Volume Title: NBER Macroeconomics Annual 2013, Volume 28

Volume Author/Editor: Jonathan A. Parker and Michael Woodford, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 978-0-226-16540-0 (cloth); 978-0-226-16540-0 (paper); 978-0-226-16554-7 (eISBN)

ISSN: 0889-3365

Volume URL: http://www.nber.org/books/park13-1

Conference Date: April 12–13, 2013

Publication Date: April 2014

Chapter Title: Reference Dependence and Labor Market Fluctuations

Chapter Author(s): Kfir Eliaz, Ran Spiegler

Chapter URL: http://www.nber.org/chapters/c12938

Chapter pages in book: (p. 159 – 200)

3

Reference Dependence and Labor Market Fluctuations

Kfir Eliaz, *Tel Aviv University and University of Michigan*
Ran Spiegler, *Tel Aviv University and University College London*

I. Introduction

Economists have long pondered over the observation that wages display downward rigidity and do not fall in recessions as much as one might expect on the basis of supply and demand analysis. An intuitive idea with a long pedigree—going back to Keynes (1936), Solow (1979), Akerlof (1982), Akerlof and Yellen (1990), Kahneman, Knetsch, and Thaler (1986), Fehr and Falk (1999), and many others—is that reciprocal-fairness considerations deter employers from cutting wages during recessions. Specifically, the theory is that the labor contract's inherent incompleteness forces employers to rely to some extent on workers' intrinsic motivation. When workers feel that they have been treated unfairly, their intrinsic motivation is dampened and their output declines. According to this "morale hazard" theory, wage cuts relative to a "reference point" have such an effect, which is why employers try to avoid them.[1]

The intuitive appeal of this argument is also reflected in survey data. Blinder and Choi (1990) and Bewley (1999) interviewed personnel managers and other labor market actors, and found overwhelming support for the morale theory. As Bewley (1999, p. 1) puts it:

> My findings support none of the existing economic theories of wage rigidity, except those emphasizing the impact of pay cuts on morale. Other theories fail in part because they are based on the unrealistic psychological assumptions that people's abilities do not depend on their state of mind and that they are rational in the simplistic sense that they maximize a utility that depends only on their own consumption and working conditions.

978-0-226-16540-0/2014/2013-0301$10.00

The morale theory of labor relations has substantial experimental support. In one prominent example, Fehr and Falk (1999) studied an experimental labor market in which wages are determined by a double auction and labor contracts are incomplete (in the sense that firms cannot monitor effort and output, and workers receive a flat wage). They found that the market-clearing wage significantly exceeds the competitive level, and hired workers reciprocate by exerting high effort. Numerous studies along similar lines are surveyed in Fehr, Goette, and Zehnder (2009).

In this paper we incorporate a morale hazard account of the labor relation into a search-and-matching (S&M) model of the labor market in which productivity fluctuates according to a log-linear AR(1) process, and explore its theoretical implications for equilibrium wage and unemployment fluctuations. Our main departure from the standard S&M model in the Mortensen-Pissarides tradition (see Pissarides [2000] and Shimer [2010] for textbook treatments) lies in the assumption that the labor contract is incomplete; workers receive a flat wage, and their normal productivity relies on "intrinsic motivation." When the worker's wage falls below his reference point, he becomes less motivated and his output falls below the normal level by a random fraction that captures the importance of morale in the production function.

In our model, the worker's reference point evolves over the course of his relationship with the firm. A formerly unemployed worker enters his first employment period only with the aspiration to be paid the lowest admissible wage (normalized to zero). At the end of the worker's first period of employment, having developed a relationship with his employer, he cultivates an aspiration to earn the expected equilibrium wage of existing workers (conditional on his current information). This aspiration will constitute the worker's reference point in the next period. Thus, the reference wage of a newly hired worker is zero, while the reference point of an existing worker at period t is equal to his expected wage, calculated according to his rational expectations at period $t - 1$. In appendix B, we present a slightly different formulation of the reference point, which *endogenizes* this distinction between newly hired and existing workers; our main results are robust to this variation.

Our lagged expectations approach to the formation of workers' reference point follows an influential model due to Köszegi and Rabin (2006). The justification for the expectation-based specification is that a given wage offer may be greeted as a pleasant surprise or as a demoralizing disappointment, depending on how it compares with the worker's former expectations. For instance, if the worker expected a big salary increase, failure to meet this expectation may hurt his morale, even if his current wage is higher than yesterday's wage. The justification for the lagged aspect is that it takes the reference point some time to adapt to changing circumstances, just as it takes people time to change a habit. This delayed adaptation will be the source of wage rigidity in our model.[2]

Before stating our main result, we wish to comment on our methodology. The standard Mortensen-Pissarides model mixes noncooperative game-theoretic modeling with the "cooperative" Nash bargaining solution. Instead, we formulate the model entirely as an extensive noncooperative game with moves of nature and study its subgame perfect equilibria (SPE), as in Rubinstein and Wolinsky (1985). We seek a complete analytical characterization of dynamic equilibria with transparent qualitative features, and this impels us to make a few simplifications. First, we focus exclusively on the labor market and leave out consumption and capital. Second, we eliminate two degrees of freedom in the standard S&M model: workers have no bargaining power (firms repeatedly make take-it-or-leave-it one-period wage offers), and their nonmarket payoff is proportional to productivity. Doing so not only simplifies the analysis, but also ensures that all wage-rigidity effects are due to the novel behavioral element. We do not add any new parameters, and our equilibrium characterization is presented for quite general reference-dependent output functions. Finally, for most of the paper, we impose a two-period exogenous separation process, which is innocuous in the reference-independent benchmark but facilitates analysis under reference dependence.[3]

Our main result is that as long as the magnitude of productivity shocks is not too large, our model generates a unique SPE, which displays the following features.

Wage rigidity and destruction of output. Equilibrium wage for existing workers displays downward rigidity w.r.t. current productivity shocks. Specifically, for intermediate noise realizations, the firm offers the reference wage, and therefore does not respond to local productivity fluctuations. At high noise realizations, the firm pays the current outside option (which we assumed to be proportional to current productivity). In certain special cases of the model, when the drop in worker productivity due to loss of morale is large, the wage is entirely rigid. At low noise realizations, the firm either lays off existing workers or pays them their outside option (in which case the workers' output declines), depending on the realized importance of "morale" in the production function. Thus, existing workers experience layoffs or demoralization in equilibrium following bad shocks.

History dependence. The fraction of existing workers' output that is destroyed as a result of wage rigidity is purely a function of current morale and productivity shocks. Since both shocks are drawn from stationary distributions, this fraction is history-independent. Thus, existing workers' observed output depends on both productivity *levels* and productivity *changes*. In particular, for a given productivity level, we may observe recession symptoms (notably layoffs) if this level is a consequence of a bad shock.

Entry-level wages. Newly matched workers are always hired in equilibrium and paid a wage below existing workers' wage. The entry-level wage is not rigid; it strictly increases with current productivity, albeit at a lower rate than in the benchmark model without reference dependence. Unlike existing workers, the equilibrium wage of new hires is purely a function of current productivity.

Increased volatility of market tightness. As in the standard S&M model, free entry implies that market tightness is determined by the firms' hiring incentive. We show that the elasticity of tightness with respect to productivity is higher than in the reference-independent benchmark. This effect is stronger for intermediate values of the AR(1) autocorrelation coefficient. The reason for this volatility effect is that existing workers' output destruction due to reference-dependence increases the weight of newly hired workers' output in the determination of the value of a vacancy. This raises the sensitivity of this value to initial conditions, since the stochastic process that governs productivity is mean-reverting.

Relation to the Shimer puzzle. In an influential paper, Shimer (2005) argued that the S&M model has shortcomings in accounting for real-life labor market fluctuations, in the sense that the wage volatility it predicts is too large and the unemployment volatility it predicts is too small. A fast-growing literature ensued. One research direction—suggested by Shimer (2004) and Hall (2005) and challenged by Pissarides (2009), Kudlyak (2009), and Haefke, Sonntag, and van Rens (2008)—has centered around the hypothetical role of wage stickiness in addressing Shimer's puzzle.

Our results can be viewed in light of this debate. As we show in section IV, our model synthesizes some *theoretical* arguments raised by the two sides in the debate, showing they need not be mutually contradictory after all. We wish to emphasize that although our model generates a volatility effect in the "right" direction, it *cannot* be viewed as an attempt to resolve Shimer's puzzle, which is quantitative in nature, whereas our paper is theoretical and qualitative in style. However, since our model provides a behavioral foundation for the association between wage rigidity and enhanced tightness volatility, it suggests a basis for future attempts to address the puzzle.

II. A Model

Consider the following complete information, infinite horizon game. There is a continuum of players: a measure one of workers and an unbounded measure of firms (the latter assumption captures free entry among firms). We break the description into the following components: search and matching, separation, wage and output determination, the agents' information, and their preferences.

Search and matching. Time is discrete. At each period t, firms and workers are matched according to the following process. An unemployed worker (including workers who lost their job at the beginning of period t, as described later) is automatically in the search pool (i.e., we abstract from questions of labor market participation). An unmatched firm (including firms that dismissed workers at the beginning of the period, as described later) decides whether to be in the search pool (i.e., post a vacancy).[4]

If there are U_t unemployed workers and V_t open vacancies at this stage, then a measure $m(U_t, V_t) \le \min\{U_t, V_t\}$ of unemployed workers are matched to vacancies at the beginning of period $t + 1$. The matching function m satisfies the standard assumptions: it is continuous, strictly increasing in each of its arguments, and exhibits constant returns to scale. The matching probabilities for workers and firms at period t are thus $\mu_t = m(U_t, V_t)/U_t$ and $q_t = m(U_t, V_t)/V_t$, respectively. Note that $\lim_{V\to\infty} m(U, V)/V = 0$. We assume that if all firms post vacancies, then $q = 0$.

Define *market tightness* at t as the ratio

$$\theta_t = \frac{V_t}{U_t} = \frac{\mu_t}{q_t}.$$

Since m exhibits constant returns to scale, θ is a strictly decreasing function of q, given by

$$m\left(1, \frac{1}{\theta}\right) = q. \tag{1}$$

From now on we will be primarily interested in market tightness as an indicator of the state of unemployment, and we will suppress U and V.

Separation and wage determination. Consider a worker who, at the end of period t, completes a tenure of $i \geq 1$ consecutive periods of employment at the same firm. We say that the worker is of type i at period t. With probability $s(i)$, the two parties will be separated by the beginning of period $t + 1$ for some unspecified exogenous reason. With probability $1 - s(i)$, the match will survive into the beginning of period $t + 1$, and the worker will turn into type $i + 1$.

When the two parties are matched at the beginning of period t, the firm first chooses whether to employ the worker. We use $r_{i,t} \in \{0, 1\}$ to denote the firm's endogenous separation decision when facing a worker of type i, where $r_{1,t} = 1$ means that the firm chooses to employ the worker at t, and $r_{i,t} = 0$ means that the firm chooses to dismiss him. Conditional on employing a worker of type i at period t, the firm makes a take-it-or-leave-it, flat-wage offer $w_{i,t} \geq 0$. This is a "spot" contract that covers period t only (put differently, the firm can renegotiate the labor contract at the beginning of every period).

The two parties are endogenously separated at period t if the firm dismisses the worker, or if the worker rejects the firm's wage offer. In this case (as well as following an exogenous separation), the worker joins the search pool of period t, while the firm chooses whether to be in the search pool of period t.

Reference-dependent output. Conditional on accepting a wage offer w_t at period t, an employed worker of type i produces an output level given by

$$y_{i,t} = \begin{cases} p_t & if \ \ w_{i,t} \geq e_{i,t} \\ \gamma_t p_t & if \ \ w_{i,t} < e_{i,t} \end{cases}, \tag{2}$$

where:

- p_t is the level of productivity that characterizes the economy at period t. We assume that p_t follows a log-linear AR(1) process with a long-run mean of 1, that is, $p_t = p_{t-1}^{\beta}\varepsilon_t$, where $\beta \in (0, 1)$ is the autocorrelation coefficient, and ε_t is i.i.d. according to a continuous, strictly increasing *cdf* $F[1/\xi, \xi]$, where $\xi > 1$. Finally, $F(\varepsilon) \equiv 1 - F(1/\varepsilon)$; that is, $\ln(\varepsilon)$, is symmetrically distributed around zero.
- $e_{i,t}$ is the worker's reference wage. We assume that a worker enters his first period of employment at a given firm with modest aspirations, in the sense that his reference point $e_{1,t}$ equals the lowest possible

wage, which is zero. On the other hand, existing workers, who were employed by the same firm at period $t - 1$, enter period t with a reference point equal to the wage they expected to earn at t conditional on being retained. Thus, at any period t, $e_{1,t} = 0$; and for every $i > 1$, $e_{i,t}$ is the expectation of $w_{i,t}$ conditional on being retained at t, given the worker's information at the end of period $t - 1$ and the continuation strategies followed by all agents.

