Slow Recoveries and Unemployment Traps: Monetary Policy in a Time of Hysteresis*

Sushant Acharya Julien Bengui Keshav Dogra Shu Lin Wee

November 14, 2018

Abstract

We analyze monetary policy in a model where temporary shocks can permanently scar the economy's productive capacity. Unemployed workers' skill losses generate multiple steady-state unemployment rates. When monetary policy is constrained by the zero bound, large shocks reduce hiring to a point where the economy recovers slowly at best – at worst, it falls into a permanent unemployment trap. Since monetary policy is powerless to escape such traps ex-post, it must avoid them ex-ante. The model quantitatively accounts for the slow U.S. recovery following the Great Recession, and suggests that lack of swift monetary accommodation helps explain the European periphery's stagnation.

JEL Classification: E24, E3, E5, J23, J64

Keywords: hysteresis, path dependence, monetary policy, multiple steady states, skill depreciation

*Acharya: FRB NY, sushant.acharya@ny.frb.org; Bengui: Université de Montréal, julien.bengui@umontreal.ca; Dogra: FRB NY, keshav.dogra@ny.frb.org; Wee: Carnegie Mellon University, shuwee@andrew.cmu.edu. We thank Jess Benhabib, Javier Bianchi, Marco Del Negro, Roger Farmer, Luca Fornaro, Jordi Gali, Marc Giannoni, Luigi Iovino, Todd Keister, Rasmus Lentz and seminar participants at Rutgers University, University of Pittsburgh, University of Wisconsin - Madison, Federal Reserve Board of Governors, Federal Reserve Bank of New York, SED 2018 and the 7th BI-CEPR conference on Low Inflation and Wage dynamics: Implications for Monetary policy and Financial stability for useful discussion and comments. The paper was previously circulated under the title "Escaping Unemployment Traps". The views expressed in this paper are those of the authors and do not necessarily represent those of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

In the aftermath of the global financial crisis, economic activity remained subdued, suggesting that the world economy may have settled on a lower growth trajectory than the one prevailing before 2007. Some observers have attributed this sluggish growth to permanent, exogenous structural changes - either permanently lower productivity growth (Gordon, 2015) or *secular stagnation*.[1] An alternative explanation is that large, temporary downturns can themselves permanently damage an economy's productive capacity. This hysteresis view, according to which changes in current aggregate demand can have a significant effect on future aggregate supply, dates back to the 1980s but recently underwent a surge of interest in the wake of the Great Recession (e.g., Yellen, 2016). While the two sets of explanations may be observationally similar, they have very different normative implications. If exogenous structural factors drive slow growth, countercyclical policy may be unable to resist or reverse this trend. In contrast, if temporary downturns themselves lead to persistently or permanently slower growth, then countercyclical policy, by limiting the severity of downturns, may have a role to play to avert such adverse developments.[2]

In this paper, we present a theory in which hysteresis might occur and countercyclical monetary policy can moderate its impact if timed appropriately. In our model, hysteresis can arise because workers lose human capital whilst unemployed and unskilled workers are costly to retrain, as in Pissarides (1992). In the presence of nominal rigidities and a zero lower bound (ZLB) constraint on monetary policy, large adverse fundamental shocks can cause recessions whose legacy is persistent or permanent unemployment. Our theory stresses nominal rigidities and constraints on monetary policy as triggers of episodes of depressed economic activity, but relies exclusively on real forces to explain the slow recovery or stagnation that may ensue. Accordingly, it crucially implies that the timing of monetary policy matters significantly for long-term outcomes. Accommodative policy early in a recession can prevent hysteresis from taking root and enable swift a recovery. In contrast, delayed monetary policy interventions may be powerless to bring the economy back to full employment.

Our formal environment is an economy with downwardly sticky nominal wages[3] and search frictions in the labor market. Our key assumption is that human capital depreciates during unemployment spells and unskilled workers are costly to retrain.[4] As in Pissarides (1992), these features generate multiple steady states. One steady state is a high pressure economy: job finding

[1]Here by *secular stagnation* we refer to the literature arguing that a chronic excess of global savings relative to investment has depressed equilibrium real interest rates. This imbalance has been variously attributed to permanent changes in either borrowing constraints, supply of safe assets, demographics, inequality or monopoly power. See for example, Eggertsson and Mehrotra (2014), Caballero and Farhi (2017), Jones and Philippon (2016) among many others.

[2]In the words of Yellen (2016): *"...hysteresis would seem to make it even more important for policymakers to act quickly and aggressively in response to a recession, because doing so would help to reduce the depth and persistence of the downturn, thereby limiting the supply-side damage that might otherwise ensue."*

[3]See for example, Fallick et al. (2016), for models where nominal wages are slow to adjust downwards but can freely adjust upwards. See also Schmitt-Grohe and Uribe (2013) for a survey of the empirical literature documenting wage rigidity.

[4]There is a large empirical literature documenting the scarring effects of unemployment; see, for example, Song and von Wachter (2014) and Wee (2016).

rates are high, unemployment is low and job-seekers are highly skilled. While tight labor markets - by improving workers' outside options - cause wages to be high, firms still find job creation attractive, as higher wages are offset by low average training costs when job-seekers are mostly highly skilled. The economy, however, can also be trapped in a low pressure steady state. In this steady state, job finding rates are low, unemployment is high, and many job-seekers are unskilled as long unemployment spells have eroded their human capital. Slack labor markets lower the outside options of workers and drive wages down, but hiring is still limited as firms find it costly to retrain these workers. Crucially, this is an *unemployment trap* - an economy near the low pressure steady state can never self-correct and return to a high pressure state.

The presence of nominal wage rigidities, together with constraints on monetary policy, allows temporary shocks to permanently move the economy from a high pressure steady state into an unemployment trap. Consider the effect of a large but temporary decline in the households' rate of time preference. This increases desired savings, pushing real interest rates below zero. Monetary policy tries to accommodate this by lowering nominal interest rates but is constrained by the ZLB. As such, current prices are forced to adjust downwards as households' demand for current consumption relative to the future declines. Under nominal downward wage rigidity, the decline in prices cause real wages to rise, and hiring to fall. This decline in hiring lengthens the average duration of unemployment and increases the incidence of skill loss, leading to a worsening in the skill composition of the unemployed. The deterioration in the average skill quality of the unemployed in turn raises the effective cost of job creation, causing the economy to take time to re-train workers and return to the high pressure steady state even after the shock has abated. In the event of a large enough shock, the economy may be pushed into an unemployment trap from which it is powerless to escape.

Importantly, the transition to an unemployment trap following a large severe shock can be avoided. If monetary policy commits to temporarily higher inflation after the liquidity trap has ended, it can mitigate both the initial rise in unemployment, and its persistent (or permanent) negative consequences. Monetary policy, however, is only effective if it is implemented early in the downturn, before the recession has left substantial scars. If the skill composition of the unemployed has significantly worsened following the shock, monetary policy cannot undo the average high cost of hiring through the promise of higher future prices. With nominal wages free to adjust upwards, any attempt to generate inflation is met by nominal wage inflation, leaving real wages unaffected. Thus, once the economy has entered into an unemployment trap or a slow recovery, monetary policy cannot engineer an escape from this trap nor hasten the recovery. In such cases, fiscal policy, in the form of hiring or training subsidies, is necessary to engineer a swift recovery.

Overall, in the presence of hysteresis, a failure to deliver stimulus early on in a recession can have irreversible costs. This contrasts with standard New Keynesian models, in which accommodative policies are equally effective at any point in a liquidity trap (Eggertsson and Woodford, 2003). In fact, these models predict that while overly tight policy may be costly in the short-run, it has no long run consequences, since temporary shocks have no permanent effects in stationary models.

Because standard New Keynesian models study stationary fluctuations around a unique steady state, the possibility that short-run disturbances can cause permanent damage is precluded, negating monetary policy's ability to influence long-run outcomes.[5] Our model instead focuses on a monetary economy with multiple steady states.[6] This allows monetary policy to affect not just fluctuations around steady state, but also the level of steady state activity.[7]

Our focus on multiple steady states also distinguishes our analysis from recent work which studies the persistent effects of recessions (Benigno and Fornaro, 2017; Schmitt-Grohe and Uribe, 2017). These papers study economies which can switch to a bad equilibrium featuring permanently low or negative inflation, binding ZLB, and high unemployment. This bad equilibrium is the result of self-fulfilling pessimistic beliefs; equally, self-fulfilling optimism can return the economy to the good equilibrium. Our analysis differs sharply in two ways. First, high unemployment can persist even after monetary policy is no longer constrained by the ZLB. Second, it features path dependence: once the economy is stuck in an unemployment trap, optimistic beliefs cannot move the economy back towards full employment.[8] This is because dynamics in our economy are driven by a slow moving state variable - the fraction of unskilled job-seekers. Even if firms anticipated a swift recovery, this would not induce them to hire and train relatively unskilled job-seekers today. In fact, firms would postpone hiring, preferring to wait until there are more skilled job-seekers. Since hiring would fall, the skill composition of job-seekers would actually worsen and firms' optimism would be self-defeating. Since self-fulfilling optimism cannot escape the trap ex-post, it is all the more important to avoid it ex-ante.

Finally, we test whether our model can quantitatively explain the slow recovery in the U.S. following the Great Recession. A calibrated version of the model suggests that allowing for a realistic degree of skill depreciation and training costs, in line with the existing literature, is sufficient to generate multiple steady states. Furthermore, this multiplicity is essential in explaining why the unemployment rate in the U.S. took 7 years to return to its pre-crisis level. In contrast, the standard search model without skill depreciation and/or training costs predicts that the U.S. economy should have fully recovered by 2011. Under our preferred calibration, the model indicates that had monetary policy been less accommodative or timely during the crisis, leading to a peak unemployment rate higher than 11 percent, the economy might have been permanently scarred and stuck in an unemployment trap. Furthermore, our model suggests that the persistently high proportion of long-term unemployed in the European periphery countries may reflect a lack of timely monetary accommodation by the European Central Bank. Additionally, our quantitative analysis also suggests that relatively modest hiring or training subsidies can hasten recoveries. In particular,

[5] Summers (2015) expounds on this criticism of New Keynesian models at length.

[6] It is important to distinguish our approach from that of Farmer (2012) who considers economies with a continuum of steady states but focuses on how beliefs cause an economy to transition between these.

[7] This does not mean that monetary policy can manipulate a long-run trade-off between inflation and unemployment. Once the economy has converged to a particular steady state unemployment rate, monetary policy is powerless to reduce unemployment below this rate.

[8] Similarly, if the economy is experiencing a slow recovery, optimistic beliefs cannot accelerate its return to full employment.

a 4% reduction in training costs would have sped up the U.S. recovery by 2 years.

The remainder of the paper is structured as follows. Next we discuss related literature. Section 2 presents the model economy. Section 3 characterizes steady states and equilibria under a flexible wage benchmark. Section 4 introduces nominal rigidities, and studies how demand shocks can cause slow recoveries or permanent stagnation. Section 5 analyses whether the mechanism studied here can account for the slow U.S. recovery since the Great Recession. Section 6 discusses potential extensions and Section 7 concludes.

Related literature A small number of recent papers study hysteresis and monetary policy in the presence of nominal frictions. Closest to our work are Laureys (2014), who studies optimal monetary policy when skill depreciates during unemployment spells,[9] and Galí (2016), who studies optimal policy in a New Keynesian model with insider-outside labor markets drawing on the earlier work of Blanchard and Summers (1986). These papers argue that monetary policy should deviate from strict inflation targeting and put more emphasis on unemployment stabilization. While hysteresis in these papers generates extra persistence, the focus is on local dynamics around a unique steady state. In contrast, we consider a model with multiple steady states where monetary policy potentially affects long run outcomes.

While we focus on hysteresis operating through the labor market, a recent literature has studied the *innovation channel* of hysteresis. Bianchi et al. (2014) find that declines in R&D during recessions can explain persistent effects of cyclical shocks on growth, while Garga and Singh (2016) study the conduct of optimal monetary policy in a model embedding this feature. Benigno and Fornaro (2017) also study an economy in which pessimism can drive the economy to the ZLB and lead to persistent or permanent slowdowns driven by a fall in innovation. A commitment to alternative monetary policy rules or subsidies to innovation can help avoid or exit such *stagnation traps*. While we study a different channel through which hysteresis might operate, our results resonate with Benigno and Fornaro (2017): a commitment to an alternative monetary policy can avoid an unemployment trap as long as it is implemented swiftly. However, if an economy is already stuck in an unemployment trap, monetary policy may be unable to engineer an exit from the trap, although fiscal policy in the form of hiring or training subsidies can still be effective.

Our analysis also contributes to the broader theoretical literature studying hysteresis, which has largely abstracted from nominal rigidities. Drazen (1985) argues that the loss of human capital due to job-loss in recessions can lead to delayed recoveries. Schaal and Taschereau-Dumouchel (2016) show that a labor search model with aggregate demand externalities can generate additional persistence in labor market variables. Similarly, in Schaal and Taschereau-Dumouchel (2015), large recessions frustrate coordination on a high-activity equilibrium, allowing temporary shocks to cause quasi-permanent recessions. Our model instead draws on Pissarides (1992), who demonstrates how skill depreciation can give rise to multiple steady states. Sterk (2016) studies a quantitative version of Pissarides (1992)'s model and argues that it can account for the behavior of job finding rates in

[9] Kapadia (2005) performs a similar exercise in the 3 equation New Keynesian model by incorporating hysteresis in output in a reduced form fashion.

the United States. Relative to our work, all these studies consider purely real models. Importantly, in our framework, hysteresis could never take root in the absence of nominal rigidities.

On the empirical side, a large literature finds evidence in support of drops in productive capacity after recessions. Dickens (1982) finds that recessions can permanently lower productivity; Haltmaier (2012) finds that trend output falls by 3 percentage points on average in developed economies four years after a pre-recession peak. Using cross-country data, Martin et al. (2014) find that severe recessions have a sustained and sizable negative impact on output. Similarly, Ball (2009) finds that large increases in the natural rate of unemployment are associated with disinflations, and large decreases with inflation.[10] Song and von Wachter (2014) find that the persistent decline in employment following job displacement is larger during recessions, suggesting that a spike in job-destruction rates can persistently affect unemployment.[11]

Aside from the literature on hysteresis, our analysis connects to a few recent developments in monetary economics. Like us, Dupraz et al. (2017) study a *plucking model* in which downward-nominal wage rigidity gives rise to asymmetric effects of monetary policy: while deflation can lead to an increase in real wages and a fall in hiring, inflation has limited ability to reduce unemployment. This asymmetry increases the costs of business cycles; but since their economy features a unique steady state, shocks have at most a temporary effect. In contrast, we show that this asymmetry becomes especially dangerous when combined with hysteresis: temporary deflation can lead to permanently higher unemployment and deterioration in the skill composition of the unemployed, which cannot be reversed by higher inflation at a later date. Thus, in our setting it is especially important for monetary policy to stabilize employment, even at the cost of compromising price stability. This result resonates with Berger et al. (2016), who find that monetary policy should prioritize employment stabilization over price stability when households are imperfectly insured against layoff risk. Our analysis provides another reason why employment fluctuations might have higher costs, and warrant more attention.

Finally, our paper also relates to the secular stagnation literature. Eggertsson and Mehrotra (2014) and Caballero and Farhi (2017) present models in which the market clearing interest rate is persistently or permanently negative, leading to persistently low output, as the ZLB prevents nominal rates from falling to clear markets. In such situations, a permanent change in fiscal or monetary policy (such as an increase in target inflation) is typically required to prevent stagnation. We share this literature's concern with long run outcomes, but consider a different mechanism: in our model *temporary* falls in market clearing interest rates have permanent effects, which *temporary* monetary accommodation can prevent.

[10] In other work, Ball (2014) finds that countries with a larger fall in output during the Great Recession experienced a larger decline in *potential output*. See also Blanchard et al. (2015) for a similar account.

[11] These papers also resonate with an older literature asking whether cyclical fluctuations could be studied independently of factors affecting longer-run outcomes. See e.g. Plosser (1989).

2 The Model Economy

We start by establishing the properties of a benchmark economy with search frictions but no nominal rigidities. We use a standard Diamond-Mortensen-Pissarides (DMP) model of labor market frictions. Time is discrete and there is no uncertainty. The only addition to the standard DMP model is that we assume that workers can lose skill following an unemployment spell.

Workers There exists a unit mass of ex ante identical workers, who are risk neutral and discount the future at a rate β. Workers can borrow and save in a nominal bond which pays a nominal return of $1 + i_t$. Workers can either be employed or unemployed. We denote the mass of employed workers as n and the mass of unemployed as $u = 1 - n$. All unemployed workers produce $b > 0$ as home-production. The stock of employed workers evolves as:

$$n_t = [1 - \delta(1 - q_t)]\, n_{t-1} + q_t u_{t-1} \tag{1}$$

where δ is the exogenous rate at which workers get separated from their current jobs and q_t is the job-finding rate. Note that equation (1) implies that a worker separated at the beginning of period t can find another job within the same period. Next, let $\mathbb{W}_t$ denote the value of an employed worker and $\mathbb{U}_t$ denote the value of an unemployed worker at time t. These can be expressed as follows:

$$\mathbb{W}_t = \omega_t + \beta\Big\{ [1 - \delta(1 - q_{t+1})]\, \mathbb{W}_{t+1} + \delta(1 - q_{t+1})\mathbb{U}_{t+1}\Big\} \tag{2}$$

$$\mathbb{U}_t = b + \beta\Big\{q_{t+1}\mathbb{W}_{t+1} + (1 - q_{t+1})\mathbb{U}_{t+1}\Big\} \tag{3}$$

where ω_t denotes the real wage at date t and b is the value of home production.

Labor market As in Pissarides (1992), we assume that a worker who gets separated from her job and is unable to transition back to employment immediately loses skill that she acquired while employed. That is, any worker unemployed for at least 1 period becomes *unskilled*. Because unskilled workers produce zero output when matched with a firm, a firm that hires an unskilled worker must pay a training cost $\chi > 0$ to use that worker in production. Once the firm trains the worker, she remains *skilled* until the next unemployment spell of at least 1 period. Let μ_t denote the fraction of unskilled workers in the pool of job-seekers (l_t) at date t and is defined as:

$$\mu_t = \frac{u_{t-1}}{l_t} \equiv \frac{u_{t-1}}{1 - (1 - \delta)(1 - u_{t-1})} \tag{4}$$

Equation (4) shows that a higher level of unemployment in the past corresponds to a higher fraction of unskilled job-seekers. As such, there is a one-to-one mapping between u_{t-1} and μ_t.

