." *Working Paper*.

Holden, Steinar, and Fredrik Wulfsberg. 2009. "How Strong is the Macroeconomic Case for Downward Real Wage Rigidity?" *Journal of Monetary Economics*, 56: 605–615.

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

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

Juhn, Chinhui, Kristin McCue, Holly Monti, and Brooks Pierce. 2018. "Firm Performance and the Volatility of Worker Earnings." *Journal of Labor Economics*, 36(S1): S99–S131.

Kahneman, Daniel, and Richard H Thaler. 1991. "Economic Analysis and the Psychology of Utility: Applications to Compensation Policy." *American Economic Review*, 81(2): 341–346.

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

Kaur, Supreet. 2019. "Nominal Wage Rigidity in Village Labor Markets." *American Economic Review*, 109(10): 3585–3616.

Øksendal, Bernt, and Agnès Sulem. 2019. *Applied Stochastic Control of Jump Diffusions. Universitext.* 3 ed., Springer Cham.

Pham, Huyên. 2009. *Continuous-time Stochastic Control and Optimization with Financial Applications. Stochastic Modelling and Applied Probability.* 1 ed., Springer Berlin, Heidelberg.

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

Quach, Simon. 2025. "Wage Hysteresis and Entitlement Effects: The Persistent Impacts of a Temporary Overtime Policy." *The Quarterly Journal of Economics*, 140(2): 1633–1680.

Revuz, Daniel, and Marc Yor. 1999. *Continuous Martingales and Brownian Motion.* Vol. 293 of *Grundlehren der mathematischen Wissenschaften.* 3 ed., Berlin, Heidelberg:Springer.

Schmitt-Grohé, Stephanie, and Martín Uribe. 2016. "Downward Nominal Wage Rigidity, Currency Pegs, and Involuntary Unemployment." *Journal of Political Economy*, 124(5): 1466–1514.

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

Sliwka, Dirk, and Peter Werner. 2017. "Wage Increases and the Dynamics of Reciprocity." *Journal of Labor Economics*, 35(2): 299–344.

Stokey, Nancy L. 2009. *The Economics of Inaction: Stochastic Control Models with Fixed Costs.* Princeton University Press.

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

Appendix A: Proofs and Derivations

Proof of Proposition 1. Throughout the proof we assume standard HJBQVI regularity conditions; that is, on the interior of $\mathcal{I}$ and $\mathcal{A}$, $\mathcal{J}$ is C^1 in both state variables so $\mathcal{J}_R$ is continuous. These conditions are indeed satisfied by the value function as proved in propositions to follow. We also work at interior points where feasibility is slack (away from the layoff or quit region) so that small wage deviations are admissible. We prove necessity and sufficiency in turn.

($\Rightarrow$) *If the inaction region $\mathcal{I}$ is non-trivial, then $\lambda > 1$.*

We proceed by contradiction. Assume that the inaction region is non-trivial, meaning that $\mathcal{I}$ is a non-empty open set and thus has positive Lebesgue measure. Assume $\lambda = 1$ for the sake of contradiction. Fix (R, Z) in the interior of $\mathcal{I}$. By definition of inaction, any immediate impulse $W \neq R$ is not profitable, so the following one-sided conditions at $W = R$ must hold:

$$M^-(R,Z) \triangleq \frac{\partial \mathcal{M}\mathcal{J}}{\partial W}\Big|_{W=R^-} \geq 0 \;\text{ and }\; M^+(R,Z) \triangleq \frac{\partial \mathcal{M}\mathcal{J}}{\partial W}\Big|_{W=R^+} \leq 0.$$

When $\lambda = 1$, the one-sided derivatives coincide since $G(.)$ is C^1 in W at $W = R$, and hence

$$\frac{\partial \mathcal{M}\mathcal{J}}{\partial W}\Big|_{W=R} = \eta Z R^{-\gamma} - 1 + \mathcal{J}_R(R,Z) = 0.$$

At the same time, $\mathcal{J}$ solves the HJB equation $(\rho+\delta+\pi)\mathcal{J} = (Z\bar{e}-R)+(\mu+\pi)Z\mathcal{J}_Z+\frac{1}{2}\sigma^2 Z^2\mathcal{J}_{ZZ}$ in the inaction region. Differentiating w.r.t. R yields

$$(\rho+\delta+\pi)\mathcal{J}_R = -1 + (\mu+\pi)Z\mathcal{J}_{RZ} + \frac{1}{2}\sigma^2 Z^2 \mathcal{J}_{RZZ}.$$

From $\mathcal{J}_R(R,Z) = 1 - \eta Z R^{-\gamma}$, we get $\mathcal{J}_{RZ} = -\eta R^{-\gamma}$ and $\mathcal{J}_{RZZ} = 0$. Substituting back gives

$$(\rho+\delta+\pi)\big(1-\eta Z R^{-\gamma}\big) = -1 - (\mu+\pi)\eta Z R^{-\gamma},$$

which cannot hold on any open set of positive Lebesgue measure. This leads to a contradiction. The hypothesis that $\lambda = 1$ must be false, and it follows that $\lambda > 1$.

($\Leftarrow$) *If $\lambda > 1$, then the inaction region $\mathcal{I}$ is non-trivial.*

Suppose that $\lambda > 1$. This introduces a concave downward kink in the function $G(.)$ at $W = R$. For any (R, Z), the one-sided derivatives at $W = R$ are

$$M^+(R,Z) = \eta Z R^{-\gamma} - 1 + \mathcal{J}_R(R,Z) \;\text{ and }\; M^-(R,Z) = \lambda\eta Z R^{-\gamma} - 1 + \mathcal{J}_R(R,Z).$$

As we prove formally below, $\mathcal{J}_R(R,Z) < 0\ \forall (R,Z)$, hence $1 - \mathcal{J}_R(R,Z) > 0$, and at the same time $M^-(R,Z) - M^+(R,Z) = (\lambda-1)\eta Z R^{-\gamma} > 0$. Now consider the set

$$\mathcal{S} \triangleq \{(R,Z) \in \mathcal{X} :\ M^+(R,Z) < 0 < M^-(R,Z)\}.$$

Since $M^{\pm}(R,Z)$ are continuous and the inequalities are strict, the set $\mathcal{S}$ is open as the intersection of preimages of open rays under continuous maps. For every $(R,Z) \in \mathcal{S}$, we have the subgradient condition at the kink $0 \in [M^{-}(R,Z), M^{+}(R,Z)]$, so $W = R$ is a local optimum among admissible wages. This establishes that $\mathcal{S} \subset \mathcal{I}$, which means that the inaction region $\mathcal{I}$ is a non-empty open set and thus has positive Lebesgue measure in $\mathcal{X}$. ■

Proof of Lemma 1. We prove this by a direct comparison argument. Let $X = \{\tau_i, W_i\}_{i\geq 1}$ denote an admissible policy consisting of a sequence of intervention times and wage impulse controls. The value function is the supremum of the expected present value of profits, call it $J(\cdot,\cdot;X)$, over all admissible policies: $\mathcal{J}(R_0,Z_0) = \sup_{x\in\mathcal{X}} J(R_0,Z_0;X)$.

Consider two initial states, (R_0,Z_0) and $(R_0+\epsilon,Z_0)$ for an arbitrarily small $\epsilon > 0$. Let X^* be the optimal policy for the firm starting in the state $(R_0+\epsilon,Z)$, so that $\mathcal{J}(R_0+\epsilon,Z) = J(R_0+\epsilon,Z;X^*)$. Now construct a feasible, but not necessarily optimal, policy $\hat{X} = X^*$, for the firm starting in the state with the lower reference wage, (R_0,Z_0). We compare the stream of flow profits generated by these two policies, subject to the *same* realization of the productivity process Z_t. Without loss of generality, assume that the realization of Z is such that the firm is in the inaction region at Z_0 (the existence of which is proved in Proposition 1) so that $\mathbb{P}(\tau_1^* > 0) > 0$ for the first stopping time τ_1^*. *For the initial inaction period, $t \in [0,\tau_1^*)$:* During this interval, the firm starting at (R_0,Z_0) under policy $\hat{X}$ earns a flow profit of $Z_t\bar{e} - R$, while the firm starting at $(R_0+\epsilon,Z_0)$ earns $Z_t\bar{e} - (R_0+\epsilon)$. The profit difference is strictly positive as $\epsilon > 0$. *For all subsequent periods, $t \geq \tau_1^*$:* Both firms reset their wage to the same values thereafter, yielding identical flow profits.

The value generated by the feasible policy $\hat{X}$ is therefore strictly greater than the value from the optimal policy X^*, as it dominates in the profit stream over a period that occurs with positive probability: $J(R_0,Z_0;\hat{X}) - J(R_0+\epsilon,Z_0;X^*) = \mathbb{E}\left[\int_0^{\tau_1^*} e^{-(\rho+\delta+\pi)t}\epsilon\, dt \,|\, \mathcal{F}_0\right] > 0$. At the same time, by definition, $\mathcal{J}(R,Z) \geq J(R_0,Z_0;\hat{X})$. Combining these results yields the strict inequality, $\mathcal{J}(R,Z) > \mathcal{J}(R+\epsilon,Z)$, hence $\mathcal{J}(R,Z)$ is strictly decreasing in R. To formally establish the property of its derivative, we rearrange the inequality as $\mathcal{J}(R+\epsilon,Z) - \mathcal{J}(R,Z) < 0$. Dividing by $\epsilon > 0$ and taking the right-hand limit yields:

$$\mathcal{J}_R(R,Z) \triangleq \lim_{\epsilon\to 0^+} \frac{\mathcal{J}(R+\epsilon,Z) - \mathcal{J}(R,Z)}{\epsilon} \leq -\mathbb{E}\left[\int_0^{\tau_1^*} e^{-(\rho+\delta+\pi)t} dt \,\middle|\, \mathcal{F}_0\right] < 0 \qquad \text{(A1)}$$

From the above follows that any (unconstrained) optimal wage must be strictly positive. ■

Proof of Proposition 2. The HJB equation governing the value function in the inaction region is: $(\rho+\delta+\pi)\mathcal{J}(R,Z) = (Z\bar{e} - R) + (\mu+\pi)Z\mathcal{J}_Z(R,Z) + \frac{1}{2}\sigma^2 Z^2 \mathcal{J}_{ZZ}(R,Z)$. The general solution to this linear second-order partial differential equation can be expressed as the sum of a particular solution and a homogeneous solution:

$$\mathcal{J}(R,Z) = \mathcal{J}^p(R,Z) + \mathcal{J}^h(R,Z).$$

First, it is easy to see that an affine ansatz of the form $\mathcal{J}^p(R,Z) = AZ - BR$, where $A, B > 0$ are constants to be determined, yields a valid and intuitive particular solution. Calculating the implied derivatives and matching coefficients yields

$$\mathcal{J}^p(R,Z) = \frac{Z\bar{e}}{\rho+\delta-\mu} - \frac{R}{\rho+\delta+\pi}. \tag{A2}$$

Moving on to $\mathcal{J}^h(R,Z)$, which solves the HJB equation with the non-homogeneous terms $(Z\bar{e} - R)$ set to zero. This is a homogeneous Cauchy-Euler linear second-order ODE in Z for *fixed* R—that is, R is effectively treated as a parameter in the inaction region—whose solution form is well known. Let $\beta_1 > 1$ and $\beta_2 < 0$ be the roots of the quadratic characteristic equation $\frac{1}{2}\sigma^2\beta(\beta-1) + (\mu+\pi)\beta - (\rho+\delta+\pi) = 0$. The homogeneous solution is then given by

$$\mathcal{J}^h(R,Z) = C_1(R)Z^{\beta_1} + C_2(R)Z^{\beta_2}, \tag{A3}$$

$$\beta_{1,2} = \frac{-(\mu+\pi-\frac{1}{2}\sigma^2) \pm \sqrt{(\mu+\pi-\frac{1}{2}\sigma^2)^2 + 2\sigma^2(\rho+\delta+\pi)}}{\sigma^2}, \tag{A4}$$

where $C_1(R), C_2(R)$ are arbitrary functions of R to be determined below. In sum, the general solution to the HJB in the inaction region and its derivative with respect to R are

$$\mathcal{J}(R,Z) = \frac{Z\bar{e}}{\rho+\delta-\mu} - \frac{R}{\rho+\delta+\pi} + C_1(R)Z^{\beta_1} + C_2(R)Z^{\beta_2}, \tag{A5}$$

$$\mathcal{J}_R(R,Z) = -\frac{1}{\rho+\delta+\pi} + C_1'(R)Z^{\beta_1} + C_2'(R)Z^{\beta_2}. \tag{A6}$$

The current challenge is to pin down the functions $C_1(R)$ and $C_2(R)$. Recall that in the action region the optimal wage is obtained by $\widetilde{W}(R,Z) = \arg\max_W\{G(W,R,Z) + \mathcal{J}(W,Z)\}$. The FOC can be expressed as

$$\mathcal{J}_R(W,Z) = 1 - \mathbb{1}_\lambda \eta Z W^{-\gamma},$$

where $\mathbb{1}_\lambda = \lambda$ if $\widetilde{W} < R$ and equal to one otherwise. The FOC reveals that $\mathcal{J}_R(W,Z)$ must be a specific function of $ZW^{-\gamma}$ at all intervention points. For the general solution to be consistent with this constraint, $\mathcal{J}_R(W,Z)$ must inherit the same scaling properties. The rest of the proof relies on this key observation that gives rise to an invariance property of the value function in order to preserve the structure of the necessary FOC. Specifically, we will show that $\mathcal{J}_R(W,Z)$ is homogeneous of degree zero in (W^γ, Z).

