Available online at www.sciencedirect.com

ScienceDirect

Review of Economic Dynamics 11 (2008) 664–678

Review of Economic Dynamics

www.elsevier.com/locate/red

Worker heterogeneity and labor market volatility in matching models

Michael J. Pries

Department of Economics, 3105 Tydings Hall, University of Maryland, College Park, MD 20742, USA

Received 1 September 2006; revised 27 October 2006

Available online 4 December 2007

Abstract

Shimer demonstrated that aggregate productivity shocks in a standard matching model cause fluctuations in key labor market statistics—such as the job-finding rate, the vacancy/unemployment ratio, and the unemployment rate—that are too small by an order of magnitude [Shimer, R., 2005. The cyclical behavior of equilibrium unemployment and vacancies. American Economic Review 95 (1) 25–49]. This paper shows that when the standard model is extended to allow for worker heterogeneity, it exhibits considerably greater volatility. In the model, marginal workers, whose productivity only slightly exceeds the value of their alternative use of time, constitute a disproportionate share of unemployment on average, and that share rises when aggregate conditions deteriorate. These composition effects cause firms to open fewer vacancies during downturns.

JEL classification: E24; E32; J63; J64

Keywords: Volatility; Amplification; Matching models; Worker heterogeneity

1. Introduction

Search and matching models such as the Mortensen and Pissarides (1994) model have become a primary tool for analyzing labor market fluctuations. In an important recent paper, however, Shimer (2005) argued persuasively that a calibrated version of a standard matching model exhibits very little volatility: fluctuations in vacancies, job-finding rates, and unemployment are too small, relative to US data, by at least an order of magnitude.

In Shimer (2005) and in the various recent attempts at resolving this apparent puzzle,[1] workers are assumed to be ex ante homogeneous. For these workers, a small increase in aggregate productivity leads to a nearly equal increase in wages. This increase in wages means that the payoff to the firms who are matched with the workers is only minimally affected by the productivity shock. Similarly, for firms that are considering opening new vacancies, the potential payoff from filling a vacancy is only slightly increased by a positive productivity shock. Thus, the response by wages dampens the response of vacancies—and thus job-finding rates and the unemployment rate as well—to productivity shocks.

E-mail address: mpries@nd.edu.

[1] See, for example, Hall (2005b), Hagedorn and Manovskii (2005), Costain and Reiter (2005), and Nagypál (2005). For a more extensive review of papers that address this puzzle, see Mortensen and Nagypál (2006).

1094-2025/$ – see front matter
doi:10.1016/j.red.2006.10.003

In this paper, I demonstrate a channel by which worker heterogeneity can increase labor market volatility in the standard model in a quantitatively important way. To understand why worker heterogeneity may affect the volatility of the vacancy/unemployment ratio, note that if firms profit more from matching with some types of workers than from matching with other types, then the firms' incentives to open vacancies will depend on the composition of the pool of unemployed from which they will potentially hire. For example, during downturns if the pool of unemployed consists of a larger than usual share of workers who are less attractive as potential hires—as the evidence on the composition of unemployment, discussed in Section 4.5, suggests—then firms have less incentive to open vacancies.

Aside from the extension to allow for worker heterogeneity, the model, like the model in Shimer (2005), is an otherwise standard matching model. For tractability, worker heterogeneity in the model is binary: there are low-type and high-type workers. The two types differ in two important ways. First, low-type workers have lower net productivity (i.e. the difference between their productivity while employed, y, and their opportunity cost from employment, z). Second, the (exogenous) separation rates for the two types of workers differ both in their average level and in their cyclicality. Low-type workers are assumed to have a higher average separation rate, which implies a higher average unemployment rate and a share of the pool of unemployment that exceeds their share in the population. In addition, the separation rate of low-type workers is assumed to vary in a more strongly countercyclical way, which implies that the share of low-type workers in the pool of unemployed rises during downturns.

Both aspects of low-type workers' share of unemployment—its high average level and the fact that it rises during downturns—may potentially contribute to the volatility of the vacancy/unemployment ratio. Consider first why a high average share of low-type unemployment may contribute to volatility. Shimer (2005) showed in his model that if net productivity $y - z$ is small, then the value of worker–firm matches is in fact sensitive to aggregate productivity, and thus so is vacancy creation. Although, as Shimer argued, a very small net productivity for the average worker is questionable,[2] for some fraction of the population it is quite plausible. Even if that fraction is relatively small, the fact that it can account for a disproportionate share of unemployment suggests that the worker–firm matches that are particularly sensitive to aggregate productivity (i.e. matches with low-type workers) would have a disproportionate impact on firms' expected value from posting a vacancy. Accordingly, incentives for firms to open vacancies would be more sensitive to aggregate conditions.

The changing composition of unemployment can also increase volatility. The value to a firm from meeting with a low-type worker is considerably lower than the value from meeting with a high-type worker. Consequently, during downturns, when the low-type share of unemployment rises, the expected value from filling a vacancy falls.[3] In response, firms open fewer vacancies.

Simulations of the model show that, in terms of volatility of the key labor market statistics, cyclical variation in the composition of the unemployed is quantitatively more important than the average composition. Thus, labor market volatility in the model is more directly driven by the separation rate shocks that change the composition of unemployment than by aggregate productivity shocks. Aggregate productivity shocks are perhaps better seen as an indirect driving force, at least to the extent that they account for the separation rate fluctuations (as in a model with endogenous separations).

For reasonable amounts of variation in the composition of unemployment, the model exhibits as much as three to four times the volatility exhibited in Shimer (2005). Nevertheless, the model's simulations still exhibit, at most, only a third of the volatility that is observed in the data; a substantial amount of volatility remains to be explained by other factors.

The remainder of the paper is organized as follows. Section 2 describes the economic environment. Section 3 characterizes and analyzes the model's equilibrium. Section 4 calibrates the model and utilizes simulations to quanti-

[2] Hagedorn and Manovskii (2005) argue that Shimer's puzzle is really not a puzzle at all; they dispute his parameterization of the model and contend that in fact a small net productivity is plausible (not just for a fraction of the population, as assumed here, but for all workers). Together with a very small bargaining share for workers, this parameterization results in very realistic labor market volatility. However, an important problem with this alternative parameterization, highlighted by Costain and Reiter (2005), is that it implies that the equilibrium unemployment rate is implausibly responsive to very small increases in z (in the form of higher unemployment benefits, for example).

[3] Hall (2005a) explores a complementary model in which firms' effective recruiting costs are countercyclical because workers self-select into better matches during periods of low unemployment. The impact of the changing composition of unemployment in the present model has a similar interpretation in that a firm's expected cost of recruiting a high-type worker is lower during periods of low unemployment.

tatively assess the significance of worker heterogeneity by way of comparison with Shimer (2005) and with US data. Finally, Section 5 offers concluding remarks.

2. The model

The model described here attempts to highlight the role of worker heterogeneity by assuming a rather simple form of heterogeneity. Possible alternative assumptions and generalizations will be discussed below.

The model is in discrete time. There are two types of infinitely lived agents: workers and entrepreneurs. There is a continuum of identical potential entrepreneurs (or 'firms'), with preferences defined by

$$\sum_{t=1}^{\infty} \beta^t (c_t - k v_t) \tag{1}$$

where β is a discount factor and c_t is consumption. The number of vacancies posted in period t is given by v_t. Although firms can costlessly create those vacancies, each period they incur a recruiting cost k per vacancy posted.

