Business Cycles and Labor Market Flows with Skill Heterogeneity

Federico Ravenna*and Carl E. Walsh†

Preliminary and Incomplete: Do Not Circulate
Dec. 2010

Abstract

We build a business cycle model where employers' screening of heterogeneous workers plays a central role in determining both the flows into and out of unemployment. The model can address how differences between the US and European labor market flows affect business cycle dynamics. It provides a novel and rich environment to study the implications of labor market structure for the goals and constraints faced by central banks and on the optimal design of monetary policy.

JEL: E52, E58, J64

1 Introduction

Monetary policy models that incorporate labor search and matching generally assume an exogenous separation rate and homogenous workers (e.g., Ravenna and Walsh 2008a, 2008b, 2009, Gertler, Sala and Trigari 2007, and Gertler and Trigari 2009). Equilibrium in these models depends on the cost of posting vacancies, the replacement ratio of unemployment benefits, and the relative bargaining power of workers and firms. This makes them useful frameworks for investigating how the monetary transmission mechanism is

*HEC Montreal, Institute of Applied Economics and Department of Economics, University of California, Santa Cruz. Email: federico.ravenna@hec.ca.

†Department of Economics, University of California, Santa Cruz. Email: walshc@ucsc.edu. This research is supported by the Fondation Banque de France grant on "European Labor Market Institutions and their Consequences for Monetary Policy".

affected by labor market institutions (e.g., Walsh 2003, 2005, Blanchard and Gali, 2007).[1] However, this family of models is unable to confront the data on important dimensions that characterize some of the striking differences between US and European labor flows such as unemployment duration, wage dispersion, exit rates from unemployment, and workers' reallocation across firms. Our aim is to study the impact of monetary policy on the business cycle within a framework that can account for this set of empirical observations in the US and in European countries.

We build a business cycle model where screening of heterogenous workers by firms plays a central role in determining both the separation rate of jobs and the exit rate from unemployment.[2] That is, we focus on the *ins* and *outs* of unemployment (Petrongolo and Pissarides 2008). Because workers are heterogeneous with respect to skill level, our framework generates a time-varying share of long term unemployed within the pool of searching workers, negative duration dependence for the job finding probability, and time-varying wage dispersion across employed workers.

In standard models with exogenous separation, the worker separation rate and the job destruction rate are by construction identical. In models with endogenous separation (Den Haan, Ramey, and Watson 2000, Walsh 2003, 2005) a worker becomes unemployed when the worker and the firm jointly decide to end a match. However, our model provides a natural framework to generate different job turnover and worker turnover rates, because the chance of hiring a higher-skill worker creates incentives for firms to separate from low-skill workers without destroying an employment position. The incentives driving the relative size of job reallocations and worker reallocations change with the level of aggregate productivity. The model provides a novel and rich environment to study the implications of the labor market structure on the goals and constraints of the policymaker, and on the optimal behavior of monetary policy.

Because there exists different skill levels, the model generates two pools of unemployed, short and long term, with different job-finding probability. Moreover, the employment wages of low and high-skill workers are also a function of the difference in

[1] Lechthaer, Merkl, and Snower (2010) model labor frictions through the assumption of hiring and firing costs and investigate the implications for macro dynamics in a model with nominal rigidities.

[2] Workers' heterogeneity in search models of the labor market has been studied by, among others, Brown, Merkl and Snower (2009), Faraglia and Esteban-Pratel (2009), and match heterogeneity plays a key role in the models of Nagypal (2007), Tasci (2007) and in models with job-to-job transitions. Our model is closest to Pries and Rogerson (2005) in providing a mechanism through which firms decide how much screening takes place prior to forming a match.

productivity, and are negatively correlated with unemployment duration (although there will also exist frictional wage dispersion because of the idiosyncratic match productivity assigned to employed low-skill workers). Finally, the model allows aggregate shocks to affect labor flows through two channels missing from other models with endogenous separation. The first channel arises because the presence of heterogenous skills among workers implies that a firm with a low-skill worker may terminate the match in hopes of finding a high-skill replacement. This leads to an increase in worker reallocation. The second channel arises because some low-skill applicants who are interviewed are not hired because the firm does not wish to forego the opportunity of finding a high-skill worker if the position is kept open. Both these margins change over the business cycle; the probability a low-skill worker is fired (hired) decreases (increases) as the pool of high-skill unemployed workers shrinks.

In the next section, we review some of the evidence on labor flows in the US and European countries.