To establish the key homogeneity property, consider the one-parameter family of transformations $T_k : (W,Z) \mapsto (k^{1/\gamma}W, kZ)$, for $k > 0$. This transformation preserves the structure of the effort function while scaling the state variables in a proportion determined by the curvature parameter γ. Notice that the FOC is actually invariant under this transformation:

$$\mathcal{J}_R(k^{1/\gamma}W, kZ) = 1 - \mathbb{1}_\lambda \eta (kZ)(k^{1/\gamma}W)^{-\gamma} = 1 - \mathbb{1}_\lambda \eta Z W^{-\gamma} = \mathcal{J}_R(W,Z) \tag{A7}$$

This property is necessary because the set of intervention points must be preserved under the

transformation that leaves the FOC constraint unchanged, hence the derivative of the general solution must also be invariant:

$$\mathcal{J}_R(k^{1/\gamma}R, kZ) = \mathcal{J}_R(R, Z),\ \forall k > 0 \tag{A8}$$

Applying this insight to equation (A6) yields the system of differential equations

$$C_1'(R) = k^{\beta_1} C_1'(k^{1/\gamma}R) \ \text{ and } \ C_2'(R) = k^{\beta_2} C_2'(k^{1/\gamma}R).$$

It is not difficult to deduce an analytical solution to these functional equations by using a general power-law ansatz of the form

$$C_1'(R) = A_1 R^{\alpha_1} \ \text{ and } \ C_2'(R) = A_2 R^{\alpha_2}, \tag{A9}$$

where A_1, A_2 are real constants depending on all of model's parameters, and the exponents α_1, α_2 are determined just below. Substituting back the ansätze gives

$$\begin{aligned} A_1 R^{\alpha_1} &= k^{\beta_1} A_1 (k^{1/\gamma}R)^{\alpha_1} = A_1 k^{\beta_1 + \alpha_1/\gamma} R^{\alpha_1}, \\ A_2 R^{\alpha_2} &= k^{\beta_2} A_2 (k^{1/\gamma}R)^{\alpha_2} = A_2 k^{\beta_2 + \alpha_2/\gamma} R^{\alpha_2}. \end{aligned}$$

The above need to hold for all $k > 0$, so it must be that

$$\begin{aligned} 1 = k^{\beta_1 + \alpha_1/\gamma} &\iff \alpha_1 = -\gamma\beta_1, \\ 1 = k^{\beta_2 + \alpha_2/\gamma} &\iff \alpha_2 = -\gamma\beta_2. \end{aligned}$$

Therefore, $C_1'(R) = A_1 R^{-\gamma\beta_1}$ and $C_2'(R) = A_2 R^{-\gamma\beta_2}$. Integrating over R yields

$$C_1(R) = \int A_1 R^{-\gamma\beta_1} dR = \frac{A_1}{1-\gamma\beta_1} R^{1-\gamma\beta_1} + B_1, \tag{A10}$$

$$C_2(R) = \int A_2 R^{-\gamma\beta_2} dR = \frac{A_2}{1-\gamma\beta_2} R^{1-\gamma\beta_2} + B_2. \tag{A11}$$

One can verify that the ansätze satisfy the functional equations $C_i'(R) = A_i R^{-\gamma\beta_i}$, for $i = 1, 2$. Setting $\Omega_1(r) \triangleq \frac{A_1}{1-\gamma\beta_1}$ and $\Omega_2 \triangleq \frac{A_2}{1-\gamma\beta_2}$ yields the stated result.

Finally we prove that the constants of integration B_1 and B_2 must both be equal to zero, by showing that otherwise the transversality condition is violated. For geometric Brownian motion, and for any β_i solving $\frac{1}{2}\sigma^2\beta(\beta-1) + (\mu+\pi)\beta - (\rho+\delta+\pi) = 0$, one can show that the following discounted moment is a martingale: $\mathbb{E}\left[e^{-(\rho+\delta+\pi)(t-s)} Z_t^{\beta_i} \,\middle|\, \mathcal{F}_s\right] = Z_s^{\beta_i}, \forall\, t \geq s$. This follows either from the explicit GBM solution or from Itô's lemma applied to Z_t^{β}. Hence the discounted contribution of the homogeneous terms $B_1 Z_t^{\beta_1} + B_2 Z_t^{\beta_2}$ is time-invariant and equals $B_1 Z_s^{\beta_1} + B_2 Z_s^{\beta_2}$. Since the TVC $\lim_{T\to\infty} \mathbb{E}\left[e^{-(\rho+\delta+\pi)(T-s)} \mathcal{J}(R_{T\wedge\tau}, Z_{T\wedge\tau}) \,\middle|\, \mathcal{F}_s\right] = 0$ must hold for all initial states (R_s, Z_s), it must be that $B_1 = B_2 = 0$. ■

Proof of Proposition 3. To derive the optimal wage policy, it is useful to define $U \triangleq \widetilde{W}^{-\gamma} Z$ if $\widetilde{W} > R$, and $L \triangleq \widetilde{W}^{-\gamma} Z$ if $\widetilde{W} < R$. This allows to express $\mathcal{J}_R(W, Z)$ (whose general form is established in Proposition 2) as $\mathcal{J}_R(W, Z) = -\frac{1}{\rho+\delta+\pi} + (1-\gamma\beta_1)\Omega_1 x^{\beta_1} + (1-\gamma\beta_2)\Omega_2 x^{\beta_2}$, for $x = U, L$. Therefore, the necessary first-order conditions for the optimal wage $\widetilde{W}(R, Z)$ solve

$$(1-\gamma\beta_1)\Omega_1 U^{\beta_1} + (1-\gamma\beta_2)\Omega_2 U^{\beta_2} + \eta U - \frac{1+\rho+\delta+\pi}{\rho+\delta+\pi} = 0, \text{ for } \widetilde{W} > R, \tag{A12}$$

$$(1-\gamma\beta_1)\Omega_1 L^{\beta_1} + (1-\gamma\beta_2)\Omega_2 L^{\beta_2} + \lambda\eta L - \frac{1+\rho+\delta+\pi}{\rho+\delta+\pi} = 0, \text{ for } \widetilde{W} < R. \tag{A13}$$

These are non-linear *algebraic* equations, one for each adjustment region, whose roots are the *real constants* U and L. Then, since $\widetilde{W}^{-\gamma} Z = U$ or $\widetilde{W}^{-\gamma} Z = L$, the implied optimal wage policy at time t takes the form

$$\widetilde{W}(R, Z) = (Z_t/U)^{1/\gamma}, \text{ for } \widetilde{W} > R, \tag{A14}$$

$$\widetilde{W}(R, Z) = (Z_t/L)^{1/\gamma}, \text{ for } \widetilde{W} < R. \tag{A15}$$

Consider now the boundaries of the inaction region, $Z_U(R)$ and $Z_L(R)$. These are the productivity cutoff values (for any given reference wage) where the firm is indifferent between staying in the inaction region—keeping the wage equal to R—and moving to the action region—adjusting the wage to $\widetilde{W}(R, Z)$. On these free boundaries, the value-matching conditions, $\mathcal{J}(R, Z_U) = \mathcal{M}\mathcal{J}(R, Z_U)$ and $\mathcal{J}(R, Z_L) = \mathcal{M}\mathcal{J}(R, Z_L)$, must hold:

$$\mathcal{J}(R, Z_U) = \mathcal{J}(\widetilde{W}, Z_U) + G(\widetilde{W}, R, Z_U), \ \forall R \in \mathcal{R},$$
$$\mathcal{J}(R, Z_L) = \mathcal{J}(\widetilde{W}, Z_L) + G(\widetilde{W}, R, Z_L), \ \forall R \in \mathcal{R}.$$

We need to substitute in the expression for $\mathcal{J}(R, Z)$ derived in Proposition 2 and evaluated it at $Z = Z_U, Z_L$, as well as the corresponding expressions evaluated at $\widetilde{W}(R, Z_U)$ and $\widetilde{W}(R, Z_L)$. After a few algebraic manipulations we get

$$\sum_{i=1}^{2} \Omega_i Z_U^{\beta_i} \left(R^{\theta_i} - \left(\frac{Z_U}{U} \right)^{\frac{\theta_i}{\gamma}} \right) = K_1 \left(R - \left(\frac{Z_U}{U} \right)^{1/\gamma} \right) + K_2 Z_U \left(\left(\frac{Z_U}{U} \right)^{\frac{1-\gamma}{\gamma}} - R^{1-\gamma} \right), \tag{A16}$$

$$\sum_{i=1}^{2} \Omega_i Z_L^{\beta_i} \left(R^{\theta_i} - \left(\frac{Z_L}{L} \right)^{\frac{\theta_i}{\gamma}} \right) = K_1 \left(R - \left(\frac{Z_L}{L} \right)^{1/\gamma} \right) + \lambda K_2 Z_L \left(\left(\frac{Z_L}{L} \right)^{\frac{1-\gamma}{\gamma}} - R^{1-\gamma} \right), \tag{A17}$$

where $\theta_i = 1 - \gamma\beta_i$, $K_1 = \frac{1+\rho+\delta+\pi}{\rho+\delta+\pi}$, $K_2 = \frac{\eta}{1-\gamma}$ are defined as such just for notational convenience. For the value-matching conditions to be satisfied for any $R \in \mathcal{R}$, it is clear that the free boundaries must take the form

$$Z_U(R) = U R^{\gamma}, \ \forall R \in \mathcal{R}, \tag{A18}$$

$$Z_L(R) = L R^{\gamma}, \ \forall R \in \mathcal{R}. \tag{A19}$$

Next we show that the smooth-pasting conditions at the boundaries $Z_U(R)$ and $Z_L(R)$ are *automatically* satisfied as long as the FOCs (A12) and (A13) hold. By taking the first derivative w.r.t. Z in equations (A16) and (A17), substituting in the boundaries $Z_U = UR^\gamma$ and $Z_L = LR^\gamma$, and rearranging, we obtain

$$(1-\gamma\beta_1)\Omega_1 U^{\beta_1} + (1-\gamma\beta_2)\Omega_2 U^{\beta_2} + \eta U - \frac{1+\rho+\delta+\pi}{\rho+\delta+\pi} = 0, \text{ for } \widetilde{W} > R,$$
$$(1-\gamma\beta_1)\Omega_1 L^{\beta_1} + (1-\gamma\beta_2)\Omega_2 L^{\beta_2} + \lambda\eta L - \frac{1+\rho+\delta+\pi}{\rho+\delta+\pi} = 0, \text{ for } \widetilde{W} < R.$$

These are *exactly* the same equations as in (A12) and (A13). Therefore, smooth-pasting at both free boundaries holds—the first derivative of the value function is always continuous at $Z_U(r)$ and $Z_L(r)$ for any R—as long as the FOCs hold.

Finally we show how to deduce unique closed-form solutions for the constants Ω_1 and Ω_2 by making use of the FOCs (A12) and (A13). The idea is that, for any U and L, the FOCs form a *linear* 2×2 system in Ω_1 and Ω_2, so we may effectively solve it using Cramer's rule. Define the following constants for notational convenience: $K \triangleq \frac{1+\rho+\delta+\pi}{\rho+\delta+\pi}$; $K_L \triangleq K - \eta\lambda L$; $K_U \triangleq K - \eta U$. Now we can compactly express the FOC system in matrix form,

$$\begin{pmatrix} L^{\beta_1} & L^{\beta_2} \\ U^{\beta_1} & U^{\beta_2} \end{pmatrix} \begin{pmatrix} \Omega_1(1-\gamma\beta_1) \\ \Omega_2(1-\gamma\beta_2) \end{pmatrix} = \begin{pmatrix} K_L \\ K_U \end{pmatrix}. \tag{A20}$$

The coefficient matrix is non-singular because its determinant is always non-zero,

$$\Delta \triangleq \det\begin{pmatrix} L^{\beta_1} & L^{\beta_2} \\ U^{\beta_1} & U^{\beta_2} \end{pmatrix} = L^{\beta_1}U^{\beta_2} - U^{\beta_1}L^{\beta_2}. \tag{A21}$$

since $\beta_1 \neq \beta_2$ and $U \neq L$. We next calculate the determinants of the replacement matrices:

$$\Delta_1 \triangleq \det\begin{pmatrix} K_L & L^{\beta_2} \\ K_U & U^{\beta_2} \end{pmatrix} = K_L U^{\beta_2} - K_U L^{\beta_2}, \tag{A22}$$
$$\Delta_2 \triangleq \det\begin{pmatrix} L^{\beta_1} & K_L \\ U^{\beta_1} & K_U \end{pmatrix} = L^{\beta_1}K_U - U^{\beta_1}K_L. \tag{A23}$$

There exists therefore a unique solution vector, (Ω_1, Ω_2), which we can obtain by Cramer's rule:

$$\Omega_1 = \frac{\Delta_1}{(1-\gamma\beta_1)\Delta} = \frac{U^{\beta_2}\left(\frac{1+\rho+\delta+\pi}{\rho+\delta+\pi} - \eta\lambda L\right) - L^{\beta_2}\left(\frac{1+\rho+\delta+\pi}{\rho+\delta+\pi} - \eta U\right)}{(1-\gamma\beta_1)\left(L^{\beta_1}U^{\beta_2} - U^{\beta_1}L^{\beta_2}\right)}, \tag{A24}$$
$$\Omega_2 = \frac{\Delta_2}{(1-\gamma\beta_2)\Delta} = \frac{L^{\beta_1}\left(\frac{1+\rho+\delta+\pi}{\rho+\delta+\pi} - \eta U\right) - U^{\beta_1}\left(\frac{1+\rho+\delta+\pi}{\rho+\delta+\pi} - \eta\lambda L\right)}{(1-\gamma\beta_2)\left(L^{\beta_1}U^{\beta_2} - U^{\beta_1}L^{\beta_2}\right)}. \tag{A25}$$

■

Proof of Proposition 4. First off we need to determine a minimally sufficient function space for the purposes of the proof. To that end we work within the admissible class $\mathcal{C}_p$ of polynomially growing continuous functions on $\mathcal{X}$; see Fleming and Soner (2006) for a rigorous treatment. We also use the notation $\mathcal{C}_p^2$ to denote the class of twice continuously differentiable polynomially growing functions. For further convenience, let us define the proper degenerate-elliptic *HJB operator* $\mathcal{H}(.)$ for any function $V(R,Z) \in \mathcal{C}_p$ as

$$\mathcal{H}\big(V, V_z, V_{zz}\big) \triangleq (\rho+\delta+\pi)V - (\mu+\pi)ZV_Z - \frac{1}{2}\sigma^2 Z^2 V_{ZZ} - (Z\bar{e} - R). \tag{A26}$$

The HJBQVI (15) can be compactly expressed as

$$\min\Big\{\mathcal{H}\big(\mathcal{J}, \mathcal{J}_Z, \mathcal{J}_{ZZ}\big),\ \mathcal{J}(R,Z) - \mathcal{M}\mathcal{J}(R,Z)\Big\} = 0, \quad \forall (R,Z) \in \mathcal{X}, \tag{A27}$$

where the intervention operator $\mathcal{M}$ is defined in (13). Note that under the model assumptions, $\mathcal{M} : \mathcal{C}_p \to \mathcal{C}_p$ is order-preserving and pointwise continuous on $\mathcal{C}_p$ (the latter follows from Berge's maximum theorem). It is straightforward to check that the candidate value function satisfies the following regularity conditions:

- $\widetilde{\mathcal{J}} \in \mathcal{C}_p$, since there exists a constant $K > 0$ and exponents $p, q \geq \max\{1/\gamma, \beta_1\}$ such that $\big|\widetilde{\mathcal{J}}(R,Z)\big| \leq K\left(1 + R^p + Z^q\right), \forall (R,Z) \in \mathcal{X}$.
- $\widetilde{\mathcal{J}} \in \mathcal{C}_p^2$ on the interior of the inaction region, $\mathcal{I}/\{Z_L(R), Z_U(R)\}$, since it solves the HJB equation $\mathcal{H}(\widetilde{\mathcal{J}}, \widetilde{\mathcal{J}}_Z, \widetilde{\mathcal{J}}_{ZZ}) = 0$ in the *classical sense*, as proved in Proposition 2.
- $\widetilde{\mathcal{J}} \in \mathcal{C}_p^2$ on the interior of the action region, $\mathcal{A}/\{Z_L(R), Z_U(R)\}$.
- $\widetilde{\mathcal{J}} \in \mathcal{C}_p^1$ on the free boundaries $Z = Z_L(R)$ and $Z = Z_U(R)$ for any $R \geq \underline{W}$; however, $\widetilde{\mathcal{J}}_{ZZ}$ is not necessarily continuously differentiable on these points.
- $\widetilde{\mathcal{J}}$ satisfies the value-matching and smooth-pasting conditions on both free boundaries, as well as value matching (but not smooth pasting) on the layoff boundary.