- $[\gamma_t \in 0, 1]$ is a random parameter representing the fraction of output loss due to worker demoralization when their wage falls below the reference point. It captures the effect of wage disappointment on workers' output (and implicitly, the extent to which the labor contract is incomplete; this interpretation is substantiated in appendix C). We assume that γ_t is i.i.d. according to a *cdf* G that has no mass point in [0,1). We also assume that $G(\gamma) < 1$ for every $\gamma < 1$.

Information. In each period $t \geq 1$, every agent observes the realizations of all exogenous random variables up to (and including) period t. In particular, ε_t and γ_t are common knowledge at the time the firm chooses its wage offer w_t. The agent also observes his own private history. Finally, whenever a firm and a worker interact, they observe the history of wage offers since they were matched. They do not observe the negotiation history in other firm-worker matches.

Preferences. All agents in the model maximize their expected discounted sum of payoffs, using the same constant discount factor δ. The payoff flow for firms at each period is as follows. A firm outside the labor market earns zero. A firm in the search pool earns $-c$, where $c > 0$ is the cost of posting a vacancy. A firm in a relationship with a worker earns a payoff that equals output minus the wage paid. An unemployed worker at period t receives a nonmarket payoff of bp_t, where $b \in (0, 1)$. An employed type-i worker gets a payoff of $w_{i,t}$. The assumption that the outside option is proportional to current productivity is made not only for simplicity, but also to ensure that in the reference-independent benchmark, equilibrium wages will be fully flexible, such that all rigidity effects will arise from the novel behavioral element.

Empirical background for our behavioral model. Our account of workers' output reflects two well-documented behavioral phenomena that involve reference dependence.

Loss aversion. This concept (due to Kahneman and Tversky 1979) means that decision makers register outcomes in terms of gains or losses relative to a reference point, and react to losses more strongly than gains. There is abundant experimental support for loss aversion, as well as field evidence that this motive is relevant in high-stakes decision problems. For example, homeowners are reluctant to lower their asking price when a boom in the real estate market is followed by a downturn (Genesove and Mayer 2001).[5] Models incorporating loss aversion have been proposed to explain phenomena such as consumer antagonism to price hikes (Anderson and Simester 2010), the Equity Premium Puzzle (Benartzi and Thaler 1995), and skewed managerial compensation schemes (de Meza and Webb 2007; Dittmann, Maug, and Spalt 2010).

Negative reciprocity. When individual i is willing to incur a cost in order to punish another individual j for choosing an action that inflicts harm on i relative to some reference point, we say that he exhibits negative reciprocity (an example is responders' tendency to reject "insultingly low" offers in the Ultimatum Game).[6] Fehr and Gächter (2000) survey the literature, and point out that negative reciprocity emerges as a stronger motive than its counterpart, positive reciprocity (the propensity to reward friendly behavior): "Whereas the positive effects of fair treatment on behavior are usually small, the negative impact of unfair behavior is often large" (Fehr, Goette, and Zehnder 2009, 366). Our model approximates this finding by assuming away positive reciprocity. Models of negative reciprocity have been used to shed light on economic phenomena such as the prevalence of incomplete labor contracts and the endogenous emergence of long-term relational contracts (see Fehr and Schmidt 2002, and Hart and Moore 2008).

Formula (2) captures these phenomena in reduced form: workers perceive a wage offer below their reference point as a loss, which triggers a negative-reciprocity response.

A. The Reference-Independent Benchmark

Let us first consider the benchmark model in which $\gamma = 1$ with probability one, where output is reference-independent. In this case, our model reduces to a standard S&M model in which firms have all the bargaining power.

Proposition 1 *Let* $\gamma = 1$. *There is a unique SPE, in which firms choose* $(r_t, w_t) = (1, bp_t)$ *at every* t *and regardless of the worker's type, and workers accept any wage offer weakly above* bp_t.

Equilibrium in the reference-independent benchmark exhibits several noteworthy features. First, equilibrium behavior is Markovian in a narrow sense: hiring/retention and wages at any period t are purely a function of p_t. Second, wages are entirely flexible, in the sense that they are proportional to productivity. Third, there is no behavioral distinction between newly matched and existing workers. Finally, there are no layoffs.

Proposition 1 determines equilibrium market tightness via a free-entry property. A firm's expected discounted benefit from posting a vacancy at period t, conditional on finding a new match at the beginning of $t + 1$, is equal to the expected discounted sum of the firm's payoffs over the duration of the employment relation. Formally, it is a function of the state at t, defined as follows:

$$\Pi(p_t) = (1 - b)\sum_{i=1}^{\infty}\delta^i \left(\prod_{0<j<i} (1 - s(j))\right)\mathbb{E}(p_{t+i}| \, p_t). \tag{3}$$

Note that Π is an increasing function. If $c > \Pi(p_t)$, then in SPE no firm posts a vacancy at t, and market tightness is infinite. If $c \leq \Pi(p_t)$, then in equilibrium firms will be indifferent between searching and not searching. The probability q_t that a searching firm will find a match at the beginning of $t + 1$ will be set such that $c = q_t\Pi(p_t)$. Market tightness is derived from q_t according to (1). Hence, equilibrium tightness at t is purely a function of p_t as well.

III. Equilibrium under a Two-Period Separation Process

We now analyze SPE in our model, under the following restriction on the exogenous job separation process: $s(1) = 0$ and $s(2) = 1$. That is, the employment relation lasts at most two periods. This could approximate industries in which firm-specific human capital depletes quickly as a result of rapid technological changes. However, we assume it mainly for tractability. We briefly discuss more complex finite-horizon separation processes at the end of this section.[7]

It is useful to make two preliminary observations. First, in SPE, all newly matched workers at any given period are treated identically; similarly, all existing workers at any given period are treated identically. The reason is that all agents on each side of the market are identical, and no firm-worker pair gets to observe the history of any pairwise interaction prior to their own match, thus preventing the emergence of history-dependent asymmetries. In what follows we often refer to the way "the worker" or "the firm" behave at a given history, with the understanding that this pertains to all firms and all workers of the same type at the same period.

Second, we can think about an equilibrium wage offer in terms of whether it satisfies a worker's individual rationality (IR) and morale hazard (MH) constraints, in analogy to IR/IC constraints in contract theory. Fix a history h following a wage offer. An SPE satisfies the *IR constraint* at h if the worker is weakly better off than if he rejects the firm's wage offer and sticks to his equilibrium strategy thereafter. An SPE satisfies the *MH constraint* at h if the wage offer at h is weakly higher than the worker's reference wage at that history.

By assumption, the MH constraint coincides with the constraint that wages are nonnegative as far as newly matched workers are concerned. Therefore, the MH constraint is only relevant for existing workers. According to the one-deviation property of SPE, the IR constraint always holds in equilibrium, and the only question is at which histories it is binding. Note that in SPE, if the IR constraint holds with slack at h, the MH constraint must be binding. The reason is simple: if the MH constraint is violated or holds with slack, the firm can slightly lower its wage and the worker will accept the offer (because if he rejects it he will join the search pool, and since this rejection is unobserved by future employers, they will be unable to adjust their wage offers to it) and his productivity will be unaffected.

The following result characterizes wage and retention policies in SPE, under a mild condition on the magnitude of the business cycle.

Proposition 2 *Let $\xi \leq (1/2)(1+\sqrt{5})$. Then, the game has a unique SPE outcome, which has the following properties.*[8]

(i) An existing worker's period-t reference point is

$$e_{2,t} = \phi \cdot b p_{t-1}^{\beta}$$

where the coefficient $\phi \in [\mathbb{E}(\varepsilon), \xi]$ is uniquely determined by the following equations:

$$\phi = \frac{\int_0^1 \int_{\varepsilon > \varepsilon^*(\gamma)} \max\{\phi, \varepsilon\} dF(\varepsilon) dG(\gamma) + \int_b^1 \int_{\varepsilon < \varepsilon^*(\gamma)} \varepsilon dF(\varepsilon) dG(\gamma)}{1 - G(b)F(\phi b)} \tag{4}$$

$$\varepsilon^*(\gamma) = \frac{b\phi}{1 - \max\{0, \gamma - b\}}. \tag{5}$$

(ii) An existing worker is dismissed at period t if and only if $\gamma_t < b$ *and* $\varepsilon_t < \phi b$. *Conditional on being retained at t, his wage is*

$$w_2(p_{t-1}, p_t) = \begin{cases} \max\{e_{2,t}, bp_t\} & \text{if } \ \varepsilon_t > \varepsilon^*(\gamma_t) \\ bp_t & \text{if } \ \gamma_t > b \text{ and } \varepsilon_t < \varepsilon^*(\gamma_t) \end{cases}. \tag{6}$$

(iii) A newly matched worker at period t is always hired; his wage at period t is

$$w_1(p_t) = b\left[p_t - \delta p_t^{\beta}\int_0^1\int_{\varepsilon^*(\gamma)}^{\phi}(\phi - \varepsilon)dF(\varepsilon)dG(\gamma)\right]. \tag{7}$$

Let us list the important qualitative features of the SPE outcome.

Wage rigidity and endogenous output destruction. Existing workers may experience wage rigidity, layoffs, or loss of morale, depending on the realizations ε_t, γ_t. When $\varepsilon_t \in (\varepsilon^*(\gamma_t), \phi)$, existing workers at period t are retained and paid their reference wage $e_{2,t}$, which is purely a function of p_{t-1} and therefore rigid in the sense that it is not responsive to productivity shocks in the range $(\varepsilon^*(\gamma_t), \phi)$. When $\varepsilon_t > \phi$, existing workers receive their participation wage, which lies above the reference wage, and therefore produce normal output. When $\varepsilon_t < \varepsilon^*(\gamma_t)$, existing workers experience destruction of output: either $\gamma < b$, in which case they are fired; or $\gamma > b$, in which case they are kept at their participation wage, which lies below their reference wage, and thus produce subnormal output due to loss of morale. Because the destruction of output experienced by an existing worker is purely a function of $(\varepsilon_t, \gamma_t)$, the expected output that a newly hired worker at period t believes he will produce at $t + 1$ is $\lambda p_t^{\beta}\mathbb{E}(\varepsilon)$, where the constant λ is given by

$$\lambda = \frac{1}{\mathbb{E}(\varepsilon)}\left[\int_0^1\int_{\varepsilon>\varepsilon^*(\gamma)}\varepsilon dF(\varepsilon)dG(\gamma) + \int_b^1\int_{\varepsilon<\varepsilon^*(\gamma)}\gamma\varepsilon dF(\varepsilon)dG(\gamma)\right]. \tag{8}$$

History dependence. The equilibrium treatment of existing workers at period t is Markovian with respect to an extended state (p_t, p_{t-1}, γ_t) (or, equivalently, $(p_t, \varepsilon_t, \gamma_t)$). Their reference wage is purely a function of p_{t-1}. Whether they receive it or the participation wage bp_t (which is purely a function of p_t) depends entirely on $(\varepsilon_t, \gamma_t)$; and so does their retention policy. Therefore, existing workers' ex ante layoff rate at any period t is $G(b)F(\phi b)$, independently of the history up to period $t - 1$. Newly hired workers' wage is purely a function of p_t.

IR and MH constraints. The IR constraint of newly matched workers is always binding, and consequently their continuation payoff at any period t is as if they earn bp_t, at every $t' \geq t$. By contrast, existing workers' IR constraint holds with slack whenever $\varepsilon_t \in (\varepsilon^*(\gamma_t), \phi)$; that is, whenever they are retained and paid their reference wage, in which case their MH constraint is binding. When existing workers are retained at their participation wage, their IR constraint is binding while their MH constraint is violated (if $\gamma_t > b$ and $\varepsilon_t < \varepsilon^*(\gamma_t)$) or satisfied with slack (if $\varepsilon_t > \phi$).