Matching technology Search is random. The number of successful matches m_t between job-seekers l_t and vacancies v_t is given by a CRS matching technology $m(v_t, l_t)$. We define market tightness θ_t as the ratio of vacancies to job-seekers. The job-finding probability of a job-seeker, q_t, and the job-filling probability of a vacancy, f_t, are then given by:

$$q(\theta_t) = \frac{m(v_t, l_t)}{l_t} \qquad \text{and} \qquad f(\theta_t) = \frac{m(v_t, l_t)}{v_t} = \frac{q(\theta_t)}{\theta_t} \tag{5}$$

Firms A representative CRS firm uses labor as an input to produce the final good. The production function is given by $y_t = An_t$ where $A > b$ is aggregate productivity and n_t is the number of employed workers in period t. A firm must incur a vacancy posting cost of $\kappa > 0$ and an additional training cost of χ for each unskilled worker hired. A firm with n_{t-1} workers at the beginning of period t chooses vacancies (taking wages as given) to maximize lifetime discounted profit:

$$\begin{aligned} \mathbb{J}_t = \quad & \max_{v_t \geq 0} (A - \omega_t) n_t - (\kappa + \chi \mu_t f_t) v_t + \beta \mathbb{J}_{t+1} \\ \text{s.t.} \quad & n_t = (1 - \delta) n_{t-1} + f_t v_t \end{aligned} \tag{6}$$

where ω_t is the wage paid to all workers.[12] This is the standard problem of a firm in search models, with one difference: the total cost of job creation depends on the skill composition of job-seekers. Since the firm pays a cost χ to train each unskilled job-seeker it hires, the effective average cost of creating a job is increasing in the fraction of unskilled job-seekers μ_t. Recall from equation (4) that μ_t depends on past unemployment rates, making the cost of job creation *increasing* in the unemployment rate. The job-creation condition is then:

$$\frac{\kappa}{f_t} + \chi \mu_t + \lambda_t = A - \omega_t + \beta(1 - \delta) \left\{ \frac{\kappa}{f_{t+1}} + \chi \mu_{t+1} + \lambda_{t+1} \right\} \tag{7}$$

where $\lambda_t f_t \geq 0$ is the multiplier on the non-negativity constraint on vacancies. Using the Envelope Theorem, the firm's value of a filled vacancy, $J_t = \partial \mathbb{J}_t / \partial n_t$, can be written as:[13]

$$J_t = A - \omega_t + \beta(1 - \delta) J_{t+1} \tag{8}$$

Resource constraint The resource constraint in the real economy can be written as:

$$c_t = An_t + b(1 - n_t) - \kappa v_t - \chi \mu_t f_t v_t$$

To close the model, we now need to specify how wages and prices are determined.

Wage and price determination While we ultimately seek to analyze the conduct of monetary policy in an environment with sticky nominal wages, it is useful to first define a *flexible wage benchmark economy*, in which wages are determined by Nash bargaining every period. Because bargaining occurs after all hiring and training costs have been paid, all workers are paid the same

[12] Firms pay the same wage to both initially skilled and unskilled hires as well as existing skilled workers. We discuss this in more detail in the section on wage determination.

[13] Using the notation J_t, the job creation condition can also be written as $f_t[J_t - \chi \mu_t] \leq \kappa$, $\theta_t \geq 0$

wage.[14] Formally, the Nash bargaining problem is:

$$\max_{\omega_t} J_t^{1-\eta} \left(\mathbb{W}_t - \mathbb{U}_t\right)^{\eta}$$

where $\eta \in [0, 1)$ denotes the bargaining power of the workers. The Nash-bargained wage is:

$$\omega_t^* \quad = \quad \eta A + (1-\eta)b + \beta(1-\delta)\eta q_{t+1} J_{t+1} \tag{9}$$

Crucially, an increase in next period's job-finding rate puts upward pressure on the Nash wage because it increases the worker's outside option. Plugging in the Nash-bargained wage (9) into the expression for J_t (8) yields:

$$J_t = a + \beta(1-\delta)(1-\eta q_{t+1})J_{t+1} \tag{10}$$

where we define $a = (1-\eta)(A-b)$. Thus, an increase in job finding rate at every future date, through its upward pressure on the Nash wage, results in a smaller profit to the firm and thus a lower J_t. Iterating forward on (10) and using equation (7), the job creation condition can be rewritten as :

$$J_t = a \sum_{s=0}^{\infty} \beta^s (1-\delta)^s \prod_{\tau=0}^{s} (1-\eta q_{t+\tau}) \leq \frac{\kappa}{f_t} + \chi\mu_t \ , \ \theta_t \geq 0 \ , \ \textit{with at least one strict equality} \tag{11}$$

In this benchmark, the classical dichotomy holds and the price level does not affect real allocations. Thus, it is not necessary to describe the conduct of monetary policy which determines prices.[15] Equilibrium in the benchmark economy is completely characterized by:

$$\begin{aligned}
\mu_{t+1} &= \frac{1-q(\theta_t)}{1+(1-\delta)[1-q(\theta_t)-\mu_t]} && (12)\\
J_t &= a + \beta(1-\delta)(1-\eta q(\theta_{t+1}))J_{t+1} && (13)\\
J_t &\leq \frac{\kappa}{f(\theta_t)} + \chi\mu_t, \theta_t \geq 0, \ \textit{at least one strict equality} && (14)
\end{aligned}$$

where (12) is derived by combining equations (4) and (6). (13) implies that the value of a filled vacancy to a firm lies in the interval: $J_t \in [J_{min}, J_{max}]$, for $J_{min} \equiv a/[1-\beta(1-\delta)(1-\eta)]$ and $J_{max} \equiv a/\left[1-\beta(1-\delta)\right]$.[16]

[14] Given the fact that the training cost is sunk at the time of bargaining and that all job-seekers have the same probability of finding a job, all workers share the same outside options.

[15] In Section 4, we describe the economy with sticky nominal wages and specify how monetary policy is conducted in that economy.

[16] J_{min} is the value achieved by the firm when labor markets are expected to be the tightest forever (i.e., $q_t = 1$ for all t). Conversely, J_{max} is the firm's value when labor markets are expected to be the slackest forever (i.e., $q_t = 0$ for all t).

3 Flexible wage benchmark

Our ultimate goal is to describe an economy in which transitory increases in unemployment can permanently scar the economy and to ask whether monetary policy can do anything about it. In our model, temporary shocks can have permanent effects because the economy features multiple steady states and a large enough shock can move the economy between these steady states.

In this section and Section 4, we assume a particular form for the matching function, $m_t = \min\{l_t, v_t\}$, which implies $q(\theta_t) = \min\{\theta_t, 1\}$, $f(\theta_t) = \min\{1/\theta_t, 1\}$. This simplifies the analysis without losing any generality.[17] In particular, it implies that the short side of the market matches with probability 1. We refer to the case with $\theta_t < 1$ as the *slack labor market regime* and the one with $\theta_t \geq 1$ as the *tight labor market regime*.

3.1 Steady states

In our model, multiplicity of steady state unemployment rates can arise naturally because workers' skills depreciate during spells of unemployment and firms must pay a cost to train unskilled workers. Consider an economy plagued by high unemployment. Since the average duration of unemployment is high, the average skill quality of the workforce is low. Consequently, firms need to spend more on training workers, which raises the effective average cost of job creation and makes firms less willing to post vacancies, even though slack labor markets lower workers' outside options and drive down wages. Thus, a high unemployment rate can be self-sustaining. Conversely, when unemployment is low, mean unemployment duration is low and the average skill of the workforce is high. While wages are high because tight labor markets improve workers' outside options, firms still find job creation attractive - as higher wages are offset by low average training costs when job-seekers are mostly highly skilled - sustaining low unemployment.

Given our Leontief matching function, the low unemployment steady state corresponds to zero unemployment.[18] The existence of this low unemployment steady state is guaranteed by the following assumption.

Assumption 1. *Vacancy posting costs are low enough: $\kappa < J_{min}$.*

Note that from the law of motion for employment (1), full employment ($n = 1$) implies $q = 1$ (and $f = 1/\theta \leq 1$). Job-seekers are on the short side of the market, and always find a job within one period. As such, skill depreciation never occurs, and the law of motion for the skill composition (12) implies $\mu = 0$. As a result, the effective cost of hiring a worker is simply κ/f. Thus, the job creation condition (11) becomes $\kappa\theta^{fe} = J_{min}$ in steady state. Assumption 1 ensures that this equation has a solution featuring $\theta > 1$, consistent with full employment.

[17] We use a more general matching function in Section 5 when we test the quantitative implications of our model.

[18] In the quantitative analysis in section 5, we use the more conventional CES matching function in which case the level of unemployment in the "low unemployment" steady state features an unemployment rate of about 5 percent instead of 0 percent.

While skill depreciation *can* generate multiple steady states, whether it in fact does so depends on the strength of the scarring effects of unemployment (measured by χ) and the sensitivity of wages to workers' outside options (measured by η). The following assumption ensures that both forces are strong enough such that in addition to the full employment steady state, there exists additional interior steady states featuring higher unemployment.

Assumption 2. *The training cost χ is neither too small nor too large, i.e., $\chi \in (\underline{\chi}, J_{max} - \kappa)$. In addition, workers' bargaining power is not too small, i.e. $\eta > \underline{\eta}$. The thresholds $\underline{\eta}$ and $\underline{\chi}$ are defined in Appendix B.*

$\kappa + \chi < J_{max}$ ensures that training costs are not too large, so that the worst steady state features a positive level of employment.[19] The remaining elements of Assumption 2 ensures that two interior steady states with unemployment exist (in addition to the full employment steady state). From the law of motion for employment (1), at any interior steady state ($n < 1$), firms are on the short side of the labor market ($q < 1$). This implies that there is some skill depreciation ($\mu > 0$), since from the law of motion for the skill composition (12), we have $q = 1 - \mu < 1$ in steady state. At an interior steady state, the job creation condition (14) becomes:[20]

$$\frac{a}{1 - \beta(1-\delta)\left[1 - \eta(1-\mu)\right]} = \kappa + \chi\mu. \tag{15}$$

The left-hand side (LHS) of (15) is the value of a filled vacancy with $q = 1 - \mu$, while its right-hand side (RHS) is the cost of creating a job. (15) describes a quadratic in μ which has at most two solutions. Appendix B shows that Assumption 2 guarantees that economically meaningful solutions to this equation exist. High bargaining power η increases the sensitivity of wages and profits to labor market conditions. When unemployment is low, wages are high because workers' outside option is relatively favorable. Firms are willing to tolerate high wages because training costs are low. When labor markets are slack and unemployment is high, workers are relatively unskilled and expensive to train; firms are willing to pay the high training costs because wages are relatively low.

Figure 1 graphically depicts the arguments above. The red curve plots the LHS of (15), while the blue line plots its RHS for different values of χ. When χ is too low, the two curves do not intersect and there are no interior steady states. When χ is too high, the blue line lies above the red curve at $\mu = 1$ and there exists a zero-employment steady state, in violation of Assumption 2. When χ is in the appropriate range, then there are two interior steady states, $\widetilde{\mu}$ and $\overline{\mu}$ (with $\widetilde{\mu} < \overline{\mu}$).[21] Finally, recall that there is always a full employment steady state at $\mu = 0$. The three steady states

[19] If, by contradiction, there was zero employment in steady state, everyone would be unskilled ($\mu = 1$) and labor markets would be completely slack ($q = 0$). Since wages are the lowest they can be, the value of a filled vacancy is J_{max}, the highest it can be. Since all workers are unskilled, the effective cost of hiring a worker is now $\kappa + \chi$. $\kappa + \chi < J_{max}$ ensures that at even such high levels of μ, a firm would find it profitable to post some vacancies, ruling out the uninteresting possibility of a zero employment steady state. Qualitatively, none of our results would change if we allowed for a zero employment steady state.

[20] In this section and in what follows, it will be convenient to work with the fraction of unskilled workers μ rather than the unemployment rate u as the state variable of interest. Notice that equation (4) defines a one-to-one map between μ_t and u_{t-1}.

[21] Since there is a one-to-one relation between μ and u, we also have $\tilde{u} < \overline{u}$.

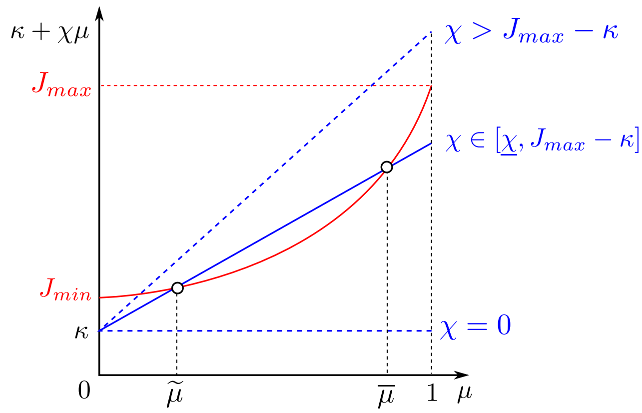


Figure 1. Multiple steady states

are associated with different degrees of market tightness, and accordingly, with different levels of wages: wages at the full employment steady state are higher than at moderate unemployment steady state $\widetilde{\mu}$, which in turn are higher than at the high unemployment steady state $\overline{\mu}$.

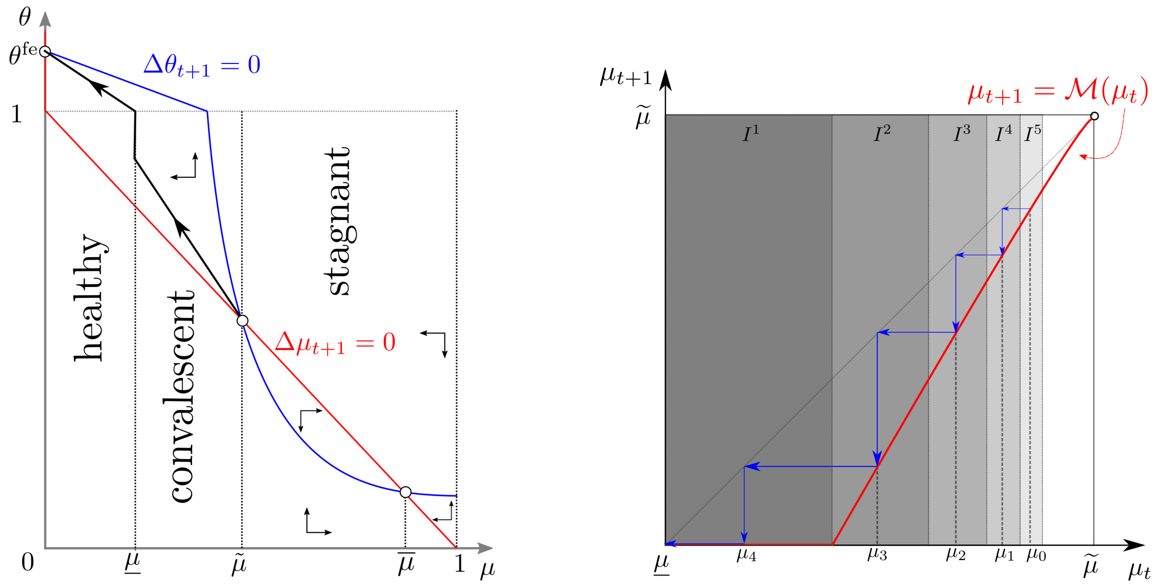
3.2 Dynamics

Next, we characterize the transitional dynamics of the economy starting from any $\mu_0 \in [0, 1]$. While we have thus far not introduced any aggregate shocks which would move the economy away from a steady state, we will do so in Section 4. For now, we can think of the experiment as studying the evolution of the economy after past shocks having moved it to a point μ_0.

The dynamics of the economy can be described by two equations in the market tightness θ and the fraction of unskilled job-seekers μ. The first one is (12), which describes how the current market tightness θ_t affects the change in the fraction of unskilled job-seekers μ_t. The second one can be obtained by substituting (12) and (14) into the firm's value from being matched (13), and describes the evolution of θ_t, given a value of μ_t:

$$\frac{\kappa}{f(\theta_t)} + \chi\mu_t = a + \beta(1-\delta)\Big(1 - \eta q(\theta_{t+1})\Big)\left[\frac{\kappa}{f(\theta_{t+1})} + \frac{\chi\Big(1 - q(\theta_t)\Big)}{1 + (1-\delta)[1 - q(\theta_t) - \mu_t]}\right]. \tag{16}$$

In equilibrium, the economy's evolution is decribed by a mapping $\mu_{t+1} = \mathcal{M}(\mu_t)$. Figure 2a describes the dynamics of the economy with μ on the horizontal axis and market tightness θ on the vertical axis. The red and blue curves respectively denote the nullclines corresponding to (12) and (16), and their intersections denote the three steady states described above. The upper-left intersection at $\left(0, \theta^{\text{fe}}\right)$ represents the full employment steady state, while the other two intersections denote the two interior steady states. As the figure shows, the state space can be partitioned into 3 regions, depending on the initial skill composition of job-seekers μ (or equivalently, the level of unemployment): (i) a *healthy region* which features low unemployment and a highly skilled workforce (low μ), (ii) a *convalescent region* which features moderate levels of unemployment and a



(a) Phase diagram **(b)** Dynamics in convalescent region

Figure 2. Global dynamics

moderately skilled workforce (intermediate level of μ) ; and finally (iii) a *stagnant region* with high unemployment and a largely unskilled workforce (high μ). Dynamics differ between these three regions, as we now describe.

Healthy region If the economy starts in the healthy region, defined as $\mu \in [0, \underline{\mu}]$ where $\underline{\mu} \equiv (J_{min} - \kappa)/\chi < \widetilde{\mu}$, then labor markets are tight and the economy immediately converges back to the full employment steady state, as formalized in the following lemma.

Lemma 1. *Suppose $\mu_0 \leq \underline{\mu}$. Then $\theta_t = (J_{min} - \chi\mu_0)/\kappa$ for $t = 0$, and $\theta_t = J_{min}/\kappa > 1$, $n_t = 1$, $\mu_t = 0$ for $t \geq 1$. Furthermore, $J_t = J_{min}$ and $\omega_t = \omega^*_{fe} \equiv \eta A + (1-\eta)b + \beta\eta(1-\delta)J_{min}$ for all $t \geq 0$.*[22]

Proof. See Appendix C. □

Intuitively, when the unemployment rate is very low, the average skill quality of job seekers is very high. Hence, low training costs make it attractive for firms to post enough vacancies to absorb all job seekers, despite the high wages associated with tight labor markets in the present and future. Consequently, unemployment duration is short (and equal to zero after the first period), and the skill quality of the workforce is high. Owing to workers' attractive outside options, wages are high and the value of a filled vacancy is low. While we have not yet introduced any shocks, one interpretation is that the full employment steady state is stable with respect to shocks which only cause small deteriorations in the average skill composition of job seekers. In particular, if μ_0 rises to

[22] The equilibrium is unique, except in the knife-edge case where $\mu = \underline{\mu}$, where there also exists other equilibria in which the economy returns to the full employment steady state in 2 periods instead of 1. Note, however, that in these equilibria the value of a filled vacancy and the real wage also satisfy $J_t = J_{min}$, $\omega_t = \omega^*_{fe}$ for all $t \geq 0$.

a level in the interval $(0, \underline{\mu}]$, the effect of the shock is immediately reversed as job seekers are still largely skilled, and firms are willing to post enough vacancies to hire and retrain all job-seekers on the spot. The equilibrium relation between θ_t and μ_t in this region is depicted by the black line in the top-left of Figure 2a. As long as $\mu_0 < \underline{\mu}$, $\theta_0 > 1$ and the economy immediately returns to full employment: $\mu_1 = \mathcal{M}(\mu_0) = 0$.