2 Evidence on labor flows

A very extensive literature has documented the differences in labor flows between US and the European countries on which we focus. Machin and Manning (1999) report that in 1995 the share of unemployed workers with unemployment spells longer than 12 months was 62.7% in Italy, 56,5% in Spain, and 45.6% in France. For the US, the corresponding figure is just 9.7%, and Esteban-Pratel and Faraglia (2009) find that the share of US unemployed with spells longer than six months, measured by the BLS, has never been higher than 25% in the 1979-2003 period, with the peak in the beginning of the 1980s. In the 1960s the figures for European countries were much closer to those in the US, and the worsening trend in Europe since the 1960s is due to a marked collapse in the exit rate from unemployment at all durations.

The large share of long term unemployed workers is a troubling issue because the composition of the unemployment pool changes with the duration of the unemployment spell. Evidence across many countries - including the US - shows that the job finding probability decreases with unemployment duration. This evidence has been explained by the loss of skill occurring for workers who are detached from employment for long spells. In fact, many authors (see references in Machin and Manning, 1999, and Villena-Roldan, 2008) find that observable and unobservable skill heterogeneity can explain nearly all of the negative duration dependence found in the data.

Elsby, Hobijn, and Sahin (2008) document striking differences in the monthly rates of inflow and outflow from unemployment among OECD countries. They find inflow and outflow rates are positively correlated, with continental European countries characterized by low rates of both inflow and outflow, consistent with the description of European labor markets as displaying sclerosis. The average of the inflow and outflow rates in France, Germany, Italy, Portugal, and Spain ranged from 4.8% (Italy) to 10.2% (Spain). By way of contrast, the rates averaged 40% in the US. Outflow rates exhibited a larger dispersion across countries, but inflow rates also differed. The estimated rate of outflow from unemployment for Spain was 1% while rates for France, Germany, Italy, and Portugal were even lower. For the US, the comparable figure was estimated to be 3.6%. Elsby, Hobijn, and Sahin argue that inflows contribute only about 20% of the time series variation of unemployment rates in Anglo-Saxon and Nordic countries, a finding consistent with Shimer (2008). However, the corresponding figure for continental European economies is 50%, suggesting a much larger relative role is played by variations in the inflow to unemployment in accounting for fluctuations in European unemployment experiences.

The important role played by fluctuations in the rate of inflow into unemployment in European economies is inconsistent with the standard assumption of most recent models of labor market frictions, business cycles and monetary policy as these models typically assume a constant and exogenous separation rate (e.g., Ravenna and Walsh 2008a, 2008b, 2009, Gertler, Sala and Trigari 2007, Gertler and Trigari 2009, Blanchard and Gali, 2010).

Rogerson and Pries (2005) suggest that hiring policies may play a large role in explaining differences in job market flows based on data on worker turnover and job turnover across countries. The levels of job creation and job destruction are similar across the US and Europe, while worker turnover, which includes both job reallocations across establishments and worker reallocation across existing jobs, is substantially greater in the US. Burgess, Lane and Stevens (2000) find that in the US about 13% of job positions are destroyed in a year, while the number of separations over the same period is roughly five times larger.

Heterogeneity in workers' skills has also been prominently suggested as an explanation for wage dispersion. The amount of wage dispersion that search models with idiosyncratic match-productivity can reasonably produce is by an order of magnitude too small compared to the data. Again, the US and European data show striking differences. Simon (2009) reports for 2002 data that the ratio of the 50th to 10th earnings percentile is

1.32 in Finland, 1.38 in Italy, 1.58 in Spain, and 1.64 in France. Hornstein, Krusell and Violante (2007) use 1990 US Census data to show that the ratio of the mean wage to the 10th percentile is 1.83 even conditioning on low-skill occupations and a set of workers with less than 10 years of experience.

The cyclical movements of unemployment rates differ by educational level, one indicator of skill differences across workers. Figure 1 shows, for the U.S. since 1992, the excess of the unemployment rate for workers with less than a high school diploma over the rate for those with a college degree. Because the average unemployment rates for these groups differ, the sample means are first subtracted from each rate. Recessions are associated with a widening gap between the two unemployment rates, consistent with economic downturns leading to a rise in unemployment of low-skill workers relative to unemployment for high-skill workers. During economic expansions, the gap narrows.

Another characteristic associated with skill levels is age. Figure 2 shows the U.S. quarterly rate for workers 16-19 years of age minus the aggregate unemployment rate. The strong counter-cyclical movements in this unemployment difference is clear.