The structure of entry-level wages. The equilibrium wage paid to new hires is both strictly positive and strictly increasing in p_t (this is ensured by our restriction on ξ), albeit at a lower rate than in the $\gamma = 1$ benchmark. In this sense, entry-level wages are partially flexible w.r.t. current productivity. Note that unlike the $\gamma = 1$ benchmark, equilibrium wages exhibit a seniority premium: existing (newly matched) workers earn wages above (below) the current outside option.

Sketch of the proof of proposition 2. First, we derive an upper bound on the rent that existing workers can get in equilibrium, which translates into a lower bound on newly hired workers' wage. This bound is above zero, such that the wage offer to newly hired workers satisfies the MH constraint with slack. Hence, their IR constraint is always binding in equilibrium. (This is the difficult part of the proof, and it relies on the assumption that $\xi < (1/2)(1 + \sqrt{5})$.) This in turn implies that newly matched workers must always be indifferent between accepting an equilibrium wage offer (and sticking to their equilibrium strategy thereafter) and being permanently unemployed. Therefore, an existing worker at period t would accept any wage above bp_t. We have thus fixed existing workers' participation wage.

For any given reference wage $e_{2,t}$, we can check, for every realization of ε_t, γ_t, which of the following three courses of action maximizes the firm's profit: (1) dismiss an existing worker, (2) keep him at his participation wage, (3) keep him at his reference wage. This enables us to write down the expression for $e_{2,t}$, which is uniquely given by equations (4) and (5). The assumption that $G(\gamma) < 1$ for every $\gamma < 1$ is instrumental in the uniqueness of the solution. Otherwise, it could be possible that existing workers' reference wage at t is strictly higher than $b\xi p_{t-1}^{\beta}$, namely the maximal outside option that is feasible given p_{t-1}, and firms would always stick to the reference wage in order to avert loss of worker morale. When γ is very close to one, firms would not have an incentive to do so, and this prevents the reference wage from being equal to the lagged-expected wage. Uniqueness can be generated by other perturbations as well (see Eliaz and Spiegler 2012).

The cutoff $\varepsilon^*(\gamma)$ is the productivity shock for which the firm is indifferent between keeping the worker at his reference wage and dismissing him or keeping him at his participation wage, depending on whether γ is below or above b. We have thus derived existing workers' equilibrium wage, and the firm's retention policy immediately follows from that. To obtain new hires' wage, we use their indifference to permanent unemployment, such that their equilibrium wage at t is equal to bp_t minus the discounted rent they expect to receive as existing workers at $t + 1$.

Two special cases. First, revisit the reference-independent benchmark by letting $G(\gamma) = 0$ for all $\gamma < 1$. Since $G(b) = 0$, existing workers are always retained. Applying formulas (4) and (5), we obtain $\phi = \mathbb{E}(\varepsilon) = \varepsilon^*(1)$; hence, an existing worker at t receives bp_t. Applying formula (7), we obtain that a newly hired worker receives the same wage. This reproduces proposition 1 for the two-period separation process.

Second, consider the limit case $G(b) \to 1$. Observe that formula (4) collapses into

$$\phi = \mathbb{E}[\max\{\phi, \varepsilon\} \mid \varepsilon > \phi b].$$

The solution to this equation is $\phi = \xi$, which implies $\varepsilon^*(\cdot) = \xi b$ with probability one. Existing workers are thus retained and paid $w_{2,t} = \xi b p_{t-1}^{\beta}$ whenever $\varepsilon_t > \xi b$, and dismissed otherwise. Existing workers' output coefficient λ is given by

$$\lambda = \frac{1}{\mathbb{E}(\varepsilon)} \int_{\xi b}^{\xi} \varepsilon dF(\varepsilon).$$

Newly hired workers earn

$$w_{1,t} = b\left[p_t - \delta p_t^{\beta} \int_{\xi b}^{\xi} (\xi - \varepsilon) dF(\varepsilon)\right]. \tag{9}$$

Existing workers' equilibrium wage in this case is absolutely rigid, in the sense that it is purely a function of productivity in the previous period. Wage rigidity here has a flavor of "grade inflation." When $G(b) \to 1$, existing workers' reference wage is the expectation of the maximum between the outside option and the reference wage itself. This means that the reference wage must always be greater than or equal to the expected outside option, which can only be true if the reference wage equals the highest possible value of the outside option. When $\gamma < b$, a firm would rather dismiss a worker than paying him a wage below his reference point. Thus, existing workers get their reference wage with probability one, conditional on being retained.

A. Volatility of Market Tightness

In order to study the equilibrium volatility of market tightness, we follow the S&M literature, and assume in this subsection that the matching function takes the following form

$$m(U_t, V_t) = kU_t^{\alpha}V_t^{1-\alpha}, \tag{10}$$

where $\alpha \in (0, 1)$ and k is sufficiently small so that match probabilities are always well-defined. This allows us to get an explicit, closed-form expression for market tightness. Let us first establish that in SPE, tightness at any period t is purely a function of p_t. The expected discounted profit generated by a vacancy opened in period t conditional on getting a new match at the beginning of period $t + 1$ is

$$\delta(1 - b)\mathbb{E}(\varepsilon)[p_t^{\beta} + \delta\lambda p_t^{\beta^2}\mathbb{E}(\varepsilon^{\beta})],$$

where λ is given by (8). This expression is an increasing function of p_t, and we denote it by $J(p_t)$. Note that in the $\gamma = 1$ benchmark, we have $\lambda = 1$, hence $J(p_t)$ is reduced to $\Pi(p_t)$, as given by (3).

Lemma 1 *In the SPE characterized by proposition 2, θ_t is a function of p_t given by the following equation:*

$$\theta_t(p_t) = \sqrt[\alpha]{\frac{kJ(p_t)}{c}},$$

as long as $c/J(p_t) < 1$. Otherwise, market tightness is zero.

To understand why equilibrium market tightness is a well-defined function of current productivity, recall that θ_t is a strictly decreasing function of q_t, the probability that a searching firm finds a match at t. Because of free entry, q_t itself is a function of $J(p_t)$. Thus, although some aspects of equilibrium behavior at t (specifically, the treatment of existing workers) depend on p_{t-1}, tightness is only a function of p_t.

To see how reference dependence affects tightness volatility, let us write down the expression for the elasticity of θ_t w.r.t. p_t (omitting the subscript):

$$\frac{\beta}{\alpha} \cdot \frac{p^{\beta} + \beta\delta\lambda p^{\beta^2}\mathbb{E}(\varepsilon^{\beta})}{p^{\beta} + \delta\lambda p^{\beta^2}\mathbb{E}(\varepsilon^{\beta})}. \tag{11}$$

It is easy to verify that this expression decreases with λ in absolute terms. Recall that $\lambda = 1$ in the reference-independent benchmark, and that $\lambda < 1$ when reference dependence is introduced. We conclude that reference dependence increases tightness volatility. The intuition for this effect is as follows. In expectation, a constant fraction $1 - \lambda$ of existing workers' normal output is destroyed, independently of the history. Therefore, the constant λ acts like an additional discount factor between the worker's first and second periods of employment. The extra discount factor increases the weight that first-period output receives in the calculation of the value of the vacancy. Because $\beta < 1$, the worker's productivity in his second period of employment is less sensitive to the value of p that prevailed at the time the firm originally posted the vacancy than his first-period productivity. Therefore, introducing the new term λ increases the sensitivity of the vacancy's value to the initial value of p. This intuition clarifies why the volatility effect disappears when $\beta \to 1$.

For a closer look at the interplay between the effects of reference dependence and the persistence of the business cycle, suppose that productivity is at the long-run average, that is, $p_t = 1$. At this point, the elasticity of tightness is

$$-\frac{\beta}{\alpha} \cdot \frac{1 + \beta\delta\lambda\mathbb{E}(\varepsilon^{\beta})}{1 + \delta\lambda\mathbb{E}(\varepsilon^{\beta})}. \tag{12}$$

When β is high, the "standard" tightness volatility—that is, the value of (12) for $\lambda = 1$—is higher. However, at such values of β, the effect of reference dependence on tightness volatility vanishes. At the other extreme, when β is low, "standard" tightness volatility is low, but the effect of reference dependence is large. The derivative of (12) w.r.t. λ is maximized at $\beta = 1/2$. Thus, the effect of reference dependence on tightness volatility is maximized at intermediate levels of persistence.

B. The Role of γ

Thus far, we have compared SPE in our model for an arbitrary G to the reference-independent benchmark $\gamma = 1$. Let us extend this comparative-statics exercise and ask more generally how the equilibrium outcome changes as output becomes more sensitive to wage drops below the worker's reference point. For simplicity, we focus on the limit case in which G assigns probability one to some particular value $\gamma < 1$. We already saw that in the range $\gamma < b$, changes in γ have no impact on the equilibrium outcome.

Let $\gamma > b$. In this case, the expression for ϕ is reduced to

$$\phi = \mathbb{E}(\varepsilon) + \int_{m\phi}^{\phi} (\phi - \varepsilon) dF(\varepsilon),$$

where

$$m = \frac{\varepsilon^*(\gamma)}{\phi} = \frac{b}{1 - \gamma + b}.$$

It is straightforward to check that as γ goes up, ϕ decreases while $\varepsilon^*(\gamma)$ rises. This means that existing workers' reference wage, as well as the range of realizations of ε_t for which they are paid this wage, shrink. As a result, newly hired workers' wage goes up and approaches their outside option. When $\gamma \to 1$, the interval $[m\phi, \phi]$ vanishes, and equilibrium wage and retention policies converge to the reference-independent benchmark. The effect on λ of raising γ is ambiguous: on one hand, the probability of subnormal output due to demoralization increases, but on the other hand, the output loss due to demoralization is lower because γ is higher. Therefore, it is not clear to us whether tightness volatility monotonically decreases with γ in the range $(b, 1)$.

C. *The Role of Contract Renegotiation*

Our model assumes that firms repeatedly offer one-period contracts. In particular, they are unable to commit at period $t - 1$ to the worker's wage at t. Does this assumption make a difference?[9] For simplicity, consider the limit case $G(b) \to 1$. In the absence of commitment, we saw that an existing worker's wage at period t is $\xi b p_{t-1}^{\beta}$. Now suppose that at the end of period $t - 1$, the firm can commit to pay a fixed wage $w^* < \xi b p_{t-1}^{\beta}$ to its worker at period t. Assume the firm cannot commit not to dismiss the worker. As a result, the worker will be retained and paid w^* when $p_t \geq w^*$, and dismissed otherwise. His reference wage will consequently be w^*, and he will produce the normal output p_t if he continues to be employed. However, the worker will reject the firm's offer at t if $b p_t^{\beta} > w^*$. If the probability that $b p_t^{\beta} > w^*$ is sufficiently low, the firm's loss from these inefficient rejections will be outweighed by the gain from the reduced wage whenever $b p_t^{\beta} \leq w^*$. In this case, the firm would benefit from commitment.

Note, however, that the worker's rejection of the fixed wage offer $b p_t^{\beta} > w^*$ is precisely the event in which the two parties would want to renegotiate the contract, because the rejection involves inefficient destruction of surplus, which can be prevented by a renegotiated flat-wage contract at period t. Therefore, the assumption that contracts are renegotiated every period seems appropriate in the present context.

D. *An Exercise: The Effect of Payroll Tax*

Suppose that a payroll tax at a constant rate τ is imposed on firms.[10] For simplicity, consider the limit case $G(b) \to 1$, where firms pay existing workers their reference wage conditional on retaining them. Existing workers' wage does not change, because their outside option at t continues to be bp_t, hence their wage continues to be ξbp_{t-1}. However, as far as the firm is concerned, imposing the payroll tax is equivalent to raising the workers' outside option coefficient from b to $b/(1-\tau)$, hence the layoff cutoff changes to $\varepsilon^* = \xi b/(1-\tau)$. By (9), new workers' wage *increases*, because their expected future rent shrinks. The rise in ε^* also implies that λ goes down, hence tightness volatility increases.

E. *General Finite-Horizon Separation*

Our analysis in this section was based on the assumption that $s(1) = 0$ and $s(2) = 1$. Let us consider a generalization of this exogenous separation process, in which $s(i) = 0$ for every $i = 1, \ldots, T-1$, and $s(T) = 1$, where $T \geq 2$. Characterization of SPE would proceed along the same lines as in proposition 2, with three key differences. First, the retention policy is more complex. In particular, a firm may prefer to retain a worker at a reference wage above his output because of a high continuation payoff. Second, in order to ensure that newly hired workers' wage is strictly positive, a stronger condition on the magnitude of the business cycle is required. Third, the expressions for the worker's wage as a function of his tenure are more cumbersome.