(Part 1) We prove that the candidate value function is a *classical solution* of the HJBQVI almost everywhere on $\mathcal{X}$—that is, it satisfies the HJBQVI pointwise everywhere in the interior of $\mathcal{X}$, except for the free boundaries and the layoff boundary—if and only if the constants U and L each satisfy a simple parameter condition.

To show under what conditions the HJBQVI is satisfied, we divide the domain into three sets, $(Z_U(R), \infty)$, $(\underline{Z}(R), Z_L(R))$, and $(0, \underline{Z}(R)$, and consider them separately.

(1a) $Z \in (Z_U(R), \infty)$. By construction of the candidate value function, $\mathcal{J}^U = \mathcal{M}\mathcal{J}$ for any $R \in \mathcal{R}$ and for any $Z \in (Z_U(R), \infty)$. What remains to be proved is

$$\mathcal{H}(\mathcal{J}^U, \mathcal{J}_Z^U, \mathcal{J}_{ZZ}^U) \geq 0, \quad \forall Z \in (Z_U(R), \infty), \forall R \in \mathcal{R}. \tag{A28}$$

To compute the derivatives we must use the envelope theorem, so we differentiate w.r.t. Z by holding W constant and then evaluate at the optimal wage policy $\widetilde{W}$:

$$\mathcal{J}_Z^U(R,Z;\widetilde{W}) = \frac{\eta}{1-\gamma}\Big(\widetilde{W}^{1-\gamma} - R^{1-\gamma}\Big) + \frac{\bar{e}}{\rho+\delta-\mu} + \sum_{i=1}^{2}\Omega_i\beta_i Z^{\beta_i-1}\widetilde{W}^{1-\gamma\beta_i}, \tag{A29}$$

$$\mathcal{J}_{ZZ}^U(R,Z;\widetilde{W}) = \sum_{i=1}^{2}\Omega_i\beta_i(\beta_i-1)Z^{\beta_i-2}\widetilde{W}^{1-\gamma\beta_i} \tag{A30}$$

Plugging the above into the HJB operator, the terms $Z\bar{e}/(\rho+\delta-\mu)$ cancel out and the Ω_i-terms vanish because β_i solve $\frac{1}{2}\sigma^2\beta(\beta-1)-(\mu+\pi)\beta-(\rho+\delta+\pi)=0$. Substituting in the optimal policy $\widetilde{W}=(Z/U)^{1/\gamma}$, the resulting expression becomes

$$\mathcal{H}(\mathcal{J}^U,\mathcal{J}_Z^U,\mathcal{J}_{ZZ}^U) = (\rho+\delta-\mu)\eta Z\frac{\widetilde{W}^{1-\gamma}-R^{1-\gamma}}{1-\gamma} - (1+\rho+\delta+\pi)(\widetilde{W}-R). \tag{A31}$$

Now consider the function $q(y)\triangleq y^{1-\gamma}$, which is increasing and concave in y. As a result, $q(\widetilde{W})-q(R)\geq q'(\widetilde{W})(\widetilde{W}-R)=(1-\gamma)\widetilde{W}^{-\gamma}(\widetilde{W}-R)$, and therefore,

$$\frac{\widetilde{W}^{1-\gamma}-R^{1-\gamma}}{1-\gamma}\geq\widetilde{W}^{-\gamma}(\widetilde{W}-R).$$

Substituting back into the expression for $\mathcal{H}$ and using $Z\widetilde{W}^{-\gamma}=U$ yields

$$\mathcal{H}(\mathcal{J}^U,\mathcal{J}_Z^U,\mathcal{J}_{ZZ}^U)\geq\Big[(\rho+\delta-\mu)\eta U-(1+\rho+\delta+\pi)\Big](\widetilde{W}-R). \tag{A32}$$

On the upward adjustment region $Z>Z_U(R)$, we have by construction $(\widetilde{W}-R)\geq 0$. Hence a sufficient condition for $\mathcal{H}(\mathcal{J}^U,\mathcal{J}_Z^U,\mathcal{J}_{ZZ}^U)\geq 0$ for all $Z>Z_U(R)$ and all $R\in\mathcal{R}$ is

$$(\rho+\delta-\mu)\eta U\geq 1+\rho+\delta+\pi. \tag{A33}$$

It is easy to show that (A33) is also a necessary condition. Fix $R>0$ and consider the Z-sequence approaching the free boundary $Z_U(r)$ from above, $Z_\varepsilon\triangleq U\,(R+\varepsilon)^\gamma$, with $\varepsilon\downarrow 0$, so that $\widetilde{W}=(Z_\varepsilon/U)^{1/\gamma}=R+\varepsilon$ and we remain in the upward adjustment region. Substituting into the expression for $\mathcal{H}$ and using $Z_\varepsilon\widetilde{W}^{-\gamma}=U$ gives

$$\mathcal{H}(\mathcal{J}^U,\mathcal{J}_Z^U,\mathcal{J}_{ZZ}^U)=\Big[(\rho+\delta-\mu)\eta U-(1+\rho+\delta+\pi)\Big]\varepsilon+o(\varepsilon).$$

If $(\rho+\delta-\mu)\eta U-(1+\rho+\delta+\pi)<0$, the right-hand side is strictly negative for some sufficiently small $\varepsilon>0$, contradicting the requirement $\mathcal{H}\geq 0$ arbitrarily close to the boundary. Therefore, the relevant necessary and sufficient condition is

$$U\geq\frac{1+\rho+\delta+\pi}{(\rho+\delta-\mu)\eta}\iff\mathcal{H}(\mathcal{J}^U,\mathcal{J}_Z^U,\mathcal{J}_{ZZ}^U)\geq 0,\ \forall Z\geq Z_U(R),\forall R\in\mathcal{R}. \tag{A34}$$

(1b) $Z \in (\underline{Z}(R), Z_L(R))$. By construction of the candidate value function, $\mathcal{J}^L = \mathcal{M}\mathcal{J}$ for any $R \in \mathcal{R}$ and for any $Z \in (\underline{Z}(R), Z_L(R))$. What remains to be proved is

$$\mathcal{H}(\mathcal{J}^L, \mathcal{J}^L_Z, \mathcal{J}^L_{ZZ}) \geq 0, \ \ \forall Z \in (\underline{Z}, Z_L(R)), \forall R \in \mathcal{R}. \tag{A35}$$

We follow exactly the same procedure as in part **(1a)** to derive

$$\mathcal{H}(\mathcal{J}^L, \mathcal{J}^L_Z, \mathcal{J}^L_{ZZ}) = (\rho + \delta - \mu)\eta\lambda Z \frac{\widetilde{W}^{1-\gamma} - R^{1-\gamma}}{1-\gamma} - (1 + \rho + \delta + \pi)(\widetilde{W} - R). \tag{A36}$$

On the downward adjustment region we have by construction $(\widetilde{W} - R) \leq 0$, so the derivations of part **(1a)** should be carried out with the inequality sign reversed. Therefore, the relevant necessary and sufficient condition is

$$L \leq \frac{1 + \rho + \delta + \pi}{(\rho + \delta - \mu)\eta\lambda} \iff \mathcal{H}(\mathcal{J}^L, \mathcal{J}^L_Z, \mathcal{J}^L_{ZZ}) \geq 0, \ \forall Z \in (\underline{Z}, Z_L(R)), \forall R \in \mathcal{R}. \tag{A37}$$

(1c) $Z \in (Z_L(R), Z_U(R))$. By construction of the candidate value function, $\mathcal{J} = \overline{\mathcal{J}}$ for any $Z \in (Z_L(R), Z_U(R))$ and for any $R \in \mathcal{R}$. Recall that $\overline{\mathcal{J}}$ solves $\mathcal{H}(\overline{\mathcal{J}}, \overline{\mathcal{J}}_Z, \overline{\mathcal{J}}_{ZZ}) = 0$ by Proposition 2. What remains to be proved is the obstacle inequality

$$\overline{\mathcal{J}} \geq \mathcal{M}\mathcal{J}, \ \ \forall Z \in (Z_L(R), Z_U(R)), \forall R \in \mathcal{R}. \tag{A38}$$

We do so by testing upward and downward impulses and propagating the boundary dominance into the interior through the degenerate-elliptic linear operator $\mathcal{H}(.)$. This technique is justified by an optional stopping argument, or equivalently, by the weak maximum principle.

For $B \in \{U, L\}$, define $\kappa^B \triangleq 1$ if $B = U$ and $\kappa^B \triangleq \lambda$ if $B = L$; $\widetilde{W}^B(Z) \triangleq (Z/B)^{1/\gamma}$; $s^B \triangleq \frac{\widetilde{W}_B}{R}$; and the secant-slope factor $\Phi(s) \triangleq s^\gamma \frac{s^{1-\gamma}-1}{(1-\gamma)(s-1)}$; so that on the inaction region we have $s^U \in (0,1)$ and $s^L > 1$, with $Z = Bs^\gamma R^\gamma$. The concavity of $q(y) \triangleq y^{1-\gamma}$ and the mean value theorem imply the useful bounds $\min\{1, s^\gamma\} \leq \Phi(s) \leq \max\{1, s^\gamma\}$. Next we define the impulse-inaction differences

$$\Delta^B(R, Z) \triangleq \mathcal{J}^B(R, Z; \widetilde{W}_B(Z)) - \overline{\mathcal{J}}(R, Z), \ \forall B \in \{U, L\}. \tag{A39}$$

By construction and value matching at the free boundaries,

$$\Delta^B(R, Z_B(R)) = 0 \ \text{ and } \ \Delta^B(R, Z_{B'}(R)) \leq 0 \ \ \forall B' \neq B. \tag{A40}$$

Using the envelope theorem for $\mathcal{J}^B$ once again, we obtain the identity

$$\begin{aligned}\mathcal{H}\left(\Delta^B_Z, \Delta^B_Z, \Delta^B_{ZZ}\right) &= \mathcal{H}\left(\mathcal{J}^B, \mathcal{J}^B_Z, \mathcal{J}^B_{ZZ}\right) - \underbrace{\mathcal{H}\left(\overline{\mathcal{J}}, \overline{\mathcal{J}}_Z, \overline{\mathcal{J}}_{ZZ}\right)}_{=0} \\ &= R(s^B - 1)\left[(\rho + \delta - \mu)\eta\kappa^B B\Phi(s^B) - (1 + \rho + \delta + \pi)\right].\end{aligned} \tag{A41}$$

Therefore, the condition

$$\mathcal{H}(\Delta^B) \geq 0, \quad \forall Z \in \big(Z_L(R), Z_U(R)\big) \tag{A42}$$

holds for both $B \in \{U, L\}$ *if and only if* the following bounds hold:

$$(\rho + \delta - \mu)\eta U \leq 1 + \rho + \delta + \pi \quad \text{and} \quad (\rho + \delta - \mu)\eta\lambda L \geq 1 + \rho + \delta + \pi. \tag{A43}$$

Now let $\bar{\tau}$ be the first exit time of Z_t from the interval $\big(Z_L(R), Z_U(R)\big)$. Applying Itô's formula to $e^{-(\rho+\delta+\pi)t}\Delta^B(R, Z_t)$ up to $\bar{\tau}$ gives

$$\mathbb{E}\left[e^{-(\rho+\delta+\pi)\tau}\Delta^B(R, Z_\tau)\right] - \Delta^B(R, Z_0) = \mathbb{E}\int_0^\tau e^{-(\rho+\delta+\pi)s}\,\mathcal{H}\left(\Delta^B_Z, \Delta^B_Z, \Delta^B_{ZZ}\right)(R, Z_s)ds. \tag{A44}$$

If (A43) holds, then by (A42) the r.h.s of (A44) is ≥ 0, so $e^{-(\rho+\delta+\pi)t}\Delta^B(R, Z_t)$ is a submartingale up to $\bar{\tau}$. Using the boundary properties (A40) together with the optional stopping theorem, we deduce that $\Delta^B(R, Z_0) < 0$ for all interior Z_0; that is,

$$\mathcal{J}^B(R, Z) < \overline{\mathcal{J}}(R, Z) \Longrightarrow \overline{\mathcal{J}} > \mathcal{M}\mathcal{J}, \quad \forall Z \in \big(Z_L(R), Z_U(R)\big), \ \forall B \in \{U, L\}, \tag{A45}$$

Conversely, suppose (A38) holds but one of the bounds in (A43) is violated. Take $Z \uparrow Z_B(R)$; then (A41) yields $\mathcal{H}(\Delta^B) < 0$ on a neighborhood of $Z_B(R)$. The r.h.s. of (A44) is then strictly negative for points sufficiently close to $Z_B(R)$, while the boundary term is ≤ 0 by (A40), forcing $\Delta^B(R, Z_0) > 0$ for some interior Z_0—this is a contradiction to (A38). Hence the conditions (A43) are also necessary.