There is also a continuum of workers, of total mass equal to one. Workers' preferences are defined by

$$\sum_{t=1}^{\infty} \beta^t \big(c_t + z(1-n_t)\big) \tag{2}$$

where β is a discount factor, c_t is consumption, and $n_t \in \{0, 1\}$ is time spent working. The parameter z measures the opportunity cost associated with working.[4]

For tractability, I assume that there are only two types of workers. The two types are distinguished by their net productivity—i.e. how much they produce when employed, y, relative to their opportunity cost of employment, z. Low-type workers constitute a fraction ν_l of the workforce and have a small net productivity, $y_l - z_l$. High-type workers make up a fraction $\nu_h = 1 - \nu_l$ of the workforce and have a larger net productivity: $y_h - z_h > y_l - z_l$. As we will see below, for the model's results it does not matter whether the differences in net productivity are attributable to heterogeneity in y_i, z_i, or both.

In addition to the type-specific components of productivity, y_l and y_h, there is an aggregate component p that is common across all workers. This aggregate component is assumed to follow a mean-zero discrete state Markov process. The vector of values is given by $\bar{p}$ and the elements of the transition matrix, Π, are given by $\pi_{ij} = \text{prob}\{p' = \bar{p}_j \mid p = \bar{p}_i\}$.

The process by which workers and firms match with each other is described by a constant-returns-to-scale function of the number (measure) of vacancies and the total number of unemployed workers of either type (they are assumed to be perfect substitutes in the search process). Specifically, the number of matches is determined by the function $n(v, u)$, where v is the number of vacancies and $u = u_l + u_h$ is total unemployment. It follows from the constant-returns-to-scale assumption that the rate at which workers match with firms depends only on the ratio of vacancies to unemployed workers, $\theta = \frac{v}{u}$. I let $f(\theta)$ denote the workers' matching rate and $q(\theta)$ the firms' matching rate.

Matched workers and firms experience exogenous separations. I allow separation rates for the two types of workers, s_l and s_h, to differ. In particular, we are interested in the case in which the average value of s_l exceeds the average value of s_h, which implies that low-type workers will on average constitute a share of unemployment that exceeds their share of the population. In order to examine the role of cyclical changes in the composition of the unemployed, I also allow the separation rates to covary negatively with aggregate productivity (for notational ease I do not express s_i as a function of p). More specifically, I allow s_l to covary (negatively) more strongly with the cycle than s_h, so that low-type workers constitute a higher share of unemployment during downturns.

The timing of events within a period is straightforward. At the beginning of the period, p is revealed. Among the workers and firms who are matched, wages are negotiated, production takes place, and wages are paid. After production, exogenous separations occur. Meanwhile, firms also decide whether to open vacancies after observing the aggregate state. The vacancies that are opened seek workers, and vice versa. The matches that result from this process must wait until the following period to produce.

[4] A wide variety of factors, such as the value attributed to leisure, the level of unemployment benefits, and spousal income, determine a worker's z. See Hall (2006) for an analysis of the determinants of z in a search model in which worker preferences depend on both consumption and leisure.

3. Equilibrium

This section describes the model's stochastic equilibrium. The model's main difference relative to the standard model—the worker heterogeneity—expands the set of state variables that are relevant to workers and firms. In addition to the exogenously determined aggregate component of productivity, p, workers and firms must also keep track of the distribution of workers across different labor market states. Workers and firms care about this distribution of workers across employment states because the composition of the pool of unemployed influences the number of vacancies that firms will open, which in turn affects workers' and firms' matching rates.[5]

Because the shares of low-type and high-type workers in the population are constant, workers and firms do not need to track the numbers of workers in all four labor market states. Instead, they must track just two, since the number of employed low-type and high-type workers are given by $e_l = \nu_l - u_l$ and $e_h = \nu_h - u_h$. I let $X = \{p, u_l, u_h\}$ denote the vector of state variables in the equations below. Although I consider separation rates that fluctuate exogenously, it is not necessary to include an additional state variable since the separation rates are assumed to correlate perfectly with p.

Aside from the slightly more complicated state space, analysis of the equilibrium is entirely standard for this class of models. Two key endogenous variables in an equilibrium are the vacancy/unemployment ratio, $\theta(X)$, and the wage, $w_i(X)$. The equilibrium value of $\theta(X)$ in each aggregate state is determined by a free-entry condition. Wages are assumed to be determined by generalized Nash Bargaining. That is, matched workers and firms agree each period on a wage that gives the worker a fraction ϕ of the match surplus that arises from the model's search frictions. As the notation suggests, this equilibrium wage $w_i(X)$ depends both on the worker's type and on the aggregate state.

Analysis of the model's equilibrium begins with the value functions associated with workers and firms—matched and unmatched. The value functions for unemployed and employed workers are given (for $i = l, h$) by

$$U_i(X) = z_i + \beta \mathbb{E}_{X'|X}\big[f\big(\theta(X)\big)M_i(X') + \big(1 - f\big(\theta(X)\big)\big)U_i(X')\big] \tag{3}$$

and

$$M_i(X) = w_i(X) + \beta \mathbb{E}_{X'|X}\big[(1 - s_i)M_i(X') + s_i U_i(X')\big]. \tag{4}$$

The expectations operator $\mathbb{E}_{X'|X}$ depends on the transition matrix Π for aggregate productivity and on the laws of motion for the measures of workers in the different labor market states (described below).

The interpretation of (3) and (4) is straightforward. In (3), the payoff in the current period for a type-i unemployed worker is z_i. With probability $f(\theta(X))$, search in the current period results in a match, yielding an employment value of $M_i(X')$ and with probability $1 - f(\theta(X))$ there is no match and the continuation value is $U_i(X')$. In (4), the employed worker earns the wage $w_i(X)$ and keeps the job in the next period with probability $1 - s_i$ and loses it with probability s_i.

The value to a firm of an open vacancy is given by

$$\begin{aligned} V(X) = -k + \beta \mathbb{E}_{X'|X}\big[&q\big(\theta(X)\big)\big(\mu(X)J_l(X') + \big(1 - \mu(X)\big)J_h(X')\big) \\ &+ \big(1 - q\big(\theta(X)\big)\big)V(X')\big] \end{aligned} \tag{5}$$

where $\mu(X)$ is the faction of the unemployed who are low-productivity workers, which can be calculated directly from the last two elements of X: i.e. $\mu = u_l/(u_l + u_h)$.

$J_i(X)$ is the firm's value from a match with a worker of type i:

$$J_i(X) = p + y_i - w_i(X) + \beta \mathbb{E}_{X'|X}\big[(1 - s_i)J_i(X') + s_i V(X')\big]. \tag{6}$$

The firm produces output $p + y_i$ and pays the wage $w_i(X)$. With probability $1 - s_i$ the match continues, and with probability s_i a separation occurs.

I denote the surplus of a match involving a worker of type i as

$$S_i(X) = J_i(X) + M_i(X) - U_i(X) - V(X). \tag{7}$$

[5] In the standard model without worker heterogeneity, workers and firms do not need to track the distribution of workers across employment and unemployment. Workers and firms only need to know their matching rates, which depend on the ratio v/u. In the standard model, that ratio in equilibrium does not depend on the numbers of employed and unemployed workers; it depends only on the aggregate component of productivity, p.