The assumption homogenous labor and a constant rate of job separation, as is common in the existing literature that has blended models of labor market search with nominal rigidities to address monetary policy issues cannot account for many of the documented differences in US and EU labor markets. We explore the implications of dropping both these assumptions by introducing a simple form of worker heterogeneity and allowing for endogenous separations. In the model we develop, the share of low-skill unemployed workers is endogenous, so the skill-weighted productivity of both the workforce and the unemployed pool changes over time. Pries (2010) finds that the composition effect of the unemployed pool has a large impact on the value of vacancies over the business cycle, and thus on the behavior of employment flows. These compositional effects will also endogenously affect the average duration of unemployment and the ratio of the duration of unemployment spells between high and low-skill workers over the business cycle. When a nominal rigidity is introduced, monetary policy will affect the dynamics of the economy, and the welfare level of the agents, by changing the composition of the unemployed pool.

3 The model

The model consists of households, wholesale and retail firms, and a monetary authority. Following the approach to labor market frictions in Walsh (2003, 2005) and Gertler and

Trigari (2009), we locate search frictions in a wholesale sector, where production requires that a firm and a worker be matched. Wholesale firms produce an homogenous good which is sold in a competitive market to retail firms, of which there are a continuum of mass one. Retail firms sell differentiated goods to households, and the retail sector is characterized by monopolistic competition and price stickiness as in standard new Keynesian models. The Appendix provides the complete set of equilibrium conditions for the model.

3.1 Overview of the labor market and labor flows

Workers are assumed to be heterogeneous with respect to skill; for simplicity, we assume workers are of two types, either high (h) or low (l) skill. Firms post vacancies to which unemployed workers apply. Firms must interview applicants to determine the worker skill type. Thus, the job search and recruitment process involves both interviewing and screening. The aggregate number of interviews per period is determined through random matching as in standard matching models of the labor market. We assume all job seekers have identical interview-finding probability regardless of skill level. At the interview, the job applicant is screened. Not all interviews result in hires. We assume that if the skill level is revealed in the interview to be h, the worker is hired and produces with probability equal to one. That is, we assume the firm is able to identify a high-skill worker in the interview and the productivity of an h worker is high enough that it guarantees a positive surplus in all states.[3]

The productivity of low-skill workers is assumed to be stochastic. Each period, regardless of whether employed or unemployed, each low-skill worker receives a new idiosyncratic productivity level $a_{i,t}$, where $a_{i,t}$ is the idiosyncratic stochastic productivity level of low-skill worker i in period t. We assume $a_{i,t}$ is serially uncorrelated and drawn from a distribution with support (0 1]. While productivity is randomly drawn in each period for a low-skill worker, the worker's skill-type, h or l, is permanently assigned.[4] While all high-skill unemployed workers who are interviewed are subsequently hired, only low-skill unemployment workers with $a_{i,t} > \bar{a}_t$ will be hired, where $\bar{a}_t$ is an endogenously

[3] This assumption is for simplicity as it will imply that endogenous separations and interviews that do not lead to hires only involve low skilled workers.

[4] We could assume match productivity is also random for high skill workers; if the support of the distribution is such that high skill workers always produce a positive surplus, the basic results of our model would be unchanged.

determined threshold level of productivity that will be shown to depend on an aggregate productivity shock and on the markup of retail over wholesale prices. In the absence of direct hiring and firing costs, $\bar{a}_t$ will also be the cut off value for determining whether an existing employed low-skill worker is retained by the firm. That is, from the perspective of the firm, the decision between retaining an existing worker with productivity $a_{i,t}$ is the same as the decision whether to hire a newly interviewed worker with productivity $a_{i,t}$.[5]

In addition to idiosyncratic productivity shocks, all employed workers are subject to an aggregate productivity shock z_t. Hence, the total productivity of low-skilled worker i at t is $z_t a_{i,t}$ while that of a high-skilled worker is z_t.

As is well know, a form of congestion externality is present in search and matching models; a firm that posts a vacancy reduces the probability other firms are able to fill their vacancies. With worker heterogeneity and endogenous separations, additional externalities arise. When a firm fails to retain a low-skill worker, the average skill-quality of the pool of job seekers is lowered, thus making it less likely a firm with a vacancy will make a hire. And as firms hire high-skill workers, they increase the probability that other firms will end up with a low-skill worker.