Conclusion. Combining the action-side conditions from parts **(1a)-(1b)** with the inaction-side conditions from part **(1c)**, yields the sharp conditions

$$L = \frac{1 + \rho + \delta + \pi}{(\rho + \delta - \mu)\eta\lambda} \quad \text{and} \quad U = \frac{1 + \rho + \delta + \pi}{(\rho + \delta - \mu)\eta}, \tag{A46}$$

which are necessary and sufficient for the candidate value function to satisfy the HJBQVI pointwise everywhere in the interior of $\mathcal{X}$, i.e., to be a classical solution almost everywhere on $\mathcal{X}$.

(Part 2) We prove that the candidate value function is a *viscosity solution* of the HJBQVI everywhere on $\mathcal{X}$. We henceforth assume that conditions (A46) hold. Obviously $\mathcal{J}$ is a viscosity solution on the interior of each region since it is a classical solution there by **(Part 1)**, it remains to verify the viscosity subsolution and supersolution properties at the free boundaries $Z_B(R) = BR^\gamma, \forall B \in \{U, L\}$. We begin by reminding the reader of three relevant definitions:

$\triangleright$ A function $V \in \mathcal{C}_p$ is a *viscosity subsolution* of the HJBQVI if for any $(R_0, Z_0) \in \mathcal{X}$ and any $\Phi \in \mathcal{C}_p^2$ such that $V - \Phi$ attains a local maximum 0 at (R_0, Z_0), we have

$$\min\Big\{\mathcal{H}\big(\Phi, \Phi_z, \Phi_{zz}\big),\, V(R_0, Z_0) - \mathcal{M}V(R_0, Z_0)\Big\} \leq 0.$$

$\triangleright$ A function $V \in \mathcal{C}_p$ is a *viscosity supersolution* of the HJBQVI if for any $(R_0, Z_0) \in \mathcal{X}$ and any $\Phi \in \mathcal{C}_p^2$ such that $V - \Phi$ attains a local minimum 0 at (R_0, Z_0), we have

$$\min\Big\{\mathcal{H}\big(\Phi, \Phi_z, \Phi_{zz}\big),\ V(R_0, Z_0) - \mathcal{M}V(R_0, Z_0)\Big\} \geq 0.$$

$\triangleright$ A function $V \in \mathcal{C}_p$ is *viscosity solution* of the HJBQVI if it is *both* a viscosity subsolution and a viscosity supersolution of the HJBQVI in $\mathcal{X}$.

(2a) *Subsolution property at $Z = Z_B(R)$.* Fix $(R_0, Z_0) = (R_0, Z_B(R_0))$ and consider a test function $\Phi \in \mathcal{C}_p^2$ touching $\mathcal{J}$ from above at (R_0, Z_0). From the value-matching matching conditions at the free boundaries, $\mathcal{J}(R, Z_B(R)) = \mathcal{J}^B(R, Z_B(R)) = \mathcal{M}\mathcal{J}(R, Z_B(R))$, we have $\mathcal{J}(R_0, Z_0) - (\mathcal{M}\mathcal{J})(R_0, Z_0) = 0$, hence

$$\min\{\mathcal{H}(\Phi, \Phi_Z, \Phi_{ZZ})(R_0, Z_0),\ 0\} \ \leq\ 0 \tag{A47}$$

holds trivially regardless of $\mathcal{H}(\Phi, \Phi_Z, \Phi_{ZZ})$. This proves that $\mathcal{J}$ is a viscosity subsolution at both $Z = Z_U(R)$ and $Z = Z_L(R)$.

(2b) *Supersolution property at $Z = Z_B(R)$.* Fix $(R_0, Z_0) = (R_0, Z_B(R_0))$ and consider a test function $\Phi \in \mathcal{C}_p^2$ touching $\mathcal{J}$ from below at (R_0, Z_0). As before, because of the value-matching matching conditions at the free boundaries, the obstacle term is zero; hence it suffices to prove $\mathcal{H}(\Phi, \Phi_Z, \Phi_{ZZ})(R_0, Z_0) \geq 0$. Let $\kappa_B = +1$ if $B = U$ and $\kappa_B = -1$ if $B = L$. Choose $\delta_\varepsilon > 0$ so that $I_B \triangleq \big\{Z :\ Z_0 < Z < Z_0 + \delta_\varepsilon\big\}$ if $B = U$, $I_B \triangleq \big\{Z :\ Z_0 - \delta_\varepsilon < Z < Z_0\big\}$ if $B = L$ lie inside the action region, on which $\mathcal{J} = \mathcal{J}^B$ and $\mathcal{J}^B$ is $\mathcal{C}_p^2$. Pick $\varepsilon_k \downarrow 0$ and define

$$Z_k \triangleq Z_0 + \kappa_B \varepsilon_k \in I_B, \quad Z_k \to Z_0 \text{ from each action side.}$$

Since Φ touches $\mathcal{J}$ from below at (R_0, Z_0) and $\mathcal{J}$ is C^1 across the free boundary (smooth pasting), the difference $\mathcal{J}(R_0, Z) - \Phi(R_0, Z)$ has a one-sided local minimum 0 at Z_0; hence

$$\Phi_Z(R_0, Z_0) = \mathcal{J}_Z(R_0, Z_0) = \mathcal{J}_Z^B(R_0, Z_0). \tag{A48}$$

Define $D(Z) \triangleq \mathcal{J}^B(R_0, Z) - \Phi(R_0, Z)$ on I_B. Then $D(Z_k) \geq 0$ since $\Phi \leq \mathcal{J}$ and $\mathcal{J} = \mathcal{J}^B$ on the action side, and $D(Z_0) = 0$. By Taylor's theorem, for each k there exist points

$$\begin{gathered}
\xi_k, \widetilde{\xi}_k \in \big(Z_0, Z_k\big) \ \text{ if } B = U, \quad \xi_k, \widetilde{\xi}_k \in \big(Z_k, Z_0\big) \ \text{ if } B = L, \ \text{ such that} \\
\mathcal{J}^B(R_0, Z_k) = \mathcal{J}^B(R_0, Z_0) + \mathcal{J}_Z^B(R_0, Z_0)\,(Z_k - Z_0) + \tfrac{1}{2}\,\mathcal{J}_{ZZ}^B(R_0, \xi_k)\,(Z_k - Z_0)^2, \\
\Phi(R_0, Z_k) = \Phi(R_0, Z_0) + \Phi_Z(R_0, Z_0)\,(Z_k - Z_0) + \tfrac{1}{2}\,\Phi_{ZZ}(R_0, \widetilde{\xi}_k)\,(Z_k - Z_0)^2.
\end{gathered}$$

Subtracting and using $\mathcal{J}^B(R_0, Z_0) = \Phi(R_0, Z_0)$ and (A48), we obtain

$$D(Z_k) = \frac{1}{2}\Big(\mathcal{J}_{ZZ}^B(R_0, \xi_k) - \Phi_{ZZ}(R_0, \widetilde{\xi}_k)\Big)\,(Z_k - Z_0)^2.$$

Since $D(Z_k) \geq 0$ and $(Z_k - Z_0)^2 > 0$, it follows that

$$\mathcal{J}^B_{ZZ}(R_0, \xi_k) \;\geq\; \Phi_{ZZ}(R_0, \widetilde{\xi}_k).$$

Because $\xi_k, \widetilde{\xi}_k \to Z_0$ from the action side and Φ_{ZZ} is continuous, we have

$$\Phi_{ZZ}(R_0, Z_0) \leq \liminf_{Z \to Z_0} \mathcal{J}^B_{ZZ}(R_0, Z). \tag{A49}$$

Now let $\hat{\rho} \triangleq \rho + \delta + \pi$ and $\hat{\mu} \triangleq \mu + \pi$. Using (A48) and (A49), as well as the degenerate ellipticity (monotonicity) property of $\mathcal{H}$ in the second derivative, i.e., $\partial_{V_{ZZ}} \mathcal{H} = \frac{1}{2}\sigma^2 Z^2 > 0$, we obtain

$$\begin{aligned}
\mathcal{H}(\Phi, \Phi_Z, \Phi_{ZZ})(R_0, Z_0) &= \hat{\rho}\Phi(R_0, Z_0) - \hat{\mu} Z_0 \Phi_Z(R_0, Z_0) + \frac{1}{2}\sigma^2 Z_0^2 \Phi_{ZZ}(R_0, Z_0) - (Z_0 \bar{e} - R_0) \\
&\geq \liminf_{k \to \infty} \Big\{ \hat{\rho}\mathcal{J}^B(R_0, \xi_k) - \hat{\mu}\xi_k \mathcal{J}^B_Z(R_0, \xi_k) \\
&\qquad\qquad + \tfrac{1}{2}\sigma^2 \xi_k^2 \, \mathcal{J}^B_{ZZ}(R_0, \xi_k) - (\xi_k \bar{e} - R_0) \Big\} \\
&= \liminf_{k \to \infty} \mathcal{H}\big(\mathcal{J}^B, \mathcal{J}^B_Z, \mathcal{J}^B_{ZZ}\big)(R_0, \xi_k) \;\geq\; 0,
\end{aligned}$$

where we have used continuity of $\mathcal{J}^B$ and $\mathcal{J}^B_Z$ together with $\xi_k \to Z_0$, as well as the fact that $\mathcal{H}(\mathcal{J}^B, \mathcal{J}^B_Z, \mathcal{J}^B_{ZZ}) \geq 0$ on the action region by **(Part 1)**. Therefore, $\mathcal{H}(\Phi, \Phi_Z, \Phi_{ZZ})(R_0, Z_0) \geq 0$. Since the obstacle term equals zero at the free boundaries by value matching, we conclude that

$$\min\{\mathcal{H}(\Phi, \Phi_Z, \Phi_{ZZ})(R_0, Z_0), 0\} \geq 0. \tag{A50}$$

This proves that $\mathcal{J}$ is a viscosity supersolution at both $Z = Z_U(R)$ and $Z = Z_L(R)$.

Conclusion. We have thus proved that $\widetilde{\mathcal{J}}$ is both a viscosity subsolution and supersolution at $Z = Z_U(R)$ and $Z = Z_L(R)$, as well as a classical (hence viscosity) solution in the interior regions. Consequently, $\widetilde{\mathcal{J}}$ is a viscosity solution of the HJBQVI everywhere on $\mathcal{X}$.

(Part 3) We prove that the viscosity solution of the HJBQVI is unique within the relevant class $\mathcal{C}_p$ using a global *comparison principle* via the *doubling-of-variables* method, a key technique for establishing uniqueness of solutions to second-order PDEs. The core of the argument follows classical references such as Crandall, Ishii, and Lions (1992).

Let $u \in \mathcal{C}_p$ be a viscosity subsolution and $v \in \mathcal{C}_p$ a viscosity supersolution of the HJBQVI (15). As explained at the start of the proof, $\mathcal{H}(.)$ is continuous and degenerate elliptic. The intervention operator is order-preserving, pointwise continuous, and translation-invariant; hence it is non-expansive in the sup-norm: $\|\mathcal{M}u - \mathcal{M}v\|_\infty \leq \|u - v\|_\infty, \;\forall u, v \in \mathcal{C}_p$. The purpose is to prove $u \leq v$ on $\mathcal{X}$. Uniqueness then follows by a global comparison principle via doubling-of-variables/Ishii's lemma, or equivalently, by the corresponding viscosity weak maximum principle, which rules out any positive interior maximum of $u - v$ and forces $u = v$ on $\mathcal{X}$.

For any $\varepsilon, \alpha > 0$, define

$$\psi(R, Z) \triangleq R^2 + Z^2$$
$$\Psi_{\varepsilon,\alpha}\big((R, Z), (R', Z')\big) \triangleq u(R, Z) - v(R', Z') - \frac{(Z - Z')^2}{2\varepsilon} - \alpha\big(\psi(R, Z) + \psi(R', Z')\big),$$
$$\Theta \triangleq \sup_{\mathfrak{X}}(u - v) \;\text{ and }\; \Theta_{\varepsilon,\alpha} \triangleq \Psi_{\varepsilon,\alpha}(X_{\varepsilon,\alpha}, Y_{\varepsilon,\alpha}).$$

By polynomial growth, $\Psi_{\varepsilon,\alpha}$ attains its maximum at $X_{\varepsilon,\alpha} = (R_{\varepsilon,\alpha}, Z_{\varepsilon,\alpha})$, $Y_{\varepsilon,\alpha} = (R'_{\varepsilon,\alpha}, Z'_{\varepsilon,\alpha})$. Moreover, as $(\varepsilon, \alpha) \to (0, 0)$, standard arguments yield

$$\Theta_{\varepsilon,\alpha} \to \Theta, \quad \frac{(Z_{\varepsilon,\alpha} - Z'_{\varepsilon,\alpha})^2}{\varepsilon} \to 0, \quad \alpha\big(|X_{\varepsilon,\alpha}|^2 + |Y_{\varepsilon,\alpha}|^2\big) \to 0. \tag{A51}$$

Define the $\mathcal{C}^2$ supertangent/subtangent test functions as

$$\phi_u(R, Z) \triangleq v(Y_{\varepsilon,\alpha}) + \frac{(Z - Z'_{\varepsilon,\alpha})^2}{2\varepsilon} + \alpha\big(\psi(R, Z) + \psi(Y_{\varepsilon,\alpha})\big), \tag{A52}$$
$$\phi_v(R', Z') \triangleq u(X_{\varepsilon,\alpha}) - \frac{(Z_{\varepsilon,\alpha} - Z')^2}{2\varepsilon} - \alpha\big(\psi(X_{\varepsilon,\alpha}) + \psi(R', Z')\big). \tag{A53}$$