Consistent with the Nash Bargaining assumption, the equilibrium wage $w_i(X)$ is such that a matched worker receives a fraction ϕ of the surplus: $M_i(X) - U_i(X) = \phi S_i(X)$ and $J_i(X) - V(X) = (1-\phi)S_i(X)$. From Eqs. (3)–(7), together with the equilibrium free-entry assumption, which implies $V(X) = 0$, it follows that the surplus of a match with a worker of type i in state X can be expressed as

$$S_i(X) = p + y_i - z_i + \beta\big(1 - s_i - \phi f\big(\theta(X)\big)\big)\mathbb{E}_{X'|X} S_i(X'). \tag{8}$$

Given this surplus function, an equilibrium of the model can be characterized in a standard way. For each possible state X, there is an equilibrium vacancy/unemployment ratio $\theta(X)$ such that the value from opening a vacancy, $V(X)$, is equal to zero. From (5), this implies the free-entry condition

$$\frac{k}{q(\theta(X))} = \beta(1-\phi)\mathbb{E}_{X'|X}\big(\mu(X)S_l(X') + \big(1-\mu(X)\big)S_h(X')\big). \tag{9}$$

Given the $\theta(X)$ that solves this equation,[6] the evolution of the state X is straightforward. The laws of motion for the last two elements of the state vector, u_l and u_h, are given by

$$u_i' = u_i\big(1 - f\big(\theta(X)\big)\big) + (\nu_i - u_i)s_i \quad i \in \{l, h\}. \tag{10}$$

The first element of the state vector X, i.e. p, evolves exogenously.

3.1. Implications of worker heterogeneity for labor market volatility

There are two potential channels by which worker heterogeneity may enhance labor market volatility relative to the standard model. One relates to the average composition of the pool of unemployed (average μ) and the other relates to variations in that composition.

Volatility in $\theta(X)$ is the primary source of volatility in other labor market indicators such as the job-finding rate $f(\theta)$ and the unemployment rate u. For this reason, it is worthwhile to focus on the determinants of the volatility in $\theta(X)$. From the free-entry condition (9), it is apparent that fluctuations in the equilibrium $\theta(X)$ will be greater the larger are fluctuations in the expected surplus of a new match (the weighted average of $S_l(X)$ and $S_h(X)$ on the right side of (9)). Intuitively, when firms' expected return from forming new matches varies significantly, so too will the number of vacancies (per unemployed worker) that firms create.

As both Shimer (2005) and Hagedorn and Manovskii (2005) showed, the surplus value of a match is more sensitive to aggregate productivity shocks when the net productivity of that match is small, because then a small productivity shock has a greater impact, in percentage terms, on net productivity and consequently on match surplus. In terms of the present model, it follows that the surplus value of a match with a low-type worker, $S_l(X)$, is more sensitive to changes in aggregate productivity than is the value of a match with a high-type worker, $S_h(X)$. This points to the first channel by which heterogeneity can affect volatility: a higher average value of μ places a larger weight on the more volatile surplus $S_l(X)$ and makes the overall expected surplus of a new match more sensitive to changes in aggregate productivity.

There is reason to suspect, however, that this first channel will actually have a limited impact on volatility. To see why, note that $S_h(X)$ is roughly proportional to $y_h - z_h$ and $S_l(X)$ is roughly proportional to $y_l - z_l$,[7] so that $S_h(X)$ will be significantly greater than $S_l(X)$ when $y_h - z_h$ is significantly greater than $y_l - z_l$. Thus, although the surplus value $S_h(X)$ is less volatile than $S_l(X)$, it is also much larger on average, so that the behavior of the overall weighted average of the two surplus values is dominated by the less volatile component, $S_h(X)$—even if the weight on $S_l(X)$, i.e. μ, is quite large.

While this difference in the magnitudes of the two surplus values limits the importance of the first channel by which worker heterogeneity can increase volatility, it actually forms the basis for the second channel. Changes in the composition of the pool of unemployed—changes in μ—result in a shift in the weights on the surplus values $S_l(X)$ and $S_h(X)$ in the free-entry condition. When the difference in the relative magnitudes of the two surplus values is large, even a small shift in the weights can have a significant impact on the overall expected surplus, and thus on

[6] The left side of the free-entry condition is increasing in θ, whereas the right side is decreasing, since the surplus values are decreasing in θ. Thus, in standard fashion, there is a unique θ that solves (9).

[7] To see this, just note that in a steady state, with $p = 0$, the surplus equation in (8) becomes $S_i = (y_i - z_i)/(1 - \beta(1 - s_i - \phi f))$.

the equilibrium vacancy/unemployment ratio. In particular, if μ rises during economic downturns, then the expected surplus of a match declines and firms respond by opening fewer vacancies. Intuitively, if firms benefit much more from meeting a high-type worker, then they will have less incentive to open vacancies when the pool of unemployed contains a higher fraction of low-type workers.

Although a high average value of μ by itself may generate only limited additional volatility, it can actually enhance the volatility that results from variation in μ. To see why, note that a higher average value of μ means a lower weighted average $\mu(X)S_l(X') + (1-\mu(X))S_h(X')$, so that for a given shift in weight from $S_h(X)$ to $S_l(X)$, due to a change in μ, the impact on the weighted average is greater in percentage terms.

It is worth discussing here the assumption that there is no directed search—i.e. there is just one matching function and thus only one free-entry condition. While in reality firms can direct their search toward workers in a particular profession or in a particular geographic region, this paper focuses on the heterogeneity in workers' net productivity that exists within these more narrowly segmented labor markets. Various other search models that feature heterogeneity in worker productivity—for example, Acemoglu (1999), Shimer (2001), and Dolado et al. (2003)—also make the random matching assumption. Acemoglu (1999) argues for the plausibility of random matching in an environment with heterogeneous skill levels by pointing out that skill is imperfectly correlated with observable characteristics, such as years of education, which makes it difficult for firms to target recruiting exclusively at workers with a particular skill level. This argument is even more forceful in the context of the current paper, since workers are heterogeneous not only in their skill or productivity, y, but also in their opportunity cost of employment, z, which is even less likely than y to be observable by firms.[8]

4. Quantitative results

This section reports various simulations of the model in order to highlight and distinguish the ways in which worker heterogeneity affects the model's volatility. The analysis and discussion is divided into several subsections. First, the chosen parameter values are discussed. The calibration follows Shimer (2005) as closely as possible, in order to allow direct comparison. Second, the role of the *average* composition of unemployment is examined. Third, the role of *variation* in the composition of the unemployed is considered. The fourth subsection provides a more comprehensive discussion of some of the implications of the simulation results and discusses some possible generalizations of the model. Finally, the quantitative importance of the results is discussed in the context of an assessment of the empirical evidence on the composition of the unemployed.

4.1. Calibration

Table 1 summarizes the parameter values used in the quantitative results discussed below. While some parameters are varied across different simulations, in order to evaluate the way in which the composition of unemployment affects volatility, several parameters are held constant across the different parameterizations. Each period represents a month and so the discount factor is set to $\beta = 0.996$ in every parameterization, which corresponds to an annual discount rate of approximately 5%. The matching function has a Cobb–Douglas specification: $n(v,u) = \gamma v^{\eta} u^{1-\eta}$. Following Shimer (2005), who used monthly data on job-finding rates and the vacancy/unemployment ratio to estimate the matching function parameters, I set $\eta = 0.28$. I also choose a value for the workers' bargaining share that matches Shimer's choice: $\phi = 0.72$. In the baseline model, this value satisfies the efficiency condition in Hosios (1990): $\phi = 1-\eta$.[9]

The Markov process for the aggregate component of productivity, p, is assumed to have 20 states and the vector of values, $\bar{p}$, and the transition matrix, Π, are chosen so that p approximates an AR(1) process with a mean of zero. As in Shimer (2005), the autocorrelation and the standard deviation of the process are chosen to match (after aggregating

[8] Though firms may be able to identify a worker's y or z once they meet, as assumed in the model, for the question of the feasibility of directed search, what matters is whether y or z can be accurately identified prior to meetings, which is less likely.