Convalescent region If the economy starts in what we label the convalescent region, defined as $[\underline{\mu}, \widetilde{\mu})$, it eventually returns to full employment, but does not do so instantaneously. The following proposition characterizes this dynamics.

Proposition 1 (Dynamics in the convalescent region)**.** *For β sufficiently close to 1, there exists a unique strictly increasing sequence $\{\mu^n\}_{n=0}^{\infty}$ with $\mu^0 \equiv \underline{\mu}$ and $\lim_{n\to\infty} \mu^n = \widetilde{\mu}$, such that if $\mu_0 \in I^n \equiv (\mu^{n-1}, \mu^n]$, the economy reaches the healthy region in n periods and reaches the full-employment steady-state in $n+2$, i.e., $\mu_n = \underline{\mu}$, $\mu_{n+1} \in (0, \underline{\mu})$ and $\mu_{n+2} = 0$. Furthermore:*

1. *Recoveries can be arbitrarily long: As $\mu_0 \to \widetilde{\mu}$, the time it takes for the economy to return to the healthy region tends to infinity.*[23]

2. *Recoveries can be arbitrarily slow: If μ_0 is close to $\widetilde{\mu}$, then μ declines very slowly early on in the recovery.*[24]

Proof. See Appendix D. □

Again, the black line in Figure 2a describes the equilibrium relation between μ_t and θ_t. Throughout the convalescent region, this black curve lies above the red curve ($\theta_t > 1 - \mu_t$), so μ_t is falling over time. Figure 2b illustrates this decline in μ_t, described in Proposition 1, by depicting the equilibrium starting from a point μ_0 in the convalescent region. The horizontal axis denotes μ_t, the vertical axis denotes μ_{t+1}, and the red curve denotes the function $\mu_{t+1} = \mathcal{M}(\mu_t)$. μ_0 is shown to lie in the interval $I^5 = (\mu^4, \mu^5]$, so it takes 5 periods for the economy to reach the healthy region, and 7 periods to reach full employment. During the transition, employment is growing over time and the proportion of unskilled individuals in the pool of job-seekers is shrinking. As long as the economy is in the convalescent region, labor markets are slack and real wages are low. And as soon as the economy reaches the interior of the healthy region, labor markets become tight and real wages reach their steady state level ω^*_{fe}.

When the fraction of unskilled job-seekers μ_0 is higher than $\underline{\mu}$, firms' expected training costs $\chi\mu_0$ are so high that firms are unable to recoup these costs if wages are high. Higher wages are supported only when labor markets going forward are expected to be tight as this improves workers' outside options. Thus, slack labor markets must persist for some time for firms to be willing to post vacancies today. In other words, in equilibrium, the labor market must experience a *slow recovery*. In fact, the speed of the recovery decreases in the initial fraction of unskilled job-seekers.

[23]Formally, for any $T \in \mathbb{N}$, there exists $\varepsilon > 0$ such that if $\mu_0 \in (\widetilde{\mu} - \varepsilon, \widetilde{\mu})$, $\mu_t > 0$ for all $t < T$.

[24]Formally, for any $\delta > 0$, $T \in \mathbb{N}$, there exists $\varepsilon > 0$ such that if $\mu_0 \in (\widetilde{\mu} - \varepsilon, \widetilde{\mu})$, $\mu_t > \mu_0 - \delta$ for all $t < T$.

Intuitively, a higher μ_t requires a lower job-finding rate for wages to be driven down and firms to be induced to post vacancies. But such a low job-finding rate in turn reduces the rate at which unskilled job-seekers are re-hired and regain skill. In the convalescent region, $\mu_t - \mu_{t+1}$ (the gap between the 45 degree line and red line in Figure 2b) is a decreasing function of μ_t: the worse the current state of the labor market, the slower it recovers. Accordingly, the economy can spend an arbitrary long time in the convalescent region before transitioning to the healthy region. When the economy starts deep in the convalescent region (μ_0 close to $\widetilde{\mu}$) the recovery takes disproportionately longer (point 1 of Proposition 1), and the rate at which the skill composition of job-seekers improves become smaller in the early stage of the recovery (point 2).

To visualize formally how a high μ_0 requires lower wages to sustain hiring, we rearrange the job-creation condition (7) to yield an expression for the real wage in the convalescent region relative to the real wage in the full employment steady state:

$$\omega_t^* = \omega_{fe}^* - \chi \left\{ [1-\beta(1-\delta)] \underbrace{(\mu_t - \underline{\mu})}_{\text{level}} + \beta(1-\delta) \underbrace{(\mu_t - \mu_{t+1})}_{\text{slope}} \right\},$$

where all terms inside the curly braces are strictly positive. This equation indicates that the real wage is necessarily lower in the convalescent region than under full employment, and expresses the deviation from that full employment steady state level as a weighted average of two terms.

The first term represents how far the economy is from the healthy region ($\mu_t - \underline{\mu}$) – the *level* effect on the wage discount. A large distance to the healthy region implies a larger fraction of job-seekers who would need to be trained if matched, lowering the value of posting a vacancy for a firm. For firms to create vacancies, they must be compensated with lower wages than in the healthy region to cover the higher training costs.

The second term relates to the speed at which labor market conditions improve ($\mu_t - \mu_{t+1}$) – the *slope* effect. A quick return to tight labor markets implies that a firm needs to pay higher wages in the near future (since tight labor markets increase job-finding rates and thus the outside options of workers). Higher future wages lower the the value of the firm in the future, discouraging vacancy posting today. To induce firms to post vacancies today, wages must therefore be even lower when a quick recovery is expected.

If the economy is deep inside the convalescent region, close to $\widetilde{\mu}$, then the level effect is the dominant force keeping wages down since the recovery is very protracted (see Proposition 1). However, if the economy is on the verge of exiting the convalescent region (μ_t close to $\underline{\mu}$), then the slope term dominates and it is the expectation of transitioning into the healthy region in the near future that keeps current wages low.

A corollary of Proposition 1 is that the wage ω_t^* is uniquely determined throughout the convalescent region. Since by Lemma 1, it is constant at $\omega_t^* = \omega_{fe}^*$ in the healthy region, it follows that it is unique and continuous throughout the convalescent and healthy regions. This wage $\omega^*(\mu_t)$, which we will refer to as the natural real wage, will play an important role in our analysis of monetary policy in Section 4.1.

Stagnant region If the economy starts in the stagnant region, defined as $[\widetilde{\mu}, 1]$, it never returns to full employment. Intuitively, when the fraction of unskilled job-seekers is this high, expected training costs are so high that they discourage firms from posting enough vacancies to bring the economy out of the stagnant region. Importantly, this is not because real wages are sticky. In the stagnant region, the high fraction of unskilled job-seekers μ_t is accompanied by depressed real wages which induce firms to post some vacancies. But such depressed real wages can only be sustained by a low job-finding rate, which in turn prevents unskilled workers from being hired and retrained in sufficient numbers for the economy to escape the stagnant region. This stagnant region is an *unemployment trap*.[25] We summarize this property in the following corollary.

Corollary 1 (Unemployment traps)**.** *If the economy is pushed into the stagnant region, i.e., $\mu \geq \widetilde{\mu}$, then it never returns to the full employment steady state.*

The right half of Figure 2a depicts dynamics in this region. As the arrows suggest, an economy that starts in the stagant region can never leave it. In principle, the policy function $\mu_{t+1} = \mathcal{M}(\mu_t)$ could describe multiple non-monotonic paths to the high unemployment steady state in this region, but this is unimportant for our purposes – what matters is that there exists *no* equilibrium in which the economy exits the stagnant region. Put differently, starting from the stagnant region, an optimistic belief that the economy will eventually return to full employment cannot be self-fulfilling. In fact, given a high μ_0, expectations of a return to full employment and a restoration of future skill composition of the workforce to high levels would lead firms to postpone hiring today in anticipation of lower training costs tomorrow. The lack of vacancy posting today in turn lengthens average unemployment duration and worsens the average skill quality amongst the pool of unemployed, entrenching the economy in the stagnant region. We discuss this in greater detail in Section 4.

Note that the same forces that trap the economy in the stagnant region are also responsible for generating a slow recovery in the convalescent region. When μ_t is high, labor markets must be slack to bring down wages and to maintain hiring, reducing the rate at which μ_t improves. In the stagnant region, this force is so severe that it prevents any permanent improvement in μ. In the convalescent region, in contrast, the situation is less dire, but even so, μ can only improve slowly.

4 Nominal rigidities

The analysis above highlighted that starting from a high level of unemployment, the economy may be unable to return to full employment. We now turn to our main objective: analyzing how temporary shocks can have permanent effects, depending on the conduct of monetary policy.

[25] To see this more formally, note that any trajectory which starts to the right of $\underline{\mu}$ and reached full employment at some date T has to be at $\underline{\mu}$ at date $T-2$. But Proposition 1 showed that *all* trajectories that reach $\underline{\mu}$ lie entirely within the convalescent region. It follows that if the economy starts in the stagnant region - defined as the set $[\widetilde{\mu}, 1]$ - it can never converge to full employment - this region is an *unemployment trap*.

Shocks We focus on the economy's response to a temporary demand shock, modeled as a temporary increase in households' patience: $\beta_0 > 1$, $\beta_t = \beta < 1$ for all $t > 0$. The New Keynesian literature has used this type of shock to capture an increase in the supply of savings which pushes real interest rates below zero.[26] In the present context, we choose to focus on a temporary demand shock (rather than, e.g., a temporary productivity shock) since it can only have persistent effects in the presence of nominal rigidities.[27]

Nominal rigidities The model specified in the previous section is characterized by the classical dichotomy and thus, monetary policy is unable to affect allocations. Since our objective is to understand whether monetary policy can prevent or moderate hystereses, we need monetary policy to have real effects. We thus break the classical dichotomy by introducing nominal rigidities in assuming that nominal wages are unable to freely adjust downwards. In particular, we suppose that at any date t the nominal wage must satisfy $W_t \geq \varphi W_{t-1}$. The parameter $\varphi \in (0, 1]$ limits how much nominal wages can fall between dates $t-1$ and t. In the spirit of Schmitt-Grohe and Uribe (2013), given the current state μ_t, we further assume that nominal wages are set to $W_t = \omega^*(\mu_t)P_t$ whenever possible, where $\omega^*(\mu_t)$ is the real wage in the flexible wage benchmark (given the state μ_t). However, if $\omega^*(\mu_t)P_t < \varphi W_{t-1}$, then $W_t = \varphi W_{t-1}$. To sum up, we postulate the wage setting rule:

$$W_t = \max \left\{ \varphi W_{t-1}, P_t \omega^*(\mu_t) \right\} \tag{17}$$

$\varphi = 1$ means that nominal wages cannot fall, while $\varphi \in (0, 1)$ implies that nominal wages can adjust downwards to some extent.[28] Note that even if nominal wages are unable to adjust downwards, *real* wages can still fall when inflation is positive: the wage setting rule (17) in real terms is given as:

$$\omega_t = \max \left\{ \frac{\varphi P_{t-1}}{P_t} \omega_{t-1}, \omega^*(\mu_t) \right\} \tag{18}$$

where $\omega_t = W_t/P_t$ denotes the prevailing real wage, which may differ from the flexible wage benchmark real wage, $\omega^*(\mu_t)$.[29]

4.1 Monetary policy

We assume that the monetary authority sets the nominal interest rate i_t subject to the ZLB constraint $i_t \geq 0$. Because of risk neutrality, the real interest rate is fixed and equal to $r_t = \beta_t^{-1} - 1$.

[26] In a richer model, such a shock could arise from a tightening of borrowing limits or an increase in precautionary savings motives. See, for example, Eggertsson and Krugman (2012), Guerrieri and Lorenzoni (2017).

[27] We discuss how productivity shocks affect the model economy in Section 6.

[28] This specification of nominal wage rigidities is not very restrictive and our characterization holds for any $\varphi \in (0, 1]$. Equation (17) just implies that nominal wages cannot *jump* downwards but can adjust downwards at some finite rate. See Remark 1 for a more in-depth discussion.

[29] The real wage can never fall below the flexible wage benchmark real wage, but it can exceed it.

Inflation then follows from the Fisher equation:

$$\frac{P_{t+1}}{P_t} = \beta_t(1+i_t) \tag{19}$$

Since nominal wages are not perfectly flexible downwards, monetary policy can affect real wages by influencing the price level. A large literature argues that monetary policy should seek to replicate allocations that would arise in an economy without nominal rigidities (Goodfriend and King, 1997; Woodford, 2003).[30] In this spirit, we assume that, whenever possible, the monetary authority sets the nominal interest rate so that, given the current state of the economy, real wages attain their flexible wage benchmark level, and nominal wages remain stable. From Section 3, we know that in the flexible wage benchmark economy, real wages are given by some function $\omega^*(\mu_t)$ of the current state μ_t, which itself evolves according to $\mu_{t+1} = \mathcal{M}(\mu_t)$.[31] Given μ_t and last period's nominal wages W_{t-1}, monetary policy attempts to implement a price level $P_t = W_{t-1}/\omega^*(\mu_t)$ which replicates allocations in the benchmark economy (starting from any initial condition μ_t) while keeping nominal wages constant.[32] Formally, monetary policy sets the nominal interest rate such that:[33]

$$P_t \leq \frac{W_{t-1}}{\omega^*(\mu_t)}, \qquad i_t \geq 0, \qquad \text{with at least one equality,} \tag{20}$$

and

$$\text{if} \qquad P_t = \frac{W_{t-1}}{\omega^*(\mu_t)}, \qquad \text{then} \qquad \mu_{t+1} = \mathcal{M}(\mu_t). \tag{21}$$

Given initial conditions μ_0 and W_{-1}, a *monetary equilibrium* is a sequence $\{W_t, P_t, i_t, J_t, \omega_t, \theta_t, \mu_{t+1}\}_{t=0}^{\infty}$ satisfying (8), (12), (14), (17), (19), (20), (21) and $\omega_t = W_t/P_t$ for all $t \geq 0$.

An immediate implication of the monetary policy rule described above is that in equilibrium nominal wages do not grow: $W_t/W_{t-1} \in [\varphi, 1]$.[34] Whenever the ZLB does not bind, monetary policy

[30] Clearly, this would be Pareto optimal if nominal rigidities were the only distortion rendering equilibrium inefficient. Such a policy may be optimal even in models with an inefficient real equilibrium. For example, the canonical New Keynesian model features inefficiencies due to both nominal rigidities and monopolistic competition.

[31] There are three technical issues worth clarifying here. First, when the economy is on the cusp of the healthy region ($\mu_t = \underline{\mu}$), there are multiple equilibria consistent with this value of μ_t. However, given a particular value of μ_{t-1}, the equilibrium is unique. Thus, to be precise, the state variables needed to describe the economy at this point are $\mu_t = \underline{\mu}$ and μ_{t-1}. This technicality is relevant for cases such as when the economy begins at $\mu_0 = \underline{\mu}$. In such cases, we adopt the convention of appending $\mu_{-1} = \underline{\mu}$, which eliminates the multiplicity. Similarly, for histories which cannot arise in the economy without nominal rigidities, such as $\mu_{t-1} = 0$ and $\mu_t = \underline{\mu}$, we follow the convention that the economy behaves as if it features the history $\mu_{t-1} = \mu_t = \underline{\mu}$. None of our results depend on this selection. Second, for some initial μ_0 in the stagnant region, there may be multiple equilibria. In this case, $\omega^*(\mu_t)$ refers to any selection from the set of equilibria. Again, none of our results depend on equilibrium selection.

[32] The fact that policy attempts to implement a flat path of nominal wages is not essential. More generally, monetary policy could also implement the same allocations if it wanted a constant rate of nominal wage inflation $\overline{\Pi} > 1$. We discuss the implications of this in Section 4.4.

[33] Our specification of monetary policy can be thought of as the limit of an interest rate rule:

$$1 + i_t = \max\left\{\beta^{-1}\left(\frac{P_t}{W_{t-1}/\omega^*(\mu_t)}\right)^{\phi_p}, 1\right\} \qquad \text{as} \qquad \phi_p \to \infty$$

[34] (17) states that $W_t = \max\{\varphi W_{t-1}, P_t\omega^*(\mu_t)\}$. This directly implies that $W_t/W_{t-1} \geq \varphi$. Combining (17) with (20), we also have $W_t \leq W_{t-1}$.

implements a price level consistent with constant nominal wages. As we will see next, a binding ZLB may cause nominal wages to fall, but these cannot decline at a rate greater than φ. Thus, in the particular case where $\varphi = 1$, nominal wages are necessarily constant in equilibrium.

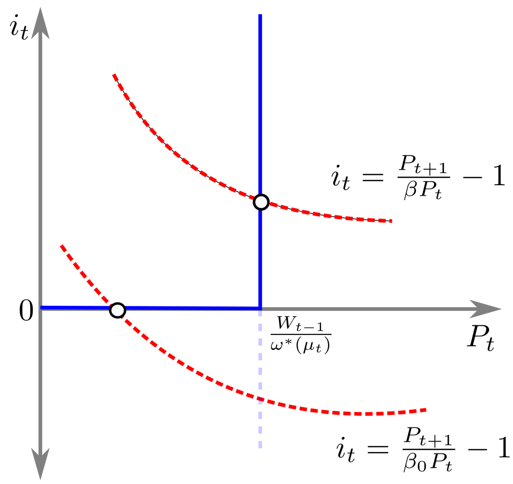


Figure 3. ZLB and the price level

Figure 3 describes the determination of the nominal interest rate and the price level at date t, given an anticipated future price level P_{t+1}. The dashed downward-sloping curve represents (19): the combinations of (i_t, P_t) consistent with bond market clearing for a given P_{t+1}. When the price level is higher, goods are more expensive today relative to tomorrow and households would rather save; a lower nominal interest rate discourages them from doing so, restoring bond market clearing. The higher dashed curve represents this locus when β is at its steady state level and the desire to save is moderate while the lower one represents the locus when $\beta_t > \beta$ and the desire to save is stronger. When the savings motive is stronger, lower nominal interest rates or prices are required to clear the bond market.