Then $u - \phi_u$ has a local maximum 0 at $X_{\varepsilon,\alpha}$ and $v - \phi_v$ has a local minimum 0 at $Y_{\varepsilon,\alpha}$. By the viscosity definitions we have

$$\min\Big\{\mathcal{H}\big(u, \partial_Z\phi_u, \partial_{ZZ}\phi_u\big)(X_{\varepsilon,\alpha}), u - \mathcal{M}u\big(X_{\varepsilon,\alpha}\big)\Big\} \le 0, \tag{A54}$$
$$\min\Big\{\mathcal{H}\big(v, \partial_Z\phi_v, \partial_{ZZ}\phi_v\big)(Y_{\varepsilon,\alpha}), v - \mathcal{M}v\big(Y_{\varepsilon,\alpha}\big)\Big\} \ge 0. \tag{A55}$$

To show why the subsolution is on the PDE branch at the near-maximum, suppose it is not. Then along a subnet $(\varepsilon, \alpha) \to (0, 0)$ there exists $\theta > 0$ with $u - \mathcal{M}u \le -\theta$ at $X_{\varepsilon,\alpha}$. Then using the non-expansiveness and pointwise continuity of $\mathcal{M}$ we have

$$u(X_{\varepsilon,\alpha}) - \mathcal{M}u(X_{\varepsilon,\alpha}) \le -\theta \Longrightarrow \mathcal{M}u(X_{\varepsilon,\alpha}) \le \mathcal{M}v(X_{\varepsilon,\alpha}) + \Theta + o(1),$$
$$\text{hence } \; u(X_{\varepsilon,\alpha}) - v(Y_{\varepsilon,\alpha}) \le \Theta - \theta + o(1),$$

contradicting $u(X_{\varepsilon,\alpha}) - v(Y_{\varepsilon,\alpha}) \ge \Theta_{\varepsilon,\alpha} \to \Theta$ by (A51). This implies that, for all small (ε, α),

$$\mathcal{H}\big(u, \partial_Z\phi_u, \partial_{ZZ}\phi_u\big)(X_{\varepsilon,\alpha}) \le 0, \;\; \text{while} \;\; \mathcal{H}\big(v, \partial_Z\phi_v, \partial_{ZZ}\phi_v\big)(Y_{\varepsilon,\alpha}) \ge 0. \tag{A56}$$

By applying Ishii's lemma to $\max \Psi_{\varepsilon,\alpha}$, there exist symmetric 2×2 matrices X, Y such that

$$\big(\partial_R\phi_u, \partial_Z\phi_u, X\big) \in \mathcal{J}^{2,+}_* u(X_{\varepsilon,\alpha}), \quad \big(\partial_{R'}\phi_v, \partial_{Z'}\phi_v, Y\big) \in \mathcal{J}^{2,-}_* v(Y_{\varepsilon,\alpha}), \tag{A57}$$

where $\mathcal{J}^{2,+}_*/\mathcal{J}^{2,-}_*$ denote what Crandall-Ishii-Lions call the limiting second-order "superjet"/"subjet",

and the following block matrix inequality holds

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq \frac{3}{\varepsilon}\begin{pmatrix} E & -E \\ -E & E \end{pmatrix} + I_4 C\alpha, \quad E \triangleq \operatorname{diag}(0,1), \tag{A58}$$

where the constant C depends on ψ but not on (ε, α). Also, by construction,

$$\partial_Z \phi_u(X_{\varepsilon,\alpha}) = \frac{Z_{\varepsilon,\alpha} - Z'_{\varepsilon,\alpha}}{\varepsilon} + \mathcal{O}(\alpha), \quad \partial_{ZZ}\phi_u(X_{\varepsilon,\alpha}) = \frac{1}{\varepsilon} + 2\alpha$$
$$\partial_Z \phi_v(Y_{\varepsilon,\alpha}) = \frac{Z_{\varepsilon,\alpha} - Z'_{\varepsilon,\alpha}}{\varepsilon} + \mathcal{O}, \quad \partial_{ZZ}\phi_v(X_{\varepsilon,\alpha}) = \frac{1}{\varepsilon} - 2\alpha$$

Using (A56) along with (A57) and (A58), we obtain

$$\mathcal{H}\big(u, \partial_Z\phi_u, \partial_{ZZ}\phi_u\big)(X_{\varepsilon,\alpha}) - \mathcal{H}\big(v, \partial_Z\phi_v, \partial_{Z'Z'}\phi_v\big)(Y_{\varepsilon,\alpha}) \geq \hat{\rho}\big(u(X_{\varepsilon,\alpha}) - v(Y_{\varepsilon,\alpha})\big) - o(1) \leq 0,$$
$$o(1) \leq C\left[\frac{Z_{\varepsilon,\alpha} - Z'_{\varepsilon,\alpha}}{\varepsilon} + \alpha\big(1 + |X_{\varepsilon,\alpha}|^2 + |Y_{\varepsilon,\alpha}|^2\big) + |Z_{\varepsilon,\alpha} - Z'_{\varepsilon,\alpha}| + |R_{\varepsilon,\alpha} - R'_{\varepsilon,\alpha}|\right] \to 0.$$

This implies $\hat{\rho}\,\Theta_{\varepsilon,\alpha} - o(1) \leq 0$. Letting $(\varepsilon, \alpha) \to (0,0)$ yields $\Theta \leq 0$, as needed.

Conclusion. We have shown $\sup_{\mathcal{X}}(u - v) \leq 0$, which implies $u \leq v$ everywhere on $\mathcal{X}$. Therefore, the viscosity solution of the HJBQVI in $\mathcal{C}_p$ is unique.

Proof of Proposition 5. We go through the steps of a formal *verification theorem* tailored to our problem. Fix an initial time s and (R_s, Z_s), setting $s = 0$ w.l.o.g. Let $\widetilde{\mathcal{J}}$ and $\widetilde{W}$be the candidate value and policy functions, and for notational ease, let $\hat{\rho} = \rho + \delta + \pi$ and $\hat{\mu} = \mu + \pi$. Throughout the proof we invoke the important properties of $\widetilde{\mathcal{J}}$ established in Proposition 4.

For each integer $n \in \mathbb{N}$, define the bounded stopping time $\tau_n \triangleq \bar{\tau} \wedge n$, where $\bar{\tau}$ is the layoff/termination time. Then $\{\tau_n\}_{n\in\mathbb{N}}$ is nondecreasing and either $\tau_n \uparrow \bar{\tau} < \infty$ or $\tau_n \uparrow \infty$ almost surely. With this localization device, optional sampling applies at the bounded stopping time $t \wedge \tau_n$. As in the main text, $\{\tau_i\}_{i\geq 1}$ denotes an admissible sequence of impulse/intervention times and $\{W_i\}_{i\geq 1}$ the associated impulse control. All expectations are conditional on $\mathcal{F}_s$. For any $t \geq 0$, define the stopped verification process

$$Y_{t\wedge\tau_n} \triangleq e^{-\hat{\rho}(t\wedge\tau_n)}\widetilde{\mathcal{J}}(R_{t\wedge\tau_n}, Z_{t\wedge\tau_n}) + \int_0^{t\wedge\tau_n} e^{-\hat{\rho}u}\Pi(R_u, R_u, Z_u)du + \sum_{\tau_i \leq t\wedge\tau_n} e^{-\hat{\rho}\tau_i}G(W_i, R_{\tau_i^-}, Z_{\tau_i}). \tag{A59}$$

The process $Y_t = Y_t(W, R, Z)$ is a supermartingale for any admissible control W_t. The goal is to prove that Y_t is a proper *martingale* under the candidate impulse control $\widetilde{W}$. We then apply the optional sampling theorem at $t \wedge \tau_n$, and finally let $t \to \infty$ and $n \to \infty$, using the TVC and the boundary condition $\widetilde{\mathcal{J}} = 0$ at layoff to conclude that our candidates are indeed the true value function and the optimal impulse control.

Step 1: We apply the Itô-Tanaka-Meyer formula[17] to the term $e^{-\hat{\rho}t}\widetilde{\mathcal{J}}$ of (A59) on $[0, \tau_n]$:

$$\begin{aligned} d\big(e^{-\hat{\rho}t}\widetilde{\mathcal{J}}(R,Z)\big) =& -\hat{\rho}\, e^{-\hat{\rho}t}\widetilde{\mathcal{J}}(R,Z)dt + e^{-\hat{\rho}t}\widetilde{\mathcal{J}}_Z(R,Z)dZ_t + \tfrac{1}{2}e^{-\hat{\rho}t}\widetilde{\mathcal{J}}_{ZZ}(R,Z)d\langle Z_t\rangle \\ &+ \frac{1}{2}e^{-\hat{\rho}t}\big[\widetilde{\mathcal{J}}_Z^+(R,Z_L(R)) - \widetilde{\mathcal{J}}_Z^-(R,Z_L(R))\big]d\Lambda_t^{Z_L(R)} \\ &+ \frac{1}{2}\,e^{-\hat{\rho}t}\big[\widetilde{\mathcal{J}}_Z^+(R,Z_U(R)) - \widetilde{\mathcal{J}}_Z^-(R,Z_U(R))\big]d\Lambda_t^{Z_U(R)}. \end{aligned} \tag{A60}$$

Since Z_t follows a GBM, we have $dZ_t = \hat{\mu}Z_t dt + \sigma Z_t dB_t$ and $d\langle Z_t\rangle = \sigma^2 Z_t^2 dt$. Also, smooth pasting at the free boundaries implies $\widetilde{\mathcal{J}}_Z^+(R,Z_L(R)) = \widetilde{\mathcal{J}}_Z^-(R,Z_L(R))$ as well as $\widetilde{\mathcal{J}}_Z^+(R,Z_U(R)) = \widetilde{\mathcal{J}}_Z^-(R,Z_U(R))$, so the local-time terms vanish. Using these results along with the definition of the generator $\mathcal{L}\widetilde{\mathcal{J}} \triangleq \hat{\mu}Z\widetilde{\mathcal{J}}_Z + \frac{1}{2}\sigma^2 Z^2\widetilde{\mathcal{J}}_{ZZ}$, equation (A60) becomes

$$d\big(e^{-\hat{\rho}t}\widetilde{\mathcal{J}}(R,Z)\big) = e^{-\hat{\rho}t}\Big[-\hat{\rho}\widetilde{\mathcal{J}}(R,Z) + \mathcal{L}\widetilde{\mathcal{J}}(R,Z)\Big]\,dt + e^{-\hat{\rho}t}\widetilde{\mathcal{J}}_Z(R,Z)\sigma Z\,dB_t \tag{A61}$$

For any non-intervention time $t < \tau_n$ such that $Z_t \in \mathcal{I}$, we have

$$dY_t = e^{-\hat{\rho}t}\Big[-\hat{\rho}\widetilde{\mathcal{J}}(R_t,Z_t) + \mathcal{L}\widetilde{\mathcal{J}}(R_t,Z_t) + \Pi(R_t,R_t,Z_t)\Big]\,dt + e^{-\hat{\rho}t}\sigma Z_t\widetilde{\mathcal{J}}_Z(R_t,Z_t)dB_t. \tag{A62}$$

Since the candidate solves the HJB equation in the inaction region, $\mathcal{H}(\widetilde{\mathcal{J}},\widetilde{\mathcal{J}}_Z,\widetilde{\mathcal{J}}_{ZZ}) = 0$ and the drift in (A62) vanishes, whereas the drift is weakly negative for any other admissible policy $W_t \in \mathcal{W}$. Moreover, the stochastic integral of the second term in (A62) is a local martingale when stopped at τ_n, and a true martingale on $[0,\tau_n]$ by square-integrability on the bounded horizon, so we have $\mathbb{E}\left[\int_0^t \sigma Z_s\widetilde{\mathcal{J}}_Z(R_s,Z_s)dB_s\right] = 0$. Therefore,

$$\mathbb{E}\,[Y_t|\mathcal{F}_s] = Y_s, \quad \forall Z_t \in \mathcal{I},\, \forall t \neq \tau_i < \tau_n, \text{ if } W_t = \widetilde{W}, \tag{A63}$$

$$\mathbb{E}\,[Y_t|\mathcal{F}_s] \leq Y_s, \quad \forall Z_t \in \mathcal{I},\, \forall t \neq \tau_i < \tau_n, \,\forall\, W_t \in \mathcal{W}. \tag{A64}$$

For any intervention time $t = \tau_i < \tau_n$ such that $Z_t \in \mathcal{A}$, the discrete jump of Y_t becomes

$$\Delta Y_t = e^{-\hat{\rho}t}\Big[\mathcal{M}\widetilde{\mathcal{J}}(W_t,Z_t) - \widetilde{\mathcal{J}}(R_t,Z_t)\Big]. \tag{A65}$$

Since the candidate solves the HJBQVI in the action region, $\widetilde{\mathcal{J}} = \mathcal{M}\widetilde{\mathcal{J}}$ and the the bracket term in (A65) vanishes, whereas the same term is weakly negative for any other admissible candidate value and policy. Therefore, the discrete jump of Y_t becomes

$$\mathbb{E}\,[\Delta Y_t|\mathcal{F}_s] = 0, \quad \forall Z_t \in \mathcal{A},\, \forall t = \tau_i < \tau_n, \text{ if } W_t = \widetilde{W}, \tag{A66}$$

$$\mathbb{E}\,[\Delta Y_t|\mathcal{F}_s] \leq 0, \quad \forall Z_t \in \mathcal{A},\, \forall t = \tau_i < \tau_n, \,\forall\, W_t \in \mathcal{W}. \tag{A67}$$

By (A63)–(A66) and (A64)–(A67), we deduce that $Y_{t\wedge\tau_n}$ is a local supermartingale for any admissible policy, and a local martingale for the candidate policy. It is easy to show that these prop-

[17] Please see the note at the end of the proof for a detailed derivation.

erties extend beyond local arguments. The process $Y_{t\wedge\tau_n}$ is bounded from below, and therefore uniformly integrable. This follows from the facts that $\widetilde{\mathcal{J}}$ is non-negative and has polynomial growth, the flow profit integral converges, and the jump process at intervention times is globally bounded. As a result, the process $Y_{t\wedge\tau_n}$ is a proper *supermartingale* under any admissible policy $W_t \in \mathcal{W}$, and a proper *martingale* under the candidate policy $\widetilde{W}$.