We neglect labor force participation decision and normalize the total workforce to equal one:

$$L^l + L^h = L = 1,$$

where L^j denotes the labor force of type j, $j = h, l$. Let $\overline{\Lambda} = L^l/L$ be the (fixed) fraction of the total labor force that is low skilled. Let S^j be the number of type j workers who are seeking jobs, and let N^j be the number of type j workers who are employed. Then the probability a worker drawn from the pool of unemployed job seekers is low skill is

$$\gamma_t = \frac{S_t^l}{S_t^l + S_t^h},$$

while the share of employed workers of skill l is

$$\xi_t = \frac{N_t^l}{N_t^l + N_t^h}.$$

The timing of activities is as follows. The stock of producing matches (filled jobs) in

[5] In models with hiring and firing costs, an existing employee with productivity $a_{i,t}$ might be retained while a new applicant of the same productivity would not be hired. See, for example, Lechthaller, Merkl, and Snower (2010).

period t is N_t of which $1-\xi_t$ are quality h and ξ_t are quality l. At the start of each period, there is an exogenous separation probability, denoted by ρ^x. Workers who are not in a match at the start of the period, or who do not survive the exogenous separation hazard, are unemployed and seek new interviews. There are

$$S_t = 1 - (1-\rho^x)\, N_{t-1}$$

such job seekers. We define the end-of-period number of unemployed workers as

$$U_t = 1 - N_t.$$

The two measures of unemployment can differ as some job seekers find employment (and produce) during the period. In search models based on a monthly period of observation, it is more common to assume workers hired in period t do not produce until period $t+1$. In this case, the number of job seekers in period t plus the number of employed workers adds to the total work force. Because we base our model on a quarterly frequency, we allow for some workers seeking jobs to find jobs and produce within the same period.

After exogenous separation occurs, all aggregate shocks are realized and observed. This allows firms to determine $\bar{a}_t$, the cutoff point for low-skill productive that will determine hiring and retention.[6]

Firms post vacancies V_t. The number of vacancies, together with the number of job seekers, determined the number of interviews I_t via a standard matching function. The probability a job seeker gets an interview is k_t^w. So $I_t = k_t^w S_t$. Firms interview $k_t^f V_t$ workers in the aggregate, where k_t^f is the probability a given vacancy receives an applicant to interview.

The time t idiosyncratic productivity shocks $a_{j,t}$ associated employed low-skill workers or low-skill workers who are interviewed are observed. A fraction $1-\rho_t^n$ type l workers receive productivity levels $a_{i,t} > \bar{a}_t$. So new hires H_t are given by number of interviewees who are high skill, all of whom are hired, plus the number of interviewees who are low skill times the fraction of these with productivity levels that exceed $\bar{a}_t$.

$$H_t = (1-\gamma_t)k_t^w S_t + (1-\rho_t^n)\,\gamma_t k_t^w S_t = (1-\rho_t^n \gamma_t)\, k_t^w S_t.$$

Note that fewer workers are hired than are interviewed: $H_t = (1-\gamma_t \rho_t^n)\, k_t^w S_t < k_t^w S_t$.

[6] We show below that $\bar{a}_t$ is the same for all firms.

The probability a randomly selected unemployed worker is screened out in the interview process (i.e., actually gets interviewed with a firm but has a $a_{i,t} < \bar{a}_t$ and so is not hired) is $\gamma_t \rho_t^n$. In standard matching models, new hires equal $k_t^w S_t$. Screening implies new hires are less than this level and depend on the average skill quality of the pool of unemployed workers γ_t and the aggregate productivity level which we show below will affect ρ_t^n.

Low-skill workers employed in existing matches that survived the exogenous separation hazard also receive a new productivity shock and are retained if and only if $a_{i,t} > \bar{a}_t$. Thus, actual employment in period t is equal to

$$\begin{aligned} N_t &= (1-\rho^x)\left[\left(1-\xi_{t-1}\right)+\xi_{t-1}(1-\rho^x)(1-\rho_t^n)\right] N_{t-1} + H_t \\ &= (1-\rho^x)\left(1-\xi_{t-1}\rho_t^n\right) N_{t-1} + H_t \end{aligned}$$

The total separation rate is $(1-\rho^x)\left(1-\xi_{t-1}\rho_t^n\right)$ and depends on the exogenous hazard ρ^x, the endogenous hazard for low-skill workers ρ_t^n, and the average skill-quality of beginning-of-period matches ξ_{t-1}. The average quality of employed workers evolves according to

$$\xi_t = (1-\rho_t^n)\left[\frac{(1-\rho^x)\xi_{t-1}N_{t-1} + \gamma_t k_t^w S_t}{N_t}\right]. \tag{1}$$

