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: Comment on "Reference Dependence and Labor Market Fluctuations"

Chapter Author(s): Giuseppe Moscarini

Chapter URL: http://www.nber.org/chapters/c12940

Chapter pages in book: (p. 209 – 222)

Comment

Giuseppe Moscarini, *Yale University and NBER*

A perennial question in macroeconomics, going back to Keynes and now more salient than ever, is why real wages fail to decline and to clear the labor market in the presence of high and persistent unemployment. Answers to this question come from four different directions.

First, from an empirical viewpoint, the main focus has been on measuring how cyclical and flexible wage rates are, both in nominal and real terms, using real-world observations from either survey or administrative data. This measurement requires controlling for the dependence of wage rates on observable and, when possible, unobservable worker and firm characteristics, in order to isolate the purely cyclical component. Part of the literature investigated the different, yet ultimately closely related question whether wage rates respond to firm-specific shocks. Many potential sources of wage rigidity apply to any kind of profitability shock that hits the firm, be it demand- or supply-driven, idiosyncratic, or aggregate. Second, the theoretical literature identified and analyzed a wide variety of potential sources of wage rigidity, originating from strategic/commitment considerations (efficiency wages, bargaining), adverse selection, and insurance motives. Third, experimental studies directly tested some of these hypotheses in the laboratory, to enjoy the benefits of controlled experimental conditions and to bypass identification problems that always plague field data. Finally, a much smaller literature, best exemplified by Bewley's 1999 book, takes an "anthropological" approach, interviewing people in charge of setting wages in real-world companies; the relevant question for our purposes is, why are these people reluctant to cut wages?

Even abstracting from the relative methodological merits of these various approaches, the question itself has profoundly changed in na-

978-0-226-16540-0/2014/2013-0303$10.00

ture in the last thirty years. The predominant view of aggregate labor markets in macroeconomics has experienced a sea-change. Rather than the traditional view of a spot labor market, where "stocks" of demand and supply meet anew every period and various frictions may impede market-clearing, the new prevailing view is the so-called "flow approach" to labor markets. Inspired by a vast body of new empirical evidence, frontier research now envisions employment, unemployment, and inactivity as persistent states, whereas wages and labor markets determine the *flows* between them: accessions, job separations, discouragement after a long job search, and so on. In this paradigm, the question is not why wages are so rigid to keep supply in excess of demand, but rather why wages of new hires do not fall in and after a recession to moderate the observed discouragement in new job creation, and why wages of incumbent workers do not fall to avoid layoffs and separations. The questions are all valid, in reverse, during expansions, but the interest has been definitely asymmetric and focused on high unemployment phases.

Within the flow approach, the search and matching (S&M) model has emerged as the canonical theoretical framework for the analysis of aggregate labor markets. This is a rich incomplete-market framework, which can accommodate in a tractable manner a wide variety of preferences, wage-setting protocols, idiosyncratic and aggregate shocks of various kinds affecting workers and firms, savings, and various forms of self- or implicit insurance against those shocks. The model is flexible enough to dovetail nicely with DSGE models of all flavors, from Neoclassical to New Keynesian. From an empirical viewpoint, the S&M model is firmly grounded in the vast body of micro-level observations on labor market transitions. Indeed, the two sides of this literature, empirical and theoretical, have developed together, and even changed the way labor market statistics are collated. For example, in 2000, the Bureau of Labor Statistics started the Job Openings and Labor Turnover Survey (JOLTS), which is expressly designed to provide timely and high-frequency empirical information about such concepts as vacancies, hires, and quits, that are central to the S&M model. JOLTS allows, in particular, to directly measure job market tightness, the vacancy/unemployment rate, which plays the same role in the S&M flow model as the labor demand/supply gap in the traditional spot market stock framework.