The solid curve depicts (20): the combination of (i_t, P_t) which monetary policy attains given the ZLB. Whenever possible, monetary policy stands ready to adjust the nominal rate so as to attain the price level $P_t = W_{t-1}/\omega^*(\mu_t)$, in order for the real wage to be at its flexible benchmark level for an unchanged nominal wage, given μ_t. When desired savings is not too high, this results in $P_t = W_{t-1}/\omega^*(\mu_t)$ and $i_t > 0$ (intersection of the solid vertical line and the upper dashed curve). In this case the flexible wage allocation is replicated, and nominal rigidities do not affect the dynamics of the economy. However, when desired savings is very high, implementing $P_t = W_{t-1}/\omega^*(\mu_t)$ would require a negative nominal interest rate (intersection of the dashed vertical line and the lower dashed curve), which is unattainable given the ZLB. Instead, the price level must fall to clear the bond market at a zero nominal interest rate (the intersection of the solid line and the lower dashed curve). Combining the Fisher equation (19) with the monetary policy rule (20) reveals that a binding ZLB imposes a lower bound on inflation, i.e., an upper bound on the date t price level, given the date $t+1$ price level: $P_t \leq P_{t+1}/\beta_t$. When the ZLB binds, the nominal return on bonds cannot be lowered anymore and the price of today's consumption must decline relative to the price of tomorrow's consumption to dissipate the excess demand for bonds. A binding ZLB thus curtails the ability of the monetary authority to implement its desired price level. The equilibrium price level is hence given by

$$P_t = \min \left\{ \frac{W_{t-1}}{\omega^*(\mu_t)}, \frac{P_{t+1}}{\beta_t} \right\} \tag{22}$$

Most crucially for our question of interest, a binding ZLB also has consequences for the labor market, since wages cannot freely adjust downwards. If the fall in prices required to clear the bond market when the ZLB binds is sufficiently severe, the nominal wage hits the constraint $W_t \geq \varphi W_{t-1}$, leading to a higher real wage than in the flexible benchmark. This lowers the value of a filled vacancy and discourages hiring. Less hiring in turn increases unemployment and lowers job-finding rates, raising the average duration of unemployment and worsening the skill quality of the job seekers relative to the flexible benchmark: $\mu_{t+1} > \mathcal{M}(\mu_t)$.[35]

4.2 Temporary shocks and permanent effects

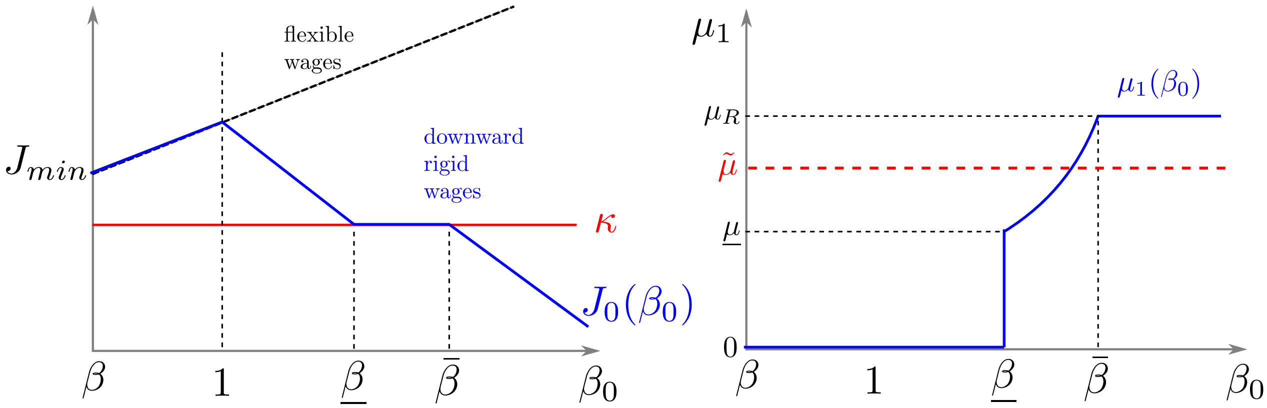
Assuming that the economy is intially at the full employment steady state ($\mu_0 = 0$), we now describe its response to a transitory demand shock modeled as a temporary increase in households' discount factor ($\beta_0 > 1$, $\beta_t = \beta < 1$ for all $t > 0$), as motivated above. If the ZLB did not prevent monetary policy from replicating the flexible wage outcome, a temporary increase in the discount factor would not raise unemployment. In fact, since a filled vacancy is a long-lived asset yielding dividends in the future and the cost of posting a vacancy is paid today, a temporary increase in the discount factor increases the net present value of vacancy posting, *encouraging* vacancy creation. However, when the ZLB binds, this *neoclassical effect* of an increase in households' desire to save can be outweighed by a deflationary *Keynesian effect*, causing a fall in the value of vacancy creation. This can lead to a decline in hiring, which in turn leads to a persistent or even permanent increase in the unemployment rate and the fraction of unskilled workers, due to the mechanisms described in Section 3. Proposition 2 formally describes the evolution of the economy following a one-period negative demand shock.

Proposition 2 (Monetary equilibrium with demand shocks). *There exists* $\underline{\beta} = \frac{A-\kappa}{\omega^*_{fe} - (1-\delta)J_{min}} > 1$ *such that:*

1. *if* $\beta_0 \in (1, \underline{\beta}]$*, there exists an equilibrium with* $\theta_0 \in (1, \theta^{FE})$ *and* $\mu_t = 0$ *for all* t*,*

2. *if* $\beta_0 > \underline{\beta}$*, there exists an equilibrium with* $\theta_0 \in [0, 1)$ *and* $\mu_1 \in (\underline{\mu}, \mu_R]$*,*

[35] More precisely, $\mu_{t+1} \in (\mathcal{M}(\mu_t), \frac{1}{1+(1-\delta)(1-\mu_t)}]$. In the worst case, real wages might rise so much that the net value of hiring an additional worker for a firm becomes negative, i.e., $J_t < \kappa + \chi\mu_t$. In this case, there is no vacancy posting, the job-finding rate is zero, and from (12), the fraction of unskilled workers next period is given by $\mu_{t+1} = \frac{1}{1+(1-\delta)(1-\mu_t)}$. This is the fastest possible rate of aggregate skill depreciation given μ_t. But even when a complete hiring freeze does not occur, higher real wages still induce a lower job-finding rate and and a higher fraction of unskilled workers μ_{t+1} than the fraction consistent with the flexible wage benchmark economy $\mathcal{M}(\mu_t)$.



(a) Value of a vacancy

(b) Skill composition of job-seekers

Figure 4. Economy's response to a demand shock.

where $\mu_R = 1/(2-\delta)$ is the rate of skill depreciation after a one-period hiring freeze, starting from full employment. Furthermore, for $\beta_0 > \underline{\beta}$, if $\mu_1 < \widetilde{\mu}$, then the economy eventually returns to full employment ($\lim_{T\to\infty} \mu_T = 0$) whereas if $\mu_1 \geq \widetilde{\mu}$, then the economy never returns to full employment ($\lim_{T\to\infty} \mu_T \geq \widetilde{\mu}$).

Proof. See Appendix F. □

To simplify the exposition of the economy's response to a temporary demand shock, it is useful to consider the limiting case where nominal wages cannot fall, i.e., where $\varphi = 1$.[36] As noted previously, in this case, the nominal wage remains constant in equilibrium. Achieving the full employment real wage at a given date would thus require implementing the price level $P_t = W_{-1}/\omega^*_{fe}$. Our temporary demand shock assumption ($\beta_0 > 1$, $\beta_t = \beta < 1$ for all $t > 0$) makes this feasible from date 1 onwards, but not at date 0. A temporary increase of the discount factor above 1 makes the ZLB bind at date 0, requiring a fall in price level today. Given constant nominal wages, such a fall in prices raises the real wage and reduces the value of vacancy creation. Thus, the demand shock has two opposite effects on vacancy creation: it increases the future value of a job (neoclassical effect) while reducing current profits (Keynesian effect). With perfectly flexible nominal wages, only the neoclassical effect would operate, generating a positive relationship between the value of a filled vacancy J_0 and size of the shock β_0, as shown by the dashed upward sloping line in Figure 4a. In contrast, with nominal rigidities, the Keynesian effect is also at work, and under Assumption 2, it always dominates the neoclassical effect, resulting in a negative relationship between J_0 and β_0 whenever $\beta_0 > 1$, as shown by the solid downward sloping line in 4a. In turn, the effect of this negative relationship onto labor market outcomes depends on whether the shock is moderate or large, as we now discuss.

[36] As explained in Remark 1 below, the same logic applies for any $\varphi \in (0, 1]$.

4.2.1 Moderate demand shock

When the demand shock is moderate, i.e., for $\beta_0 \in (1, \underline{\beta}]$, despite a binding ZLB and an accordingly higher real wage at date 0, the value of a filled vacancy does not decline enough to reduce vacancy posting substantially. As a consequence, the economy remains at full employment at date 0 and thereafter. To see this, suppose that the economy leaves date 0 with full employment ($\mu_1 = 0$). Then at date 1, monetary policy is no longer constrained by the ZLB, and the date 1 price level is given by $P_1 = W_{-1}/\omega^*_{fe}$. However, at date 0, the ZLB must bind and the increase in desired savings causes prices to fall to clear the bond market: $P_0 = P_1/\beta_0 < P_{-1}$. This deflation in turn increases the date 0 real wage $\omega_0 = \beta_0\omega^*_{fe} > \omega^*_{fe}$, which tends to reduce the value of a filled vacancy:

$$J_0 = \underbrace{A - \beta_0\omega^*_{fe}}_{\text{current profit}} + \underbrace{\beta_0(1-\delta)J_{min}}_{\text{valuation of future profits}}. \tag{23}$$

The two opposite effects of a higher β_0 onto the value of a filled vacancy are visible in (23): a lower current profit due to a higher current real wage (Keynesian effect) and an increased valuation of future profits (neoclassical effect). Under Assumption 2, the Keynesian effect dominates, i.e., $\partial J_0/\partial\beta_0 = -\omega^*_{fe} + (1-\delta)J_{min} < 0$.[37] As long as β is not too large, the value of a filled vacancy J_0 might fall, but it remains above the vacancy posting cost κ, implying that firms are still willing to post enough vacancies to keep the economy at full employment. Consequently, the economy remains at $\mu = 0$, as shown by Figure 4b, validating our earlier conjecture that it leaves date 0 with full employment. However, when $\beta_0 > \underline{\beta} \equiv \frac{A-\kappa}{\omega^*_{fe}-(1-\delta)J_{min}}$, the value of a vacancy would fall below the vacancy posting cost, and firms would not be willing to post enough vacancies to maintain full employment. We discuss this case next.

4.2.2 Large demand shock

Large enough demand shocks ($\beta_0 > \underline{\beta}$) cause a persistent or even permanent increase in unemployment. To see why, recall that when $\beta_0 > \underline{\beta}$, if firms expected full employment (tight labor markets) in the future, they would be unwilling to post vacancies at date 0 since we would have $J_0 < \kappa$. Thus, in equilibrium a large enough demand shock must move the economy away from the full employment steady state. Firms' incentive to post vacancies depends on real wages both today and in the future. If deflation increases real wages at date 0, firms would only be willing to post vacancies if they anticipate lower real wages in the future. Lower real wages in the future can only be an equilibrium outcome if labor markets in the future are slack, i.e. if the economy is away from full employment for a sustained period. Thus, unemployment must increase following a large demand shock.

Formally, the value of a filled vacancy at date 0 is then given by

$$J_0 = \underbrace{A - \beta_0\omega^*(\mu_1)}_{\text{current profit}} + \underbrace{\beta_0(1-\delta)(\kappa + \chi\mu_1)}_{\text{valuation of future profits}} \quad \text{where } \mu_1 \in [\underline{\mu}, \mu_R]$$

[37] See Appendix E for the derivation.

For firms to post vacancies at date 0 despite a high β_0, future profits must rise just enough to make the value of a filled vacancy J_0 in the above equation at least as large as the vacancy posting cost κ. For this to be the case, the skill composition must deteriorate enough to generate a context of low real wages even after the shock has abated and the ZLB no longer binds: the economy must enter the convalescent or stagnant region, i.e., $\mu_1 > \underline{\mu}$ (see Figure 4b). This adverse labor market development induces a stabilizing feedback effect: a higher μ_1 reduces date 1's real wage in the flexible wage economy, so monetary policy targets a higher price level $P_1 = W_{-1}/\omega^*(\mu_1) > W_{-1}/\omega^*_{fe}$, which tends to mitigate the decline in prices required to clear the bond market at date 0. As a result, it is possible to have $J_0 = \kappa$ even though $\beta_0 > \underline{\beta}$ (see Figure 4a). For extremely large demand shocks ($\beta_0 > \overline{\beta}$)[38] these stabilizing mechanisms are overwhelmed by the deflationary forces, resulting in zero hiring and increasing the fraction of unskilled workers at date 1 to $\mu_1 = \mu_R$ (see Figure 4b).[39]

Whether the economy is only persistently or permanently affected by the temporary shock depends on the equilibrium value of μ_1.

Slow recovery When the reduction in hiring drives the economy into the convalescent region, i.e., when $\mu_1 \in (\underline{\mu}, \widetilde{\mu})$, the recovery takes time but full employment is ultimately restored. Following an initial deterioration in the skill composition of the workforce due to a hiring slump, the economic forces underlying this slow recovery are essentially the ones outlined in Section 3. Faced with a higher likelihood of meeting unskilled applicants and hence higher expected training costs, firms only post vacancies if they are compensated by lower real wages. In turn, the only way low wages can be an equilibrium outcome is if job-finding rates are depressed for a period of time, keeping workers' outside option low. As a result, the unemployment rate and the fraction of unskilled job-seekers only decline gradually, but the economy ultimately converges back to the full employment steady state.

Permanent stagnation When the date 0 hiring slump takes the economy into the stagnant region, i.e., when $\mu_1 \geq \widetilde{\mu}$, the outcome is permanent stagnation. Again, conditional on the initial deterioration of the skill composition, the forces behind the ensuing stagnation dynamics are not nominal but real. The fraction of unskilled job-seekers is so high that real wages must be very low for firms to post *any* vacancies. Such low real wages can only be sustained if slack labor markets are expected to persist forever. In this scenario the economy converges to the low-pressure steady state with a high fraction of unskilled job-seekers $\overline{\mu}$ and high unemployment. In this steady state, even though high unemployment depresses wages, firms are reluctant to post vacancies because the average job-seeker is likely to be unskilled and costly to retrain. These low vacancy posting rates support high unemployment. Thus, even a transitory demand shock can permanently depress employment and output.

What determines whether the economy experiences a slow recovery or permanent stagnation

[38] See Appendix F for the definition of $\overline{\beta}$.

[39] In our economy with only one period shock, μ_R is the maximum damage that can be inflicted on the skill composition of the workforce during the ZLB episode, starting from full employment. With longer duration shocks, naturally there is no limit on the havoc that demand shocks can wreak.

in response to a large temporary demand shock? Unsurprisingly, one key element is the size of the demand shock. The larger the shock, the larger the decline in hiring and the worse the skill deterioration, making it more likely for the economy to land in the stagnant region. More interestingly, another key element relates to the strength of the forces generating multiplicity of steady states. The higher the training cost χ, the lower $\widetilde{\mu}$, and the more likely it is that the economy enters the stagnant region following a large temporary demand shock.[40] Figure 5 illustrates all these possibilities. The vertical axis denotes the size of the shock β_0 while the horizontal axis shows the magnitude of training costs χ. When $\beta_0 < \underline{\beta}$, the economy always returns to full employment. When $\beta_0 > \underline{\beta}$, the economy can experience either slow recovery or permanent stagnation; the latter outcome is more likely when β_0 and χ are larger. Figure 4b depicts the case in which $\mu_R > \widetilde{\mu}$, so a large enough shock that causes a hiring freeze at date 0 always drives the economy to the stagnant region. More generally, whether this is the case depends on the strength of the forces generating multiplicity, as Figure 5 illustrates.

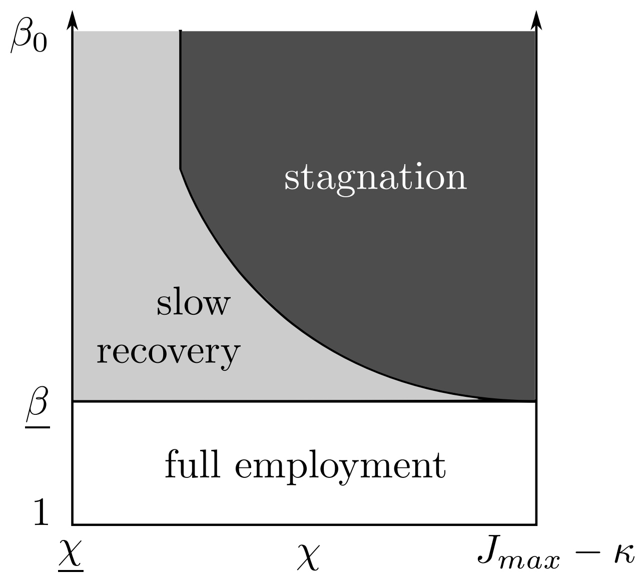


Figure 5. Parameters for which demand shocks lead to slow recoveries or stagnation.

Remark 1 (Paradox of flexibility)**.** *The characterization above holds more generally for any* $\varphi \in (0, 1]$. *Even when nominal wages can decline to some extent, deflationary forces drive them down to the lowest allowable level at date* 0*, i.e.,* $W_0 = \varphi W_{-1}$*. Since monetary policy is unconstrained at date* 1*, nominal wages remain constant between date* 0 *and* 1*, i.e.,* $W_1 = W_0 < W_{-1}$*. As a result, to implement the full employment steady state natural real wage* ω^*_{fe} *at date* 1*, the monetary authority must target a lower price level than the one prevailing initially:*

$$P_1 = \frac{W_0}{\omega^*_{fe}} < \frac{W_{-1}}{\omega^*_{fe}} = P_{-1}.$$

[40] More generally, it would also be true that the longer the duration of the shock, the more likely that the economy ends up in the stagnant region. We focus on one period shocks to emphasize that even very transitory recessions can have permanent effects.

Anticipating lower prices at date 1*, households are even less willing to consume today, further depressing the demand for current production. Thus, even lower prices must prevail at date* 0 *to clear goods markets. Lower prices in turn neutralize the benefits of the fall in nominal wages. This is a version of the* paradox of flexibility *(Eggertsson and Krugman, 2012).*

4.3 Monetary policy and hysteresis

The persistent or permanent effects of temporary shocks detailed in the previous sub-section makes our model stand out from the existing literature in monetary economics. First, and most obviously, it is in stark contrast to standard models with nominal rigidities in which transitory shocks do not affect long run outcomes. Second, it is also distinct from the recent literature on secular stagnation (e.g., Eggertsson and Mehrotra, 2014; Caballero and Farhi, 2017), where the focus is on the permanent effects of permanent changes in the environment.