Step 2. Before we move on with the verification, we show that the candidate $\widetilde{\mathcal{J}}$ satisfies the relevant transversality condition, provided that the parameter condition $\rho+\delta+\pi > \frac{\mu+\pi}{\gamma}+\frac{1-\gamma}{2\gamma^2}\sigma^2$ is satisfied, which we have already assumed in (8). For analytical convenience, we rewrite $\mathcal{J}^U(R_t, Z_t)$ in the upward adjustment region—since this is the only potentially explosive region along non-terminated paths—by grouping terms based on Z_t and R_t as

$$\mathcal{J}^U = (Z_t/U)^{1/\gamma} X_U + Z_t \left(\frac{\bar{e}}{\rho+\delta-\mu} - \frac{\eta}{1-\gamma} R_t^{1-\gamma} \right) + R_t, \tag{A68}$$

$$\text{where } X_U = \frac{\eta}{1-\gamma} U + \Omega_1 U^{\beta_1} + \Omega_2 U^{\beta_2} - \frac{1+\rho+\delta+\pi}{\rho+\delta+\pi}. \tag{A69}$$

Recall that the TVC for the problem is given by

$$\lim_{t\to\infty} \mathbb{E}\left[e^{-\hat{\rho}(t-s)} \mathcal{J}(R_{t\wedge\bar{\tau}}, Z_{t\wedge\bar{\tau}}) \,|\, \mathcal{F}_s \right] = 0.$$

where τ is a random termination (layoff or quit) stopping time. If $t \wedge \bar{\tau} = \min\{t, \bar{\tau}\} = \bar{\tau} < \infty$, then obviously the TVC holds because the conditional expectation is guaranteed to converge to zero given the absorbing convention $\mathcal{J} = 0$ for every $t > \bar{\tau}$. The case of interest is $\bar{\tau} = \infty$, which arises from the realization of some Z_t sequence such that the controlled state never enters the termination set and the firm-worker match is never endogenously dissolved.

To ensure that the discounted continuation value always vanishes in expectation, we first need to determine the dominant term in $\mathcal{J}$ as $t \to \infty$. Let's analyze the growth of each term. The first term $(Z_t/U)^{1/\gamma} X_U$ grows asymptotically as $Z_t^{1/\gamma}$; the second term grows as Z_t; the third term grows at rate less than $Z_t^{1/\gamma}$ because by the nature of reference wage formation the magnitude of R_t is naturally constrained by $R_t < (Z_t/U)^{1/\gamma}$. The dominant term is $Z_t^{1/\gamma}$, and thus the value function grows asymptotically at most as $\mathcal{O}(Z_t^{1/\gamma})$. Next we need to be able to evaluate the conditional expectation of $Z_t^{1/\gamma}$ for an arbitrary power p. Recall that the strong SDE solution for a geometric Brownian motion starting from Z_s at time s is

$$Z_t = Z_s \exp\left(\left(\mu + \pi - \frac{1}{2}\sigma^2 \right)(t-s) + \sigma(B_t - B_s) \right).$$

Let $\hat{\tau} = t - s$ be the time elapsed. The increment of the Wiener process $\Delta B_{\hat{\tau}} = B_t - B_s$ is distributed as $\Delta B_{\hat{\tau}} \sim \mathcal{N}(0, \hat{\tau})$. Therefore,

$$\mathbb{E}[Z_t^{1/\gamma} \,|\, \mathcal{F}_s] = \mathbb{E}\left[\left(Z_s \exp\left(\left(\mu + \pi - \frac{1}{2}\sigma^2 \right)\hat{\tau} + \sigma \Delta B_{\hat{\tau}} \right) \right)^{1/\gamma} \,|\, \mathcal{F}_s \right]$$

Since Z_s is $\mathcal{F}_s$-measurable, we can pull $Z_s^{1/\gamma}$ out of the expectation,

$$\mathbb{E}[Z_t^{1/\gamma} \mid \mathcal{F}_s] = Z_s^{1/\gamma} \mathbb{E}\left[\exp\left(\frac{1}{\gamma}\left(\mu + \pi - \frac{1}{2}\sigma^2\right)\hat{\tau} + \frac{1}{\gamma}\sigma \Delta B_{\hat{\tau}}\right)\right].$$

Let $A = (1/\gamma)\left(\mu + \pi - \frac{1}{2}\sigma^2\right)\hat{\tau}$ and $Y = (1/\gamma)\sigma\Delta B_{\hat{\tau}}$. The expectation can be rewritten as $Z_s^{1/\gamma} e^A \mathbb{E}[e^Y]$. The random variable Y has mean $\mathbb{E}[Y] = (1/\gamma)\sigma\mathbb{E}[\Delta B_{\hat{\tau}}] = 0$ and variance $\mathrm{Var}(Y) = ((1/\gamma)\sigma)^2\hat{\tau}$. Substitute back to derive the formula for the conditional expectation:

$$\mathbb{E}[Z_t^{1/\gamma} \mid \mathcal{F}_s] = Z_s^{1/\gamma} \exp\left(\left((1/\gamma)(\mu + \pi) + \frac{1}{2}(1/\gamma)((1/\gamma) - 1)\sigma^2\right)\hat{\tau}\right). \tag{A70}$$

Given that $\mathcal{J}(R_t, Z_t) = \mathcal{O}(Z_t^{1/\gamma})$ and using (A70), the TVC will always hold as long as

$$X_U Z_s^{1/\gamma} \lim_{t\to\infty} \exp\left(\left(-(\rho + \delta + \pi) + (1/\gamma)(\mu + \pi) + \frac{1}{2}(1/\gamma)((1/\gamma) - 1)\sigma^2\right)\hat{\tau}\right) = 0.$$

For the limit to converge to zero, the exponent multiplying $\hat{\tau}$ must be strictly negative. This means that the TVC along non-terminated paths holds if

$$\rho + \delta + \pi > \frac{\mu + \pi}{\gamma} + \frac{1}{2}\frac{\sigma^2(1-\gamma)}{\gamma^2}.$$

This is exactly the condition we assumed in (8), which ensures that the maximum expected growth rate of the value function is sufficiently low relative to the effective discount rate.

Step 3. Establishing the (super)martingale property of $Y_{t\wedge\tau_n}$ in Step 1 is key to the verification proof, as it is the essential precondition to use an important tool: the *optional sampling theorem*. Since $t \wedge \tau_n$ is a bounded stopping time for each fixed n, the optional sampling theorem for (super)martingales yields

$$\mathbb{E}\left[Y_{t\wedge\tau_n} \mid \mathcal{F}_s\right] = Y_s = \widetilde{\mathcal{J}}(R_s, Z_s), \text{ if } W_t = \widetilde{W}, \tag{A71}$$

$$\mathbb{E}\left[Y_{t\wedge\tau_n} \mid \mathcal{F}_s\right] \leq Y_s = \widetilde{\mathcal{J}}(R_s, Z_s), \ \forall\, W_t \in \mathcal{W}. \tag{A72}$$

Next, we let $t \to \infty$ while keeping n fixed. Note that $\tau_n \leq n$ is bounded and $t \wedge \tau_n \uparrow \tau_n$ a.s. when $t \to \infty$, so by the right-continuity of the sample paths together with the integrability ensured by the polynomial growth bound of $\widetilde{\mathcal{J}}$, we know that $\left|Y_{t\wedge\tau_n}\right|$ is globally bounded, and $Y_{t\wedge\tau_n} \to Y_{\tau_n}$ a.s. when $t \to \infty$. Therefore, by the *dominated convergence theorem*, we have

$$\mathbb{E}\left[Y_{\tau_n} | \mathcal{F}_s\right] = Y_s = \widetilde{\mathcal{J}}(R_s, Z_s), \text{ if } W_t = \widetilde{W}, \tag{A73}$$

$$\mathbb{E}\left[Y_{\tau_n} | \mathcal{F}_s\right] \leq Y_s = \widetilde{\mathcal{J}}(R_s, Z_s), \ \forall\, W_t \in \mathcal{W}, \tag{A74}$$

Finally, we take the limit $n \to \infty$. Using the same arguments as right above, we may pass the

limit $n \to \infty$ through the conditional expectations in each term of Y_{τ_n} to obtain

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}\left[e^{-\hat{\rho}(\tau_n - s)}\widetilde{\mathcal{J}}(R_{\tau_n}, Z_{\tau_n}) \,|\, \mathcal{F}_s\right] &= 0, \\
\lim_{n\to\infty} \mathbb{E}\left[\int_s^{\tau_n} e^{-\hat{\rho}(t-s)}\Pi(R_t, R_t, Z_t)dt \,|\, \mathcal{F}_s\right] &= \mathbb{E}\left[\int_s^{\infty} e^{-\hat{\rho}(t-s)}\Pi(R_t, R_t, Z_t)dt \,|\, \mathcal{F}_s\right], \\
\lim_{n\to\infty} \mathbb{E}\left[\sum_{\tau_i \le \tau_n} e^{-\hat{\rho}(\tau_i - s)} G\left(\widetilde{W}_i, R_{\tau_i^-}, Z_{\tau_i}\right) | \, \mathcal{F}_s\right] &= \mathbb{E}\left[\sum_{i=1}^{\infty} e^{-\hat{\rho}(\tau_i - s)} G\left(\widetilde{W}_i, R_{\tau_i^-}, Z_{\tau_i}\right) | \, \mathcal{F}_s\right],
\end{aligned}$$

where the first equality holds because of the TVC from Step 2. Therefore, we conclude that

$$\begin{aligned}
\widetilde{\mathcal{J}}(R_s, Z_s) &= \mathbb{E}\left[\int_s^{\infty} e^{-\hat{\rho}(t-s)}\Pi(R_t, R_t, Z_t)dt \,|\, \mathcal{F}_s\right] + \mathbb{E}\left[\sum_{i=1}^{\infty} e^{-\hat{\rho}(\tau_i - s)} G\left(\widetilde{W}_i, R_{\tau_i^-}, Z_{\tau_i}\right) | \, \mathcal{F}_s\right] \\
&= \mathcal{J}(R_s, Z_s). \qquad \text{(A75)}
\end{aligned}$$

This means that $\widetilde{\mathcal{J}}$ coincides with the value function and $\widetilde{W}$ is indeed optimal:

$$\widetilde{\mathcal{J}}(R_s, Z_s; \widetilde{W}) = \sup_{W \in \mathcal{W}} \mathcal{J}(R_s, Z_s; W) = \mathcal{J}(R_s, Z_s), \ \forall (R_s, Z_s) \in \mathcal{X}. \qquad \text{(A76)}$$

(Step 4) To find the optimal U and L that maximize the value function, start by rewriting the value function in the upward action region under the optimal policy $\widetilde{W} = (Z/U)^{(1)/\gamma}$ as

$$\mathcal{J}^U(R, Z) = C_U(R, Z) + Z^{1/\gamma}\,\varphi_U(U; L), \qquad \text{(A77)}$$

$$C_U(R, Z) \triangleq -\frac{\eta Z}{1-\gamma} R^{1-\gamma} + R + \frac{Z\bar{e}}{\rho + \delta - \mu}, \qquad \text{(A78)}$$

$$\varphi_U(U; L) \triangleq \frac{\eta}{1-\gamma} U^{\,1-1/\gamma} - K\,U^{-1/\gamma} + \Omega_1(L, U)\,U^{\,\beta_1 - 1/\gamma} + \Omega_2(L, U)\,U^{\,\beta_2 - 1/\gamma}. \qquad \text{(A79)}$$

where $K \triangleq \frac{1+\rho+\delta+\pi}{\rho+\delta+\pi}$. Since the term $C_U(R, Z)$ does not depend on U and since $Z^{1/\gamma} > 0$, maximizing $\mathcal{J}^U$ with respect to U is equivalent to maximizing φ_U with respect to U. Differentiating φ_U while holding L fixed gives

$$\begin{aligned}
\varphi_U'(U; L) = &-\frac{\eta}{\gamma} U^{-1/\gamma} + \frac{K}{\gamma} U^{-1/\gamma - 1} + U^{\beta_1 - 1/\gamma}\Omega_{1,U} + \left(\beta_1 - \frac{1}{\gamma}\right)\Omega_1 U^{\beta_1 - 1/\gamma - 1} \\
&+ U^{\beta_2 - 1/\gamma}\Omega_{2,U} + \left(\beta_2 - \frac{1}{\gamma}\right)\Omega_2 U^{\beta_2 - 1/\gamma - 1}, \qquad \text{(A80)}
\end{aligned}$$

where $\Omega_{1,U} \triangleq \frac{\partial \Omega_1}{\partial U}$ and $\Omega_{2,U} \triangleq \frac{\partial \Omega_2}{\partial U}$. Multiplying by $\gamma U^{1+1/\gamma}$ yields

$$\gamma U^{1+1/\gamma}\varphi_U'(U; L) = K - \eta U + (\gamma\beta_1 - 1)\Omega_1 U^{\beta_1} + (\gamma\beta_2 - 1)\Omega_2 U^{\beta_2} + \gamma U^{\beta_1 + 1}\Omega_{1,U} + \gamma U^{\beta_2 + 1}\Omega_{2,U}. \qquad \text{(A81)}$$

Now notice that the FOC condition for the optimal wage (A12) shows up in the above equation

(multiplied by minus one), which means that

$$K - \eta U + (\gamma\beta_1 - 1)\Omega_1 U^{\beta_1} + (\gamma\beta_2 - 1)\Omega_2 U^{\beta_2} = 0. \tag{A82}$$

Then (A81) simplifies immediately to

$$\varphi'_U(U; L) = U^{-1-1/\gamma}\Big(U^{\beta_1+1}\Omega_{1,U} + U^{\beta_2+1}\Omega_{2,U}\Big). \tag{A83}$$