The S&M model was used predominantly as a tool to understand long-run issues, such as the determinants of the natural rate of unemployment and the effect of labor market policies. But the business cycle side of the research was always active. In this vein, Robert Shimer recently offered two seminal contributions. Shimer (2012) shows that, in the United States, unemployment rises in recessions mostly because job losers, who exist in large numbers even in good times, can no longer find a new job quickly, rather than (as common wisdom held since Keynes) due to waves of layoffs. Recessions are times of increasing unemployment duration more than of increased risk of separation. Conversely, expansions are times when job losers, who are always there, can regain employment quickly. Part of Shimer's contribution was technical, proving that a particular type of time aggregation bias affected previous empirical inference. In light of Shimer's finding, the question is indeed why wages fail to decline in order to stimulate new hires, *not* to prevent layoffs. This provocative finding spurred a lively debate on the relative importance of the cyclicality of job finding and separations in explaining that of the unemployment rate (Elsby, Michaels, and Solon 2009, Fujita and Ramey 2009). It is fair to conclude that the former explains at least half of unemployment fluctuations. So Shimer's finding has been qualified, but it still changed our view of labor market fluctuations. Indeed, after this debate unfolded, the Great Recession only reinforced his conclusion: the layoff rate observed in JOLTS spiked briefly, while the job finding rate calculated from the Current Population Survey, using either gross flows from matched files or short-term unemployment stocks as in Shimer (2012), dropped at least as dramatically, and then never recovered. Fujita and Moscarini (2013) update the evidence to the Great Recession and show that this phenomenon is even more pronounced if one focuses on new hires and excludes recalls of previous employees, who should not be mediated by the frictional matching mechanism central to the S&M model.

The second part of Shimer's contribution pertains to the S&M model itself, in the canonical form with Nash Bargaining wages that earned Peter Diamond, Dale Mortensen, and Chris Pissarides the 2010 Nobel Prize. Shimer (2005) derives an exact algorithm to compute the stochastic equilibrium of this model in the presence of shocks to average labor productivity. He shows that a reasonable calibration of the model generates fluctuations in job market tightness that are a full order of magnitude smaller than in US data. This theoretical finding spurred an even livelier debate.

A proliferation of proposed solutions to this puzzle ensued. One simple fix proposed by Hagedorn and Manovskii (2008) is to calibrate market and nonmarket activity as close substitutes. In this respect, the S&M model does not provide a solution to the main failure of a competitive DSGE model, which is to explain employment volatility without resorting to implausibly high elasticity of labor supply. Since this fix makes unemployment barely costly, thus, ultimately, irrelevant, and is also in contrast with other evidence, the majority of the literature has attacked the Nash Bargaining assumption for not providing sufficient wage rigidity. More precisely, because hiring requires an up-front investment, the volatility of job market tightness depends directly on the volatility of the present value of rents that a firm expects to earn from a new hire. In turn, this volatility is inversely related to the cyclical volatility of the present value of wages that a new hire expects to earn. Therefore, this present value is the relevant notion of "wages" that researchers have focused on.

Eliaz and Spiegler (henceforth, ES) join the debate at this juncture. They introduce a new source of wage rigidity, based on reference-dependent behavior, a phenomenon that has been amply documented by the experimental literature and that is consistent with the answers to Bewley's interviews. Their departure from the standard S&M model is intentionally minimal, and they take great care in eliminating other sources of volatility in job market tightness. They provide an elegant analysis of noncooperative equilibrium of a repeated wage bargaining game, where the firm makes all the offers and yet the worker may earn rents ex post, because of reference dependence: if the wage offer is lower than expected, the worker cannot help reducing unobservable effort. The Subgame Perfect Equilibrium has a flavor of efficiency wages, but is not based on shirking/monitoring, rather on a primitive aspect of people's preferences. Importantly, it is unique, so it does not generate wage rigidity through indeterminacy (as in Hall 2005).