For illustration, let us construct an SPE in which newly hired workers' IR constraint is always binding, under the following parametric restrictions: $T = 3$, $\beta = 1$, $\xi < \sqrt{2}$, and $G[b - (1-b)\mathbb{E}(\varepsilon)] \to 1$ (the latter restriction implies $b > 1/2$). Consider first a worker of type $i = 3$ at period t. This worker is essentially equivalent to a worker of type 2 in the two-period model: his wage equals the maximal outside option at period t conditional on p_{t-1}, that is, $w_{3,t} = e_{3,t} = \xi bp_{t-1}$, and he is retained if and only if $\varepsilon_t \geq \xi b$. Since $\gamma < b$ with probability one, the worker almost surely receives his reference wage conditional on being retained.

Next, consider a worker of type $i = 2$ at the same period t. His *participation* wage, denoted $\bar{w}_{2,t}$, is the same as newly hired workers' equilibrium wage in the two-period model, that is, $\bar{w}_{2,t} = \hat{b}p_t$, where

$$\hat{b} = b\left[1 - \delta\int_{\xi b}^{\xi}(\xi - \varepsilon)dF(\varepsilon)\right] < b.$$

Let us guess that type 2 workers receive their reference wage with probability one conditional on being retained. Therefore, by the same reasoning as in the case of type 3 workers, we obtain $w_{2,t} = e_{2,t} = \xi\hat{b}p_{t-1}$. To confirm that the guess is correct, we need to verify that the firm's expected discounted sum of profits from keeping the worker at a wage below his reference point is almost surely negative, that is,

$$\gamma_t p_t - \hat{b}p_t + \delta p_t\int_{\xi b}^{\xi}(\varepsilon - \xi b)dF(\varepsilon) < 0$$

for almost all realizations of γ_t. Our assumption on G ensures that this is the case. It follows that the firm retains type 2 workers at period t if and only if ε_t is above a cutoff ε^*, that is given by

$$p_{t-1}\varepsilon^* - \xi\hat{b}p_{t-1} + \delta p_{t-1}\varepsilon^*\int_{\xi b}^{\xi}(\varepsilon - \xi b)dF(\varepsilon) = 0,$$

hence

$$\varepsilon^* = \frac{\xi\hat{b}}{1 + \delta\int_{\xi b}^{\xi}(\varepsilon - \xi b)dF(\varepsilon)} < \xi b.$$

It remains to derive the wage of type 1 workers at period t and verify that it is strictly positive. Because we are asserting a binding IR constraint for new hires, these workers should be indifferent to permanent unemployment. Therefore, their wage is equal to their outside option minus the expected discounted sum of rents they accumulate as existing workers:

$$w_{1,t} = bp_t - \delta\int_{\varepsilon^*}^{\xi}\left[\hat{b}p_t(\xi - \varepsilon_{t+1}) + \delta bp_t\varepsilon_{t+1}\int_{\xi b}^{\xi}(\xi - \varepsilon_{t+2})dF(\varepsilon_{t+2})\right]dF(\varepsilon_{t+1})$$

The assumption that $\xi < \sqrt{2}$ ensures that $w_{1,t} > 0$. Hence, newly hired workers' MH constraint holds with slack, implying that their IR constraint is binding.

Observe that in this equilibrium, all existing workers at t are paid a fully rigid wage conditional on being retained, which is purely a function of p_{t-1}, whereas newly hired workers' wage at t is a function of p_t. Wages exhibit a "seniority premium": $w_{i,t}$ increases with i (it is obvious that $w_{2,t} < w_{3,t}$; verifying that $w_{1,t} < w_{2,t}$ is less immediate, and ensured by the restriction that $b > 1/2$). Finally, workers with a longer tenure are more likely to be dismissed.

IV. Discussion

This section discusses two features of our model: its relation to the Shimer puzzle, and the model's ability to capture persistent effects of productivity shocks on wages, inflows into unemployment, and tightness.

A. Wage Rigidity and the Shimer Puzzle

The enhanced tightness volatility discussed in section III, subsection A naturally brings Shimer's puzzle to mind. We reemphasize that the following discussion is not an attempted resolution of the puzzle, but a clarification of theoretical arguments that were raised in response to it. Also, this section is not intended to be a survey of recent attempts to resolve the Shimer puzzle. We focus on a small number of approaches that are straightforward to compare to ours, and a number of important works on the subject (such as Hall and Milgrom [2008], Gertler and Trigari [2009], and Kennan [2009]) are not mentioned because of the difficulty of comparison.

Shimer (2004) and Hall (2005) proposed to resolve the puzzle by assuming that wages are only *partially responsive* in the sense that they remain fixed as long as they do not violate the participation constraint of either the firm or the worker. The latter requirement is made to avoid the "Barro critique" that rational parties would find a way to renegotiate when wage rigidity threatens to destroy surplus (Barro 1977). The Shimer-Hall approach imposes wage rigidity a priori, without deriving it from explicit behavioral or institutional considerations. In contrast, our model generates wage rigidity from workers' reference-dependent behavior. In addition, the Shimer-Hall approach does not distinguish between newly hired and existing workers; whichever force that generates wage rigidity applies to all workers, regardless of their tenure.

Incorporating such a distinction into S&M models seems important since casual observation, as well as the evidence from Bewley (1999) and Fehr, Goette, and Zehnder (2009), suggest that the psychological forces that give rise to wage rigidity have a greater impact on existing relationships. The Shimer-Hall approach could be modified to accommodate this distinction, simply by imposing wage rigidity only on existing workers (this would not change implications for aggregate wage data because new hires constitute a small fraction of the total stock of employed workers at any given point in time).

However, this modification *cannot* generate increased tightness volatility. As pointed out by Pissarides (2009) (as well as Kudlyak [2009], and Haefke, Sonntag, and van Rens [2012]), a newly matched pair fully internalizes all future rigidities into their negotiation, such that the equilibrium wage offsets all future departures from the "normal" surplus-division rule. As a result, the firms' hiring incentives are unaffected by the anticipated rigidity of existing workers' wage.

The previous discussion highlights the difficulty of constructing a model that can simultaneously accommodate the following ingredients: a distinction between newly hired and existing workers in which only the latter experience wage rigidity, increased tightness volatility relative to a benchmark without wage rigidity, and robustness to the Barro critique. Our framework offers a way of accomplishing this. The key is the incompleteness of the labor contract and the workers' time-changing reference point. The standard S&M model, as well as the Shimer-Hall modification, assume complete contracts. When complete contracts are feasible, the rule for dividing the surplus does not affect the size of the surplus. This independence breaks down in our model. When a firm violates an existing worker's MH constraint by paying him a wage below his reference point, the bargaining set effectively shrinks due to the worker's loss of morale, potentially to the point where all gains from mutual agreement are dissipated. As a result, the value of a new firm-worker match is not neutral to anticipated wage rigidity.

Does the fact that our model gives rise to inefficient output destruction (in the second period of the firm-worker relationship) mean that it is vulnerable to the Barro critique? An implicit assumption behind the critique is that the labor contract is complete: the firm and the worker can always reach a contract that shares any available surplus (recall that Barro [1977] predated the rich literature on incomplete contracts). Our model rests on the assumption that the labor contract is incomplete: the interaction between the two parties involves events outside the contract's scope, which arise from the change in the worker's reference point. Inefficient output destruction in our model is a consequence of this contractual incompleteness. Thus, in a deep sense, our model is not vulnerable to Barro's critique.[11]

The claim that new hires' wage is more flexible than existing workers' wage has some empirical support. Most recently, Haefke, Sonntag, and van Rens (2008) constructed a time series for wages of new hires using micro-data on earnings and hours worked from the Current Population Survey (CPS) outgoing rotation groups. They found that the wage for newly hired workers is relatively responsive to productivity (however, their estimate has a rather large standard error). Our paper is of course not the place for empirical evaluation of claims as to the relative flexibility of the wage of existing workers and new hires, and our discussion here focuses on the *theoretical* aspects of the debate. Furthermore, it is important to distinguish between *flexible determination* of new hires' wage and the *responsiveness* of this wage to productivity. In particular, in our model new hires' wages are only *partially* responsive to productivity, even though they are flexibly determined (in the sense that the MR constraint holds with slack).[12]

It is also useful to think about our theoretical argument in relation to another well-known response to Shimer's puzzle. Hagedorn and Manovskii (2008) argued that the Shimer Puzzle can be resolved under a different calibration strategy that assigns more bargaining power to the firm and implies a less procyclical outside option for workers. Note that while this approach can account for rigid wages, it does not distinguish between new hires and existing workers, and thus implies that the former earn rigid wages as well. Our model, of course, takes the hagedorn-Manovskii assumption regarding the distribution of bargaining power to the extreme. However, it shows that even if the workers' outside option is maximally procyclical, one obtains patterns of rigid wages for existing workers and partially flexible wages for new hires.

B. Comparison with Other Classes of Models

In this subsection we discuss our results in comparison with alternative S&M models of the labor market. We will demonstrate that the combination of effects that our model generates—wage rigidity for existing workers, flexible entry-level wages, a seniority premium, endogenous job destruction that is sensitive to changes in productivity, and enhanced volatility of market tightness—cannot be reproduced by these alternative models.

Idiosyncratic Shocks and Endogenous Job Destruction

Since endogenous destruction of output plays a major part in our tightness volatility result, it is natural to ask whether other mechanisms of endogenous job destruction would generate similar patterns. The most well-known S&M model that exhibits endogenous job destruction, due to Mortensen and Pissarides (1994)—referred to as the MP model henceforth—generates this effect through idiosyncratic productivity shocks. To create an MP-like model that is as comparable to ours as possible, modify the benchmark model as follows. First, assume the same two-period separation process as in section III: $s(1) = 0$, $s(2) = 1$. Assume further that the output of an *existing* worker is subjected to a random idiosyncratic shock, such that an employed worker produces an output of $p_t\nu_t$, where ν_t is i.i.d. with $\mathbb{E}(\nu) = 1$ across firms and periods.

SPE wage offers in this model are exactly as in the benchmark: workers are always offered bp_t when they are employed. Hiring and retention decisions are as follows: $r_t = 1$ if and only if $\nu_t \geq b$. Thus, in each period, a constant fraction of firms will choose to fire existing workers—just like our model. However, this variation *lowers* the volatility of market tightness—the exact opposite of our effect. To see why, note that we could reinterpret the benchmark model as an MP model in which firms make their hiring/retention decisions *before* learning the realization of their idiosyncratic shock. Because $\mathbb{E}(\nu) = 1 > b$, firms will always choose $r = 1$. When $\nu_t < b$, the firm's precommitment to play $r = 1$ is inefficient ex post; if the firm could delay its decision until *after* it has learned its idiosyncratic shock, it would efficiently fire the worker.

This is a simple value-of-information argument: enabling firms to move after learning their idiosyncratic shock increases expected profits. But this means that when we switch from the benchmark model to the MP-like model of this subsection, this is equivalent to introducing a *premium* factor $\lambda > 1$ to the firm's profit in the second period of its relationship with the worker. In other words, the MP-like variation *increases* the importance of the second period in determining the value of the vacancy, thereby *reducing* its sensitivity to initial conditions.

This comparison highlights the feature that endogenous separations in our model *destroy* value. The worker's changing reference point and the firm's inability to offer a complete labor contract imply that vacancies will be closed even though the two parties would have agreed ex ante that it would be efficient to keep them. In contrast, vacancies in the MP model are closed if and only if it is efficient to do so. This difference translates to tightness volatility effects in opposite directions.

Moral Hazard and Efficiency Wages

Our model is essentially an efficiency-wage model: in equilibrium, firms pay (existing) workers a wage above their reservation value, in order to induce higher output. The mechanism that generates this effect is based on reciprocal fairness considerations, in the tradition of Akerlof (1982), but there could be others. Shapiro and Stiglitz (1984) assume that when a worker shirks, he is caught and fired with some probability. In order for the worker to have an incentive to exert effort, the firm must offer him a wage above his outside option.

Mortensen (1989), and more recently, Costain and Jansen (2010) and Malcolmson and Mavroeidis (2010) incorporated the Shapiro-Stiglitz efficiency wage model into an S&M model. To illustrate the similarities and differences between such a model and ours, we briefly analyze the following modification of the benchmark model. After a worker accepts a wage offer at period t, he makes an effort decision $x_t \in \{0, 1\}$ and products an output of $p_t[\gamma_t + (1 - \gamma_t)x_t]$. Assume $\gamma < b$, and suppose that the firm can observe x_t with probability χ. An employed worker's payoff at period t is $w_t - \kappa x_t$, where κ is his cost of effort.