Hence the necessary first-order condition becomes

$$U^{\beta_1+1}\Omega_{1,U} + U^{\beta_2+1}\Omega_{2,U} = 0. \tag{A84}$$

Recall that we have obtained closed-form solutions for Ω_1 and Ω_2 given by (26) and (27). Direct differentiation with respect to U gives

$$\Omega_{1,U}(L, U) = -\frac{L^{\beta_2}}{U\,(1 - \gamma\beta_1)\,D(L,U)^2}\,H_U(L, U), \tag{A85}$$

$$\Omega_{2,U}(L, U) = \frac{L^{\beta_1}}{U\,(1 - \gamma\beta_2)\,D(L,U)^2}\,H_U(L, U),, \tag{A86}$$

where we define

$$D(L, U) \triangleq L^{\beta_1}U^{\beta_2} - U^{\beta_1}L^{\beta_2}, \tag{A87}$$

$$\begin{aligned} H_U(L, U) \triangleq\ & K\Big[\beta_1 L^{\beta_2}U^{\beta_1} - \beta_2 L^{\beta_1}U^{\beta_2} + (\beta_2 - \beta_1)U^{\beta_1+\beta_2}\Big] \\ & + \eta\Big[(1 - \beta_1)L^{\beta_2}U^{\beta_1+1} + (\beta_2 - 1)L^{\beta_1}U^{\beta_2+1} - \lambda(\beta_2 - \beta_1)LU^{\beta_1+\beta_2}\Big]. \end{aligned} \tag{A88}$$

Substituting (A85)–(A86) into (A84) yields the factorization

$$\frac{(1 - \gamma\beta_1)L^{\beta_1}U^{\beta_2} - (1 - \gamma\beta_2)L^{\beta_2}U^{\beta_1}}{D(L,U)^2(1 - \gamma\beta_1)(1 - \gamma\beta_2)}\,H_U(L, U) = 0. \tag{A89}$$

Thus the first-order condition splits into two branches:

$$H_U(L, U) = 0, \tag{A90}$$

$$(1 - \gamma\beta_1)L^{\beta_1}U^{\beta_2} - (1 - \gamma\beta_2)L^{\beta_2}U^{\beta_1} = 0. \tag{A91}$$

It is clear that the branch (A91) is not the relevant first-order condition because of a structural reason. If $\lambda = 1$, the relevant (nondegenerate) branch must connect continuously to the symmetric limit $U = L$. Evaluating the factor in (A91) at ($U = L$ gives

$$(1 - \gamma\beta_1)U^{\beta_1+\beta_2} - (1 - \gamma\beta_2)U^{\beta_1+\beta_2} = \gamma(\beta_2 - \beta_1)U^{\beta_1+\beta_2} \neq 0 \qquad (\beta_1 \neq \beta_2). \tag{A92}$$

Hence (A91) does not pass through the symmetric limit and therefore cannot describe the cor-

rect condition. On the other hand, evaluating (A90) at $U = L$ and $\lambda = 1$ gives

$$\begin{aligned} H_U(U,U;\lambda=1) &= A\Big[\beta_1 U^{\beta_2}U^{\beta_1} - \beta_2 U^{\beta_1}U^{\beta_2} + (\beta_2-\beta_1)U^{\beta_1+\beta_2}\Big] \\ &\quad + \eta\Big[(1-\beta_1)L^{\beta_2}U^{\beta_1+1} + (\beta_2-1)U^{\beta_1}L^{\beta_2+1} - (\beta_2-\beta_1)U\ U^{\beta_1+\beta_2}\Big] \\ &= AU^{\beta_1+\beta_2}\big[\beta_1-\beta_2+(\beta_2-\beta_1)\big] \\ &\quad + \eta U^{\beta_1+\beta_2+1}\big[(1-\beta_1)+(\beta_2-1)-(\beta_2-\beta_1)\big] = 0 \end{aligned}$$

Therefore, the relevant first-order condition is given by (A90). We then go through the same steps to derive the condition for the optimal L. Maximizing $\mathcal{J}^L$ with respect to L and repeating the same procedure as above yields the following (relevant) first-order condition for L:

$$H_L(L,U) = 0, \tag{A93}$$

$$\begin{aligned} H_L(L,U) &\triangleq K\Big[\beta_2 L^{\beta_2}U^{\beta_1} - \beta_1 L^{\beta_1}U^{\beta_2} + (\beta_1-\beta_2)L^{\beta_1+\beta_2}\Big] \\ &\quad + \eta\Big[\lambda(\beta_1-1)L^{\beta_1+1}U^{\beta_2} + \lambda(1-\beta_2)L^{\beta_2+1}U^{\beta_1} + (\beta_2-\beta_1)L^{\beta_1+\beta_2}U\Big]. \end{aligned} \tag{A94}$$

Before we move on, it is also easy to show that the FOCs (A88)–(A88) characterizing the optimal U and L that maximize $\mathcal{J}^U$ and $\mathcal{J}^L$ coincide with the FOCs that maximize the value function in the inaction region. Differentiating $\overline{\mathcal{J}}$ with respect to U and L, respectively, and using the expressions we derived above yields

$$\overline{\mathcal{J}}_U(R,Z) = \frac{H_U(L,U)}{U\,D(L,U)^2}\left[-\frac{L^{\beta_2}}{1-\gamma\beta_1}Z^{\beta_1}R^{1-\gamma\beta_1} + \frac{L^{\beta_1}}{1-\gamma\beta_2}Z^{\beta_2}R^{1-\gamma\beta_2}\right], \tag{A95}$$

$$\overline{\mathcal{J}}_L(R,Z) = \frac{H_L(L,U)}{L\,D(L,U)^2}\left[\frac{U^{\beta_2}}{1-\gamma\beta_1}Z^{\beta_1}R^{1-\gamma\beta_1} - \frac{U^{\beta_1}}{1-\gamma\beta_2}Z^{\beta_2}R^{1-\gamma\beta_2}\right]. \tag{A96}$$

Consequently,

$$\overline{\mathcal{J}}_U(R,Z) = 0 \iff H_U(L,U) = 0 \quad \text{and} \quad \overline{\mathcal{J}}_L(R,Z) = 0 \iff H_L(L,U) = 0 \tag{A97}$$

Next, we need to show that the FOCs above actually identify strict maxima, so that the second-order conditions at these critical points are strictly negative. Since maximizing the upward/downward action branches of the value function with respect to U and L is equivalent to maximizing the value function in the inaction region, we shall work with $\overline{\mathcal{J}}$ for analytical convenience. From (A95), differentiating once more with respect to U and then evaluating at the critical point satisfying $H_U(L,U) = 0$ yields

$$\overline{\mathcal{J}}_{UU}(R,Z) = \frac{H_{U,U}(L,U)}{U\,D(L,U)^2}\left[-\frac{L^{\beta_2}}{1-\gamma\beta_1}Z^{\beta_1}R^{1-\gamma\beta_1} + \frac{L^{\beta_1}}{1-\gamma\beta_2}Z^{\beta_2}R^{1-\gamma\beta_2}\right], \tag{A98}$$

Since the bracket in (A98) is strictly positive, the sign of $\overline{\mathcal{J}}_{UU}$ coincides with the sign of $H_{U,U}(L,U)$.

On the relevant branch $H_U(L,U)=0$, direct differentiation gives

$$H_{U,U}(L,U)=\frac{L^{\beta_1}U^{\beta_2}-L^{\beta_2}U^{\beta_1}}{U}\Big[K\beta_1\beta_2-\eta U(\beta_1-1)(\beta_2-1)\Big]. \tag{A99}$$

Given that $U>L,\beta_1>1,\beta_2<0$, we have

$$L^{\beta_1}U^{\beta_2}-L^{\beta_2}U^{\beta_1}=L^{\beta_1+\beta_2}\left[\left(\frac{U}{L}\right)^{\beta_2}-\left(\frac{U}{L}\right)^{\beta_1}\right]<0.$$

It remains to determine the sign of the second factor in (A99). To do so, solve $H_U(L,U)=0$ for λ and substitute the result into $H_L(L,U)=0$. After rearrangement one obtains

$$\begin{aligned}&\Big[K\beta_1\beta_2-\eta U(\beta_1-1)(\beta_2-1)\Big]\left[\left(\frac{U}{L}\right)^{\beta_1}-\left(\frac{U}{L}\right)^{\beta_2}\right]\\&=K\left(\frac{U}{L}\right)^{\beta_1+\beta_2}\left[\beta_1\left(\left(\frac{U}{L}\right)^{-\beta_2}-1\right)+(-\beta_2)\left(1-\left(\frac{U}{L}\right)^{-\beta_1}\right)\right].\end{aligned} \tag{A100}$$

The right-hand side of (A100) is strictly positive because $U/L>1,\beta_1>1,-\beta_2>0$, and the factor in the square bracket on the left-hand side is also strictly positive because $(U/L)^{\beta_1}>(U/L)^{\beta_2}$. Therefore $[K\beta_1\beta_2-\eta U(\beta_1-1)(\beta_2-1)]>0$. Combining this with (A99) yields

$$H_{U,U}(L,U)<0,\quad\text{hence}\quad\overline{\mathcal{J}}_{UU}(R,Z)<0. \tag{A101}$$

Thus the U-critical point is a strict local maximum. We next turn to L. From (A96), differentiating once more with respect to L and evaluating at a point satisfying $H_L(L,U)=0$ yields

$$\overline{\mathcal{J}}_{LL}(R,Z)=\frac{H_{L,L}(L,U)}{LD(L,U)^2}\left[\frac{U^{\beta_2}}{1-\gamma\beta_1}Z^{\beta_1}R^{1-\gamma\beta_1}-\frac{U^{\beta_1}}{1-\gamma\beta_2}Z^{\beta_2}R^{1-\gamma\beta_2}\right]. \tag{A102}$$

The bracket in (A102) is strictly negative, therefore the sign of $\overline{\mathcal{J}}_{LL}$ is opposite to the sign of $H_{L,L}(L,U)$. On the relevant branch $H_L(L,U)=0$, direct differentiation gives

$$H_{L,L}(L,U)=KL^{\beta_1+\beta_2-1}\left[\beta_1\left(\left(\frac{U}{L}\right)^{\beta_2}-1\right)-\beta_2\left(\left(\frac{U}{L}\right)^{\beta_1}-1\right)\right]. \tag{A103}$$

To show that the bracket in (A103) is strictly positive, note that for any $x>1$, the function $\beta\mapsto\frac{x^\beta-1}{\beta}$ is strictly increasing, because

$$\frac{d}{d\beta}\left(\frac{x^\beta-1}{\beta}\right)=\frac{x^\beta(\beta\ln x-1)+1}{\beta^2}=\frac{y\ln y-y+1}{\beta^2}>0,\quad y\triangleq x^\beta>0, \tag{A104}$$

where the strict inequality follows from $y \ln y - y + 1 > 0$ for $y \neq 1$, thus

$$\frac{(U/L)^{\beta_1} - 1}{\beta_1} > \frac{(U/L)^{\beta_2} - 1}{\beta_2}. \tag{A105}$$

Multiplying both sides by $\beta_1(-\beta_2) > 0$ yields

$$-\beta_2 \left(\left(\frac{U}{L} \right)^{\beta_1} - 1 \right) + \beta_1 \left(\left(\frac{U}{L} \right)^{\beta_2} - 1 \right) > 0, \tag{A106}$$

which is exactly the bracket in (A103). Therefore

$$H_{L,L}(L, U) > 0, \quad \text{hence} \quad \overline{\mathcal{J}}_{LL}(R, Z) < 0. \tag{A107}$$

Thus the L-critical point is also a strict local maximum.

Moving on, we can make further progress by reducing the 2×2 system of coupled non-linear equations (A88)–(A88) into a one-dimensional non-linear equation. Let's define the ratio $x \triangleq U/L \geq 1$ (equal to 1 if and only if $\lambda = 1$). Substituting $U = xL$ into the FOCs (A88) and (A94) and manipulating through gives two equations that are linear in L:

$$K M_U(x) + \eta L N_U(x) = 0, \tag{A108}$$
$$K M_L(x) + \eta L N_L(x) = 0, \tag{A109}$$

$$\text{where} \quad M_U(x) \triangleq \beta_1 x^{\beta_1} - \beta_2 x^{\beta_2} + (\beta_2 - \beta_1) x^{\beta_1 + \beta_2}, \tag{A110}$$
$$M_L(x) \triangleq \beta_2 x^{\beta_1} - \beta_1 x^{\beta_2} + (\beta_1 - \beta_2), \tag{A111}$$
$$N_U(x) \triangleq (1 - \beta_1) x^{\beta_1 + 1} + (\beta_2 - 1) x^{\beta_2 + 1} - \lambda(\beta_2 - \beta_1) x^{\beta_1 + \beta_2}, \tag{A112}$$
$$N_L(x) \triangleq (\beta_2 - \beta_1) x + \lambda \Big((\beta_1 - 1) x^{\beta_2} + (1 - \beta_2) x^{\beta_1} \Big). \tag{A113}$$

Now we want to eliminate L from (A108)–(A109) in order to get a single scalar equation for x. We solve for $L = -\frac{K M_U(x)}{\eta N_U(x)}$ and $L = -\frac{K M_L(x)}{\eta N_L(x)}$ and equate the two expression to get

$$M_U(x) N_L(x) - M_L(x) N_U(x) = 0. \tag{A114}$$

Factoring out (A114) and performing further algebraic manipulations yields a (well-behaved and numerically stable) scalar non-linear equation that pins down the ratio x:

$$\Lambda(x) \triangleq x \frac{P(x)}{Q(x)} - \lambda = 0, \tag{A115}$$
$$\text{where} \quad P(x) \triangleq \beta_2(\beta_1 - 1) x^{\beta_1} - \beta_1(\beta_2 - 1) x^{\beta_2} + (\beta_2 - \beta_1), \tag{A116}$$
$$Q(x) \triangleq \beta_1(\beta_2 - 1) x^{\beta_1} - \beta_2(\beta_1 - 1) x^{\beta_2} + (\beta_1 - \beta_2) x^{\beta_1 + \beta_2}. \tag{A117}$$

Once x is uniquely determined by (A115), we can use (A108)–(A109) along with the identity

$U = xL$ to recover explicit expressions for the optimal $L(x)$ and $U(x)$. These are given by

$$U(x) = \frac{K\Big[\beta_1(\beta_2-1)x^{\beta_1} - \beta_2(\beta_1-1)x^{\beta_2} + (\beta_1-\beta_2)x^{\beta_1+\beta_2}\Big]}{\eta\,(\beta_1-1)(\beta_2-1)\big(x^{\beta_1}-x^{\beta_2}\big)}, \tag{A118}$$

$$L(x) = \frac{K\Big[\beta_1(\beta_2-1)x^{\beta_1} - \beta_2(\beta_1-1)x^{\beta_2} + (\beta_1-\beta_2)x^{\beta_1+\beta_2}\Big]}{x\,\eta(\beta_1-1)(\beta_2-1)\big(x^{\beta_1}-x^{\beta_2}\big)}. \tag{A119}$$

Now that we have obtained the expressions for $U(x)$ and $L(x)$, it is not hard to also derive in closed-form the constant U^* in the case of $\lambda = 1$ (so that $x = 1$ and $U = L = U^*$). Evaluating either (A118) or (A119) at $x = 1$ and simplifying gives

$$U^* = \frac{K\beta_1\beta_2}{\eta\,(\beta_1-1)(\beta_2-1)},$$

Since $\beta_{1,2}$ solve a quadratic characteristic equation, we know by Vieta's formulae that $\beta_1+\beta_2 = 1 - \frac{2(\mu+\pi)}{\sigma^2}$ and $\beta_1\beta_2 = -\frac{2(\rho+\delta+\pi)}{\sigma^2}$. Using these along with the definition of K yields

$$U^* = \frac{1+\rho+\delta+\pi}{(\rho+\delta-\mu)\eta}, \quad \text{if and only if } \lambda = 1. \tag{A120}$$

Proving that U and L are uniquely determined. In order to prove that U and L given by the explicit expressions (A118)–(A119) are indeed uniquely determined, we need to prove that the non-linear equation $\Lambda(x) = 0$ in (A115) has a unique solution in $x \in [1,\infty)$. We shall thus show that $\Lambda(x)$ is strictly increasing on $x \in [1,\infty)$ and crosses zero exactly once.