My comment revolves around two aspects of ES analysis. In the first section, I show that reference-dependent behavior mentioned by respondents to Bewley's 1999 interviews may arise, even with standard preferences, from asymmetric information between employer and employee. I introduce a model where a wage cut discourages worker effort not because of some psychological impact, but rather because it reveals to the worker the opinion (and information) that the firm has about the worker's ability. If ability and effort are complementary, a wage cut leads the worker to revise downward his perceived marginal productivity of effort, thus, rationally, his work effort. If one is not good at something, why even try? If effort is already suboptimal to begin with, because of standard moral hazard considerations, this "discouragement" effect may be detrimental to the firm, who may prefer to keep wages high and conceal the negative information. While this new model may not explain the experimental evidence on reference dependence, it fits well the tone of Bewley's interviews. Next, ES illustrate the qualitative effects of their model of wage determination on equilibrium job market tightness, but do not explore its quantitative implications. In the second section I argue that their model cannot solve the quantitative puzzle, for reasons that are instructive about the S&M model and that show once again how deep this puzzle remains. I conclude by revisiting other predictions of the ES model, related to wages rather than to unemployment and vacancies, which suggest potentially more promising applications.

A Rational Model of Reference-Dependent Behavior

Fang and Moscarini (2005) introduced in economics the notion of morale hazard (as opposed to the different notion in the insurance industry), to capture the insights of Bewley's survey. They studied the effect of wage comparisons across workers. Indeed, human resource managers responding to Bewley's survey explicitly stated that wage cuts are much more acceptable by their workers when they happen across the board, possibly involving management also, hence credibly signaling that the whole company is in trouble. By contrast, they find selective wage cuts particularly damaging to worker "morale" and avoid them in favor of layoffs. Accordingly, morale hazard is the detrimental effect on worker effort due to the inference that he rationally draws, from the wage/contract offer that he receives, about what the firm thinks/knows he can accomplish. Simply put, if a wage cut, especially relative to other employees, tells the worker that he is not very good at performing his task, then the worker will rationally give up trying, exacerbating preexisting incentive problems.

ES focus on a different aspect, which is corroborated more by experimental evidence, namely the comparison made over time by the same worker between the wage he expects to receive and the one he actually receives. I modify in this direction the morale hazard model of Fang and Moscarini (2005), to provide an alternative microfoundation for the two new aspects of the ES model—both reference-dependent behavior and the specific “lagged expectations” model of reference formation—without introducing any behavioral assumptions, but rather appealing to plain-vanilla private information.

A worker is hired by a firm and spends unobservable effort $n \in [0, 1]$ at utility cost $n^2/2$ and, with probability na, produces two units of output, while with probability $1 - na$ he produces only one unit. The parameter a captures ability and takes two values, 0 and 1, with $\Pr(a = 1) = p_0$. The firm owns any output that gets produced. Worker ability is not directly observable by either party, and can only be learned from output performance. The worker has an outside option equal to U when deciding not to work for the firm, and the firm has an outside option normalized to zero by free entry. Both parties are risk-neutral. Before producing, the firm and the worker may receive additional, possibly private, information that turns their common prior belief p_0 into posteriors p for the firm and q for the worker.

In the first-best allocation, worker and firm cooperatively choose effort and share output, and are symmetrically informed, so they share belief, say, π. The efficient allocation yields

$$Q(\pi) = \max_{n \in [0,1]} 1 + \pi n - \frac{n^2}{2}.$$

The FOC yields $n^* = \pi$, which is in [0,1] as required, so the FOC is also sufficient. The optimal joint value is

$$Q(\pi) = 1 + \frac{\pi^2}{2}.$$

To avoid trivialities, we assume $Q(p_0) > U$: there are positive gains from trade.

In the second best, the worker cannot commit to provide effort, while the firm can commit to deliver an output-contingent payment. WLOG, an incentive scheme, is a zero wage in case of failure and a wage w in case the second unit of output is produced. The offer w may signal some or all of the firm's private information about the worker's ability, and change the worker's belief from p_0 to q, while the firm updates its belief from p_0 to p using its private signal.