The persistent or permanent effects of temporary shocks in our model are the result of a combination of nominal and real factors. First, nominal rigidities and the ZLB imply that large shocks can trigger hiring slumps, which take the economy away from full employment. Second, skill depreciation during unemployment – a purely real factor – is responsible for either a slow recovery or a permanent stagnation, once the economy finds itself away from full employment. This property has some noteworthy implications. First, nominal rigidities and the ZLB need not bind for very long for large adverse shocks to have persistent or permanent effect in our model. Provided the ensuing jump in unemployment is large enough, hysteresis can take root if downward wage rigidities and the ZLB bind for as little as one model period. This stands in contrast to much recent work on permanent effects of liquidity traps (Benigno and Fornaro, 2017; Schmitt-Grohe and Uribe, 2017; Acharya and Dogra, 2018). These models can generate permanently higher unemployment following a shock which drives the economy to the zero lower bound – but only if the ZLB and downward nominal wage rigidity also bind forever.

Second, we bring the idea of path dependence into the literature on liquidity traps and secular stagnation. Again, this yields starkly different predictions from Benigno and Fornaro (2017) and Schmitt-Grohe and Uribe (2017). These models feature multiple equilibria, one of which features deflation and high unemployment. Thus persistent unemployment can be an equilibrium outcome because agents' pessimistic beliefs are self-reinforcing. But if agents woke up one morning and expected the economy to return to full employment, the economy would indeed return to full employment. In contrast, our economy is not trapped in the high unemployment steady state because of self-fulfilling beliefs. In fact, starting from the high unemployment steady state, if (off equilibrium) firms anticipated a return to full employment, this would make them *less* willing to hire workers today, since they would anticipate a more skilled workforce and lower costs of job creation tomorrow. Lack of hiring today would further cement the skill deterioration in the workforce and the presence of high unemployment rates. Thus, persistently high unemployment arises in our model not because of self-fulfilling beliefs, but because our economy features an endogenous slow-moving state variable, namely the skill composition of job-seekers. Once hysteresis has taken its toll, it takes more than wishful thinking to reverse the damage.

Finally, it is worth highlighting that the ZLB works through a different channel in our model than in the standard New Keynesian model. In the standard model, the lower bound on nominal rates, together with an upper bound on expected inflation, make real interest rates too high, and consumption growth too large. If long run consumption is pinned down (by the assumption that the economy returns to the unique steady state), high consumption *growth* requires consumption to fall today. In contrast, in our model with linear utility, real rates are fixed at $1/\beta_t$, and the ZLB requires positive inflation to clear the bond market in response to negative demand shocks. If long run prices are pinned down (by monetary policy), high inflation requires prices to fall today. With downward sticky nominal wages, this fall in prices raises real wages, further discouraging vacancy creation and reinforcing the forces that support the onset of hysteresis. Critically, this effect of prices on job-creation has implications for the conduct of monetary policy, as we discuss next.

4.4 Can unconventional policies prevent hysteresis?

Having shown that temporary demand shocks can lead to permanent stagnation, we now ask whether monetary policy can act to avoid such outcomes. Under our baseline monetary policy specification, where the central bank tries to implement the natural allocation while ensuring zero nominal wage inflation, the ZLB binds at date 0, forcing prices to fall and real wages to rise. Therefore, as is common in the liquidity trap literature, a higher inflation target $\overline{\Pi} > 1$ could relax the ZLB constraint, making it bind only when $\beta_0 > \overline{\Pi}$ (rather than when $\beta_0 > 1$). However, such a policy would not altogether insulate the economy from the persistent effects of demand shocks. Even with $\overline{\Pi} > 1$, sufficiently large demand shocks would still cause the ZLB to bind, and potentially generate a persistent or permanent downturn as described above.

More interestingly, as emphasized by a large literature (Eggertsson and Woodford, 2003; Werning, 2011), avoiding adverse outcomes in liquidity trap episodes can be achieved by commitments to temporarily more accommodative policy after the trap is over. In our economy, this policy prescription takes the form of temporarily higher wage inflation at date 1, i.e. a permanently higher nominal wage level from date 1 onwards. The proposition below describes how such a commitment can succeed at keeping the economy at full employment following an adverse shock.[41]

Proposition 3 (Hysteresis-proof policy)**.** *Suppose that $\beta_0 > \underline{\beta}$ and the economy starts at full employment, i.e. $\mu_0 = 0$ and $W_{-1} = \omega^*_{fe} P_{-1}$. Then, the hysteresis-proof monetary policy implements a price sequence given by $P_0 = P_{-1}$ and $P_t = \beta_0 P_{-1} > P_{-1}$ for $t > 0$. The unique equilibrium consistent with this price path features full employment for all t and a nominal wage path $W_t = \omega^*_{fe} P_t$.*

Proof. Date 1 inflation equals $P_1/P_0 = \beta_0 > 1$, which clearly satisfies the ZLB constraint. Real wages are given by equation (18), which implies that $W_t/P_t = \omega^*_{fe}$ along the price path defined above, since $\varphi P_{t-1}/P_t \leq 1$ for all $t \geq 0$. $\square$

[41] Note that such a policy does not require a commitment to higher trend inflation.

Commitment to higher date 1 prices - which involves deviating from nominal wage stability even after the shock has abated - prevents prices from falling at date 0. Even with nominal rates stuck at zero, higher future prices discourage households from saving in nominal bonds, propping up the goods demand at date 0 and preventing deflation. Without a fall in prices, real wages do not rise, job creation remains profitable, and unemployment remains low. Thus, the economy never enters the convalescent or stagnant regions and hysteresis is averted.

Remark 2 (Forward guidance and the intertemporal substitution channel)**.** *The policy just described involves commitment to expansionary policy once the liquidity trap has abated. In this regard, it resembles* forward guidance *policies discussed in the recent literature.*[42] *That literature argues that committing to keep interest rates low in the future can alleviate demand driven recessions. In the context of standard New Keynesian models, this policy works through an intertemporal consumption smoothing channel. There is an ongoing debate about whether the strength of forward guidance in New Keynesian models is realistic*[43] *and about the strength of the intertemporal substitution channel.*[44] *Importantly, the power of commitment to higher future prices in our framework does not depend on the intertemporal substitution channel, which is entirely absent. Instead, a commitment to a higher future price level implies that current prices need not fall in order to deliver a given level of inflation going forward. As such, real wages do not rise and hiring does not contract. Since our hysteresis-proof policy operates through the effect of the price level on firms' hiring decisions, rather than via the time path of real interest rates, it does not depend on the strength of the intertemporal substitution channel.*

4.5 Escaping unemployment traps?

The hysteresis-proof monetary policy just described can prevent adverse shocks from causing a recession, but requires a commitment to higher prices in the period immediately following the shock. What if instead of adopting such a wise and effective policy, the central bank has allowed unemployment to rise severely enough that the economy enters the stagnant region? Can it reverse course and return the economy to full employment? Unfortunately, in our model, the answer is negative: monetary policy cannot engineer an escape from an unemployment trap. In the stagnant region, employers are unwilling to create more vacancies despite prevailing low real wages. Temporary wage cuts would incentivize hiring, but given that nominal wages are fully flexible upwards, monetary policy *cannot* reduce real wages below their natural level. Similarly, even if the recession generates a slow recovery rather than a permanent stagnation, monetary policy cannot speed up the recovery: temporarily lower real wages would in principle stimulate hiring, but monetary policy cannot reduce real wages. This result is a natural reflection of the observation that the slow recovery and permanent stagnation phenomena in our model are purely driven by real factors.

[42] To be clear though, the policy described here is not a commitment to keep nominal rates at zero for an extended period of time.

[43] See, for example, Del Negro et al. (2015) and McKay et al. (2016).

[44] See, for example, Kaplan et al. (2018).

Our analysis thus delivers the stark prediction that monetary policy can prevent any increase in unemployment if it acts at date 0, but is completely powerless to improve outcomes if it waits until date 1. The reason is that in between these two dates, absent timely action by the monetary authority, an adverse dynamic in a key endogenous state variable – unemployment or the skill composition – has set in and cannot be reversed. This contrast between the power of monetary policy at date 0 and date 1 is arguably excessively stylized. For instance, monetary policy can avert the recession altogether at date 0 only because we assume full commitment and there are no costs associated with permanently higher nominal prices and wages.[45] Such trade-offs would make a commitment to higher prices time-inconsistent. Similarly, monetary policy is ineffective after date 0 only because it cannot raise employment above its natural level. If instead it had some ability to create temporary hiring booms, then it could reverse some of the damage done to the economy.[46] With these caveats in mind, the notion that the timing of monetary policy actions is crucial for preventing hysteresis would remain true in a more general environment. Prompt policy intervention can limit the long-term damage caused by temporary shocks whereas delayed action may be ineffective. By contrast, in standard models such as Eggertsson and Woodford (2003), such policies are equally effective at any point in a crisis. Delaying monetary accommodation in such models is costly, but the costs are only temporary. Unlike the existing literature which generally finds that policy should put little weight on output or employment stabilization, our model suggests that in the presence of hysteresis, failing to stabilize employment can be extremely costly as it sets the economy down a path where it can fail to return its original low unemployment steady state.

Finally, it is worth noting that even within our model, fiscal policy, such as hiring or training subsidies, could potentially counter the effects of hysteresis when monetary policy is powerless. Committing to compensate firms for each worker they train would be equivalent to lowering the private training cost χ, speeding up recovery or even lifting the economy out of the stagnant region. Subsidizing job-creation could have similar effects.

5 Hysteresis since the Great Recession

We now ask whether the mechanisms described in the paper can help quantitatively explain the sluggish economic recovery following the Great Recession.

While the particular form of the matching function assumed above facilitates analytical results, it has the counterfactual prediction that the "high pressure" steady state has 0 percent unemployment. In what follows, we use a more standard matching function $m(v, l) = vl/\left(v^{\iota} + l^{\iota}\right)^{\frac{1}{\iota}}$.[47] This allows us

[45] If nominal wages were inflexible in both directions, implementing a higher future price level may involve a deviation from natural allocations and welfare losses. At the full employment SS, lower real wages than the natural level lead to welfare losses. Lower wages encourage firms to create more vacancies. With the economy already at full-employment, higher market tightness only leads to higher vacancy posting costs, and thus lower consumption, without creating more jobs. This is the manifestation of the Hosios (1990) externality in our setting.

[46] So long as nominal wages are not fully flexible upwards, targeting higher prices once the shock has passed would stimulate hiring by temporarily lowering real wages - potentially escaping the stagnant region.

[47] Notice that as $\iota \to \infty$, this matching function converges to the $\min\{\cdot,\cdot\}$ matching function employed in the previous sections.

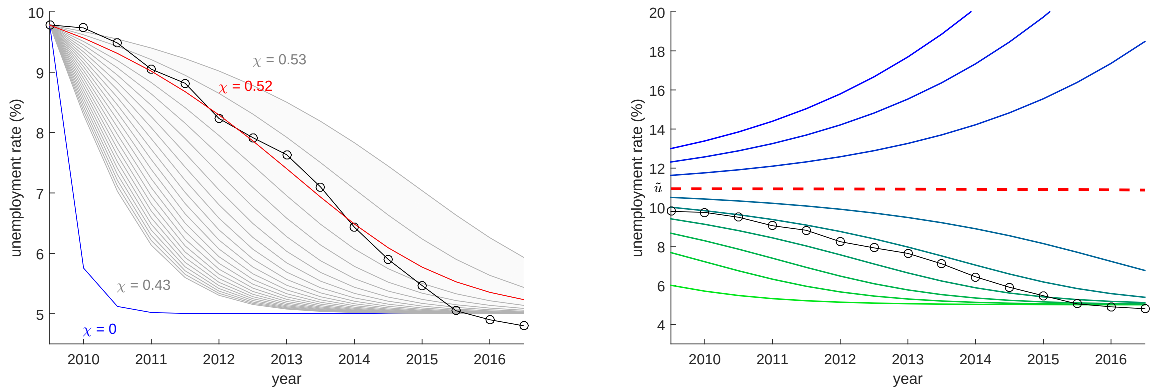
to consider a high pressure economy with an empirically plausible unemployment rate. We calibrate the model to the U.S. economy. In our model, unemployed workers lose skill after one period. We calibrate one period to six months, so "unskilled" workers correspond to those unemployed for 27 weeks or more, i.e., the long-term unemployed. Using resume audit studies, Kroft et al. (2013) find that the likelihood of receiving a call-back declines significantly after an unemployment spell, with most of the decline occurring within the first 8 months. Similarly, Ghayad (2014) finds a sharp decline in the call-back rate after 6 months. It therefore seems reasonable to posit that most of the skill loss upon losing a job occurs within the first 6 months.[48] We set $\beta = 0.98$, implying a 4 % annualized steady state real interest rate. A is normalized to 1. We set $\iota = 0.5$ following Menzio and Shi (2011) and $\eta = 0.7$ (Shimer, 2005). b is chosen to imply a steady state replacement ratio of 70 percent (Hall, 2009). We set $\delta = 0.2105$ so that the 5 percent steady state unemployment is consistent with 20 percent of job seekers being long-term unemployed in steady state as observed in the U.S. before 2008. This leaves us with two parameters, κ and χ, to target one remaining moment. We consider a range of values for these parameters.

After the Great Recession, U.S. unemployment peaked at close to 10 percent in the second half of 2009 before beginning a slow decline, not returning to 5 percent for another 6 years, as shown by the black line in Figure 6a. We have already seen analytically (Proposition 1) that our model can in principle generate an arbitrarily slow recovery. To see whether a slow recovery is plausible quantitatively, we solve our model for a range of (κ, χ) combinations consistent with multiple interior steady states, starting from the unemployment rate of 9.8% observed in the second half of 2009. These trajectories are shown by the gray lines in Figure 6a. Moving outwards from the origin, as we increase χ and decrease κ, the forces generating multiplicity become stronger and the recovery becomes slower. Quantitatively, the model can match the sluggish recovery observed in the data. The red line in Figure 6a indicates our preferred calibration, $\chi = 0.52$, which fits the data most closely. While direct evidence on training costs is hard to come by, this lies well within the empirical estimates found in the literature.[49]

The blue line in Figure 6a shows the trajectory of unemployment when $\chi = 0$, i.e. a model without training costs - essentially the standard DMP model, which has a unique steady state. Absent any further shocks, this model predicts a rapid recovery with unemployment returning to 5

[48]Strictly speaking, in our simple model, skilled and unskilled job-seekers have the same job-finding rates and are paid the same wages conditional on employment: skill depreciation only shows up in the training costs faced by employers. For our purposes, all that matters is that the findings of these studies is consistent with a bulk of the skill depreciation occurring within 6 months of job-loss. Our rate of skill depreciation is also consistent with Ljungqvist and Sargent (1998); their calibration implies a 95% probability of losing skill after a 6 month unemployment spell.

[49]Barron et al. (1989) find that, on average, a new hire spends 151 hours on training in the first 3 months of the job. If only unskilled workers require training, as assumed in our model, this implies an upper bound of $\chi = 151/(0.2 * 1043.5) = 0.72$(since the average fraction of unskilled workers in the US prior to 2008 was 20 percent, and assuming 2087 hour work-year, as is standard. Barron et al. (1989) also find that the median worker spends 81 hours in training. If we instead calibrate χ to match the difference between training costs between unskilled and skilled (median) worker, we get $\chi = (151 - 81)/(0.2 * 1043.5) = 0.34$.) The American Society for Training and Development (Paradise, 2009) estimated the average annual learning expenditure to be 2.24% of total annual payroll in 2008. In our model total training expenditures equal $\chi\mu\delta(1-u)$ in steady state while payroll equals $w(1-u)$ implying $\chi = 0.48$. Our preferred value of $\chi = 0.52$ lies comfortably within this range.



(a) Model vs data

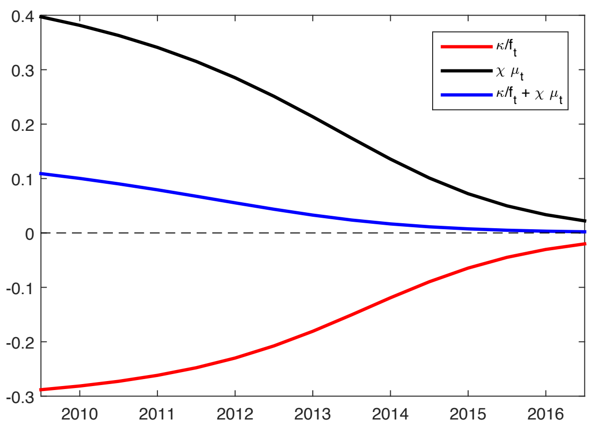
(b) Different starting points

Figure 6. Trajectories of unemployment

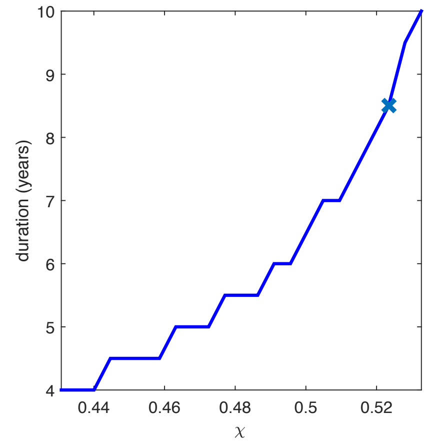
percent by the end of 2010. This suggests that in the absence of persistent shocks, our mechanism is *necessary* to match the sluggish recovery observed in the data. As highlighted in Pissarides (2009), when firms post fewer vacancies due to poor aggregate conditions, competition for workers amongst recruiters declines, shortening the average duration of a vacancy. When the only costs associated with job creation are vacancy posting costs, the decline in vacancy duration lowers the average effective cost of job creation, mitigating the recession's adverse impact on hiring. Training costs, χ, undo this phenomenon, creating a protracted recovery. Unlike the average cost of vacancy creation, κ/f, which is pro-cyclical and rises when there is more competition amongst recruiters, training costs are counter-cyclical and rise when the composition of job-seekers tilts towards the unskilled. Figure 7a highlights the inverse relationship between the average expected vacancy posting cost κ/f (red line) and the expected training costs $\chi\mu$ (black line) in the years following the Great Recession. Notably, the increase in job creation costs (blue line) solely stemmed from the large spike in training costs. The sharp rise in training costs more than counteracted the benefit of lower average expected vacancy posting costs following the Great Recession, depressing job growth and stalling the recovery.