[9] The Hosios condition does not hold here. In the standard model, searching workers and firms inflict search externalities on each other. The Hosios condition identifies the worker's share of the surplus, ϕ, that would cause the workers' and firms' externalities to offset in an optimal way. In the model with two worker types, in addition to the standard externalities the two types of workers also cause externalities for the other type, and optimality can no longer be reduced to a simple condition on the bargaining parameter ϕ.

Table 1
Parameter values

Parameter	Description	Value
β	Monthly discount factor	0.996
ϕ	Worker's bargaining share	0.72
γ	Matching function coef.	0.45
η	Matching function elasticity	0.28
k	Recruiting cost	0.136–0.211
ν_l	Proportion low-type workers	0.2
z_l, z_h	Value of leisure	0.4
y_l	Low-type productivity	0.45
y_h	High-type productivity	1.0–1.14
s_l s_h	Exogenous separation rate	(see text)
$\bar{p}, \Pi$	Process for agg. prod.	(see text)

to a quarterly frequency) the autocorrelation (0.878) and standard deviation (0.02) of quarterly US data on real average output per worker in the non-farm business sector.

The parameters c and γ are guided by targets for the average equilibrium values of θ and of the worker's job-finding rate. As in Shimer (2005), I target an average monthly job-finding rate for workers of 0.45. Reliable evidence on an appropriate target value for the average of θ is scarce. Moreover, given a target for the average job-finding rate, the average value of θ has no independent significance for the results of the model. Thus, as in Shimer (2005), I simply target an average value for θ of 1.0. Given these targets, I set $\gamma = 0.45$ in all of the parameterizations and adjust c in each parameterization in order to achieve the target for θ.

Because of the assumed coarseness of the heterogeneity, with just two worker types, there is no straightforward empirical counterpart that allows me to directly calibrate the share of low-productivity workers, ν_l, the productivity parameters y_l and y_h, or the opportunity costs z_l and z_h. I assume that low-type workers constitute 20% of the workforce. From the perspective of the model's implications for labor market volatility, whether heterogeneity in net productivity $y - z$ comes from y or from z, or from both, is not important. Nevertheless, I still must specify values for all four parameters, as opposed to just setting values for the two net productivities, because of the need to pin down the average level of productivity. Thus, for simplicity, I assume that $z_l = z_h$ so that all heterogeneity comes from y.

In choosing the productivity and opportunity cost parameters, I aim to make choices that are in some way comparable to the calibration of the homogeneous worker modeled in Shimer (2005), which set $y = 1$ and $z = 0.4$. Given this objective, I set $z_h = z_l = 0.4$. There are a couple possible approaches for setting the productivity parameters y_h and y_l. One approach would be to parameterize the high-type workers so that they are identical to the agents in Shimer (2005) ($y_h = 1$) and then parameterize the low-type workers so that their productivity is only slightly higher than their opportunity cost (for example, $y_l = 0.45$). An alternative approach would be to again make low-type workers only slightly more productive than their opportunity cost ($y_l = 0.45$), and then adjust y_h so that the average productivity of the two types of workers (taking into account that their employment rates differ endogenously) is equal to Shimer's assumed labor productivity (i.e. 1). I choose the latter approach.

For the separation rates, s_l and s_h, I consider several different parameterizations. In each of them, however, the values for s_l and s_h are chosen so that the average aggregate separation rate in the simulations is 0.033 (consistent with the quarterly separation rate of 0.1 in Shimer, 2005).

4.2. *The role of the* average *composition of unemployment*

Table 2 presents results from simulations of the model that demonstrate how labor market volatility is affected by the *average composition* of the unemployed. For these simulations, variation in the composition of unemployment is negligible.[10] The first column shows the standard deviations of several key labor market statistics in US data (these

[10] Specifically, there is no variation in μ attributable to differential volatility in the two separation rates. It is possible, at least in principal, that the composition of unemployment still fluctuates, since movements in the job-finding rate will affect the two unemployment rates in different ways when their average levels differ. However, in practice fluctuations in μ in these simulations are virtually non-existent.

Table 2
The role of *average* μ

	US data	Baseline	$\bar{\mu} = 0.20$	$\bar{\mu} = 0.35$	$\bar{\mu} = 0.50$
prod.	0.020	0.020	0.020	0.020	0.020
		(0.002)	(0.002)	(0.002)	(0.002)
$f(\theta)$	0.118	0.009	0.009	0.011	0.013
		(0.001)	(0.001)	(0.001)	(0.002)
$\theta = v/u$	0.382	0.031	0.033	0.039	0.046
		(0.004)	(0.004)	(0.005)	(0.006)
u	0.190	0.007	0.008	0.010	0.011
		(0.001)	(0.001)	(0.001)	(0.002)
v	0.202	0.024	0.025	0.030	0.035
		(0.003)	(0.003)	(0.004)	(0.005)

Note: The table reports standard deviations of the key labor market indicators, computed after taking natural logs and removing a Hodrick–Prescott trend, with smoothing parameter 10^5. The numbers in parenthesis indicate the bootstrapped standard errors. The various columns report results for the US data, for the benchmark (Shimer) model, and, in the last three columns, for parameterizations of the model with different *average* compositions of the unemployed.

statistics are taken from Table 1 in Shimer, 2005). The next four columns give results for four different parameterizations of the model. To create each set of statistics, the equilibrium described in the previous section was solved by numerical methods (details of the computational algorithm are provided in Appendix A). Then, given this solution, 1000 samples of 736 observations were simulated. The first 100 observations of each sample were discarded to eliminate sensitivity to initial conditions. For each sample, the remaining 636 observations were aggregated up to quarterly averages, yielding 212 quarters of 'data' per sample. As in Shimer (2005), the natural logs of the quarterly data were detrended using a Hodrick–Prescott filter with smoothing parameter 10^5. The table presents the mean (and, in parenthesis, the bootstrapped standard errors) of the samples' standard deviations.

The 'baseline' model in the second column is precisely Shimer's model and calibration, with no worker heterogeneity. That is, I set $\nu_l = 0$, $y_h = 1$, and $s_h = 0.033$.[11] A quick comparison of this baseline model with the US data makes apparent the dramatic lack of amplification emphasized by Shimer (2005). Although the productivity shocks that drive the model's fluctuations have the same serial correlation and standard deviation as in the data, the baseline (Shimer) model's key labor market statistics all exhibit standard deviations that are too small by a factor of at least 10.

The third column in Table 2 shows the results for a parameterization of the model in which 20% of the workforce are low-type workers ($\nu_l = 0.2$) whose low net productivity ($y_l - z_l = 0.05$) means that the surplus value of their matches are sensitive to small changes in aggregate productivity. Matches with the other 80%, by contrast, are relatively insensitive to small changes in aggregate productivity. The two types of workers have identical separation rates ($s_l = s_h = 0.033$) and job-finding rates, and thus have identical unemployment rates as well. Consequently, the low-type workers' share of unemployment μ is constant and equal to their share in the workforce, $\nu_l = 0.2$.[12]

The results for this parameterization show essentially no additional amplification relative to the (Shimer) baseline. Following the discussion above in Section 3.1, this result is not surprising. Although the weight on $S_l(X)$ in the free-entry condition is 0.2, the surplus value of a high-type worker $S_h(X)$, which exhibits very little responsiveness to aggregate conditions, dominates the weighted average of the two types of workers because it has a much higher average value. With minimal extra variation in the expected surplus of a match, θ is essentially no more volatile than if all workers were high-type workers.