Since $\gamma < b$, the incentive constraint that induces workers to exert effort must hold in order for firms to earn positive profits. In SPE, both this constraint and the IR constraint will be binding. As a result, equilibrium wage at period t will be $bp_t + \kappa/(1 - \chi)$. Firms will therefore choose $r_t = 1$ if and only if $p_t \geq \kappa/(1 - \chi)(1 - b)$. This means that separation will be more frequent when productivity is low. As a result, efficiency wages will have an adverse effect on the incentive to hire new workers at low values of p, such that the effect on tightness volatility will be roughly in the same direction as in our model. However, the equilibrium wage is linear in p, as in the benchmark model, which means that the model does not generate wage rigidity.

Long-Term Contracts and Consumption Smoothing

An alternative theory of wage rigidity is based on the idea (dating back to Azriadis 1975 and Beaudry and DiNardo 1989) that employers can commit to long-term wage contracts that enable liquidity-constrained workers to smooth consumption across periods. When productivity fluctuates, a risk-neutral employer with no liquidity constraints can essentially offer insurance to a risk-averse worker with limited access to savings. By risk aversion, the worker would be willing to take a pay cut in return for a stream of flat wages. Thus, entry wages would fluctuate with productivity, whereas ongoing wages would be rigid because of the long-term commitment to pay the same wage in each period.

To investigate the effect of risk sharing in our framework, let $\gamma = 1$, and assume that the worker is risk-averse and that the firm can commit to a two-period labor contract. For the sake of illustration, assume separable CARA utility from streams of wage earnings (an analogous argument would hold under CRRA). The risk premium that a new hire in period t would be willing to pay for a constant-wage scheme for periods t and $t + 1$ is independent of p_t. Hence, a firm's expected discounted benefit from posting a vacancy at period t is equal to $\Pi(p_t)$ plus a constant. As in section III, subsection B, assume the matching function is Cobb-Douglas. Lemma 1 implies that the ratio $\theta(p')/\theta(p)$ for $p' < p$ (and hence, volatility of market tightness) is *lower* than in the benchmark.[13]

C. Persistent Effects of Productivity Shocks

In our model, the SPE outcome at any period t is a function of the extended state (p_{t-1}, p_t)—or, equivalently, (p_t, ε_t). Longer lags have no effect; and certain aspects of the equilibrium (new hires' wage, market tightness) are exclusively a function of p_t. These are artefacts of two features of our model: (1) the stochastic process governing productivity is log-linear AR(1); (2) the reference point at t is a function of workers' expectations at $t - 1$. In this subsection we discuss how relaxing these assumptions may lead to longer-lasting effects of productivity shocks.

Suppose that the process governing productivity is $p_t = \Psi(p_{t-1}, \ldots, p_{t-K})\varepsilon_t$, where $K > 1$ is an integer, Ψ is some deterministic function, and ε_t is i.i.d. according to $F[1/\xi, \xi]$, as in our basic model. For simplicity, let us focus on the limit case $G(b) \to 1$. It can be shown that if ξ is sufficiently close to one, SPE characterization is essentially the same as in the basic model. In particular, the characterization of ϕ is the same, such that existing workers at period t earn $\xi b\Psi(p_{t-1}, \ldots, p_{t-K})$ and are retained as long as $\varepsilon_t > \xi b$. Newly hired workers' wage at t and market tightness are exclusively a function of p_t. Lagged realizations of p have no effect on the equilibrium outcome, once we control for p_t, ε_t.

Thus, it would appear that our model cannot generate long-lasting effects of shocks, even under more general productivity processes. However, this is a result of the "purely multiplicative" specification of our model: the process governing p_t is log-linear, and the workers' output and outside option at t are proportional to p_t. This specification was motivated by tractability and methodological considerations. It is not essential for the broad features of our analysis in section III. Yet, as we saw, it rules out long-lasting effects of shocks.

To illustrate this point, suppose $p_t = \Psi(p_{t-1}, \ldots, p_{t-K}) + \varepsilon_t$, where Ψ is some deterministic function and ε_t is i.i.d. according to some symmetric density over $[-\xi, \xi]$. Continue to assume $G(b) \to 1$. It can be shown that if ξ is small enough, there is a unique SPE outcome, which has the following properties. Existing workers are retained at t as long as $\varepsilon_t \geq b - (1 - b)\Psi(p_{t-1}, \ldots, p_{t-K})$, and paid $b[\Psi(p_{t-1}, \ldots, p_{t-K}) + \xi]$. Note that this means layoff rates at t are sensitive to p_{t-K}, even if we control for $p_t, \ldots, p_{t-K+1}$. In this sense, K-lagged productivity shocks have an effect on current inflows into unemployment, even after controlling for current productivity. Newly hired workers' equilibrium wage at t depends on the period-t outside option bp_t, as well as on the period-$(t + 1)$ retention threshold $b - (1 - b)\Psi(p_t, \ldots, p_{t+1-K})$. Hence, past productivity shocks will also have a persistent effect on the equilibrium wage of new hires. Recall that tightness at t is determined by the value of a created vacancy conditional on filling it at the beginning of $t + 1$. Since layoff rates at $t + 2$ depend on $(p_{t+1}, \ldots, p_{t-K+1})$, it follows that tightness at t is a function of $(p_t, \ldots, p_{t-K+1})$. Thus, when we depart from the purely multiplicative specification of our model, long-lasting effects of productivity shocks emerge.

Long-lasting effects of productivity shocks can also be generated by assuming instead that existing workers' reference wage at t is defined as the expected wage conditional on the history at the end of period $t - K$. The treatment of existing workers will then be sensitive to shocks at period $t - K$, even when we control for $p_t, \ldots, p_{t-K+1}$.

V. Concluding Remarks

Our objective in this paper was to formalize the idea that morale considerations affect the labor market's response to macroeconomic fluctuations, in the context of an S&M model. In our model, as in Akerlof (1982), workers' productivity is damaged when their wage falls below a reference point. Following Köszegi and Rabin (2006), we assumed that existing workers' reference point is a function of their lagged wage expectations. The equilibrium predictions of the model are that existing workers' wages display downward rigidity with respect to macroeconomic shocks, while entry-level wages are lower and more flexible, and market tightness is more volatile than in a reference-independent benchmark. The main open problem is to provide a complete characterization of SPE under general exogenous separation processes and arbitrary business-cycle magnitudes. Extending the model to other bargaining protocols is an additional interesting avenue for future research.

We believe that the model is capable of producing additional insights, some of which were suggested informally by Bewley (1999) on the basis of his survey. Here we make do with a brief description.

Part-time jobs. Suppose that a firm's hiring/retention decision is not binary, but a real number $r \in 0, 1]$, such that an interior r corresponds to a part-time job. Suppose further that wages are stated for full-time positions, such that an employed worker's total wage earnings are rw. It makes sense to assume that an existing worker's reference point will be based on w rather than on rw. This means that if a firm shifts its worker from full- to part-time employment without cutting w below its lagged-expected value, this will not be construed as unfair behavior, and the worker will produce normal output. It follows that after a bad productivity shock, a firm may prefer this option to the alternative of keeping the worker at full-time employment while lowering his wage, even when the latter option would have been optimal in a reference-independent model.

The role of inflation. Discussions of wage rigidity often involve a distinction between real and nominal wages and the mitigating role of inflation. In a model with reference dependence, this distinction is traced to an assumption as to whether the reference point is formed in nominal or real terms. If the reference point is stated in terms of (lagged-expected) nominal wages, it should come as no surprise that unexpected inflation can have real (yet temporary) effects on the labor market, by lowering the reference point in real terms, thereby making the MH constraint less likely to be binding.

We would like to conclude the paper with a discussion of alternative reference-point formation rules. In appendix B, we examine a close variation on our model, in which the reference point of workers of *any* type is equal to their lagged-expected wage *earnings*, thus endogenizing the distinction between newly matched and existing workers. The main qualitative results of our model are reproduced.

Another variant would abandon the lagged-expectation component, and assume that an existing worker's reference point at period t is equal to his *actual* wage at period $t - 1$. This alternative formulation could also generate persistent effects of productivity shocks: in a model with a long-horizon separation process, the treatment of existing workers would be sensitive to the value of p when they were originally hired by the firm, even after controlling for the values of p ever since.

Finally, the reference point that conditions the worker's effort decision could be a function of variables other than the worker's own (expected) wage. For instance, it could be the wage earned by his peers, or what he considers to be a fair share of his output. Analyzing the model under such alternative reference-point rules is important not only as a robustness check, but also because this may generate new insights into other aspects of labor relations over the business cycle.

Appendix A
Proofs

Let us first introduce some notation that will serve us in several proofs. Fix an SPE.

Unemployed workers' payoff. Recall that for a given firm-worker pair, the only observable aspect of the history prior to their match is the sequence of realizations of p. In particular, it does not matter whether the worker's unemployment at t is due to a matching failure, a firm's decision not to hire him, or his own decision to reject a wage offer. Therefore, we can denote an unemployed worker's equilibrium continuation payoff at t by $W_0(p_0, \ldots, p_t)$, without loss of generality.

Employed workers' payoff. Let h_t be the information set of a given firm-worker matched pair at period t, where the worker is of type i at t. Let (h_t, w_t) denote the immediate concatenation in which the firm hires/retains the worker and makes the offer w_t. Let $W_i(h_t, w_t)$ denote the worker's equilibrium continuation payoff at (h_t, w_t), where the subscript i clarifies the worker's type at t. Let $W_0(h_t)$ denote his reservation payoff at h_t, namely the continuation payoff if he rejects the wage offer that the firm makes and thus becomes unemployed at t. By definition, $W_i(h_t, w_t) \geq W_0(h_t)$.

Employed workers' rent. We define two types of rents. First, let $R_i(h_t, w_t) = W_i(h_t, w_t) - W_0(h_t)$ be the difference between the worker's equilibrium continuation payoff at (h_t, w_t) and his reservation payoff at this history. Second, let $B(p)$ denote a worker's continuation payoff from the strategy of rejecting all wage offers when the current state is p, and define $Q_i(h_t, w_t) = W_i(h_t, w_t) - B(p_t)$. By revealed preferences, $Q_i(h_t, w_t) \geq R_i(h_t, w_t) \geq 0$. In addition, $Q(\cdot)$ is bounded from above because firms will never make offers that generate negative profits.

Proof of Proposition 1

Consider some SPE of the game. Define Q^* as the maximum of $Q_i(h, w)$ over all histories (h, w) and agent types i in this SPE. In general, the maximum need not be well-defined, and complete rigor demands it to be replaced with the sup. However, this would complicate our analysis in a way we find superfluous. Thus, to simplify exposition, we deal with the case in which Q^* is well-defined and attained in some finite history (h_t^*, w_t^*) by a worker of some type i^*.

If $Q^* = 0$ we are done, and so assume that $Q^* > 0$. Note that $Q^* = w_t^* - bp_t + \delta Q_{i^*}(h_{t+1}, w_{t+1})$. Suppose that $w_t^* = 0$ and i^* accepts the wage offer so that the nonnegativity constraint is binding at (h_t^*, w_t^*). Since $w_t^* - bp_t < 0$ and $Q^* > 0$ we have that $Q_{i^*}(h_{t+1}, w_{t+1}) > Q^*$, a contradiction. Thus, the nonnegativity constraint of a wage offer to worker i^* must hold with slack at (h_t^*, w_t^*). Similarly, it cannot be the case that worker i^* rejects the wage offer w_t^*. It follows that the IR constraint of i^*'s contract (see definition at the beginning of section III) is binding at (h_t^*, w_t^*)—otherwise, the firm can slightly lower the worker's wage without changing his subsequent behavior (the reason is that the worker cannot benefit from rejecting the offer; if he does, he will join the search pool and by assumption, future employers will be unable to adjust their behavior on this rejection because they do not observe it).