Figure 6a suggests that, given the magnitude of the shock that hit the U.S. during the Great Recession, monetary policy was accommodative enough to avert permanent stagnation, but not enough to prevent a slow recovery. The model allows us to evaluate how the economy would have responded had shocks been larger or policy less accommodative. Figure 6b shows the trajectory of unemployment under our preferred calibration given different initial unemployment rates (in the second half of 2009).[50] Again, the black line plots data. The green lines show trajectories starting from lower initial unemployment; light blue lines indicate trajectories starting from higher unemployment. The red-dashed line shows the trajectory starting from $\tilde{u} = 10.9\%$ (the unstable steady state) which divides the regions of slow recovery and permanent stagnation. The figure shows that, had monetary policy been more accommodative after the initial shocks and kept the initial rise in unemployment below 8 percent, unemployment would have returned to 6% two years

[50] These initial combinations refer to different combinations of shocks and policy.



(a) Components of job-creation costs



(b) Duration as function of χ

Figure 7. The role of χ in driving costs and duration

earlier, in 2012. More strikingly, had policy been less accommodative allowing unemployment to rise to 12%, the economy would have been unable to return to full-employment (absent fiscal policy).

Figure 6b suggests that had monetary policy not been so quick to respond following the shocks in 2007, the U.S. economy could have fallen into permanent stagnation. In this regard, Europe presents a cautionary tale. Figure 8 shows the fraction of long-term unemployed in Ireland, Greece, Spain, the Euro area and the U.S. from 2008 to 2016. While the fraction of long-term unemployed increased in the U.S. following 2007, timely monetary policy accommodation ensured that this increase was temporary. In contrast, the fraction of long-term unemployed increased following the European recessions of 2008 and 2011 and has since remained elevated.[51] Many commentators[52] have argued that the European Central Bank's response was insufficient from the point of view of these economies or came too late. The model suggests that this delayed or insufficient monetary policy response could explain why long-term unemployment has remained persistently high in these economies. From this perspective, ECB President Draghi's "whatever it takes" speech in July 2012 - which can be seen as a commitment to very accommodative policy - may have come too late to reverse the effects of hysteresis. Thus, to prevent hysteresis, monetary stimulus must not just be large, but also timely.

Finally, even when monetary policy cannot ameliorate the scarring effects of recessions, fiscal policy - hiring or training subsidies - can help. Figure 7b shows the duration of the recovery (starting from 9.8% unemployment) as a function of χ with the cross denoting our preferred value of χ. The effect of χ on duration is highly non-linear. A modest subsidy which lowers the per worker training cost χ by 4% from 0.52 to 0.5 could have hastened the recovery by over 2 years.

[51] Of course, experiences differed widely across European countries; for example, the fraction of long -term unemployed actually declined in Germany over this period. Here we focus on those countries most severely affected during the European crisis, from whose perspective the ECB's monetary policy was arguably insufficiently accommodative. Having said that, despite the heterogeneous experiences of different countries, on average, the Euro area did experience an increase in the fraction of long-term unemployed.

[52] See for example Kang et al. (2015).

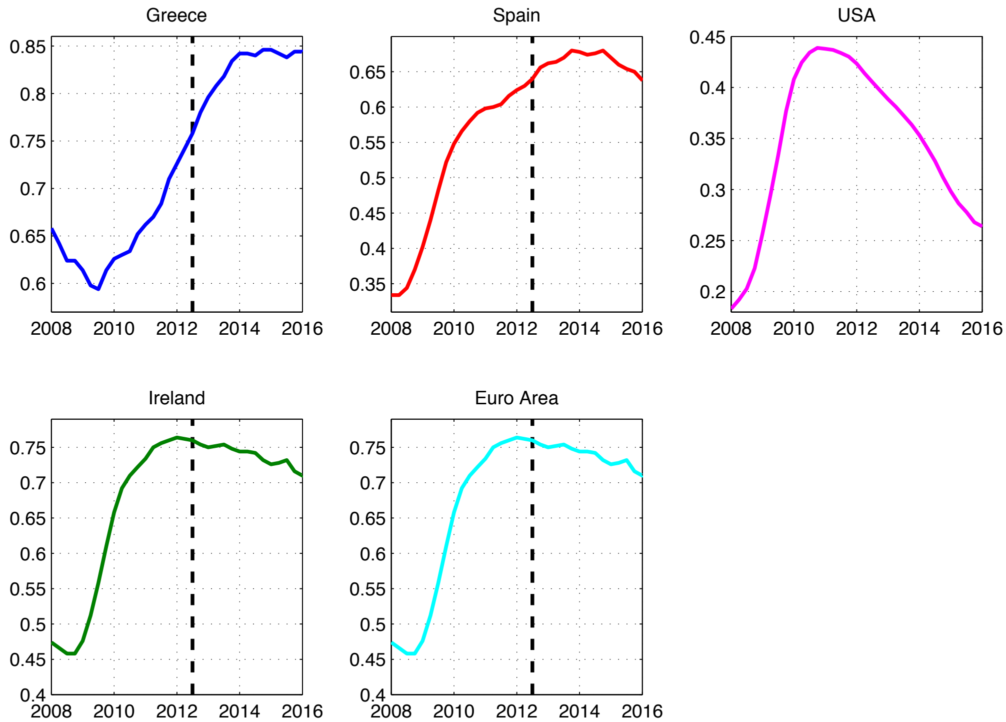


Figure 8. Fraction of Long-term unemployed (>27 weeks) in select countries. The figure plots five quarter moving averages of quarterly data. The dashed-line indicates the timing of Draghi's "whatever it takes" speech. Source: Eurostat and FRED.

6 Some extensions

Market segmentation A key assumption in our analysis has been that both skilled and unskilled workers search in the same markets. In particular, firms cannot post vacancies targeted to either unskilled or skilled workers. In reality, employers might be able to discern whether a worker requires training or not based on observable characteristics - in particular, their duration of unemployment. One might therefore wonder whether allowing firms to observe a worker's type would prevent slow recoveries and permanent stagnation from arising in our economy.

If instead skilled and unskilled workers searched in separate markets, the economy would still be characterized by hysteresis, but it would take a different form. There are two possibilities to consider. Under our maintained assumption that $J_{max} > \kappa + \chi$, the firm's share of the surplus from hiring an unskilled worker, net of training costs, is large enough to compensate firms for posting vacancies in the unskilled labor market - provided that this market is sufficiently slack. Thus, after a temporary recession which increases the fraction of unskilled job-seekers, it can take a long time for these workers to be reabsorbed into employment. Firms prefer to post vacancies in the market for skilled job-seekers rather than the market for unskilled job-seekers in order to avoid paying a training cost. With fewer vacancies posted for them, unskilled job-seekers face a lower job-finding rate and thus, the outflow from the pool of unskilled job-seekers is low. In contrast, the skilled unemployment rate recovers rapidly - in fact, faster than in the baseline model with a single labor market. However, the recovery is even slower for unskilled job-seekers relative to the baseline model.

If instead $J_{max} < \kappa + \chi$, the segmented labor markets economy could experience permanent stagnation, rather than a slow recovery, but again this stagnation would take a different form, relative to the baseline model. In this case, unskilled workers are unemployable, since firms are unwilling to pay the cost of hiring and training these workers. Thus unskilled workers effectively drop out of the labor force. Again, with perfectly segmented labor markets, unskilled workers exert no externality on the skilled labor market, which recovers rapidly. However, this economy experiences a fall is labor market participation.

Wage bargaining We have assumed that firms pay training costs before bargaining over the wage. This is not essential for our results. If instead firms bargained over the wage before paying training costs, our results would be broadly unchanged except that the effective training cost faced by firms would be given by $(1-\eta)\chi$.

Productivity shocks We have focused on how a temporary negative demand shock can permanently damage the economy's productive capacity by altering the skill composition of workers in the presence of nominal rigidities and insufficient monetary policy accommodation. Our results would similarly apply if the economy were beset by temporary negative productivity shocks instead.[53] That is, large enough falls in aggregate productivity cause unemployment to rise and average skill quality to deteriorate significantly, resulting in permanent scarring of the workforce. But since large enough negative productivity shocks drive unemployment up even in the absence of a binding ZLB, we chose to focus on demand shocks.[54]

Rate of skill depreciation In order to limit the number of state variables and retain analyticsl tractability, we have assumed that workers lose their skill after being unemployed for 1 period only. Relaxing this assumption would not qualitatively change our results. Allowing workers to lose their skill probabilistically would merely slow down the deterioration in the skill quality of the unemployed. As long as large adverse shocks reduce hiring enough and significant large deteriorations in the skill quality of job-seekers, the economy cannot observe a swift reversal back to full employment and all our results follow through.

7 Conclusion

We presented a model designed to study the positive and normative implications of hysteresis. Skill depreciation, nominal rigidities and constraints on monetary policy together allow temporary shocks to create slow recoveries or even permanent stagnation. Aggressive countercyclical policy may be

[53] In an earlier version, we looked at the impact of productivity shocks and found broadly similar results: temporary productivity shocks can have persistent real effects. The results on productivity shocks are available upon request.

[54] As emphasized in Section 4, in the absence of a binding ZLB, a negative demand shock has no sustained deleterious effect on the economy. But when monetary policy is constrained by the ZLB, timely and reactive monetary policy may be necessary to stymie a deterioration in the skill composition of job-seekers and curtail any onset of hysteresis.

able to avoid these outcomes, but only if enacted in a timely manner. Once the rot has set in, monetary policy cannot rescue the economy from an unemployment trap.

Hysteresis gained renewed attention during the recent recovery, as policymakers debated whether to begin tightening policy. Yellen (2016) discusses the argument that in the presence of hysteresis, monetary policy should "run the economy hot" to reverse the damage caused by high unemployment. Our analysis suggests that such monetary accommodation may be less effective when implemented late in the recovery, rather than early on in the recession.

While we have focused on skill depreciation, more generally recessions may damage productive capacity through multiple channels - reducing capital accumulation, reducing labor force participation, slowing productivity growth, and so on. Many of these effects may also be hard or even impossible to reverse. For example, Wee (2016) shows that recessions can permanently change young workers' search behavior, causing them to stay in careers in which they have a comparative disadvantage but have accumulated sufficient specific human capital, causing permanent misallocation. Whenever such mechanisms are operative, it is all the more important for countercyclical policy to nip recessions in the bud; the damage from failing to do so may be irreversible. In a world vulnerable to hysteresis, prevention is better than cure.

References

Acharya, Sushant and Keshav Dogra, "The Side Effects of Safe Asset Creation," Staff Reports 842, Federal Reserve Bank of New York March 2018.

Ball, Laurence, "Long-term damage from the Great Recession in OECD countries," *European Journal of Economics and Economic Policies: Intervention*, September 2014, *11* (2), 149–160.

Ball, Laurence M., "Hysteresis in Unemployment: Old and New Evidence," NBER Working Papers 14818, National Bureau of Economic Research, Inc March 2009.

Barron, John M, Dan A Black, and Mark A Loewenstein, "Job Matching and On-the-Job Training," *Journal of Labor Economics*, January 1989, *7* (1), 1–19.

Benigno, Gianluca and Luca Fornaro, "Stagnation Traps," *Review of Economic Studies*, 2017, *85* (3), 1425–1470.

Berger, David, Ian Dew-Becker, Konstantin Milbradt, Lawrence Schmidt, and Yuta Takahashi, "Layoff risk, the welfare cost of business cycles, and monetary policy," Technical Report 2016.

Bianchi, Francesco, Howard Kung, and Gonzalo Morales, "Growth, Slowdowns, and Recoveries," Working Paper 20725, National Bureau of Economic Research December 2014.

Blanchard, Olivier and Lawrence Summers, "Hysteresis and the European Unemployment Problem," in "NBER Macroeconomics Annual 1986, Volume 1," National Bureau of Economic Research, Inc, 1986, pp. 15–90.

_ **, Eugenio Cerutti, and Lawrence Summers**, "Inflation and Activity Two Explorations and their Monetary Policy Implications," Working Paper 21726, National Bureau of Economic Research November 2015.

Caballero, Ricardo J and Emmanuel Farhi, "The Safety Trap," *Review of Economic Studies*, 2017, *85* (1), 223–274.

Del Negro, Marco, Marc Giannoni, and Christina Patterson, "The forward guidance puzzle," Staff Reports 574, Federal Reserve Bank of New York 2015.

Dickens, William T., "The productivity crisis: Secular or cyclical?," *Economics Letters*, 1982, *9* (1), 37 – 42.

Drazen, Allan, "Cyclical Determinants of the Natural Level of Economic Activity," *International Economic Review*, June 1985, *26* (2), 387–97.

Dupraz, Stephane, Emi Nakamura, and Jon Steinsson, "A Plucking Model of Business Cycles," Technical Report 2017.

Eggertsson, Gauti B. and Michael Woodford, "The Zero Bound on Interest Rates and Optimal Monetary Policy," *Brookings Papers on Economic Activity*, 2003, *34* (1), 139–235.

_ **and Neil R. Mehrotra**, "A Model of Secular Stagnation," NBER Working Papers 20574, National Bureau of Economic Research, Inc October 2014.

_ **and Paul Krugman**, "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach," *The Quarterly Journal of Economics*, 2012, *127* (3), 1469–1513.

Fallick, Bruce C., Michael Lettau, and William L. Wascher, "Downward Nominal Wage Rigidity in the United States During and After the Great Recession," Finance and Economics Discussion Series 2016-1, Board of Governors of the Federal Reserve System (U.S.) January 2016.

Farmer, Roger E. A., "Confidence, Crashes and Animal Spirits," *Economic Journal*, 03 2012, *122* (559), 155–172.

Galí, Jordi, "Insider-outsider labor markets, hysteresis and monetary policy," Technical Report January 2016.

Garga, Vaishali and Sanjay R. Singh, "Output Hysteresis and Optimal Monetary Policy," Technical Report November 2016.

Ghayad, Rand, "The Jobless Trap," Technical Report 2014.

Goodfriend, Marvin and Robert King, "The New Neoclassical Synthesis and the Role of Monetary Policy," in "NBER Macroeconomics Annual 1997, Volume 12" NBER Chapters, National Bureau of Economic Research, Inc, February 1997, pp. 231–296.

Gordon, Robert J., "Secular Stagnation: A Supply-Side View," *American Economic Review*, May 2015, *105* (5), 54–59.

Guerrieri, Veronica and Guido Lorenzoni, "Credit Crises, Precautionary Savings, and the Liquidity Trap," *The Quarterly Journal of Economics*, 2017, *132* (3), 1427–1467.

Hall, Robert E, "Reconciling Cyclical Movements in the Marginal Value of Time and the Marginal Product of Labor," *Journal of Political Economy*, 2009, *117* (2), 281–323.

Haltmaier, Jane, "Do recessions affect potential output?," International Finance Discussion Papers 1066, Board of Governors of the Federal Reserve System (U.S.) 2012.

Hosios, Arthur J., "On The Efficiency of Matching and Related Models of Search and Unemployment," *Review of Economic Studies*, 1990, *57* (2), 279–298.

Jones, Callum and Thomas Philippon, "The Secular Stagnation of Investment?," Technical Report 2016.

Kang, Dae Woong, Nick Ligthart, and Ashoka Mody, "The European Central Bank: Building a shelter in a storm," CFS Working Paper Series 527, Center for Financial Studies (CFS) 2015.

Kapadia, Sujit, "Optimal Monetary Policy under Hysteresis," Economics Series Working Papers 250, University of Oxford, Department of Economics October 2005.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante, "Monetary Policy According to HANK," *American Economic Review*, March 2018, *108* (3), 697–743.

Kroft, Kory, Fabian Lange, and Matthew J. Notowidigdo, "Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment," *The Quarterly Journal of Economics*, 2013, *128* (3), 1123–1167.

Laureys, Lien, "Optimal Monetary Policy in the Presence of Human Capital Depreciation during Unemployment," Discussion Papers 1415, Centre for Macroeconomics (CFM) June 2014.

Ljungqvist, Lars and Thomas J. Sargent, "The European Unemployment Dilemma," *Journal of Political Economy*, June 1998, *106* (3), 514–550.

Martin, Robert F., Tenyanna Munyan, and Beth Anne Wilson, "Potential Output and Recessions : Are We Fooling Ourselves?," IFDP Notes 2014-11-12, Board of Governors of the Federal Reserve System (U.S.) November 2014.

McKay, Alisdair, Emi Nakamura, and Jn Steinsson, "The Power of Forward Guidance Revisited," *American Economic Review*, October 2016, *106* (10), 3133–3158.

Menzio, Guido and Shouyong Shi, "Efficient Search on the Job and the Business Cycle," *Journal of Political Economy*, 2011, *119* (3), 468 – 510.

Paradise, Andrew, "Learning Remains Steady During the Downturn," https://www.td.org/Publications/Magazines/TD/TD-Archive/2009/11/Learning-Remains-Steady-During-the-Downturn November 2009.

Pissarides, Christopher A., "Loss of Skill During Unemployment and the Persistence of Employment Shocks," *The Quarterly Journal of Economics*, 1992, *107* (4), 1371–1391.

_ , "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?," *Econometrica*, 09 2009, *77* (5), 1339–1369.

Plosser, Charles I, "Understanding Real Business Cycles," *Journal of Economic Perspectives*, Summer 1989, *3* (3), 51–77.

Schaal, Edouard and Mathieu Taschereau-Dumouchel, "Coordinating Business Cycles," 2015 Meeting Papers 178, Society for Economic Dynamics 2015.

_ **and** _ , "Aggregate Demand and the Dynamics of Unemployment," Working Paper 2016.

Schmitt-Grohe, Stephanie and Martin Uribe, "Downward Nominal Wage Rigidity and the Case for Temporary Inflation in the Eurozone," *Journal of Economic Perspectives*, August 2013, *27* (3), 193–212.

_ **and** _ , "Liquidity Traps and Jobless Recoveries," *American Economic Journal: Macroeconomics*, January 2017, *9* (1), 165–204.

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

Song, Jae and Till von Wachter, "Long-Term Nonemployment and Job Displacement," Jackson Hole Economic Symposium Conference Proceedings, Kansas City Federal Reserve Board 2014.

Sterk, Vincent, "The Dark Corners of the Labor Market," Discussion Papers 1603, Centre for Macroeconomics (CFM) January 2016.

Summers, Lawrence, "Advanced economies are so sick we need a new way to think about them," http://larrysummers.com/2015/11/03/reconstructing-keynesian-macroeconomics/ 2015.

Wee, Shu Lin, "Delayed Learning and Human Capital Accumulation: The Cost of Entering the Job Market During a Recession," Technical Report 2016.

Werning, Ivan, "Managing a Liquidity Trap: Monetary and Fiscal Policy," NBER Working Papers 17344, National Bureau of Economic Research, Inc August 2011.

Woodford, M., *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, 2003.