The worker solves

$$\max_{n \in [0,1]} qnw - \frac{n^2}{2}.$$

We ignore the constraint $n \in [0, 1]$ and later check that it is slack. The worker chooses effort $n(w) = qw$. The firm's problem is then

$$J = \max_{w} 1 + pn(w)(1 - w),$$

subject to $n(w) = qw$, with solution

$$w = \frac{1}{2} \text{ and } J(p, q) = 1 + \frac{pq}{4}.$$

The worker receives in expectation

$$S(q) = qn(w)w - \frac{n^2(w)}{2} = \frac{q^2}{8}.$$

Now suppose that, just like in ES, employment relationships last two periods, and what we just described is the second and last period. In the first period, firm and worker start symmetrically informed with prior $p_0 = \Pr(a = 1)$, but output is neither observable by nor verifiable to the worker; thus the firm cannot commit to an output-contingent payment. The worker puts effort n_1. After observing the output outcome, the firm updates its prior p_0 by Bayes rule to the posterior p, which is the prior with which it enters the second period. Suppose that only one unit of output is produced in the first period, and accordingly the firm becomes pessimistic and updates beliefs to $p = \underline{p}$ where

$$\underline{p} = \frac{p_0(1 - n_1)}{p_0(1 - n_1) + 1 - p_0} \leq p_0.$$

If instead the firm observes two units of output, it learns that the worker is of high ability, and $p = 1$. If the firm credibly reveals to the worker the outcome, the worker also updates his belief to $q = p$. If the firm can credibly reveal nothing, the worker keeps the prior p_0 into the second period. Suppose:

$$S(\underline{p}) = \frac{\underline{p}^2}{8} < U < \frac{p_0^2}{8} = S(p_0).$$

Then the worker would accept the second-period contract, with either wage $w = 1/2$ or dismissal depending on output outcome, only if his prior perception p_0 of his own ability did not change between periods. Otherwise, upon learning that his first-period performance was poor, he would be "discouraged," update belief to $\underline{p} < p_0$, and exacerbate the moral hazard problem of effort underprovision, making the relationship not viable, as his expected value $\underline{p}^2/8$ from the second-period contract would fall short of his outside option U. This may be possible even if the relationship was still viable under commitment, $Q(p) = 1 + (\underline{p}^2/2) > U$ and the firm would gain from continuing, $J(\underline{p}, \underline{p}) = 1 + (\underline{p}^2/4) > 0$. So the source and cost of discouragement is worker lack of commitment. In this case, at the end of the first period, the firm may prefer to reveal nothing about first-period performance. Fang and Moscarini (2005) labeled morale hazard this effect of the agent's discouragement on the payoffs of the principal.

How may the firm credibly reveal its information? It can do it either through the wage effectively paid at the end of the first period, or through the wage offered in case of good performance in the second period. Suppose the firm and the worker are myopic and maximize, at the beginning of each period, their static payoffs; but, at the end of the first period, after observing output, the firm cares about the continuation. In ES terminology, in the first period the worker has a reference wage $w = 1/2$, the wage he expects to earn conditional on being retained. If first-period output is low, the firm may prefer to conceal the failure, pay $w = 1/2$ anyway, trigger no updating in worker's beliefs, and earn $J(\underline{p}, p_0) = 1 + (\underline{p}p_0/4)$ in expectation in the second period.

Based on these considerations, we find a unique Bayesian Perfect Equilibrium of this game. At the end of the first period the firm plays a pooling strategy, hides first-period performance, and pays the wage $w = 1/2$ that the worker expected at the end of the first period, as well as promises $w = 1/2$ in case of (verifiable) success in the second period. The pessimistic firm who observed low output in the first period pools with the optimistic firm type and pays $w = 1/2$ nonetheless, because the continuation value of not discouraging the worker and keeping the relationship viable is worth the wage cost, and beats paying nothing at the end of the first period and losing the worker:

$$0 < -w + J(\underline{p}, p_0) = \frac{1}{2} + \frac{\underline{p}p_0}{4}.$$

By backward induction, at the beginning of the first period, the worker knows that he will receive at the end of the first period a wage $w = 1/2$ independently of the outcome of his first period's effort, because a zero payment would discourage him beyond rescue, and the firm would lose from the resulting destruction of output. Thus, in the first period, the worker puts zero effort $n_1 = 0$, the firm learns nothing ($\underline{p} = p_0$), neither does the worker ($q = p_0$), the firm pays the reference wage

($w = 1/2$), and the relationship continues into the second period. This equilibrium is supported by simple off-path beliefs. Any offer that might credibly signal to the worker low output in the first period discourages the worker to the point that the relationship breaks down. In that case, the pessimistic firm, which observed low output, would pay nothing, ex post. But it would lose on net, as shown earlier, and would rather mimic an optimistic one, and pay any wage below 1 at the end of the first period and (subject to a good final outcome) the second period. Any such wage path keeps the worker morale high and yields a positive net continuation value, as opposed to losing the discouraged worker.