First, differentiate (A115) with respect to x and rearrange to get

$$\Lambda'(x) = -\frac{(\beta_1-1)(\beta_2-1)x^{\beta_1+\beta_2}}{Q(x)^2}\Big(\beta_1x^{\beta_2} - \beta_1 - \beta_2x^{\beta_1} + \beta_2\Big)\Big(\beta_1x^{-\beta_1} - \beta_1 - \beta_2x^{-\beta_2} + \beta_2\Big). \tag{A121}$$

To determine the sign of $\Lambda'(x)$, consider the function

$$\phi_y(\beta) \triangleq \frac{y^\beta - 1}{\beta} = \int_1^y t^{\beta-1}\,dt, \quad \text{for any } y > 0 \text{ with } y \neq 1. \tag{A122}$$

Differentiating under the integral sign yields

$$\partial_\beta\phi_y(\beta) = \int_1^y t^{\beta-1}\log t\,dt > 0. \tag{A123}$$

Hence ϕ_y is strictly increasing in β, since if $y > 1$ then $t^{\beta-1}\log t > 0$ on $[1,y]$, and if $0 < y < 1$ then $t^{\beta-1}\log t < 0$ on $[y,1]$, and the reversed orientation of the integral makes $\partial_\beta\phi_y(\beta) > 0$. Now rewrite the first bracket in (A121) as

$$\beta_1x^{\beta_2} - \beta_1 - \beta_2x^{\beta_1} + \beta_2 = \beta_1\beta_2\left(\frac{x^{\beta_2}-1}{\beta_2} - \frac{x^{\beta_1}-1}{\beta_1}\right) = \beta_1\beta_2\big(\phi_x(\beta_2) - \phi_x(\beta_1)\big). \tag{A124}$$

Since $x \geq 1$, ϕ_x is strictly increasing, and since $\beta_1 > \beta_2$, we have $\phi_x(\beta_2) - \phi_x(\beta_1) < 0$. Also, since $\beta_1\beta_2 < 0$, it holds that

$$\beta_1 x^{\beta_2} - \beta_1 - \beta_2 x^{\beta_1} + \beta_2 > 0 \tag{A125}$$

So we have proved that the first bracket of $\Lambda'(x)$ is positive. Now exactly the same argument applies to the second bracket of (A121), because $(1/x \in (0, 1]$. Indeed,

$$\beta_1 x^{-\beta_1} - \beta_1 - \beta_2 x^{-\beta_2} + \beta_2 = \beta_1\beta_2\big(\phi_{1/x}(\beta_2) - \phi_{1/x}(\beta_1)\big) > 0. \tag{A126}$$

Since the terms (A125) and (A126) are positive, and since $(\beta_1 - 1)(\beta_2 - 1) < 0$ so that the first term of (A115) is positive, we conclude that

$$\Lambda'(x) > 0 \ ; \ \text{for all } x \geq 1. \tag{A127}$$

What remains is to determine the behavior of $\Lambda(x)$ at the endpoints of the domain. At $x = 1$, we have $P(1) = Q(1) = 0$ from equations (A116) and (A117). Moreover, $P'(1) = \beta_1\beta_2(\beta_1 - \beta_2)$ and $Q'(1) = \beta_1\beta_2(\beta_1 - \beta_2)$. Hence the first-order Taylor expansions around $x = 1$ are

$$P(x) = \beta_1\beta_2(\beta_1 - \beta_2)(x - 1) + \mathcal{O}\big((x - 1)^2\big), \tag{A128}$$
$$Q(x) = \beta_1\beta_2(\beta_1 - \beta_2)(x - 1) + \mathcal{O}\big((x - 1)^2\big). \tag{A129}$$

Therefore,

$$\lim_{x\downarrow 1} x\frac{P(x)}{Q(x)} = 1, \quad \text{hence} \quad \lim_{x\downarrow 1} \Lambda(x) = 1 - \lambda \leq 0. \tag{A130}$$

Then, as $x \to \infty$, since $\beta_2 < 0$ the lower-order terms satisfy $x^{\beta_2} = o(1)$ and $x^{\beta_1+\beta_2} = o(x^{\beta_1})$. Therefore the asymptotic dominant terms in $P(.)$ and $Q(.)$ are

$$P(x) \overset{a}{\sim} \beta_2(\beta_1 - 1)x^{\beta_1}, \tag{A131}$$
$$Q(x) \overset{a}{\sim} \beta_1(\beta_2 - 1)x^{\beta_1}. \tag{A132}$$

It follows that $x\frac{P(x)}{Q(x)} \overset{a}{\sim} \frac{\beta_2(\beta_1-1)}{\beta_1(\beta_2-1)}x$. Since $\beta_2 < 0, \beta_1 > 1, (\beta_2 - 1) < 0$, the coefficient $\frac{\beta_2(\beta_1-1)}{\beta_1(\beta_2-1)}$ is strictly positive. Therefore,

$$\lim_{x\uparrow\infty} x\frac{P(x)}{Q(x)} = \infty, \quad \text{hence} \quad \lim_{x\uparrow\infty} \Lambda(x) = \infty. \tag{A133}$$

Since Λ is continuous and strictly increasing on $[1, \infty)$, and at the same time satisfies $\Lambda(1) \leq 0$ and $\lim_{x\uparrow\infty} \Lambda(x) = +\infty$, we can conclude that $\Lambda(x) = 0$ has exactly one solution in $[1, \infty)$. The constants L and U are thus uniquely determined for any set of parameter configurations.

Conclusion. By going through the steps of a verification theorem, we have proved that the candidate value function $\widetilde{\mathcal{J}}$ under the candidate control $\widetilde{W}$ coincides with the value function of the problem for any initial initial state vector. Therefore we have obtained the value function

of the problem and the unique optimal impulse control. ■

Supplement to Proof of Proposition 5: Here we explain how to derive the simplified version of the Itô-Tanaka-Meyer (ITM) formula we used in (A60). The tools we use to get to the final formula (A138) can be found in stochastic calculus textbooks such as Revuz and Yor (1999). We start by stating a general version of the ITM formula in (A134).

Let X_t be a càdlàg semimartingale and let $f : \mathbb{R} \to \mathbb{R}$ so that f' exists a.e. and is of locally bounded variation. Then, for all $t \geq 0$,

$$f(X_t) = f(X_0) + \int_0^t f'_-(X_s)\, dX_s + \frac{1}{2}\int_{\mathbb{R}} \Lambda_t^a\, f''(da), \tag{A134}$$

where $f''(da)$ is the measure-valued distributional (weak) second derivative of f; Λ_t^a is the *local time* of X_t at level a; and f'_- denotes the left-derivative.

Now, as in our problem, suppose f is C^2 except at k finitely many points a_k, where the one-sided derivatives $f'_\pm(a_k)$ exist. Using integration by parts on each smooth interval between the points a_k together with Lebesgue's decomposition theorem for the signed measure, the distributional derivative $f''(da)$ decomposes as

$$f''(da) = f''(a)da + \sum_k \big(f'_+(a_k) - f'_-(a_k)\big)\delta_{a_k}(da), \tag{A135}$$

where δ_{a_k} is the Dirac measure on a_k. The next step is to use the *occupation times formula*, $\int_{\mathbb{R}} g(a)\, \Lambda_t^a\, da = \int_0^t g(X_s)\, d\langle X\rangle_s$. We apply this formula to the last term in in (A134), taking $g = f''$ and $f''(da) = f''(a)da$, together with (A135), to obtain

$$\frac{1}{2}\int_{\mathbb{R}} \Lambda_t^a\, f''(da) = \frac{1}{2}\int_0^t f''(X_s)\, d\langle X\rangle_s + \frac{1}{2}\sum_k \big(f'_+(a_k) - f'_-(a_k)\big)\Lambda_t^{a_k}. \tag{A136}$$

Putting it all together, we get the simplified ITM formula in integral form:

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\, dX_s + \frac{1}{2}\int_0^t f''(X_s)\, d\langle X\rangle_s + \frac{1}{2}\sum_k \big(f'_+(a_k) - f'_-(a_k)\big)\, \Lambda_t^{a_k}, \tag{A137}$$

or equivalently, the simplified ITM formula in differential form:

$$\mathrm{d}f(X_t) = f'(X_t)\, dX_t + \frac{1}{2}\, f''(X_t)\, d\langle X\rangle_t \;+\; \frac{1}{2}\sum_k \big(f'_+(a_k) - f'_-(a_k)\big)\, d\Lambda_t^{a_k}. \tag{A138}$$

Proof of Proposition 6. For this proof it is more convenient to work with the fundamental state variable $X_t \triangleq \dfrac{Z_t}{R_t^\gamma}$. Since Z_t follows a geometric Brownian motion and R_t remains constant up the next (potential) adjustment time, X_t also follows a geometric Brownian motion:

$$dX_t = (\mu + \pi)X_t dt + \sigma X_t dB_t \tag{A139}$$

Now consider a discrete time interval of length $\Delta > 0$. Starting from the initial condition $X_t = x$ at time t, the state at time $t + \Delta$ is obtained by the strong solution to the above SDE:

$$X_{t+\Delta} = X_t \exp\left[\left(\mu + \pi - \frac{\sigma^2}{2}\right)\Delta + \sigma(B_{t+\Delta} - B_t)\right] \tag{A140}$$

The Wiener increment $(B_{t+\Delta} - B_t)$ is normally distributed with mean zero and variance Δ, $(B_{t+\Delta} - B_t) \sim \mathcal{N}(0, \Delta)$, so we can rewrite it as $(B_{t+\Delta} - B_t) = \sqrt{\Delta}\,\varepsilon$, where $\varepsilon \sim \mathcal{N}(0, 1)$ is a standard normal random variable. Substituting into equation (A140) yields

$$X_{t+\Delta} = X_t \exp\left[\left(\mu + \pi - \frac{\sigma^2}{2}\right)\Delta + \sigma\sqrt{\Delta}\,\varepsilon\right] \tag{A141}$$

To derive the probability of an upward wage adjustment at time $t + \Delta$, $\theta_U(x, \Delta)$, we use the fact that adjusting upwards occurs when $X_{t+\Delta} > U$. Specifically,

$$\begin{aligned}
\theta_U(x, \Delta) &\triangleq \mathbb{P}\left(X_{t+\Delta} > U \mid X_t = x\right) \\
&= \mathbb{P}\left(x \cdot \exp\left[\left(\mu + \pi - \frac{\sigma^2}{2}\right)\Delta + \sigma\sqrt{\Delta}\varepsilon\right] > U\right) \\
&= \mathbb{P}\left(\varepsilon > \frac{\ln(U/x) - (\mu + \pi - \sigma^2/2)\Delta}{\sigma\sqrt{\Delta}}\right)
\end{aligned} \tag{A142}$$

Since $\varepsilon \sim \mathcal{N}(0, 1)$, we have that $\mathbb{P}(\varepsilon > a) = 1 - \Phi(a)$ for any $a \in \mathbb{R}$, where $\Phi(\cdot)$ is the standard normal cdf. As a result

$$\theta_U(x, \Delta) = 1 - \Phi\left(\frac{\ln(U/x) - (\mu + \pi - \sigma^2/2)\Delta}{\sigma\sqrt{\Delta}}\right) \tag{A143}$$

and using the symmetry property $1 - \Phi(a) = \Phi(-a)$ and the logarithm property $-\ln(U/x) = \ln(x/U)$, it follows that

$$\theta_U(x, \Delta) = \Phi\left(\frac{\ln(x/U) + (\mu + \pi - \sigma^2/2)\Delta}{\sigma\sqrt{\Delta}}\right) \tag{A144}$$

To derive the probability of a downward wage adjustment at time $t + \Delta$, $\theta_L(x, \Delta)$, we note that adjusting downwards occurs when $X_{t+\Delta} < L$. Following the same steps as above we obtain

$$\theta_L(x, \Delta) = \Phi\left(\frac{\ln(L/x) - (\mu - \sigma^2/2)\Delta}{\sigma\sqrt{\Delta}}\right) \tag{A145}$$

Finally, the probability of no wage adjustment equals the probability that $X_{t+\Delta} \in [L, U]$, thus

$$\begin{aligned}
\mathbb{P}(L \leq X_{t+\Delta} \leq U) &= 1 - \mathbb{P}(X_{t+\Delta} < L) - \mathbb{P}(X_{t+\Delta} > U) \\
&= 1 - \theta_L(x, \Delta) - \theta_U(x, \Delta)
\end{aligned} \tag{A146}$$

■