Christopher Foote is an assistant professor of economics at Harvard University.

Commentary

Christopher Foote

This stimulating paper by den Haan, Ramey, and Watson (DRW) seeks to contribute to two literatures. First, it helps answer the microeconomic question of when firms will use performance pay rather than salaries to motivate workers. Second, it examines the microeconomic issue of why employment relationships appear so fragile with respect to aggregate shocks. The overarching goal is to show that these two literatures are related, or, as the authors put it, "the particular form of imperfections that are present in the con-tracting environment can have major implications for economic outcomes." The central result of the paper is that employment contracts often are robust with respect to aggregate shocks only if the effort choices of both the worker and the firm are verifiable and contractable. The paper is exceptionally clear and insightful; even though the model is simple, the authors can analyze a surprisingly large number of special cases. As a result, it is a great paper for a volume that seeks an integrated study of the macroeconomics of the labor market. I will comment first on the struc-ture of the model and then discuss its micro and macro implications.

The nicely parsimonious model is based around three possible exchanges between the worker and the firm. The two parties begin the period by negotiating the contract and making an *initial* exchange (s_0^k), which depends on the aggregate state k. In the negotiation process, they promise to make *intermediate* exchanges (s_A and s_B) if one of the parties shirks. Finally, if both parties provide the required levels of high effort, production completes satis-factorily and the firm receives the output, making a *final* exchange (s_C^k) to the worker.

The authors contend that long-term employment relationships can survive even in bad aggregate states if the agents can successfully commit themselves not to shirk, but this requires low effort to be detectable by a court so that it can bring about the required intermediate exchanges. The setup reminded me of an unemployed worker in the Shapiro-Stiglitz model; such a worker would like to be hired at the pre-vailing wage, but his offer is not accepted because he cannot make a credible promise not to shirk. Here, in the case of perfect verifiability, the offer is accepted even when the production value is low because shirking is detected with probability one and can be proven in court. This brings about the required intermediate transfer from the worker to the firm. (Another difference between this model and the Shapiro-Stiglitz framework is that in this model the workers can post bonds for jobs via negative values of the initial payment s_0^k, a point to which I will return below). When verification of effort is not possible in the DRW model, the intermediate-pay-ment strategy for the robust employment contract fails. The constant benefit of shirking is too large and tempting for either the worker or the firm, relative to the low outcome from production.

The model has several interesting microeconomic implications. But even though the possibility of initial, intermediate, and final payments makes the model very versatile, I am not sure that it is general enough to speak definitively on the issue of salary versus performance pay in an optimal compensation policy. The applica-bility of the model for this issue hinges on whether a negative payment to a worker in the initial period ($s_0^k < 0$) can be thought of as a salary, or rather, as some sort of nega-tive bond reflecting the *ex ante* terms of trade in the labor market. Because such bonds can eliminate involuntary unemploy-ment when effort elicitation is a problem, such bonds have featured prominently in

work of this kind. Yet the authors claim that in their model, a positive value s_0^k is like a salary, while the final payment, s_C^k, resembles performance-based pay. Therefore, they suggest that their model can speak to the performance-pay versus salary issue. Yet the DRW model does not allow the worker both to post a bond (which would require $s_0^k < 0$) and receive a salary (which would require $s_0^k > 0$). As a result, I am not sure that the model can be considered a general case of the previous models in the effort-elicitation literature. An alternative interpretation of the salary-vs.-performance-pay distinction in the DRW model would be that some average value of the end-of-production payment (which could be denoted s_C) would be more like a salary, while the time-varying payments (s_C^k) are similar performance-based pay. The point is that a salary is generally a constant payment received by the worker regardless of how much output is produced, yet if production does not take place, no salary is received.

A larger microeconomic goal of the paper is to investigate the effects of nonverifiability of effort in employment relationships. To make the model tractable, the authors make some simplifying assumptions expressed in inequality 1. One implication of the assumptions is that the promise of continuing the job is generally not enough to preserve a current match in the bad state today, unless there is some verifiability of effort. Basically, firms and workers will not stick it out through the bad times in order to enjoy the good times again later on. As the paper shows, however, the continuation value of the job in the different states (the g's) are themselves functions of the common discount factor (β), the probability of the bad state (ρ), and the relative levels of the productive outcomes (z^G and z^B). Therefore, underlying the assumptions in inequality 1 are implicit assumptions about preferences and the stochastic properties of aggregate shocks. Indeed, firms and workers may want to endure the bad times even without verifiability—if the discount rate is small enough

or bad aggregate shocks are rare enough—but particular values of the underlying parameters, which imply that verifiability is required for a robust contract, are not obvious.

A third comment on the micro-structure of the model involves the relationship of this model to work involving specific investments in the employment relationship. Papers by Caballero and Hammour stress the "fundamental transformation" that employment relationships undergo when searching workers and firms find one another, or when either side of the employment relationship invests in capital that is specific to the match. Both of these phenomena transform the employment relationship into a bilateral monopoly where the *ex ante* terms of trade may not carry over to the *ex post* Nash bargain over the surplus in the match. In reading this paper, I was curious to know whether there was a simple mapping between its shirking interpretation and the specific capital basis of other work. It may be that the shirking model is a particularly strong form of the specific capital model. For example, consider the type of specific capital that is created simply when a searching worker finds a specific job. Match capital is created because the firm no longer has to pay the costs of posting the vacancy and the worker can start earning wages rather than spend time looking for a job. The outside alternatives of the worker, the firm, and the exogenously determined bargaining weights (the π's) determine how the surplus is divided in both models; though in the DRW setup, the worker or the firm receives an additional reward (the benefit to shirking) if the match breaks up. In the specific capital model, a party who leaves the match does not receive this type of benefit. It would be interesting to know if there is a simple way to link both the shirking and specific-capital interpretations of the employment relationship.