By the definition of Q^*, $W_{i^*}(h_t^*, w_t^*) \geq W_1(h_t^*, w_t^*)$, and

$$W_1(h_t^*, p_{t+1}, w_{t+1}) - B(p_{t+1}) \leq W_{i^*}(h_t^*, w_t^*) - B(p_t), \tag{13}$$

for any realization of p_{t+1} and a wage offer w_{t+1} made to a newly matched worker at $t + 1$. Observe that

$$W_0(h_t^*) = bp_t + \delta[\mu_t \cdot \mathbb{E}W_1((h_t^*, p_{t+1}, w_{t+1})|p_t) + (1 - \mu_t) \cdot \mathbb{E}W_0((h_t^*, p_{t+1})|p_t)], \tag{14}$$

where μ_t is the probability that an unemployed worker at t finds a match. The determinants of μ_t are immaterial for our purposes. Since $W_0(h_t^*, p_{t+1}) \leq W_1(h_t^*, p_{t+1}, w_{t+1})$, we obtain from (14) that

$$W_0(h_t^*) \leq bp_t + \delta\mathbb{E}(W_1(h_t^*, p_{t+1}, w_{t+1})|p_t).$$

Because the IR constraint of i^*'s contract is binding at (h_t^*, w_t^*), $W_{i^*}(h_t^*, w_t^*) = W_0(h_t^*)$. Using (13) we may therefore conclude that

$$W_{i^*}(h_t^*, w_t^*) = W_0(h_t^*) \leq bp_t + \delta W_{i^*}(h_t^*, w_t^*) + \delta\mathbb{E}B(p_{t+1}|p_t) - \delta B(p_t).$$

Since $bp_t + \delta\mathbb{E}B(p_{t+1}|p_t) = B(p_t)$, we have $W_{i^*}(h_t^*, w_t^*) \leq B(p_t)$, hence $Q^* = 0$.

By the definition of Q^*, it follows that for any worker type i and any (h_t, w_t) along the equilibrium path, $W_i(h_t, w_t) = B(p_t)$. Thus, if the worker accepts the wage offer, we have

$$W_i(h_t) = w_t + \delta\mathbb{E}(W_{i+1}(h_t, p_{t+1}, w_{t+1})|p_t) = bp_t + \delta\mathbb{E}B(p_{t+1}|p_t),$$

and this implies $w_t = bp_t$. Finally, there cannot be a SPE in which a worker rejects an offer of bp_t at some period t because the firm could profitably deviate by slightly raising the wage.

Proof of Proposition 2

We first prove a pair of lemmas that will serve us in several proofs. In particular, they hold for any γ. Define Q^{**} as the maximum of $Q(h, w)$ over all histories (h, w) in which a *newly matched worker* responds to a wage offer.

Lemma 2 *Let* (h_t, w_t) *be a history in which a newly matched worker responds to a wage offer, for which* $Q(h_t, w_t) = Q^{**}$. *If the IR constraint is binding at* (h_t, w_t), *then* $Q^{**} = 0$.

Proof. By the definition of Q^*, $W_1(h_t, p_{t+1}, w_{t+1}) - B(p_{t+1}) \leq W_1(h_t, w_t) - B(p_t)$. The proof that $Q^{**} = 0$ reproduces exactly the same steps that led us to conclude that $Q^* = 0$ in the proof of Proposition 1. (Note that here we simply assume that IR is binding at (h_t, w_t), rather than deriving this property.) ■

Let $\bar{w}_i^t$ denote the participation wage of a worker of tenure $i = 1, 2$ at period t (implicitly, given the history)—that is, the lowest wage offer they will accept given that all agents conform to their equilibrium continuation strategies. The next lemma shows that a worker's equilibrium wage at any history cannot exceed the highest participation wage he could get given the previous-period history.

Lemma 3 *In SPE,* $w_{i,t} \leq \max \bar{w}_i^t | h_{t-1}$.

Proof. Assume the contrary, that is, $e_{2,t} > \max \bar{w}_i^t | h_{t-1}$. Since $e_{2,t}$ is a weighted average of $e_{2,t}$ and realizations of $\bar{w}_i^t$ that are feasible given h_{t-1}, it is equal to $e_{2,t}$ only if the firm pays $w_{2,t} = e_{2,t}$ with probability one, conditional on retaining the worker. However, since $e_{2,t} > \bar{w}_i^t$ with probability one, there exists a value of γ_t sufficiently close to one, such that for any h_t, $p_t - e_{2,t} < \gamma_t p_t - \bar{w}_i^t$, in which case the firm can profitably deviate from $w_{2,t} = e_{2,t}$ to $\bar{w}_i^t$. By assumption, such realizations of γ_t occur with positive probability, a contradiction. ■

Lemma 4 *In SPE,* $w_{1,t} > 0$ *at any period* t.

Proof. Recall that W_0^t is independent of the worker's type at t, and that R_i^t is the rent (i.e., excess payoff above his reservation payoff) that a worker of type i gets at period t. If the worker is unemployed at t, we write $R_i^t = 0$. The following equations hold, by the definition of these objects:

$$\bar{w}_2^t + \delta \mathbb{E}(W_0^{t+1} | h_t) = W_0^t$$

$$\bar{w}_1^t + \delta \mathbb{E}(W_0^{t+1} | h_t) + \delta \mathbb{E}(R_2^{t+1} | h_t) = W_0^t.$$

Therefore,

$$\bar{w}_1^t = \bar{w}_2^t - \delta \mathbb{E}(R_2^{t+1}|h_t). \tag{15}$$

Moreover, since

$$W_0^t = bp_t + \delta \mathbb{E}(W_0^{t+1}|h_t) + \delta \mu_t \mathbb{E}(R_1^{t+1}|h_t),$$

we obtain

$$\bar{w}_2^t = bp_t + \delta \mu_t \mathbb{E}(R_1^{t+1}|h_t) \tag{16}$$

$$\bar{w}_1^t = bp_t + \delta \mu_t \mathbb{E}(R_1^{t+1}|h_t) - \delta \mathbb{E}(R_2^{t+1}|h_t). \tag{17}$$

If the IR constraint of a wage offer to a newly matched worker is binding at t, then his period-t wage is equal to his period t reservation wage and $R_1^t = 0$. If his MH constraint is binding at t, then the actual wage at t is zero, and $R_1^t = -\bar{w}_1^t$. If $\bar{w}_1^t < 0$ ($\bar{w}_1^t > 0$), then the MH (IR) constraint is binding. Therefore, $R_1^t = \max\{0, -\bar{w}_1^t\}$.

Let R^* and R_* denote the maximum and minimum values that R_1^t can attain at any t. By definition, $R_* \geq 0$. Assume that $R^* > 0$. Let w_* denote the minimum value that $\bar{w}_1^t$ may obtain at any t. Then $R^* = -w_*$, where $w_* < 0$. From (17) it follows that

$$\bar{w}_1^t = bp_t + \delta \mu_t \mathbb{E}(R_1^{t+1}|h_t) - \delta \mathbb{E}(R_2^{t+1}|h_t).$$

Observe that $R_2^{t+1}|p_t = w_{2,t+1} - \bar{w}_2^{t+1}$. By lemma 3, $w_{2,t+1} \leq \max \bar{w}_2^{t+1}|h_t$. Therefore, $\delta \mathbb{E}(R_2^{t+1}|h_t)$ is smaller or equal to the sum

$$\delta[\max_{h_{t+1}|h_t}(bp_{t+1} + \delta \mu_{t+1} \mathbb{E}(R_1^{t+2}|h_{t+1})) - \mathbb{E}(bp_{t+1} + \delta \mu_{t+1} \mathbb{E}(R_1^{t+2}|h_{t+1})|h_t)],$$

which in turn is lower or equal to

$$\delta b(p_t^\beta \xi - p_t^\beta \mathbb{E}(\varepsilon)) + \delta^2 \max_{h_{t+1}|h_t}[\mu_{t+1} \mathbb{E}(R_1^{t+2}|h_{t+1})] - \delta^2 \mathbb{E}[\mu_{t+1} \mathbb{E}(R_1^{t+2}|h_{t+1})|h_t].$$

Note that

$$\max_{h_{t+1}|h_t}[\mu_{t+1} \mathbb{E}(R_1^{t+2}|h_{t+1})] \leq R^*$$

$$\mathbb{E}[\mu_{t+1} \mathbb{E}(R_1^{t+2}|h_{t+1})|h_t] \geq 0$$

$$\mu_t \mathbb{E}(R_1^{t+1}|h_t) \geq 0$$

$$p_t^\beta \xi - p_t^\beta \mathbb{E}(\varepsilon) \leq p_t^\beta(\xi - 1).$$

Hence, for any t,

$$\bar{w}_1^t \geq b[p_t - \delta p_t^\beta(\xi - 1)] - \delta^2 R^*. \tag{18}$$

Since $R^* = -w^*$, inequality (18) holds for every t only if it holds at the lowest possible value of $\bar{w}_1^t$, that is, only if

$$w_* \geq b[p_t - \delta p_t^{\beta}(\xi - 1)] - \delta^2(-w_*),$$

which implies

$$w_* \geq \frac{b[p_t - \delta p_t^{\beta}(\xi - 1)]}{1 - \delta^2}.$$

Recall that p_t follows a log-linear AR(1) process where shocks take values in $[1/\xi, \xi]$. A simple calculation shows that since $\xi < (1/2)(1 + \sqrt{5})$, the numerator of the R.H.S. is strictly positive. But this contradicts our assumption that $w_* < 0$. It follows that $R^* = 0$, and this establishes the result. ■

The rest of the proof proceeds in two steps. First, we use the above lemmas to derive the retention decision, reference point, and equilibrium wages for existing workers. Second, we compute the hiring decision and equilibrium wages for newly matched workers. Since by assumption $e_{1,t} = 0$, lemma 4 implies that the MH constraint of a wage offer to newly matched workers holds with slack after every history. Therefore, their IR constraint must be binding after every history.

Step 1: Existing Workers

Let us first show that an existing worker at period t will accept a wage offer $w_{2,t}$ if and only if $w_{2,t} \geq bp_t$. This is his last period of employment. If he rejects the firm's offer, he will be unemployed and earn a payoff of bp_t at t. We have seen that newly matched workers' IR is binding after every history. By lemma 2, it follows that the worker's equilibrium continuation payoff from period $t + 1$ onwards is the same as if he were to receive bp_s in every period $s \geq t + 1$. Therefore, the existing worker's participation constraint at t will be binding if he receives a payoff of bp_t.

It follows that if $bp_t \geq e_{2,t}$, the firm will choose $r_t = 1$ and $w_t = bp_t$ in equilibrium. Let us turn to the case of $bp_t < e_{2,t}$. Conditional on playing $r_t = 1$, the firm will offer $w_t \in \{e_{2,t}, bp_t\}$ because IR or MH are binding. Retaining the worker at $w_t = bp_t$ generates a profit of $\pi = \gamma_t p_t - bp_t$. If $\gamma_t < b$ ($\gamma_t > b$), then $\pi < 0$ ($\pi > 0$); and since this is the last period of the worker's employment, the firm will choose $r_t = 0$ ($r_t = 1$). It follows that when $\gamma_t < b$, the firm will play $r_t = 0$ if $p_t - e_{2,t} < 0$ and $r_t = 1$, $w_t = e_{2,t}$ if $p_t - e_{2,t} > 0$. And when $\gamma_t > b$, the firm will play $r_t = 1$, and $w_t = e_{2,t}$ (or $w_t = bp_t$) if $p_t - e_{2,t} > \gamma_t p_t - bp_t$ (if $p_t - e_{2,t} < \gamma_t p_t - bp_t$).

We are now able to provide an expression for existing workers' reference wage at period t, which is equal to their expected wage conditional on being retained, according to their information at the end of period $t-1$. We use the abbreviated notation $e = e_{2,t}$, $p = p_{t-1}$:

$$e = \frac{G(b)\int_{\varepsilon>\varepsilon^*} \max\{e, bp^{\beta}\varepsilon\}dF(\varepsilon) + \int_b^1 \left[\int_{\varepsilon<\varepsilon^{**}} bp^{\beta}\varepsilon dF(\varepsilon) + \int_{\varepsilon>\varepsilon^{**}} \max\{e, bp^{\beta}\varepsilon\}dF(\varepsilon)\right] dG(\gamma)}{G(b)(1 - F(\varepsilon^*)) + 1 - G(b)}, \quad (19)$$

where the productivity shock cutoffs ε^* and ε^{**} are given as follows:

$$p^{\beta}\varepsilon^* = e$$

$$p^{\beta}\varepsilon^{**} - e = \gamma p^{\beta}\varepsilon^{**} - bp^{\beta}\varepsilon^{**}.$$

It is clear from (19) that $e_{2,t} \geq bp_{t-1}^{\beta}\mathbb{E}(\varepsilon)$. By lemma 3, $e_{2,t} \leq b\xi p_{t-1}^{\beta}$, namely the highest outside option that is feasible at period t given p_{t-1}. Our task now is to establish that the equation (19) has a unique solution e in the interval $[bp_{t-1}^{\beta}\mathbb{E}(\varepsilon), b\xi p_{t-1}^{\beta}]$. Rewrite the equation as $eB(e) - A(e) = 0$, where the functions A and B are the numerator and denominator of the R.H.S. of (19), respectively. The L.H.S. of this equation is a continuous function of e. Moreover, it is negative for $e = 0$ and positive for $e > b\xi p_{t-1}^{\beta}$. Differentiating w.r.t. e, we obtain $[eB(e) - A(e)]' > 0$ for all e in the relevant domain. Therefore, (19) has a unique solution. Let us guess that the solution has the form $\phi \cdot bp_{t-1}^{\beta}$, where ϕ is a constant that is a function of F, G, and b. Plugging this expression into (19) and simplifying, we obtain (4) through (5). In particular, $\varepsilon^* = \varepsilon^*(\gamma)$ for $\gamma < b$, and $\varepsilon^{**} = \varepsilon^*(\gamma)$ for $\gamma > b$. This system has a solution, by the same reasoning that ensured a solution for (19). Therefore, this solution gives us the unique solution for (19). We have thus fully characterized the equilibrium retention and wage policies for existing workers.