Yellen, Janet, "Macroeconomic Research After the Crisis," https://www.federalreserve.gov/newsevents/speech/yellen20161014a.htm October 2016.

Appendix

A Wages and Nash bargaining

Recall that the value of an employed worker and of an unemployed worker, respectively, are defined by the recursions (2) and (3). Also, the value of a filled vacancy to a firm is given by equation (8). We can then define the surplus of a match between a worker and a firm as:

$$\mathcal{S}_t = J_t + \mathbb{W}_t - \mathbb{U}_t$$

Wages are determined by Nash bargaining. Denoting workers' bargaining power by η, wages solve

$$\max_{w_t} J_t^{1-\eta}(W_t - U_t)^\eta$$

implying

$$\eta J_t = (1-\eta)(W_t - U_t)$$

Notice that the match surplus can be rewritten as:

$$\begin{aligned}
\mathcal{S}_t &= J_t + \mathbb{W}_t - \mathbb{U}_t \\
&= A - b + \beta(1-\delta)J_{t+1} + \beta(1-\delta)(1-q_{t+1})(\mathbb{W}_{t+1} - \mathbb{U}_{t+1}) \\
&= A - b + \beta(1-\delta)(1-q_{t+1})S_{t+1} + \beta(1-\delta)q_{t+1}J_{t+1}
\end{aligned}$$

Using the fact that $W_t - U_t = \eta S_t$ in the equation above, we have:

$$\omega_t = \eta A + (1-\eta)b + \beta(1-\delta)\eta q_{t+1}J_{t+1}$$

B Existence of multiple steady states

Define:

$$\underline{\eta} = \max\left\{\frac{1-\beta(1-\delta)}{\beta(1-\delta)}, \frac{1-\delta}{2-\delta}, \frac{(1-\delta)(\kappa+\chi)-b}{(1-\delta)(\kappa+\chi)+a-b}\right\} \tag{24}$$

Steady states μ satisfies $\frac{a}{1-\beta(1-\delta)[1-\eta(1-\mu)]} = \kappa + \chi\mu$. Dividing through by J_{min}, this becomes

$$(1 - e\mu)^{-1} = k + x\mu \tag{25}$$

where $k = \kappa/J_{min}$, $x = \chi/J_{min}$ and $e = \frac{\beta\eta(1-\delta)}{1-\beta(1-\delta)(1-\eta)}$. For future reference, we define $\underline{\chi} = e\left[2 - k + 2\sqrt{1-k}\right]$. Assumptions 1 and 2 impose that $k < 1$ and $(1-e)^{-1} > k + x$. Since $e \in (0, 1)$, $(1 - e\mu)^{-1}$ is an increasing, strictly convex function. Starting from $x = 0$, as we increase x, either the intersection of these two functions first occurs at $\mu \in (0, 1)$, in which case a slightly higher x would give us multiplicity, or the first intersection has $\mu \geq 1$. Consider the knife edge case in which the first intersection of these two curves is at $\mu = 1$. Then the curves must be tangent and equal to each other at $\mu = 1$, i.e.

$$\frac{e}{(1-e)^2} = k \qquad \text{and} \qquad \frac{1}{1-e} = k + x$$

which implies $k = (1 - 2e)(1 - e)^{-2}$.

In order to have multiple intersections in $(0, 1)$, there must exist some $\mu \in (0, 1)$ such that $(1 - e\mu)^{-1} = k + x\mu$ and $e(1 - e\mu)^{-2} > x$ (at the larger of the two intersections, this convex function must intersect the linear function from below). If $k < (1 - 2e)(1 - e)^{-2}$, then this cannot be the case. A smaller k implies a larger x, increasing the slope of the linear function; $\mu < 1$ decreases the slope of the convex function. Thus, we must have $k > (1 - 2e)(1 - e)^{-2}$.The assumption that $\eta \geq \underline{\eta}$ is sufficient (but not necessary) to ensure this, since it implies that $e > 0.5$. If this is true, and if x is just large enough that there is a single slack steady state, then (25), which is quadratic in μ, has a unique solution, i.e. its discriminant equals zero: $x^2 - 2e(2 - k)x + e^2k^2 = 0$.

Considered as a function of x, this equation has two real solutions since its discriminant is positive: $4(e^2(2 - k)^2 - e^2k^2) = 16e^2(1 - k) > 0$. This will have two solutions x^*, the larger of which corresponds to $\mu \in (0, 1)$. To see this, consider the following graphical argument. Fix e and $k < 1$ and start with $x = \infty$, so that the $k + x\mu$ line is vertical at $\mu = 0$. Then the two curves intersect at exactly one point, $\mu = 0$. Decreasing x rotates the straight line clockwise, increasing the smallest value of μ at which the two curves intersect from 0 to some positive number. Eventually, for low enough x, the straight line is tangent to the convex curve at some $\mu > 0$. Next, start with $x = 0$, so the straight line $k + x\mu$ is horizontal at k and intersects the convex curve at some $\mu = e^{-1}(1 - k^{-1}) < 0$. Gradually increasing x rotates the straight line counter-clockwise, lowering the first value at which the curves intersect. For x large enough, the two curves are tangent at some $\mu < 0$. Clearly, the second case corresponds to a lower value of x. Thus, the larger value of x corresponds to the economically sensible case where $\mu \in (0, 1)$. Choosing this value, we have

$$x^* = e(2 - k) + \sqrt{e^2(2 - k)^2 - e^2k^2} = e[2 - k + 2\sqrt{1 - k}]$$

Thus there will be multiple steady states if $x > x^*$.

C Proof of Lemma 1

Suppose $\mu_0 = 0$. Then, note that $\mu_t = 0$ (which implies $n_t = 1$) is consistent with (12), since in the tight labor market regime $q_t = 1$, and $n_t = 1, \mu_{t+1} = 0$. Next we show that we cannot have $\theta_0 < \theta^{fe}$ given $\mu_0 = 0$. (Since $\theta^{fe} \geq 1$ by Assumption 1, this implies in particular that we cannot have $\theta_0 < 1$.) In any equilibrium, (11) must be satisfied:

$$J_{min} \leq a \sum_{t=0}^{\infty} \prod_{\tau=0}^{t} \beta(1-\delta)(1-\eta \min\{\theta_\tau, 1\}) \leq \kappa \max\{\theta_t, 1\}$$

where the first inequality holds because the LHS is decreasing in θ_τ. Since we know that $J_{min} > \kappa$ from Assumption 1, it is immediate that this inequality can only be satisfied if $\theta_t \geq \theta^{fe} \geq 1$. Finally, we show that we cannot have $\theta_0 > \theta^{fe}$. We have shown that

$$a \sum_{t=0}^{\infty} \prod_{\tau=0}^{t} \beta(1-\delta)(1-\eta \min\{\theta_\tau, 1\}) = \kappa \theta_0$$

in any equilibrium, and that this expression is satisfied by $\theta_t = \theta^{fe}$, $\forall t \geq 0$. If $\theta_0 > \theta^{fe}$, it follows that $\theta_t < \theta^{fe}$ for some $t > 0$. Let T be the first date at which this is true. Then up to that date, since the labor market has been tight, $\mu_T = 0$. This is a contradiction, since we have already shown that if $\mu_T = 0$, $\theta_T \geq \theta^{fe}$. It follows that the unique equilibrium has $\theta_t = \theta^{fe}$ for all $t \geq 0$. The proof for any $\mu_0 \in (\mu_0, \underline{\mu})$ is similar and follows from the fact that $q_0 = 1$ which implies that all workers are employed in period 0. Before characterizing the case when $\mu = \underline{\mu}$, the following result is useful:

Lemma 2. *If $J_t = J_{min}$, then $q_{t+1} = 1$, i.e. $\theta_{t+1} \geq 1$ and $J_{t+1} = J_{min}$.*

Proof. We have $J_t = a + \beta(1-\delta)(1-\eta q_{t+1})J_{t+1}$. The only way to attain $J_t = J_{min}$ is $q_{t+1} = 1$ and $J_{t+1} = J_{min}$, since $q_{t+1} \leq 1$, $J_{t+1} \geq J_{min}$, and the expression is decreasing in q_{t+1} and increasing in J_{t+1}. □

For $\mu_0 = \underline{\mu}$, there exist a continuum of equilibria indexed by $\theta_0 \in [1-\underline{\mu}, 1]$. In all these equilibria, the value of an employed worker for a firm is given by J_{min}. To see this, notice that $J_0 \leq \kappa + \chi\underline{\mu}$ as long as labor markets are slack, $\theta_0 \leq 1$. In this case, by definition, $J_0 \leq J_{min}$ and by definition this relationship has to hold with equality. If labor markets are tight, $\theta_0 > 1$, then $J_0 = \kappa\theta_0 + \chi\underline{\mu} > J_{\min}$ since $\theta_0 > 1$. This is a contradiction since if $\theta_0 > 1$, $\mu_1 = 0$ from Lemma 1 and $J_0 = J_{min}$ from Lemma 1. Furthermore, from Lemma 2, it follows that $J_1 = J_{min}$ and $\theta_1 \geq 1$.

The contradiction above shows that $\theta_0 \leq 1$. We now need to show that $\theta_0 > 1 - \underline{\mu}$. Suppose that $\theta_0 < 1 - \underline{\mu}$. Then μ_1 is given by:

$$\mu_1 = \frac{1-\theta_0}{1+(1-\delta)[1-\theta_0-\underline{\mu}]} > \underline{\mu}$$

This is a contradiction since

$$J_1 = J_{min} = \kappa + \chi \underline{\mu} < \kappa + \chi \mu_1$$

which requires that $\theta_1 = 0$. Thus, we have shown that $\theta_0 \in [1 - \underline{\mu}, 1]$. From (4) and the earlier part of this proof, it follows that $\mu_1 = \frac{1-\theta_0}{1+(1-\delta)[1-\theta_0-\mu_0]} \leq \mu_0$ and $\theta_1 = (J_{min} - \chi\mu_1)/\kappa \geq 1$. As mentioned in footnote 22, we select the equilibrium in which $\theta_0 = 1$ implying that $\mu_1 = 0$, $\theta_1 = \theta^{fe}$.

D Proof of Proposition 1

Definition 1. *Define the functions $\Theta^1 : I^1 \to [0,1]$, $F^1 : I^1 \to \mathbb{R}_+$, $M^1 : I^1 \to \{\underline{\mu}\}$ as:*

$$\begin{aligned}
\Theta^1(\mu_{T-1}) &:= 1 - \frac{\underline{\mu}}{1-(1-\delta)\underline{\mu}}(1-(1-\delta)\mu_{T-1}) \\
F^1(\mu_{T-1}) &:= \frac{1}{\chi}\left[a - \kappa + \beta(1-\delta)(1-\eta\Theta^1(\mu_{T-1}))(\kappa + \chi\mu_{T-1})\right] \\
M^1(\mu_{T-1}) &:= \underline{\mu}
\end{aligned}$$

where $I^1 = [\underline{\mu}, \mu^1]$ and $\mu^1 := F^1(\underline{\mu})$.

Intuitively, at any date t, for any $\mu_t \in I^1$, $\Theta^1(\mu_t)$ describes the job-finding rate that ensures that the economy reaches $\underline{\mu}$ at date $t+1$. $F^1(\mu_t)$ describes the unique value that μ_{t-1} can have in period $t-1$ such that $\mu_t \in I^1$ and also $\mu_{t+1} = \underline{\mu}$. In other words, given market tightness at date t, $\Theta^1(\mu_t)$, one can compute the value of a filled vacancy at date $t-1$ and zero and by no-arbitrage, this pins down the value of μ_{t-1} for which firms would have been willing to post the requisite number of vacancies. $M^1(\mu)$ is just a constant function which by definition describes where any $\mu \in I^1$ ends up.

Corollary 2. *It must be true that $\mu^1 < \widetilde{\mu}$.*

By the definition of μ^1, it must be true that

$$\begin{aligned}
\mu^1 &= \frac{1}{\chi}\left[a - \kappa + \beta(1-\delta)\left[1 - \eta(1-\underline{\mu})\right](\kappa + \chi\underline{\mu})\right] \\
&< \frac{1}{\chi}\left[a - \kappa + \beta(1-\delta)\left[1 - \eta(1-\widetilde{\mu})\right](\kappa + \chi\widetilde{\mu})\right] \\
&= \widetilde{\mu}
\end{aligned}$$

Lemma 3. *For β sufficiently close to 1, F^1 is increasing in μ for $\mu \in [\underline{\mu}, \widetilde{\mu})$*

Proof. Since $F^1(\mu)$ is composed of constants and a concave part, it suffices to consider the concave polynomial $\xi(\mu) = \left[1 - \eta\Theta^1(\mu)\right](\kappa + \chi\mu)$. This function is increasing in μ for

$$\mu < \frac{1}{2}\left[\frac{\left(1-(1-\delta)\underline{\mu}\right)(1-\eta)}{\eta(1-\delta)\underline{\mu}} + \frac{1}{(1-\delta)} - \frac{\kappa}{\chi}\right] \tag{26}$$

It is thus sufficient to show that $\widetilde{\mu}$ satisfies this inequality. Before proceeding further, it is convenient to work with a quasi-value function of the firm defined in terms of μ as opposed to J_t. Define the quasi-value function $\mathcal{Q}(\mu)$ as:

$$\mathcal{Q}(\mu) = \frac{a}{1-\beta(1-\delta)\left[1-\eta(1-\mu)\right]}$$

By construction, $\mathcal{Q}(\mu)$ is the value of the firm as long as the job-finding rate is $1-\mu$ forever. Note that $\mathcal{Q}'(\mu) > 0$ and $\mathcal{Q}''(\mu) > 0$.

Under this quasi-value function and given free entry, $\widetilde{\mu}$ satisfies

$$\mathcal{Q}(\widetilde{\mu}) = \frac{a}{1-\beta(1-\delta)\left(1-\eta+\eta\widetilde{\mu}\right)} = \kappa + \chi\widetilde{\mu}$$

Since the left hand side is convex and the right hand side linear, since $\widetilde{\mu}$ is the smaller of two solutions to this equation, then

$$\mathcal{Q}'(\widetilde{\mu}) = \frac{a\beta(1-\delta)\eta}{\left[1-\beta(1-\delta)\left(1-\eta+\eta\widetilde{\mu}\right)\right]^2} < \chi$$

In other words, the LHS cuts the RHS from above. Next, dividing the first equality by the second inequality, we have

$$\widetilde{\mu} < \frac{1}{2}\left[\frac{1-\beta(1-\delta)}{\beta(1-\delta)\eta} + 1 - \frac{\kappa}{\chi}\right] \tag{27}$$

Define:

$$\Xi = \frac{1}{2}\left\{\frac{(1-(1-\delta)\underline{\mu})(1-\eta)}{\eta(1-\delta)\underline{\mu}} - \frac{1-\beta(1-\delta)}{\eta\beta(1-\delta)} + \frac{1}{(1-\delta)} - 1\right\}$$

Assuming that $\beta > \frac{\widetilde{\mu}}{\eta\widetilde{\mu}+1-\eta}$, it can be shown that $\Xi > 0$.[55] Thus, as required:

$$\widetilde{\mu} < \frac{1}{2}\left[\frac{1-\beta(1-\delta)}{\beta(1-\delta)\eta} + 1 - \frac{\kappa}{\chi}\right] + \Xi = \frac{1}{2}\left[\frac{\left(1-(1-\delta)\underline{\mu}\right)(1-\eta)}{\eta(1-\delta)\underline{\mu}} + \frac{1}{(1-\delta)} - \frac{\kappa}{\chi}\right]$$

□

It was already clear that given a $\mu_{t+1} \in I^1$, there exists a unique μ_t which could have led there. In addition, this Lemma shows that given any μ_t, there exists at most one $\mu_{t+1} \in I^1$ is consistent with equilibrium.

Corollary 3. *Let $I^2 = F^1(I^1)$ and let $M^2(\mu)$ be the inverse of this function. Then $M^2(\mu^1) = M^1(\mu^1) = \underline{\mu}$.*

[55] Note that this assumption is a condition on an endogenous variable, $\widetilde{\mu}$ and can be rewritten as $\widetilde{\mu} < \frac{1-\eta}{\beta^{-1}-\eta}$. Nonetheless, it is a weak condition: for any $\widetilde{\mu} < 1$, it is satisfied for β sufficiently close to 1.

Since F^1 is increasing and continuous, its inverse M^2 exists and is increasing and continuous. Consequently, $F^1(I^1)$ maps into an interval $(\mu^1, \mu^2]$. Further since $\mu^1 = F^1(\underline{\mu})$, then $M^2(\mu^1) = \underline{\mu}$.

Lemma 4. $\mu^2 = F^1(\mu^1) < \widetilde{\mu}$

Proof. Since $\Theta^1(\underline{\mu}) = 1 - \underline{\mu}$ and Θ^1 is increasing, we have $\Theta^1(\mu^1) > 1 - \underline{\mu} > 1 - \widetilde{\mu}$. It follows that:

$$\frac{1}{\chi}\left[a - \kappa + \beta(1-\delta)(1 - \eta\Theta^1(\mu^1))(\kappa + \chi\widetilde{\mu})\right] \quad < \quad \frac{1}{\chi}\left[a - \kappa + \beta(1-\delta)(1 - \eta(1-\widetilde{\mu}))(\kappa + \chi\widetilde{\mu})\right]$$

Then, from Corollary 2, since $\mu^1 < \widetilde{\mu}$:

$$\begin{aligned} F^1(\mu^1) &= \frac{1}{\chi}\left[a - \kappa + \beta(1-\delta)(1 - \eta\Theta^1(\mu^1))(\kappa + \chi\mu^1)\right] \\ &< \frac{1}{\chi}\left[a - \kappa + \beta(1-\delta)(1 - \eta(1-\widetilde{\mu}))(\kappa + \chi\widetilde{\mu})\right] \\ &= \widetilde{\mu} \end{aligned}$$

□

Lemma 5. *Define $\Theta^2(\mu) : I^2 \to [0,1]$ as:*

$$\Theta^2(\mu) := 1 - M^2(\mu)\frac{1-(1-\delta)\mu}{1-(1-\delta)M^2(\mu)}$$

Then,

$$\frac{\partial\Theta^2(\mu)}{\partial\mu} \leq \frac{(1-\delta)M^2(\mu)}{1-(1-\delta)M^2(\mu)}$$

Proof.

$$\begin{aligned} \frac{\partial\Theta^2(\mu)}{\partial\mu} &= M^2(\mu)\frac{(1-\delta)}{1-(1-\delta)M^2(\mu)} - \frac{\partial M^2(\mu)}{\partial\mu}\left[1 + \frac{(1-\delta)\left(1-(1-\delta)\mu\right)M^2(\mu)}{\left[1-(1-\delta)M^2(\mu)\right]^2}\right] \\ &\leq M^2(\mu)\frac{(1-\delta)}{1-(1-\delta)M^2(\mu)} \end{aligned}$$

where the inequality comes because $M^2(\mu)$ is increasing and the expression in square brackets is positive. □

We are now ready to characterize equilibrium in the entire convalescent region.