One way that *ex ante* terms of trade can be reflected *ex post* in the employment match is when the workers post a bond. One of the cases discussed by the authors is that of "limited liquidity," which argues

that low liquidity may prevent workers from paying a bond. The inability to pay a bond opens up the possibility that efficiency wages may be paid in order to elicit effort. Several economists, however, have pointed out that the *utility cost* of a bond payment may vary inversely to the worker's liquid assets. Workers with low liquidity may find it hard to post a bond, but it is precisely these workers who will be quite averse to shirking and losing their bond if they get caught shirking. Of course, the power of very small bonds to motivate liquidity-constrained workers depends on marginal utility going to negative infinity as assets go to zero, so the outside benefits, *b*, may prevent this from occurring in the DRW setup.

I now turn to the macroeconomic implications of the model. One of the key questions in cyclical macroeconomics is why so many employment relationships break up during recessions. Pioneering work by Davis, Haltiwanger and Schuh showed that (at least in manufacturing) the drop in employment that occurs when recessions begin comes not from a decline in the creation of new jobs, but rather, a large spike in the destruction of existing jobs. Several authors have suggested that the spike in job destruction at the onset of recessions may be linked to the economy's amplification mechanism, by which moderate innovations to productivity or aggregate demand may bring about large movements in employment and output. Not surprisingly, a number of theories to explain the burst in job destruction have been advanced. One branch of the literature stresses "cleansing" effects of recessions. Large numbers of jobs are destroyed in recessions because a large number of jobs are typically close to the margin of unprofitability at any point in time. Convex job-creation costs for the aggregate economy mean that it is more efficient for these older jobs to be destroyed than for the rate of new job creation to drop. A second branch of the job destruction literature stresses the "pit stop" role of recessions. Just as the real business cycle literature contends that recessions are a

good time to enjoy leisure, that pit-stop view suggests that recessions are good times to reorganize production.

This paper can be placed in a branch of the economic literature that contends that the increase in job destruction is a primary result of some imperfection or friction in the labor market; here, the friction is nonverifiability. Other papers of this kind suggest that job destruction is high during recessions because wages cannot fall. Two potential causes of wage rigidity are the suppression of wage renegotiation (since bargaining is costly and may encourage opportunistic behavior) and efficiency wages (the need to motivate workers provides a floor through which wages cannot fall). The DRW model suggests that nonverifiability, rather than the suppression of negotiations or efficiency wages, can better explain the cyclical dynamics of the labor market. The authors implicitly argue against the suppression of the renegotiation model by having the worker and firm bargain at the start of every production period. The suppressed renegotiation models imply that the firm and the worker would like to renegotiate and stay together, but doing so would result in redoing the employment contract and thereby violate some social norm. The DRW paper suggests, in contrast, that firms and workers are not averse to renegotating, but they prefer to separate in bad times because they cannot make a credible promise not to shirk without external verifiability.

The paper argues against the efficiency-wage-model-with-verifiability more explicitly with the experiments displayed in DRW's Figure 2. Recall that efficiency wages arise in the DRW framework if workers are unable to post a bond and the value of the final payment they can receive at the end of production is not large enough to prevent them from shirking. In this case, the firm must raise the worker's *total* compensation (here, just the final payment) above the level implied by the worker's bargaining weight, π^w. In this way, the worker is encouraged not to shirk and an efficiency-wage trade-off arises. This is due to the

inverse relationship between the total compensation of the worker and his incentive to shirk. The firm is prevented from shirking because of the verifiability assumption. Should it fail to provide high effort, the court will assess a payment to the worker of s_A. The efficiency-wage cases analyzed in Figure 2 are, therefore, ones of robust contracts, because the efficiency wage and the monitoring of the firm by the court combine to keep the parties honest. Figure 2 shows that the three verifiable contracts result in little amplification of shocks, which are modeled as an increase in the breakup parameter ρ. Even the two contracts that imply efficiency wages are robust, in part because firm behavior can be verified. On the other hand, the nonverifiable ("severance payment only") contract results in substantial amplification of the increase in ρ. Employment relationships break up because no mechanism exists to ensure high effort.

My main concern with this result is that I am not sure it portrays efficiency wages in the most familiar way. The efficiency-wage model in this paper is essentially backloaded compensation, which is paid only when the worker does not shirk, an event that is detected with probability one. (Of course, there may be an incentive for the firm to shirk when compensation is backloaded, but the courts are assumed to regulate the firm's behavior since verification is assumed in these cases.) Another interpretation of efficiency wages, however, might affirm that they arise when worker misbehavior is detected only imperfectly and, therefore, the workers may consider the shirking decision differently. The differences in the two interpretations for macro behavior of the two views of efficiency wages are not obvious immediately. Another question I have is how general equilibrium effects work in the macro simulations of DRW's Figure 2. The general equilibrium is important because the previous work by Caballero and Hammour has shown that high unemployment during a recession can result from wage rigidity engendered by specific-capital

investments. The high unemployment is an equilibrium response of the economy to wage rigidity, as it disciplines the wage demands of insiders. I am curious to know whether a similar effect is operating here.

All in all, I found this paper well worth the time and effort it took to study the subject carefully. And I hope the graduate students enrolled in my next "Macroeconomics of the Labor Market" course do so as well.