A few simple calculations help to illustrate this point. For these simulations reported in the third column, the mean and standard deviation of $S_l(X)$ are 0.14 and 0.07, respectively. $S_l(X)$ is clearly very volatile in the sense that a one standard deviation increase, starting from its mean, corresponds to a 50% increase. By contrast, the mean and standard deviation of $S_h(X)$ are 2.05 and 0.05, so that one standard deviation increase in $S_h(X)$ corresponds to merely a 2.4%

[11] Given the identical calibration strategy for this baseline case, it is reassuring that the numbers reported here are extremely close to the numbers reported by Shimer. The very minor differences in results—Shimer gets standard deviations for $f(\theta)$, θ, and u of 0.010, 0.035, and 0.009, respectively—are likely due to differences in solution methods (he solves the continuous time version of the model).

[12] Similarly, low-type workers, with productivity $y_l = 0.45$ account for a share 0.2 of employment, and so y_h is set to 1.1375 in order to achieve, as described above in the calibration discussion, an (employment weighted) average productivity of $0.2(0.45) + 0.8(1.1375) = 1$.

increase. In terms of the weighted average of the two surplus values, a one standard deviation increase in both $S_l(X)$ and $S_h(X)$ raises the weighted average from 1.67 to 1.72. This represents a 3.0% increase, which is only minimally greater than the 2.4% increase that would occur if all workers were high-type. So even though low-type workers account for 20% of unemployment, they clearly have a negligible impact on the volatility of the expected surplus value, relative to a model with no heterogeneity.

The fourth and fifth columns examine what happens as $\bar{\mu}$ increases. A higher value of μ can result either from an increase in ν_l or because, holding ν_l fixed, low-type workers represent a share of unemployment that is large relative to their share in the population. While these two possibilities are associated with slightly different interpretations of the nature of heterogeneity, they differ minimally in terms of the implications for volatility.[13] In the results in the fourth and fifth columns, I choose to hold $\nu_l = 0.20$ and raise $\bar{\mu}$ by increasing the s_l relative to s_h and thus increasing u_l relative to u_h. Specifically, in the fourth column the separation rate for low-types s_l (0.062), is raised relative to s_h (0.0265) so that the average share of low-type workers among the unemployed is $\bar{\mu} = 0.35$. In the fifth column, the difference between s_l and s_h is increased further—$s_l = 0.093$ and $s_h = 0.02$—so that the pool of unemployed is even more heavily populated by low-type workers, with $\bar{\mu} = 0.5$.[14]

While the results in these last two columns do show some additional volatility, the difference is small. As in the calculations discussed above, even when low-type workers account for half of unemployment (the fifth column), the expected surplus of a match is only slightly more volatile than $S_h(X)$. For that parameterization, a one standard deviation increase in $S_h(X)$ represents a 2.7% increase, while a one standard deviation increase in both $S_l(X)$ and $S_h(X)$ results in a 5.2% increase in the weighted average of the surplus values. The fact that the weighted average varies slightly more than $S_h(X)$ accounts for the slightly greater labor market volatility of this case relative to the baseline. Nevertheless, these results suggest that a very high average value of μ (and thus a very small weight on $S_h(X)$) would be necessary in order for the volatility of $S_l(X)$ to have a substantial impact on overall volatility.[15]

4.3. The role of variation *in the composition of unemployment*

Whereas the last section showed that the *average* composition of the pool of unemployed has a limited impact on the degree of amplification of shocks, this section will show that *variation* in the composition of the pool of unemployed can have a substantially greater impact on the degree of labor market volatility.

The first two columns of Table 3 again report the US data and simulation results for the baseline parameterization without heterogeneity. The remaining columns show results when the share of low-type workers among the unemployed, μ, fluctuates countercyclically. These fluctuations in μ arise because the increase in separation rates during downturns is greater for low-type workers. In particular, for simplicity s_l is assumed to be perfectly negatively correlated[16] with aggregate productivity, while s_h does not fluctuate.[17]

In the third and fourth columns of Table 3, the average value of μ is 0.35, and its standard deviation is 0.02 and 0.04. Even for a small standard deviation of μ—just 0.02, or 6% of the average value—the standard deviation of θ and of $f(\theta)$ is more than double that of the baseline. With a slightly higher standard deviation (0.04) for $\sigma(\mu)$, the volatility of θ and $f(\theta)$ is more than tripled. In both cases, the increase in the variation of the unemployment rate is even greater than the increase in variation of $f(\theta)$ or of θ, since the countercyclical separation rate (which is absent in the baseline) also contributes to movements in u.

As explained in Section 3.1, the free-entry condition implies that variation in μ generates volatility in θ (and thus in $f(\theta)$) by way of its impact on the expected surplus value of a match, $\mu(X)S_l(X') + (1 - \mu(X))S_h(X')$. Small

[13] If $\bar{\mu}$ is higher due to a higher s_l, then the surplus value $S_l(X)$ is also affected by the change in s_l, as can be seen in Eq. (8).

[14] Given these separation rates, y_h is set at 1.128 for the $\bar{\mu} = 0.35$ case and at 1.121 for the $\bar{\mu} = 0.5$ case in order to attain an employment weighted average productivity of 1.0.

[15] This result is not surprising in light of the observation by Hall (2006) that net productivity $y - z$ affects labor market volatility in a highly non-linear way, with volatility increasing substantially only when net productivity is very small. Here, only when $\bar{\mu}$ is very large does the *average* net productivity (and thus the average match surplus) of new matches become sufficiently small to generate substantial volatility.

[16] To allow s_l to be imperfectly correlated with p would require another state variable.

[17] The means of s_l and s_h, are the same as they were in the previous section and y_h is again set to achieve an employment weighted average productivity of 1.0. That is, $E(s_l) = 0.062$, $E(s_h) = 0.0265$, and $y_h = 1.128$ for the cases with $\bar{\mu} = 0.35$; $E(s_l) = 0.093$, $E(s_h) = 0.02$, and $y_h = 1.121$ for the cases with $\bar{\mu} = 0.5$. The standard deviation of s_l is approximately 0.01 for the cases with $\sigma(\mu) = 0.02$ and is 0.016 for the cases with $\sigma(\mu) = 0.04$.

Table 3
The role of *variation* in μ

	US data	Baseline	$\bar{\mu} = 0.35$		$\bar{\mu} = 0.50$	
			$\sigma(\mu) = 0.02$	$\sigma(\mu) = 0.04$	$\sigma(\mu) = 0.02$	$\sigma(\mu) = 0.04$
prod.	0.020	0.020	0.020	0.020	0.020	0.020
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$f(\theta)$	0.118	0.009	0.021	0.031	0.027	0.039
		(0.003)	(0.004)	(0.004)	(0.004)	(0.005)
$\theta = v/u$	0.382	0.031	0.076	0.110	0.096	0.139
		(0.010)	(0.015)	(0.015)	(0.014)	(0.018)
u	0.190	0.007	0.052	0.089	0.065	0.113
		(0.007)	(0.013)	(0.013)	(0.010)	(0.016)
v	0.202	0.024	0.027	0.026	0.035	0.033
		(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
sep. rate	0.075	–	0.036	0.068	0.045	0.086
			(0.005)	(0.009)	(0.006)	(0.011)

Note: The table reports standard deviations of the key labor market indicators, computed after taking natural logs and removing a Hodrick–Prescott trend, with smoothing parameter 10^5. The numbers in parenthesis indicate the bootstrapped standard errors. In addition to the US data and the benchmark model in the first two column, the last four columns presents results for four combinations of average μ and standard deviation of μ.

changes in μ lead to changes in this expected surplus value that are proportional to the difference between $S_l(X)$ and $S_h(X)$. A few simple calculations can demonstrate the quantitative nature of this effect. For the case with $\bar{\mu} = 0.35$ and $\sigma(\mu) = 0.04$, reported in the fourth column of Table 3, the average values of $S_l(X)$ and $S_h(X)$ in the simulations are 0.13 and 2.17, respectively. Thus, the weighted average of those two values is $0.35(0.13) + 0.65(2.17) = 1.456$. Holding the two surpluses values at their averages, so as to focus on the impact of a change in μ, a one standard deviation increase in μ—an increase from 0.35 to 0.39—decreases the weighted average of the two surplus values to 1.374. This change represents a 6% decline. This additional variation in the expected surplus value that is due to changes in μ increases the overall volatility of the expected surplus beyond that which can be attributed solely to the variation in the two surplus values (discussed in the previous subsection). Moreover, these results highlight the fact that, in terms of the impact of variation in μ on variation in the vacancy/unemployment ratio, a large difference between $y_h - z_h$ and $y_l - z_l$ (which accounts for the large difference in surplus values) is more important than a small value of $y_l - z_l$ by itself.