Lemma 6 (Induction Step)**.** *Suppose the functions $\Theta^n(\mu)$, $M^n(\mu)$ are defined on some interval $I^n = [\mu^{n-1}, \mu^n]$ and $M^{n-1}(\mu_{T-n+1})$ is defined on an interval $I^{n-1} = [\mu^{n-2}, \mu^{n-1}]$, with $\underline{\mu} < \mu^{n-2} <$*

$\mu^n < \widetilde{\mu}$, *and that these functions satisfy*

$$\begin{aligned}
\Theta^n(\mu) &= 1 - M^n(\mu)\frac{1-(1-\delta)\mu}{1-(1-\delta)M(\mu)} \\
\frac{\partial \Theta^n(\mu)}{\partial \mu} &< \frac{(1-\delta)M^n(\mu)}{1-(1-\delta)M^n(\mu)} \\
M^n(I^n) &= I^{n-1} \\
M^n(\mu^{n-1}) &= M^{n-1}(\mu^{n-1}) = \mu^{n-2}
\end{aligned}$$

Then, for β sufficiently close to 1, we have the following results:

1. *The function*
$$F^n(\mu) := \frac{1}{\chi}\left[a - \kappa + \beta(1-\delta)(1-\eta\Theta^n(\mu))(\kappa + \chi\mu)\right]$$
is monotonically increasing in μ for $\mu \le \widetilde{\mu}$.

2. *Let $I^{n+1} = F^n(I^n)$ and let $M^{n+1}(\mu)$ be the inverse of this function. Then $M^{n+1}(\mu^n) = M^n(\mu^n) = \mu^{n-1}$.*

3. *$I^{n+1} = [\mu^n, \mu^{n+1}]$ with $\mu^{n+1} < \widetilde{\mu}$.*

4. *Define $\Theta^{n+1}(\mu)$ on I^{n+1} by*
$$\Theta^{n+1}(\mu) = 1 - M^{n+1}(\mu)\frac{1-(1-\delta)\mu}{1-(1-\delta)M^{n+1}(\mu)}$$
The derivative of this function satisfies
$$\frac{\partial \Theta^{n+1}(\mu)}{\partial \mu} < \frac{(1-\delta)M^{n+1}(\mu)}{1-(1-\delta)M^{n+1}(\mu)}$$

Proof. (1.) The derivative of $F^n(\mu)$ is

$$\begin{aligned}
\frac{\partial F^n(\mu)}{\partial \mu} &= \frac{\beta(1-\delta)}{\chi}\left[-\eta\frac{\partial \Theta^n(\mu)}{\partial \mu}(\kappa + \chi\mu) + \chi(1-\eta\Theta^n(\mu))\right] \\
&> \frac{\beta(1-\delta)}{\chi}\left[-\eta\frac{(1-\delta)M^n(\mu)}{1-(1-\delta)M^n(\mu)}(\kappa + \chi\mu) + \chi(1-\eta\Theta^n(\mu))\right]
\end{aligned}$$

Substituting in the definition of Θ^n and rearranging, we see that this expression will be positive provided that

$$\mu < \frac{1}{2}\left[\frac{1-\eta}{\eta}\frac{(1-(1-\delta)M^n(\mu))}{(1-\delta)M^n(\mu)} + \frac{1}{(1-\delta)} - \frac{\kappa}{\chi}\right]$$

By the same logic as in Lemma 3, for β sufficiently close to 1, this is satisfied for any $\mu \le \widetilde{\mu}$, since we have $M^n(\mu) \le \widetilde{\mu}$. So $F^n(\mu)$ is increasing, and hence invertible, for $\mu < \widetilde{\mu}$. Let $M^{n+1}(\mu)$ be the inverse of this function.

(2.) We have

$$\begin{aligned}
M^n(\mu^{n-1}) &= M^{n-1}(\mu^{n-1}) \\
\Theta^n(\mu^{n-1}) &= \Theta^{n-1}(\mu^{n-1}) \\
F^n(\mu^{n-1}) &= F^{n-1}(\mu^{n-1}) = \mu^n \text{ by definition of } \mu^n \\
M^{n+1}(\mu^n) &= M^n(\mu^n)
\end{aligned}$$

(3.) Since F^n is a continuous, increasing function, the image of the interval $[\mu^{n-1}, \mu^n]$ under F^n must be an interval $[\mu^n, \mu^{n+1}]$. (We have already shown that $F^n(\mu^{n-1}) = \mu^n$.) We need to show that $\mu^{n+1} = F^n(\mu^n) < \widetilde{\mu}$. We know that $\widetilde{\mu} \geq M^n(\mu^n)$. Then, it must be true that

$$\begin{aligned}
1 - \widetilde{\mu} &< 1 - M^n(\mu^n) \\
&= 1 - M^n(\mu^n)\frac{1-(1-\delta)\mu^n}{1-(1-\delta)\mu^n} \\
&< 1 - M^n(\mu^n)\frac{1-(1-\delta)\mu^n}{1-(1-\delta)M^n(\mu^n)} \\
&= \Theta^n(\mu^n)
\end{aligned}$$

Then, by the same logic as in Lemma 4 we have $F^n(\mu^n) < \widetilde{\mu}$. So we have shown that $I^{n+1} \subset [\underline{\mu}, \widetilde{\mu}]$.

(4.) The bound on the derivative is established in the same way as Lemma 5.

□

Lemma 7. *$\lim_{n\to\infty} \mu^n \to \widetilde{\mu}$.*

Proof. We have shown that $\{\mu^n\}$ is an increasing sequence bounded above by $\widetilde{\mu}$; thus by the Monotone Convergence Theorem, its limit μ^∞ exists, and $\mu^\infty \leq \widetilde{\mu}$. Suppose by contradiction that $\mu^\infty < \widetilde{\mu}$. Then μ^∞ must be a steady state. But by definition, $\widetilde{\mu}$ is the smallest slack steady state. So we must have $\mu^\infty = \widetilde{\mu}$. □

Finally, we prove that recoveries can be arbitrarily slow, i.e. for any $T \in \mathbb{N}$, there exists $\varepsilon > 0$ such that if $\mu_0 \in (\widetilde{\mu} - \varepsilon, \widetilde{\mu})$, $\mu_t > 0$ for all $t < T$. Fix $\delta > 0, T \in \mathbb{N}$ and let n be the smallest integer such that $\mu^n \geq \widetilde{\mu} - \delta$ (this exists, since $\mu^n \to \widetilde{\mu}$ and $\delta > 0$. Set $\varepsilon = \widetilde{\mu} - \mu^{n+T}$. Take any $\mu_0 \in (\widetilde{\mu} - \varepsilon, \widetilde{\mu}) = (\mu^{n+T}, \widetilde{\mu})$. Then $\mu_0 \in (\mu^{m-1}, \mu^m]$ for some $m > n + T + 1$. We know from 1 that $\mu_T \in (\mu^{m-T-1}, \mu^{m-T}]$. In particular,

$$\mu_T > \mu^{m-T-1} > \mu^n \geq \widetilde{\mu} - \delta > \mu_0 - \delta$$

Finally, since $\{\mu_t\}$ is monotonically decreasing, we have $\mu_t > \mu_0 - \delta$ for all $t < T$, as claimed. Next, note that the first part of the lemma is a special case of the second part with $\delta = \widetilde{\mu}$.

E Properties of $J_0(\beta_0)$

Suppose that the economy remains at full employment steady state even after the shock $\beta_0 > 1$. There are two cases to consider. First, suppose that the ZLB does not bind at date 0. Then monetary policy is unconstrained in all periods, and nominal wages and prices remain constant. From (19), we have $1 + i_t = \frac{P_1}{P_0 \beta_0} = \frac{1}{\beta_0}$. When $\beta_0 > 1$, this would imply a negative nominal interest rate, violating the ZLB. Thus, when $\beta_0 > 1$, monetary policy is constrained at date 0 and we have $P_0 = \frac{P_1}{\beta_0}$. Since the economy returns to full employment after date 0, real wages will equal ω^*_{fe} at all dates $t \geq 1$. Iterating forward (22), it follows that prices and nominal wages remain constant thereafter and the ZLB does not bind after date 0. In particular, since $W_1 = W_0$, we have:

$$\omega_0 = \frac{W_0}{W_1} \frac{P_1}{P_0} \omega_1 = \beta_0 \omega^*_{fe}$$

Using this in the expression for J_0 we have:

$$J_0 = A - \beta_0 \omega^*_{fe} + \beta_0 (1 - \delta) J_{min}$$

The full employment steady state Nash wage equals

$$\omega^*_{fe} = \frac{\eta}{1 - \beta(1-\delta)(1-\eta)} A + \frac{[1 - \beta(1-\delta)](1-\eta)}{1 - \beta(1-\delta)(1-\eta)} b$$

So

$$\frac{\partial J}{\partial \beta_0} = -\omega^*_{fe} + (1-\delta) J_{min} = -\frac{\eta}{1 - \beta(1-\delta)(1-\eta)} A - \frac{[1 - \beta(1-\delta)](1-\eta)}{1 - \beta(1-\delta)(1-\eta)} b + (1-\delta) \frac{(1-\eta)(A-b)}{1 - \beta(1-\delta)(1-\eta)}$$

which is negative provided that $A\left[1 - \delta - \frac{\eta}{1-\eta}\right] - [2 - \delta - \beta(1-\delta)]b < 0$. By Assumption 2, both terms are negative, so this condition is satisfied.

F Proof of Proposition 2

First we show that a one-period hiring freeze takes the economy either to the convalescent or to the stagnant region.

Lemma 8. *Starting from full employment, a one period hiring freeze takes the economy out of the healthy region:* $\mu_R = \dfrac{1}{2-\delta} > \underline{\mu}$.

Proof. We prove the Lemma by proving the contrapositive. The first thing to note is that $\mu_R := \frac{1}{2-\delta} > 0.5$ since $0 < 1 - \delta < 1$. Recall that $\underline{\mu} = \frac{J_{min} - \kappa}{\chi}$. Suppose $\underline{\mu} \geq \mu_R$. This implies that $\underline{\mu}$ must also be greater than 0.5. In this case, no interior steady state can exist. Recall that any interior steady state solves:

$$
\begin{aligned}
\kappa + \chi\mu &= \mathcal{Q}(\mu) \\
&= \frac{a}{1-\beta(1-\delta)[1-\eta(1-\mu)]} \\
&= \frac{a}{1-\beta(1-\delta)(1-\eta)} \frac{1-\beta(1-\delta)(1-\eta)}{1-\beta(1-\delta)[1-\eta(1-\mu)]} \\
&= J_{min} \frac{1}{1-e\mu}
\end{aligned}
$$

where, as before $e = \frac{\beta(1-\delta)\eta}{1-\beta(1-\delta)(1-\eta)}$.

Thus interior steady states solve:

$$
\Omega(\mu) := \frac{J_{min}}{1-e\mu} - \kappa - \chi\mu = 0
$$

We show that this is not possible if $\underline{\mu} > \mu_R$. In particular, we have $\Omega(\mu) > 0$ for all $\mu \in [0,1]$. First, we show that $e > 1/2$ and $\chi < 2(J_{min} - \kappa)$. Notice that e can also be rewritten as:

$$
e = \frac{1}{1 + \frac{1-\beta(1-\delta)}{\beta(1-\delta)\eta}} > \frac{1}{1+\frac{\eta}{\eta}} = \frac{1}{2}
$$

where the inequality follows since $\eta > \frac{1-\beta(1-\delta))}{\beta(1-\delta)}$ by Assumption 2. Thus, $e > \frac{1}{2}$. To see that $\chi < 2(J_{min} - \kappa)$, note that from the definition of $\underline{\mu}$:

$$
\chi = \frac{J_{min} - \kappa}{\underline{\mu}} < 2(J_{min} - \kappa)
$$

since $\underline{\mu} > 0.5$ by assumption.

Fix $\kappa \in [0, J_{min})$, $\mu \in [0,1]$. Even though we have shown above that $e > 1/2$ and $\chi < 2(J_{min} - \kappa)$, for a moment, set $e = 1/2$, $\chi = 2(J_{min} - \kappa)$. We claim that

$$
\mathcal{Q}(\mu) = \frac{J_{min}}{1-e\mu} \geq \kappa + \chi\mu = \kappa + 2(J_{min} - \kappa)\mu
$$

with strict inequality unless $\kappa = 0$ and $\mu = 1$, in which case the expression holds with equality. When $\kappa = 0$, the RHS becomes $2J_{min}\mu$, and the LHS and RHS are only equal for $\mu = 1$. For any $\mu < 1$, the LHS is larger. When $\kappa > 0$, the RHS is strictly lower for any $\mu > 1/2$. Thus for any $\kappa \in [0, J_{min}]$, the inequality holds for all $\mu \in [0, 1)$. Finally, for any $\mu \leq 1/2$, the inequality clearly holds since the LHS is greater than J_{min}, and the RHS smaller than J_{min}.

Next, suppose $e > 1/2$ and $\chi < 2(J_{min} - \kappa)$. If $\mu = 0$, this does not change the inequality, which still holds strictly (since $\mu \neq 1$). If $\mu > 0$, this strictly increases the LHS and strictly decreases the RHS. Thus the expression is still satisfied with strict inequality. Thus we have $\Omega(\mu) > 0$ for all $\mu \in [0, 1]$, and there is no interior steady state. Since we have shown that $\underline{\mu} \geq \mu_R$ implies there exists no interior steady state, it follows that if there exist multiple interior steady states, we must have $\underline{\mu} < \mu_R$. □

Next we need to prove two lemmas. The first states that wages are lower in the convalescent region than at full employment. We need this result to show that prices will be higher in the convalescent region.

Lemma 9. *$\omega^*(\mu_t) < \omega^*_{fe}$ if $\mu_t \in (\underline{\mu}, \tilde{\mu})$.*

Proof. We know that $M(\mu_t) < \mu_t$ if $\mu_t \in (\underline{\mu}, \tilde{\mu})$.

$$\begin{aligned}
\omega^*(\mu_t) &= A - (\kappa + \chi\mu_t) + \beta(1-\delta)[\kappa + \chi M(\mu_t)] \\
&= A - \beta(1-\delta)\chi(\mu_t - M(\mu_t)) - (1-\beta(1-\delta))(\kappa + \chi\mu_t) \\
&< A - (1-\beta(1-\delta))(\kappa + \chi\mu_t) \\
&< A - (1-\beta(1-\delta))(\kappa + \chi\underline{\mu}) = \omega^*_{fe}
\end{aligned}$$

□

Lemma 10. *Under Assumption 2, $\frac{W_t}{P_t} > (1-\delta)[\kappa + \chi\mu_t]$.*

Proof. We know that $\frac{W_t}{P_t} \geq \omega^*(\mu_t)$ by definition, so it suffices to show that $\omega^*(\mu_t) > (1-\delta)[\kappa + \chi\mu_t]$. In the flexible wage benchmark we have

$$\omega_t = \eta A + (1-\eta)b + \beta(1-\delta)q_{t+1}J_{t+1} \geq \eta A + (1-\eta)b > (1-\delta)(\kappa + \chi\mu_t)$$

for any $\mu_t \in [0,1]$, given assumption 2. □

Finally, we need to characterize dynamics of the economy starting at date 1, once the shock has abated. Under neutral monetary policy, if the ZLB never binds, allocations are (by definition) equal to those in the flexible wage benchmark. The following is immediate.

Lemma 11. *If $\mu_1 \geq \widetilde{\mu}$, the economy never returns to the full employment steady state.*

Proof. If the ZLB never binds, allocations are equivalent to those in the flexible wage benchmark, and we know that the economy never returns to steady state. It only remains to show that the ZLB can never help the economy converge to the full employment steady state. Suppose by contradiction that the economy converges to the full employment steady state. Let μ_t^R, μ_t^N denote allocations in the flexible wage benchmark and in the nominal economy, respectively, given the initial condition $\mu_1 \geq \widetilde{\mu}$. Let $T \geq 1$ be the first date at which $\mu_t^N < \mu_t^R$ (there must be some such date, since in the long run $\mu_t^N = 0, \mu_t^R > 0$, by assumption). Then we have

$$\begin{aligned}
J_{T-1}^N = \kappa + \chi\mu_{T-1}^N = \kappa + \chi\mu_{T-1}^R = J_{T-1}^R \\
J_T^N = \kappa + \chi\mu_T^N < \kappa + \chi\mu_T^R = J_T^R
\end{aligned}$$

This implies that real wages are higher at date $T-1$ in the flexible wage benchmark than in the nominal economy:

$$\begin{aligned} J^N_{T-1} &= J^R_{T-1} \\ A - \omega^N_T + \beta(1-\delta)J^N_T &= A - \omega^R_T + \beta(1-\delta)J^R_T \\ \omega^R_t - \omega^N_t &= \beta(1-\delta)(J^R_T - J^N_T) > 0 \end{aligned}$$

This is a contradiction - given the downward nominal wage rigidities, wages are always weakly higher than in the flexible wage benchmark. Thus the economy cannot converge to the full employment steady state. □

We are now ready to prove Proposition 2. Part 1. follows for the same reasons as in the previous lemmas. Define the function

$$B(\mu) = \frac{A-\kappa}{\omega(\mu) - (1-\delta)(\kappa + \chi\mu)}$$

on $(\underline{\mu}, \mu_R]$, where $\omega(\mu_1)$ denotes the prevailing real wage at date 1 as a function of μ_1. It is straightforward to show that $\omega(\mu_1)$ is continuous, and thus B is continuous. We have $B(\underline{\mu}) = \underline{\beta}$. Define $\overline{\beta} := B(\mu_R)$. If $\beta_0 > \overline{\beta}$, then if $\mu_1 = \mu_R$, we have

$$J_0 = A - \beta_0\omega(\mu_R) + \beta_0(1-\delta)(\kappa + \chi\mu_R) < \kappa$$

thus $\theta_0 = 0$, which is consistent with $\mu_1 = \mu_R$. If instead $\beta_0 \in (\underline{\beta}, \overline{\beta})$, then there exists $\mu \in (\underline{\mu}, \mu_R)$ such that $B(\mu) = \beta_0$, and a corresponding $\theta_0 = 1 - \frac{\mu_1}{1-(1-\delta)\mu_1}$. Then we have

$$J_0 = \kappa = A - \beta_0\omega(\mu_1) + \beta_0(1-\delta)(\kappa + \chi\mu_1)$$

and firms are indifferent between posting any number of vacancies; thus $\theta_0 \in [0,1]$ can indeed be an equilibrium. Finally, the fact that the economy does not return to full employment if it is thrown into the stagnant region follows from Lemma 11.