In this pooling equilibrium, the first period wage is “rigid” because of morale hazard. Should the firm pay the worker at the end of the first period a wage below the reference wage $w = 1/2$ that the worker expects conditional on retention, the worker would interpret the signal as bad news about his ability to generate high output, thus about his marginal productivity of effort. Hence, the worker would reduce his effort in the second period, to the point of making the relationship no longer viable, even if gains from trade would still exist under commitment. While this is a theory of wage rigidity to idiosyncratic shocks, it is not difficult to envision how to extend it to aggregate shocks that the firm and the worker cannot immediately tell apart from idiosyncratic ones.

Cyclical Volatility in the Labor Market

The ES model leaves intentionally open the question of whether it can explain *quantitatively* amplification of TFP (total factor productivity) shocks on unemployment and amend the shortcomings of the baseline S&M model that were pointed out by Shimer (2005). A more narrow question that emerged from the debate is whether the S&M model can explain the observed large fluctuations in job market tightness, hence in the incentives to create new jobs, independently of whether those explain 50 percent or 95 percent of unemployment volatility. We start from the second question. To do so, we first summarize the equilibrium equations of the S&M model used by ES.

Following Brügemann and Moscarini (2010), the steady state equilibrium of the S&M model is fully characterized by three simple equations. Using the ES notation, these are the free entry condition in vacancy creation

$$c = q(\theta)\Pi(p),$$

the Bellman equation of the unemployed worker

$$W_0(p) = bp + \delta W_0(p) + \delta\theta q(\theta)R_1(p),$$

and the definition of new hires' rents $R_1(p)$ as the present value of wages on the job minus the value of unemployment. In any model, such as ES, where the worker returns to unemployment after two periods of work, rents are

$$R_1(p) = w_1 + \delta\lambda W_0(p) + \delta(1-\lambda)[w_2 + \delta W_0(p)] - W_0(p), \tag{1}$$

where λ is the expected layoff rate between periods, and w_1 and w_2 are the wages to be determined. A final restriction is the adding-up constraint that the sum of gross values $\Pi + R_1 + W_0$ do not exceed the expected present value of output in the relationship. These three equations and the adding-up constraint pin down a solution Π, R_1, W_0, and θ as functions of p. For convenience, we will refer to p as TFP.

Shimer (2005) derived a stochastic version of this model, where TFP is subject to mean-reverting Markovian shocks. Under the assumption of Nash bargaining wages, which are privately efficient and entirely forward-looking, in the stochastic equilibrium job market tightness θ_t is only a function of current TFP p_t, and follows a simple stochastic difference equation induced by TFP shocks. Thus θ_t has no transitional dynamics.

ES introduce a richer, noncooperative model of wage determination that features history-dependence. In the unique Subgame Perfect Equilibrium (SPE), the wage in the second period $w_2(p_t, p_{t+1})$ is independent of the Bellman values, new hires receive no rents, so $R_1(p_t) = 0$, the value of unemployment $W_0(p_t)$ equals the expected present value of leisure $B(p_t)$, which is a primitive of the model, equation (1) pins down the initial wage $w_1(p_t)$ given expected $w_2(p_t, p_{t+1})$, the firm receives residual output after paying wages, capitalized into $\Pi(p_t)$, and free entry pins down job market tightness $\theta(p_t)$.