While the previous subsection showed that even a relatively high average μ by itself generates very little additional volatility relative to the baseline, the final two columns of Table 3 show that a higher average value of μ serves to enhance the volatility that results from variation in μ. With $\bar{\mu} = 0.5$ and a standard deviation of 0.04, variations in $f(\theta)$ and θ are over four times greater than in the baseline (though the fluctuations still fall short of the variation observed in US data by a factor of about three). Another simple calculation helps to demonstrate why a high average μ reinforces the volatility due to changes in μ. For the results shown in the last column of Table 3, the average values of $S_l(X)$ and $S_h(X)$ are 0.12 and 2.21, respectively. These values imply a weighted average of $0.5(0.12) + 0.5(2.21) = 1.165$. A one standard deviation increase in μ (from 0.5 to 0.54) reduces that weighted average by 0.084 to 1.081—a 7.8% decline. This decline in the weighted average is greater, in percentage terms, due to the fact that the relatively large weight (0.5) on 0.12 reduced the initial weighted average.

In addition to the statistics for unemployment, vacancies, and job-finding rates, the bottom row of Table 3 also shows the standard deviation of the (log) aggregate separation rate for the different parameterizations of the model. Comparing these with the value for the US data suggests that the parameterization with $\bar{\mu} = 0.5$ and with $\sigma(\mu) = 0.04$ probably establishes the upper bound on the amount of amplification that can be attributed to variation in the composition of the unemployed, at least in terms of variation in μ that can be attributed to variation in the separation rates of low-type workers.[18]

[18] Indeed, Shimer (2005) argues that the standard deviation of the separation rate has declined during the last two decades, so that the standard deviation for the entire period, 0.075, may overstate the current volatility of the separation rate.

4.4. Discussion

Taken together, the above simulation results point to some implications that deserve further discussion. One implication relates to the driving forces that account for volatility. The more volatile fluctuations in θ and in $f(\theta)$ in the simulations just discussed are for the most part not driven, at least not in a direct sense, by shocks to aggregate productivity, p. Instead, the more direct driving force behind the enhanced volatilities are the fluctuations in separation rates (in particular, s_l) that cause the countercyclical changes in μ. This distinction between the driving forces is somewhat obfuscated in the model by the assumption that separation rates are perfectly negatively correlated with aggregate productivity. Of course, to the extent that fluctuations in separation rates result from movements in aggregate productivity (as they would in a model with endogenous separations), the volatility that stems from fluctuations in μ can still by viewed as indirectly driven by aggregate productivity.

The prominent role of separation rates as the driving force behind fluctuations contrasts rather sharply with results in Shimer (2005). In the simulations in which he considers fluctuations driven by separation rate shocks, Shimer finds almost no volatility in θ and $f(\theta)$. This is not surprising, however, since in his model there is no changing composition of unemployment associated with the movements in the separation rate. In addition to generating negligible volatility in θ and $f(\theta)$, fluctuations driven by separation rate shocks in Shimer (2005) imply—counterfactually—a positively sloped Beveridge Curve. To understand why this occurs in his model, note that while the fluctuations in the separation rate cause significant fluctuations in u, the fact that $\theta = v/u$ fluctuates very little implies that v must covary positively with u (and the Beveridge Curve slopes upward). In the current model, by contrast, the simulations reveal a negatively sloped Beveridge Curve that is more consistent with the data. For the simulations shown in Table 3, the correlation between u and v was about -0.85 for the cases with $\sigma(\mu) = 0.02$ (and -0.75 for the cases with $\sigma(\mu) = 0.04$), and the average slope of the Beveridge Curve was about -0.45 for the cases with $\sigma(\mu) = 0.02$ (and -0.22 for the cases with $\sigma(\mu) = 0.04$). The negatively sloped Beveridge Curve arises because vacancies are much more responsive to aggregate conditions. That is, during a downturn the decline in θ results not only from an increase in u, but also from a decrease in v.

It is also worthwhile to discuss the volatility of wages, since in Shimer's baseline model the minimal labor market volatility results from the fact that productivity movements pass through to wages almost entirely, leaving little additional incentive for firms to open more vacancies when productivity is high. As evidence of this, for the simulation of the baseline model reported in Table 2, the (log) standard deviation of wages is 0.019, which is very close to the standard deviation of the driving force p (0.02). The additional labor market volatility in the model with heterogeneous workers is not the result of significantly less responsive wages. This can be seen by comparing the baseline with the simulation that generated the greatest labor market volatility (the case with $\bar{\mu} = 0.5$ and $\sigma(\mu) = 0.04$, reported in the last column of Table 3). The wage of low-type workers in that case is modestly more volatile than wages in the baseline, with a (log) standard deviation of 0.041, while the wage of high-type workers is slightly less volatile, with a standard deviation of 0.017. Due to composition effects (low-type workers' share of employment falls during downturns), the standard deviation of average (employment weighted) wages is slightly lower than both, with a standard deviation of 0.016.

While the approach taken in this paper has been to keep the model relatively simple in order to highlight the effects of worker heterogeneity, it is worth commenting briefly on some potentially interesting generalizations or alternative assumptions. While separations have been assumed to occur exogenously in the model, it is possible to endogenize separations by incorporating a stochastic match-specific component of productivity, as in Mortensen and Pissarides (1994).[19] In such a setting, low-type workers' smaller (average) net productivity $y - z$ would place them generally closer to the destruction margin and would make them more likely to draw a match-specific productivity below the separation threshold. Thus, low-type workers' separation rates would be consistently higher than the separation rates of high-type workers, just as I assumed above in order to generate higher values of $\bar{\mu}$

Separation rates in such a model would also vary countercyclically, as I have assumed above, because the fraction of matches with match-specific productivity that falls below the separation threshold increases when the aggregate

[19] I choose not to introduce endogenous separations in this way since it would complicate the model dramatically by turning the distributions of low-type and high-type workers over match-specific productivities into state variables. This occurs because potential firms need to know those distributions in order to forecast the composition of the unemployed, which in turn affects the value of a vacancy. As a result, numerical solutions of the model run into the "curse of dimensionality." See Pries (2006), nevertheless, for a solution to such a model.

component of productivity falls.[20] Furthermore, the separation rates of the workers who are generally closer to the separation threshold—i.e. low-type workers—would vary more strongly with aggregate productivity than the separation rates of high-type workers. This too is consistent with the assumptions made above.