Step 2: Newly Matched Workers

A newly matched worker at period t expects to earn the discounted sum of payoffs in periods t and $t+1$:

$$w_{1,t} + \delta\mathbb{E}[r_{2,t+1}w_{2,t+1} + (1 - r_{2,t+1})bp_{t+1}|p_t]. \quad (20)$$

We have already noted that a new worker's SPE continuation payoff is as if he receives bp_t in every period t. Hence, in any SPE, the expected, discounted sum in (20) must equal $bp_t + \delta\mathbb{E}(bp_{t+1}|p_t)$. Expression (7) for $w_{1,t}$ thus follows from our characterization of $r_{2,t+1}$ and $w_{2,t+1}$. To see why $r_{1,t} = 1$ regardless of the history, note that in the second period of the interaction between the firm and the worker, the firm necessarily earns nonnegative profits. The newly matched worker at t produces the normal output p_t because as we saw, his MH constraint holds (with slack). Since he is paid at most bp_t, the firm earns strictly positive profits, and therefore would always prefer to hire the worker.

Appendix B
Endogenous Distinction between Worker Types

Thus far, we have assumed a reference-point formation rule that imposed an exogenous distinction between newly matched and existing workers. One could argue that there are endogenous reasons for such a distinction. In particular, they have different employment prospects: the probability that a newly hired worker at t is employed at $t + 1$ is a function of his employer's equilibrium retention policy, while the probability that an unemployed worker at t is employed at $t + 1$ is a function of market tightness at t and firms' hiring policy.

In this appendix, we modify the reference-point formation rule in order to capture this consideration and endogenize the distinction between the reference points of workers of different types. For simplicity, we restrict attention to the two-period exogenous separation process. Assume that at any period t and for any worker type $i = 1, 2$, the worker's reference point is equal to his *expected wage earnings* conditional on his information at the end of period $t - 1$. Specifically, the period-t reference points for newly matched and existing workers are

$$\begin{aligned} e_{1,t} &= \mu_{t-1} \cdot \mathbb{E}(r_{1,t} w_{1,t}) \\ e_{2,t} &= \mathbb{E}(r_{2,t} w_{2,t}), \end{aligned}$$

where the expectation over $r_{i,t}$, $w_{i,t}$ is conditional on the worker's information set at the end of $t - 1$.

This reference point formation rule puts newly matched and existing workers on the same footing a priori. However, their different employment prospects translate into different reference points. In particular, if an unemployed worker at period $t - 1$ faces a low match probability μ_{t-1}, his reference wage if matched at the beginning of t is close to zero.

Our main result in this appendix is that when the matching friction is sufficiently high and the magnitude of the business cycle is not too large, there exists an SPE that mimics the qualitative features of the unique SPE obtained in section III.

Proposition 3 *If*

$$m(1,1) < \min\left\{\frac{c}{\Pi^0(\sqrt[1-\beta]{\xi})}, 1 - \xi + \frac{1}{\xi}\right\}, \tag{21}$$

the game has a SPE with the following properties.

(i) An existing worker's period-t reference point $e_{2,t}$ is

$$e_{2,t} = \phi \cdot bp_{t-1}^{\beta}$$

where ϕ is uniquely given by

$$\phi = \int_0^1\int_{\varepsilon > \varepsilon^*(\gamma)} \max\{\phi, \varepsilon\} dF(\varepsilon)dG(\gamma) + \int_b^1\int_{\varepsilon > \varepsilon^*(\gamma)} \varepsilon dF(\varepsilon)dG(\gamma)$$

$$\varepsilon^*(\gamma) = \frac{\phi b}{1 - \max\{0, \gamma - b\}}.$$

(ii) An existing worker is dismissed at period t if and only if $\gamma_t < b$ and $\varepsilon_t < \phi b$. Conditional on being retained at t, his wage is

$$w_2(p_{t-1}, p_t) = \begin{cases} \max\{e_{2,t}, bp_t\} & \text{if } \varepsilon_t > \varepsilon^*(\gamma_t) \\ bp_t & \text{if } \gamma_t > b \text{ and } \varepsilon_t < \varepsilon^*(\gamma_t) \end{cases} \tag{22}$$

(iii) A newly matched worker at period t is always hired; his wage at period t is given by (7).

Proof. Our method of proof is as follows. First, we construct a unique SPE under the assumption that $\bar{w}_{1,t} > e_{1,t}$ at any period t, regardless of the history—that is, newly matched workers' participation wage exceeds their reference wage. Then, we show that this assumption holds under (21). Many of the steps in the proof have analogues in the proof of proposition 2, and are therefore described briefly.

Step 1: Existing Workers

By assumption, newly matched workers' IR constraint is binding after every history in equilibrium. Therefore, existing workers' participation wage at any period t is exactly the same as in the basic model, namely bp_t. For a given reference wage $e_{2,t}$, the firm's retention and wage policy in SPE is the same as in the basic model. Specifically, when $\gamma_t < b$, the firm will retain an existing worker at period t if and only if $p_t \geq e_{2,t}$, and pay $w_{2,t} = \max\{e_{2,t}, bp_t\}$ conditional on retention. And if $\gamma_t > b$, the firm will always retain an existing worker, and pay him $w_{2,t} = e_{2,t}$ when $p_t - \gamma_t p_t \geq e_{2,t} - bp_t \geq 0$, and $w_{2,t} = bp_t$ otherwise. Therefore, an existing worker's reference wage at period t is given by the following equation:

$$e = G(b)\int_{\varepsilon>\varepsilon^*} \max\{e, bp^\beta\varepsilon\}dF(\varepsilon)$$
$$+ \int_b^1 \left[\int_{\varepsilon<\varepsilon^{**}} bp^\beta \varepsilon dF(\varepsilon) + \int_{\varepsilon>\varepsilon^{**}} \max\{e, bp^\beta\varepsilon\}dF(\varepsilon)\right] dG(\gamma)'$$

where the productivity shock cutoffs ε^* and ε^{**} are given as follows:

$$p^\beta\varepsilon^* = e$$

$$p^\beta\varepsilon^{**} - e = \gamma p^\beta\varepsilon^{**} - bp^\beta\varepsilon^{**}.$$

To establish existence of a solution to this equation, note first that the R.H.S. is a continuous function of $e_{2,t}$. Second, the R.H.S. cannot take values above $b\xi p^\beta$, hence we can view the R.H.S. as a continuous mapping from $[0, b\xi p^\beta]$ to itself. By Brouwer's fixed-point theorem, this mapping has a fixed point. To see that this fixed point is unique, differentiate both sides of the equation w.r.t. e. The derivative of the L.H.S. w.r.t. e is 1, while the derivative of the R.H.S. w.r.t. e is strictly below 1. Therefore, there can be at most one point in which the functions on the two sides of the equation intersect, hence precisely one fixed point.

Step 2. Newly Matched Workers

The derivation is exactly the same as in the basic model.

Step 3: Verifying That Newly Matched Workers' MH Holds with Slack

By the expression for newly matched workers' wage,

$$w_{1,t} > b[p_{t-1} - p_{t-1}^\beta(\xi - \mathbb{E}(\varepsilon))].$$

On the other hand, by the same expression and the definition of newly matched workers' reference point,

$$e_{1,t} \leq \mu_{t-1} \cdot bp_{t-1}^\beta \mathbb{E}(\varepsilon).$$

In order to prove the result, it suffices to show that the lower bound on $w_{1,t}$ is always higher than the upper bound on $e_{1,t}$. A bit of algebra gives us the following sufficient condition (using the facts that $\mathbb{E}(\varepsilon) > 1$ and $p^{1-\beta} \geq 1/\xi$):

$$\mu_{t-1} < 1 - \xi + \frac{1}{\xi}.$$

The highest value that p_t can get is $\sqrt[1-\beta]{\xi}$. By the free entry assumption, the following inequality holds in any equilibrium:

$$q_{t-1} \geq \frac{c}{\Pi^0(\sqrt[1-\beta]{\xi})}.$$

By the assumption that m satisfies constant returns to scale, $q_{t-1} > m(1,1)$ if and only if $\mu_{t-1} < m(1,1)$. Therefore, if $m(1,1)$ satisfies condition (21), newly matched workers' participation wage exceeds their reference wage after every history. ■

Because the exact equation that describes the reference point for existing workers differs from its specification in the basic model, the SPE constructed here does not exactly replicate the SPE in the basic model. However, the qualitative features of firms' retention and wage policies for newly hired and existing workers are preserved. Note that the restriction on ξ required to obtain this result is more severe than in the basic model, but this difference vanishes as $m(1,1)$ gets closer to zero. Given the specification of ϕ, we are able to obtain the output constant λ just as in the basic model, and use it to replicate (qualitatively) the tightness volatility effect.

Appendix C
Reference-Dependent Worker Preferences

Reference dependence of output in our model is interpreted in terms of worker motivation. This suggests that our model may be viewed as a reduced form of a larger model in which workers' *preferences* are reference-dependent. Indeed, Akerlof (1982) formulated his model of the labor relation in terms of reference-dependent worker preferences that dictate their choice of unobserved effort, such that when their wage falls below the reference point, their subjective cost of effort increases. In this appendix we construct such a model, which can be viewed as a foundation for the reduced form, reference-dependent output function assumed in the main text.

The search, matching, separation, and bargaining components, as well as the firms' preferences, are exactly the same as in the basic model. The only differences are in the description of workers' output and their preferences. Suppose that conditional on accepting an offer, an employed worker is committed to a minimal level of effort. On top of that, he chooses a level of discretionary effort $x_t \in \{0, 1\}$. We refer to $x = 1$ as "normal effort." This effort decision is not observed by the firm. The worker's output is $y_t = p_t[\gamma_t + (1 - \gamma_t)x_t]$. Under this formulation, γ_t is interpreted as an indicator of the *completeness of the labor contract*, such that $1 - \gamma_t$ captures the importance of discretionary effort in the output function.

Workers maximize expected discounted payoffs. Employed workers' payoff flow is modified as follows:

$$w_t - x_t \cdot \mathbf{1}[w_t < e_t], \tag{23}$$

whereas the basic model assumed only the first term. The interpretation is that when the worker's wage is below his reference point, he perceives this as unfair treatment; his intrinsic motivation is damaged, and he strictly prefers not to exert his normal effort. Otherwise, the worker is indifferent between $x = 0$ and $x = 1$, and we assume that he chooses the latter.

Given the assumption that the worker's discretionary effort is unobserved, the worker's choice of x is entirely myopic in any SPE. At any period t in which he accepts a wage offer w_t, he will play $x_t = 1$ if and only if $w_t \geq e_t$. As a result, the worker will respond to wage offers as if he maximizes the discounted sum of expected wage and nonmarket earnings, just as in the basic model. This is the reason that this larger model collapses to our basic model in equilibrium.

The value of recasting our model in terms of reference-dependent preferences is that it clarifies the interpretation of the random parameter γ_t. It also suggests new extensions and raises interpretational questions. We discuss some of them in the following.