Volatility of Job Market Tightness

In principle, to understand the quantitative property of the S&M model with ES reference-dependent wages, we should calibrate and simulate a fully dynamic stochastic version. Fortunately, in the canonical S&M model, for many wage-setting protocols, including ES game, equilibrium remains entirely forward-looking, in the sense that the free entry condition ties job market tightness θ_t only to the current level of the aggregate impulse p_t, say TFP. In these cases, a common quantitative finding is that the relative volatility of θ_t and p_t is bounded above by the steady-state elasticity of θ to p. The reason is twofold: the comparative statics response is the same as the effect of a permanent shock to p, because θ has no transitional dynamics by the forward-looking behavior of the model and constant returns to scale in production and hiring; and, the effect of a given shock to p on θ is increasing in its persistence and decreasing in its volatility, so the full persistence case provides an upper bound, and comparative statics an even higher upper bound. Brügemann and Moscarini (2010) build on this insight to construct a simple test: any model of wage-setting $\{\Pi, R_1\}$ that satisfies three steady-state comparative statics properties, once embedded in the canonical S&M model, generates at most 23 percent of the observed variance of job market tightness in Shimer's 2005 calibration. Setting the all-important value of leisure to 71 percent of average labor productivity, as in Hall (2009), makes little difference for this conclusion.

ES characterize in their equation (12) the steady-state elasticity of θ to p in the equilibrium of their model. Simple inspection shows that the elasticity is maximized when TFP shocks are fully persistent ($\beta = 1$), because δ, λ, $\mathbb{E}(\varepsilon^{\beta})$ are all nonnegative. In this case reference dependence plays no role in amplifying shocks, and the elasticity of job market tightness to such shocks equals the inverse of α, the elasticity of the matching function. This formula is familiar from the literature (e.g., Brügemann and Moscarini 2010). Intuitively, if α is high, unemployment is very important in creating vacancies at the margin, and vacancies are not. Therefore, when vacancies rise and job market tightness θ declines, the job-filling rate $k\theta^{-\alpha}$ declines quickly: more vacancies create much congestion and few additional matches. We conclude that external congestion effects are strong, returns to hiring investment diminish fast, and any increase in p and incentives to create vacancies die out quickly. Indeeed, for plausible calibrations of the value of α, job market tightness responds no more than fourfold to TFP shocks, much less in Shimer's calibration, while the empirical target exceeds tenfold. Hence, it seems unlikely that the ES model can answer this quantitative question.

The equilibrium of the ES model gives zero rents to new hires, because firms have all the bargaining power and the first reference wage for new hires is lower than the value of leisure. In this case of firms' full bargaining power, Brügemann and Moscarini (2010) show that, given the calibration, average profits earned by firms on new hires are quite large, thus so is average job market tightness, relative to the size of cyclical fluctuations in aggregate driving forces, such as average labor productivity, or TFP. Therefore, the impulse is tiny and firms barely respond to it with their job creation. Getting around this problem would require implausibly large hiring costs, which find no support in the available empirical evidence. Brügemann and Moscarini (2010) identify in countercyclical worker rents the key to generate volatility. In Hall (2005) the wage for new hires and thereafter is fixed and, as the value of unemployment is procyclical, worker rents are countercyclical. This is why the model can generate high volatility. In ES, worker rents are zero, thus acyclical, which Brügemann and Moscarini (2010) shows is not enough to amplify aggregate shocks.

If TFP shocks are mean reverting ($\beta < 1$), then reference dependence will raise the elasticity of θ to p, but only relative to a low reference-independent benchmark. The upper bound (the expression in [12] with $\beta = 1$) that it can attain is still too low.