It is also worth mentioning two other factors, not considered here, that could contribute to cyclical movements in μ. First, job-finding rates for the two types of workers could differ in how they vary over the cycle. If low-type workers' job-finding rates fell more during downturns, relative to high-type workers, their share of unemployment would rise. One would expect this to occur in a matching model with a match-specific component of productivity and a match acceptance decision.[21] That is, if the initial match-specific component of productivity were drawn randomly, then even if the two types of workers meet vacancies at the same rate the low-type workers would have a smaller probability—particularly when aggregate productivity is lower—of receiving a draw above their acceptance threshold.

Second, suppose that a worker's net productivity is not fixed, but instead can change either from high to low, or vice versa. Evidence on the scarring effects of job displacement, such as in Jacobson et al. (1993), suggests that highly productive workers indeed often emerge from a job loss looking much more like low-productivity workers, due to loss of human capital, for example. If this kind of job loss were more likely during downturns, due perhaps to increased sectoral restructuring, then clearly μ would rise when aggregate conditions deteriorated.

4.5. Assessment of quantitative importance of heterogeneity

In the above simulations, I illustrate an important channel through which worker heterogeneity can increase the volatility of labor market fluctuations. Relatively small cyclical changes in the composition of unemployment can have a significant effect on firms' vacancy creation decisions, and thus on job-finding rates and the unemployment rate. The results also show that worker heterogeneity by itself is not enough to generate significant labor market volatility—even with a relatively high average μ, but no variation in μ, there is little volatility. However, when the composition of unemployment does vary, the average composition of unemployment can further enhance volatility.

While Table 3 provides results for parameterizations with a range of values for the average and standard deviation of μ, a more precise assessment of the actual quantitative importance of worker heterogeneity, in terms of its impact on the amplification of shocks, would require compelling evidence on those two moments. To get that evidence would require data that allowed workers to be classified according to their type, i.e. their $y - z$.

Unfortunately, several issues hinder this kind of classification of workers by type. First, workers' observable characteristics, such as years of education or work experience, account for only a limited portion of cross-sectional differences in productivity, making it difficult to accurately identify workers' productivity. Second, heterogeneity in z, which according to the model is potentially as significant as heterogeneity in productivity, is even more difficult to identify. Third, in reality a worker's y and z change over time, though for simplicity the model has assumed that they are fixed. Finally, although the model assumes just one labor market (and one matching function), in reality there are more narrowly segmented markets related to different geographic regions, professions, etc. Accordingly, the relevant measure of worker heterogeneity that we would like to obtain would be the degree of worker heterogeneity within these more narrowly defined markets. This, however, is not feasible with the available data.

While these obstacles impede a more precise calibration of the model, a somewhat crude look at the existing empirical work that has attempted to classify workers according to their productivity can nevertheless still provide us with a rough assessment of the extent of worker heterogeneity, particularly as it pertains to the composition of unemployment. In terms of the *average* composition of the unemployed, the evidence suggests very clearly that low-productivity workers account for a disproportionate share of unemployment. For example, using education as a proxy for productivity, Nagypál (2001) examines the March supplement of the CPS for the years 1970–1997 and finds that workers with less than a high school diploma have unemployment rates that are generally three to four times as high as unemployment rates among college educated workers. Nickell and Bell (1995) also examine, using panel data on unemployment rates for a cross section of OECD countries, the unemployment rates for workers in different education

[20] However, in the model with endogenous separations, the correlation with aggregate productivity is not -1. Instead, in that model higher separations occur in short spikes as aggregate productivity initially deteriorates. Accordingly, the separation rates in the endogenous separation rate model also exhibit lower serial correlation. Nevertheless, the implications for the composition of the pool of unemployed, which is the matter of primary interest, are largely the same as for the model with exogenous separation rate movements.

[21] Indeed, this is the case in the model with match-specific productivity that is solved in Pries (2006).

categories. With only one exception (Japan in the 1970s) they find that workers in their 'low education' category have unemployment rates at least twice as high as the unemployment rates among 'high education' workers. In several instances, the ratio of unemployment rates for the two types of workers exceeds five. For the US, the average ratio is greater than four.

These studies also indicate that when overall unemployment rises, it rises much more for workers with lower productivity. Nagypál (2001) shows that the unemployment rate among workers with less than a high school diploma more than tripled between 1973 and 1983, whereas for workers with a college degree, it merely doubled. In terms of unemployment rate volatility for the entire 1970–1997 period, the coefficient of variation for the unemployment rate is 0.284 for workers with less than a high school diploma, compared to 0.219 for workers with a college degree. The evidence in Nickell and Bell (1995) also suggests greater volatility in unemployment for workers with less education. This is especially true in their results for the US. For example, the ratio of the low-education unemployment rate to the high-education unemployment rate in the US rose from 3.1 in the early 1970s to 4.5 in the early 1980s. Of course, interpretation of this evidence must bear in mind that it is difficult to identify how much of the variation represents a truly cyclical component and how much represents low-frequency trends.

As suggested above, comparisons of the model's simulations with these various pieces of evidence must also be mindful of the fact that this evidence conveys information about economy-wide worker heterogeneity, whereas the more appropriate evidence would be information about the heterogeneity of workers within some appropriately defined labor submarkets. With these caveats in mind, Table 4 reports the average unemployment rates of the two types of workers, as well as the standard deviation of u_l, for the four different parameterizations of the model that were reported in Table 3. (For each parameterization of the model, the standard deviation of u_h is less than half a percentage point.) For the two parameterizations with $\bar{\mu} = 0.35$, the average unemployment rate of low-type workers is slightly above 0.12, or a little over two times the average unemployment rate of high-type workers (0.056). For the two parameterizations with $\bar{\mu} = 0.5$, the average u_l (about 0.175) is about four times the average for u_h (0.043). Again, the evidence cited above suggested a ratio between three and four, suggesting that the parameterizations with $\bar{\mu} = 0.50$ are near the upper bound for $\bar{\mu}$ (at least, in terms of heterogeneity stemming only from y, since accounting for heterogeneity in z could yield an even greater disparity between the average unemployment rates of different worker types).

Another way to evaluate these average values of u_l implied by the model's different parameterizations is to compare them with the average unemployment rates of demographic groups that are likely to contain a high share of low net productivity workers. For example, for males aged 16–24, the average unemployment rate between 1972–2003 was 0.131. For male and female African-Americans, aged 20–24, the average unemployment rate was 0.210. The values of $\bar{u}_l$ in the simulations lie in the neighborhood of these values.

Table 4 also provides some insight into whether the *variation* in the composition of unemployment that is observed in the simulations is in line with the available evidence. Some simple calculations show the impact of a one-standard deviation increase in u_l (and no increase in u_h, since its standard deviation in the simulations is very small) on the ratio of low-type unemployment to high-type unemployment. For the parameterization with $\bar{\mu} = 0.35$ and $\sigma(\mu) = 0.04$, a one standard deviation increase in u_l (starting from its average value) causes an increase in the u_l/u_h ratio from $0.123/0.056 = 2.20$ to $(0.123 + 0.024)/0.056 = 2.63$. For the parameterization with $\bar{\mu} = 0.5$ and $\sigma(\mu) = 0.04$,

Table 4
Average unemployment rates and the standard deviation of u_l

	$\bar{\mu} = 0.35$		$\bar{\mu} = 0.50$	
	$\sigma(\mu) = 0.02$	$\sigma(\mu) = 0.04$	$\sigma(\mu) = 0.02$	$\sigma(\mu) = 0.04$
$\bar{u}_h$	0.056	0.056	0.043	0.043
	(0.001)	(0.001)	(0.001)	(0.001)
$\bar{u}_l$	0.122	0.123	0.173	0.175
	(0.006)	(0.010)	(0.008)	(0.013)
$\sigma(u_l)$	0.013	0.024	0.017	0.030
	(0.002)	(0.003)	(0.002)	(0.004)

Note: This table reports the mean unemployment rates for the two types of workers, along with the standard deviation of u_l, for the four simulations reported in the final columns of Table 3. The standard deviation of u_h in each simulation is negligible, and thus not reported. The numbers in parenthesis indicate the bootstrapped standard errors.

a one standard deviation increase in u_l would cause the ratio to increase from $0.175/0.043 = 4.07$ to $(0.175 + 0.03)/0.043 = 4.76$. These increases are modest relative to the magnitude of the changes reported in Nickell and Bell (1995) and discussed above, though again the changes in Nickell and Bell (1995) may contain both trend and cyclical components.