The equilibrium concept. Because workers' preferences in this extended model depend on their expectations (both of the moves of nature and of the players' strategies), this is not, strictly speaking, a conventional game, but rather an example of an extensive-form "psychological game" (after Geanakoplos, Pearce, and Stacchetti 1989).[14] In general, extending standard game-theoretic solution concepts to this class of games may involve subtleties. However, in the present case, the standard concept of subgame perfect equilibrium (SPE) is defined and analyzed in a completely standard way.[15]

Contractual incompleteness. In our model, firms offer flat-wage contracts and do not observe workers' effort. The latter assumption may appear strange, because we assume that firms observe p_t, γ_t, hence assuming that x_t is unobservable is tantamount to assuming that output is unobserved, which may seem odd. However, recall that although the model is presented in terms of one-to-one matching, this assumption is purely expositional and the entire analysis is valid for one-to-many matching where production is separable across vacancies. It is entirely realistic to assume that while the firm can only observe its aggregate output with some noise, it cannot monitor the contribution of any individual worker.

Even under this limited monitoring, one could argue that flat-wage contracts are too restrictive, and that firms could incentivize effort by conditioning the workers' compensation on the noisy signal, namely aggregate output. However, as the literature on moral hazard in teams has demonstrated (starting with Holmstrom 1982), such incentives are limited in their ability to induce team effort. When these considerations are combined with reference-dependent worker preferences that exhibit loss aversion, the limitations are exacerbated and may lead firms to choose flat contracts (see Herweg, Müller, and Weinschenk 2010; Herweg and Mierendorff 2013).

Furthermore, incentivizing team performance may exacerbate morale problems for reasons that are not captured by our preference model, because it punishes individual workers for a drop in output that is due to chance or other workers' effort decisions. Similar issues arise when the worker has multiple tasks and the firm can only monitor a subset of those (see Fehr, Goette, and Zehnder 2009). Thus, morale considerations and limited monitoring of workers' effort complement each other in motivating firms to prefer flat-wage contracts to elaborate incentive schemes.

Endnotes

Spiegler acknowledges financial support from the European Research Council, Grant no. 230251. We thank Gadi Bar-Levy, Martin Dufwenberg, Erik Eyster, Robert Hall, Elhanan Helpman, Giuseppe Moscarini, Christopher Pissarides, Thijs van Rens, Ariel Rubinstein, Ron Siegel, Balazs Szentes, Eran Yashiv, Michael Woodford, Alwyn Young, and numerous seminar participants for helpful comments. For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see http://www.nber.org/chapters/c12938.ack.

1. The notion of "morale hazard" has a completely different meaning in the insurance industry.

2. See Crawford and Meng (2011) for an empirical implementation of the "lagged expectations" approach to reference dependence.

3. Kuang and Wang (2010) conduct a quantitative analysis of an S&M model with a reduced-form fair-wage equation, which includes past wages as some of the independent variables. Dufwenberg and Kirchstegier (2000) study a static model with one firm and

two workers, in which firms refrain from exploiting competition between workers to cut wages due to reciprocal-fairness considerations.

4. For expositional simplicity, we conventionally assume that each firm can post at most one vacancy.

5. See Odean (1998), Haigh and List (2005), and Pope and Schweitzer (2011) for other instances.

6. As Fehr and Gächter (2000, p. 160) emphasize, "Reciprocity is fundamentally different from 'retaliatory' behavior in repeated interactions. These behaviors arise because actors expect future material benefits from their actions; in the case of reciprocity, the actor is responding to a negative action even if no material gains can be expected."

7. We have been able to derive analytic solutions for a stationary infinite-horizon separation process when $\beta = 0$ or $\beta = 1$. However, since these cases are degenerate in terms of the effect of wage rigidity on tightness volatility, we chose not to include this analysis in the paper.

8. We ignore the firm's behavior at zero-probability cutoff events.

9. We are grateful to Giueseppe Moscarini for suggesting this point.

10. In principle, tax incidence might matter through framing effects on the reference point.

11. In contrast, the version of the model with commitment discussed in section III, subsection C *is* vulnerable to the Barro critique because it implies inefficient output destruction that can be avoided by a renegotiated *incomplete* contract.

12. Hall and Milgrom (2008) argue that the partially flexible wages they obtain from the alternating-offers bargaining model (without distinguishing between new and existing workers) can resolve the Shimer puzzle quantitatively.

13. In a recent paper, Rudanko (2011) assumes that the employer is also risk-averse but has better access to capital markets than the employee. She then shows that the equilibrium generates higher tightness volatility compared to a benchmark in which employees can use the capital market to smooth their consumption. Recall that our model abstracts from consumption, thus implicitly assuming that workers spend their wage earnings instantaneously.

14. Rabin (1993) was the first to use the framework of psychological games to model reciprocity considerations.

15. This is due to the fact that in our model, workers incorporate reciprocity considerations into their effort decision in a myopic way. Dufwenberg and Kirchsteiger (2004) and Battigalli and Dufwenberg (2009) develop tools to deal with more complicated dynamic settings, where reciprocity considerations may be sensitive to off-equilibrium events and higher-order beliefs.

References

Akerlof, George A. 1982. "Labor Contracts As Partial Gift Exchange." *Quarterly Journal of Economics* 97:543–69.

Akerlof, George A., and Janet L. Yellen. 1990. "The Fair-Wage Effort Hypothesis and Unemployment." *Quarterly Journal of Economics* 195:255–84.

Anderson, Eric T., and Duncan I. Simester. 2010. "Price Stickiness and Customer Antagonism." *Quarterly Journal of Economics* 125:729–65.

Azariadis, Costas. 1975. "Implicit Contracts and Underemployment Equilibria." *Journal of Political Economy* 83:1183–202.

Barro, R. 1977. "Long-Term Contracting, Sticky Prices, and Monetary Policy." *Journal of Monetary Economics* 3:305–16.

Battigalli, Pierpaolo, and Martin Dufwenberg. 2009. "Dynamic Psychological Games." *Journal of Economic Theory* 144:1–35.

Beaudry, Paul, and John DiNardo. 1991. "The Effect of Implicit Contracts on the

Movement of Wages over the Business Cycle: Evidence from Micro Data." *Journal of Political Economy* 99:665–88.

Benartzi, Shlomo, and Richard H. Thaler. 1995. "Myopic Loss Aversion and the Equity Premium Puzzle." *Quarterly Journal of Economics* 110:73–92.

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

Blinder, Alan S., and Don H. Choi. 1990. "A Shred of Evidence on Theories of Wage Stickiness." *Quarterly Journal of Economics* 105:1003–15.

Costain, James, and Marcel Jansen. 2010. "Employment Fluctuations with Downward Wage Rigidity: The Role of Moral Hazard." *Scandinavian Journal of Economics* 112:782–811.

Crawford, Vincent, and Juanjuan Meng. 2011. "New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income." *American Economic Review* 101:1912–32.

de Meza, David, and David C. Webb. 2007. "Incentive Design under Loss Aversion." *Journal of the European Economic Association* 5:66–92.

Dittmann, Ingolf, Ernst Maug, and Oliver Spalt. 2010. "Sticks or Carrots? Optimal CEO Compensation When Managers Are Loss Averse." *Journal of Finance* 65:2015–50.

Dufwenberg, Martin, and Georg Kirchsteiger. 2000. "Reciprocity and Wage Undercutting." *European Economic Review* 44:1069–78.

———. 2004. "A Theory of Sequential Reciprocity." *Games and Economic Behavior* 47:268–98.

Eliaz, Kfir, and Ran Spiegler. 2012. "Reference Dependence and Labor-Market Fluctuations." Center for Economic and Policy Research (CEPR) Discussion Paper 8997.

Fehr, Ernst, and Armin Falk. 1999. "Wage Rigidity in a Competitive Incomplete Contract Market." *Journal of Political Economy* 107:106–34.

Fehr, Ernst, and Simon Gächter. 2000. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* 14:159–81.

Fehr, Ernst, Lorenz Goette, and Christian Zehnder. 2009. "A Behavioral Account of the Labor Market: The Role of Fairness Concerns." *Annual Review of Economics* 1:355–84.

Fehr, Ernst, and Klaus M. Schmidt. 2002. "Theories of Fairness and Reciprocity—Evidence and Economic Applications." In *Advances in Economics and Econometrics—8th World Congress, Econometric Society Monographs*, edited by M. Dewatripont, L. Hansen, and St. Turnovsky, 208–57. Cambridge: Cambridge University Press.

Geanakoplos, John, David Pearce, and Ennio Stacchetti. 1989. "Psychological Games and Sequential Rationality." *Games and Economic Behavior* 1:60–79.

Genesove, David, and Christopher Mayer. 2001. "Loss Aversion and Seller Behavior: Evidence from the Housing Market." *Quarterly Journal of Economics* 116:1233–60.

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

Haefke, Christian, Marcus Sonntag, and Thijs van Rens. 2008. "Wage Rigidity and Job Creation." Discussion Paper 3714, IZA.

Hagedorn, M., and I. Manovskii. 2008. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited." *American Economic Review* 98:1692–706.

Haigh, Michael S. and John A. List. 2005. "Do Professional Traders Exhibit Loss Aversion? An Experimental Analysis." *Journal of Finance* 60: 523–34.
Hall, Robert E. 2005. "Employment Fluctuations with Equilibrium Wage Stickiness." *American Economic Review* 95:50–65.
Hall, Robert E., and Paul R. Milgrom. 2008. "The Limited Influence of Unemployment on the Wage Bargain." *American Economic Review* 98:1653–674.
Hart, Oliver, and John Moore. 2008. "Contracts As Reference Points." *Quarterly Journal of Economics* 123:1–48.
Herweg, Fabian, and Konrad Mierendorff. 2013. "Uncertain Demand, Consumer Loss Aversion, and Flat-Rate Tariffs." *Journal of the European Economic Association* 11 (2): 399–432.
Herweg, Fabian, Daniel Müller, and Philipp Weinschenk. 2010. "Binary Payment Schemes: Moral Hazard and Loss Aversion." *American Economic Review* 100:2451–77.
Holmstrom, Bengt. 1982. "Moral Hazard in Teams." *Bell Journal of Economics* 13:324–40.
Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler. 1986. "Fairness As a Constraint on Profit Seeking: Entitlements in the Market." *American Economic Review* 76:728–41.
Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47:263–92.
Kennan, John. 2009. "Private Information, Wage Bargaining and Employment Fluctuations." *Review of Economic Studies* 77:633–64.
Keynes, John M. 1936. *The General Theory of Employment, Interest and Money*. Basingstoke: Palgrave Macmillan.
Köszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics* 121:1133–65.
Kuang, Pei, and Tong Wang. 2010. "Labor Market Dynamics with Searching Friction and Fair Wage Considerations." Discussion Paper, University of Oxford.
Kudlyak, Marianna. 2009. "The Cyclicality of the User Cost of Labor with Search and Matching." Richmond Fed Working Paper 09–12.
Malcomson, James M., and Sophocles Mavroeidis. 2010. "Nash Bargaining, Credible Bargaining and Efficiency Wages in a Matching Model for the US." Economics Series Working Paper 511, University of Oxford.
Mortensen, Dale T. 1989. "The Persistence and Indeterminacy of Unemployment in Search Equilibrium." *The Scandinavian Journal of Economics* 91: 346–70.
Mortensen, Dale T., and Christopher Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *Review of Economic Studies* 61:397–415.
Odean, T. 1998. "Are Investors Reluctant to Realize Their Losses?" *Journal of Finance* 53:1775–98.
Pissarides, Christopher A. 2000. *Equilibrium Unemployment Theory*. Cambridge, MA: MIT Press.
———. 2009. "The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer?" *Econometrica* 77:1339–69.
Pope, Devin G. and Maurice E. Schweitzer. 2011. "Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes." *American Economic Review* 101: 129–57.
Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics." *American Economic Review* 83:1281–302.

Rubinstein, Ariel, and Asher Wolinsky. 1985. "Equilibrium in a Market with Sequential Bargaining." *Econometrica* 53:1133–50.

Rudanko, Leena. 2011. "Aggregate and Idiosyncratic Risk in a Frictional Labor Market." *American Economic Review* 101:2823–43.

Shapiro, Carl, and Joseph E. Stiglitz. 1984. "Equilibrium Unemployment As a Worker Discipline Device." *American Economic Review* 74:433–44.

Shimer, Robert. 2004. "The Consequences of Rigid Wages in Search Models." *Journal of the European Economic Association (Papers and Proceedings)* 2: 469–79.

———. 2005. "The Cyclical Behavior of Equilibrium Unemployment and Vacancies." *American Economic Review* 95:25–49.

———. 2010. *Labor Markets and Business Cycles*. Princeton, NJ: Princeton University Press.

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