Volatility of Unemployment

A separate question is whether the ES model may provide a partial explanation to the high cyclical volatility of unemployment through its prediction for productivity and separations, independently of its (in) ability to generate cyclical job creation. The effect of negative TFP shocks is amplified by "excessive" worker expectations about their wages. When the economy slumps, an existing employee receives a lower offer than expected, which leads him to either withdraw effort and destroy output, or to reject and enter unemployment. The first outcome amplifies the decline in average labor productivity. This means that the standard deviation of the underlying TFP impulse must be smaller than the 2 percent observed variation in average labor productivity, chosen by Shimer as his measure of the impulse, because part of the observed 2 percent follows from internal amplification via effort reduction. This effect is common to many models of endogenous productivity, and works against solving the puzzle. Even more problematic for this hypothesis is the recent observation that average labor productivity in the United States became acyclical or countercyclical since the mid-1980s, when recoveries became jobless (Berger 2012). The Great Recession is a case in point, as both TFP and average labor productivity rose sharply in late 2008. As for dismissals, as ES explain very clearly, in their model the layoff probability is stationary and uncorrelated with the initial level of economic activity. Thus, it is only weakly correlated with the contemporaneous level. So the model will have to be modified to make the layoff probability strongly countercyclical, if it aims to explain unemployment fluctuations through this channel.

Additional Predictions

The ES model provides, among others, two intriguing predictions that offer novel alternatives to lively empirical debates. These are interesting in their own right, independent of the issue of cyclical volatility. Therefore, it appears worth exploring further the implications of reference dependence on these two dimensions, to distinguish it from other hypotheses that have already survived some empirical scrutiny.

First, the ES model predicts that the wages of new hires are more cyclical than those of incumbent workers, and seniority commands a premium. Both are well-established facts that the model generates robustly. Reference dependence provides an alternative to human capital, backloading, and insurance theories. Second, the model offers a novel rationale, the lagged-expectation reference formation, for history-dependence of wages. Observed real wage rates and earnings do exhibit history-dependence. Specifically, the unemployment rate at the time of the hire and the lowest unemployment rate observed since then have explanatory power for wages above and beyond the contemporaneous unemployment rate. This finding has been traditionally interpreted as evidence in favor of implicit employment contracts, where the firm insures the risk-averse worker against wage and employment risk and in exchange earns a risk premium through a reduced average wage (Beaudry and DiNardo 1991). More recently, Hagedorn and Manovskii (2013) challenged this interpretation. They argue that these observed effects are caused by selection and competition, not by insurance. Low values of initial and intervening unemployment rates proxy for tight labor markets, that facilitate outside offers. These, in turn, raise wages permanently, either directly through wage renegotiation, or by triggering quits by the lower paid workers. In fact, in the same data, even stronger explanatory power for wages lies in the sum of the values of job market tightness since the beginning of the current employment spell, which summarizes the history of labor demand intensity, hence of possible outside offers that the worker may have received or generated. Reference dependence offers an entirely different explanation.

Endnote

For acknowledgments, sources of research support, and disclosure of the author's material financial relationships, if any, please see http://www.nber.org/chapters/c12940.ack.

References

Beaudry, P., and J. 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 (4): 665–88.

Berger, D. 2012. "Countercylical Restructuring and Jobless Recoveries." PhD Dissertation. Yale University.

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

Brügemann, B., and G. Moscarini. 2010. "Rent Rigidity, Asymmetric Information, and Volatility Bounds in Labor Markets." *Review of Economic Dynamics* 13 (3): 575–96.

Elsby, M., R. Michaels, and G. Solon. 2009. "The Ins and Outs of Cyclical Unemployment." *American Economic Journal: Macroeconomics* 1:84–110.

Fang, H., and G. Moscarini. 2005. "Morale Hazard." *Journal of Monetary Economics* 52 (4): 749–77.

Fujita, S., and G. Ramey. 2009. "The Cyclicality of Separation and Job Finding Rates." *International Economic Review* 50 (2): 415–30.

Fujita, S., and G. Moscarini. 2013. "Recall and Unemployment." Unpublished Manuscript. Yale University.

Hagedorn, M., and I. Manovskii. 2008. "The Cyclical Behaviour of Equilibrium Unemployment and Vacancies Revisited." *American Economic Review* 98 (4): 1692–706.

Hagedorn, M., and I. Manovskii. 2013. "Job Selection and Wages over the Business Cycle." *American Economic Review* 103 (2): 771–803.

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

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

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

———. 2012. "Reassessing the Ins and Outs of Unemployment." *Review of Economic Dynamics* 15:127–48.