We can also compare the standard deviation of u_l with the standard deviation of unemployment for demographic groups that are likely to contain a high share of low net productivity workers. In Table 4, the standard deviations of u_l for the different parameterizations of the model range from 0.013 to 0.030. For comparison, the standard deviation of the unemployment rate for males aged 16–24 was 0.023; for male and female African-Americans, aged 20–24, it was 0.045.

Though the evidence discussed here does not allow for a precise calibration of worker heterogeneity in the model, at a minimum it suggests that, in terms of both the average composition of unemployment and variation in the composition, the magnitudes evaluated in the simulations are reasonable ones to consider. Moreover, to the extent that the evidence discussed here only relates to heterogeneity in productivity, and fails to take into account heterogeneity in z that we cannot observe, the parameterizations considered here might be viewed as conservative.

5. Concluding remarks

Fluctuations in the unemployment rate are large and persistent. There is a rather broad consensus that fluctuations in unemployment are the outcome of an economic environment with pervasive frictions that is constantly buffeted by shocks. When shocks hit the economy that require that workers be reallocated across firms, industries, and regions, the frictions impede the reallocation. This can explain, at least qualitatively, why the unemployment rate jumps following a shock, and thereafter adjusts downward very slowly. Nevertheless, the search and matching models that economists use to describe and analyze this process of reallocation struggle quantitatively to account for the observed amplitude and persistence of fluctuations.

This paper has shown that a standard search and matching model modified to allow for worker heterogeneity can generate greater (relative to the standard model without heterogeneity) cyclical movements in key labor market statistics—the vacancy/unemployment ratio, the job-finding rate, and the unemployment rate. The composition of the pool of unemployed lies at the heart of this enhanced volatility. The pool of unemployed consists disproportionately of marginal workers with low net productivity, particularly during economic downturns. These composition effects influence firms' incentives to create vacancies, resulting in a more volatile vacancy/unemployment ratio.

The quantitative exercises suggest that the model with worker heterogeneity can plausibly generate up to three or four times the amplification of the baseline model without heterogeneity, examined in Shimer (2005). However, it still falls short, by a factor of at least three, of explaining the level of volatility observed in US data; thus, part of the puzzle uncovered by Shimer (2005) remains to be explained by other factors.

Acknowledgments

I am especially grateful to Richard Rogerson, the editor, for suggestions and discussion that improved the paper significantly. I also thank an anonymous referee, Robert Hall, Robert Shimer and seminar participants at the 2006 NBER Summer Institute, the University of Maryland, and the Federal Reserve Bank of San Francisco for helpful comments.

Appendix A. Computational algorithm

This appendix provides details of the algorithms used to compute the model's equilibrium and to generate simulated data. As discussed in the text, the aggregate component of productivity, p, is a discrete-state Markov process. I assume a 20-state process, whose vector of values and transition matrix are chosen to approximate an AR(1) process with autocorrelation $\rho(p)$ and standard deviation $\sigma(p)$. The other two aggregate state variables, u_l and u_h, are inherently continuous and must be discretized. I assumed a uniform grid of 50 values for both, so that the state space has dimension $20 \times 50 \times 50$. The minimum and maximum values of the grids for u_l and u_h vary with each different parameter set considered.

Given starting values for the $(20 \times 50 \times 50)$ matrix $\theta(X)$, the two surplus functions $S_l(X)$ and $S_h(X)$ are jointly solved by value function iteration. For a given point in the state space, the laws of motion for u_l and u_h, given in Eq. (10), determine the next period's values of those values, $u_l{}'$ and $u_h{}'$. Because those values will generally lie between gridpoints, the values of the surplus function at $u_l{}'$ and $u_h{}'$ are computed by two-dimensional linear interpolation.

Given solutions for the value functions, the values of $\theta(X)$ that solve the free-entry condition (for each point in the state space) are computed. The matrix of values for $\theta(X)$ is updated by taking a convex combination of the values that solve the free-entry condition and the previous iterate's values. The surplus functions are re-solved and the process is repeated until the values in the $\theta(X)$ matrix converge (i.e. the free-entry condition is satisfied).

To simulate each sample of 736 observations, a Markov chain with 736 values of p is generated. In the cases in which separation rates covary with p, the Markov chain implies a sequence of separation rates as well. Starting values for u_l and u_h are set equal to the unemployment rates associated with the $p = 0$ steady state. With initial values for $X = \{p, u_l, u_h\}$, the initial values for θ and $f(\theta)$ are calculated, using interpolation, from the optimal matrix $\theta(X)$ found in the solution just discussed. Given the value for $f(\theta)$, the subsequent period's values for u_l and u_h are given by their respective laws of motion. These steps are repeated until forward iteration yields 736 observations.

References

Acemoglu, D., 1999. Changes in unemployment and wage inequality: An alternative theory and some evidence. American Economic Review 89 (5), 1259–1278.

Costain, J.S., Reiter, M., 2005. Business cycles, unemployment insurance, and the calibration of matching models. Mimeo, Universidad Carlos III.

Dolado, J.J., Jansen, M., Jimeno, J.F., 2003. On-the-job search in a matching model with heterogeneous jobs and workers. Discussion paper No. 886, IZA.

Hagedorn, M., Manovskii, I., 2005. The cyclical behavior of equilibrium unemployment and vacancies revisited. Mimeo, University of Pennsylvania.

Hall, R.E., 2005. The amplification of unemployment fluctuations through self-selection. Mimeo, Stanford University.

Hall, R.E., 2005b. Employment fluctuations with equilibrium wage stickiness. American Economic Review 95 (1), 50–65.

Hall, R.E., 2006. Sources and mechanisms of cyclical fluctuations in the labor market. Mimeo, Stanford University.

Hosios, A.J., 1990. On the efficiency of matching and related models of search and unemployment. Review of Economic Studies 57, 279–298.

Jacobson, L.S., La Londe, R.J., Sullivan, D.G., 1993. Earnings losses of displaced workers. American Economic Review 83 (4), 685–709.

Mortensen, D.T., Nagypál, E., 2006. More on unemployment and vacancy fluctuations. Mimeo, Northwestern University.

Mortensen, D.T., Pissarides, C.A., 1994. Job creation and job destruction in the theory of unemployment. Review of Economic Studies 61 (3), 397–415.

Nagypál, E., 2001. Learning in the workplace. PhD thesis, Stanford University.

Nagypál, E., 2005. Labor-market fluctuations, on-the-job search, and the acceptance curse. Mimeo, Northwestern University.

Nickell, S., Bell, B., 1995. The collapse in demand for the unskilled and unemployment across the OECD. Oxford Review of Economic Policy 11 (1), 40–62.

Pries, M.J., 2006. Solving a matching model with heterogeneous workers and matches. Mimeo, University of Maryland.

Shimer, R., 2001. The impact of young workers on the aggregate labor market. Quarterly Journal of Economics 116 (3), 969–1008.

Shimer, R., 2005. The cyclical behavior of equilibrium unemployment and vacancies. American Economic Review 95 (1), 25–49.