Job seekers at t who are of quality l equal the total number of low-skilled workers minus the number of matches of quality l that survive the exogenous separation hazard, dissolving matches that were productive at $t-1$. Hence,

$$\gamma_t = \frac{L^l - (1-\rho^x)\xi_{t-1}N_{t-1}}{S_t}. \tag{2}$$

In deriving eqs. (1) and (2) we assume workers who suffer exogenous separations can search within the same period, while those who experience endogenous separation, which occurs after shocks are realized during the period, cannot search until the following period.[7]

Endogenous separations happen as in a model without skills heterogeneity, and the random productivity shock is interpreted as the skill-dependent productivity of the worker. Since $a_{i,t}$ is *i.i.d.*, the model does not generate any endogenous distribution of skill-related

[7] Combining eqs. (1) and (2), it can be seen that job seekers at t who are of quality l arise from three sources: low-skilled workers who were searching for jobs in $t-1$ and failed to be hired; those employed in $t-2$ who survived the exogenous separation hazard but were endogenously terminated; and those employed in $t-1$ but who suffer the exogenous hazard at the start of period t.

productivity (each l worker may be more or less productive in every period), and an l worker can become less productive even if already in a match. But the share of low-skill workers in the unemployment pool, γ_t, is endogenous, so the skill-weighted productivity of both the workforce and the pool of unemployed changes over time. In particular, a burst of separations raises the average productivity of surviving matches and lowers the average skill level of the pool of unemployed job seekers.

Matches that end endogenously do so because they have a non-positive surplus. This is not true, however, of matches that experience exogenous separation. Following den Haan, Ramey, and Watson (2000), assume that firms with matches that end through the exogenous separation hazard immediately repost the position. Job destruction in period t is then defined as the number of exogenous separations occurring at the start of the period ($\rho^x N_{t-1}$) plus the number of workers who produced in $t-1$, survived the exogenous separation hazard, and then had productivity too low to survive the endogenous separation process (there are $\rho_t^n(1-\rho^x)\xi_{t-1}N_{t-1}$ such workers, recalling that only low-skill workers are at risk of endogenous separation) minus the number of the exogenous separation induced vacancies that get refilled within period t and so produce in period t (of which there are $\rho^x N_{t-1}k_t^f(H_t/I_t) = \rho^x(H_t/V_t)N_{t-1}$). Hence,

$$\begin{aligned} jd_t &= \rho^x N_{t-1} + \rho_t^n(1-\rho^x)\xi_{t-1}N_{t-1} - \rho^x(H_t/V_t)N_{t-1} \\ &= \left[\rho^x + \rho_t^n(1-\rho^x)\xi_{t-1} - \rho^x(H_t/V_t)\right]N_{t-1}. \end{aligned}$$

The ratio H_t/V_t of hires to vacancies corresponds to the probably a firm fills a match. In the present model, $H_t/V_t = (H_t/I_t)(I_t/V_t)$ depends on the fraction of interviews that result in hires (and so on the average skill level of job seekers) and on the number of interviews the firm receives per vacancy. In standard matching models, I_t/V_t is also the hiring rate as all interviews (in our terminology) result in hires.

Job creation in period t is equal to the number of new hires (H_t) minus the number of the new hires that go into positions made vacant by the exogenous separation hazard ($\rho^x(H_t/V_t)N_{t-1}$). Hence

$$jc_t = H_t - \rho^x N_{t-1}(H_t/V_t).$$

This definition is consistent with with Davis, Haltiwanger, Schuh (1996) who define job creation as net employment gains in establishments expanding their labor force. Since

$H_t = k_t^f \left(H_t/I_t\right) V_t$, job creation can also be written as

$$jc_t = k_t^f \left(H_t/I_t\right)\left(V_t - \rho^x N_{t-1}\right).$$

With our notation and timing, the net employment loss across the economy is equal to

$$jd_t - jc_t = \left[\rho^x + \rho_t^n \left(1-\rho^x\right)\xi_{t-1}\right] N_{t-1} - H_t,$$

which is gross separations minus total hires, and the aggregate separation rate is

$$\rho_t = \rho^x + (1-\rho^x)\rho_t^n \xi_{t-1}.$$

Finally, we define job turnover as the sum of job creation and destruction, $jturn_t = jc_t + jd_t$, while worker turnover as the sum of all separations and all hires in the period, $wturn_t = \rho_t N_{t-1} + H_t$, where $\rho_t = \rho_t^x + (1-\rho^x)\rho_t^n$. By construction, worker turnover is larger than job turnover.

If the aggregate productivity shock is low, $\bar{a}_t$ will rise, lowering the fraction of low-skill unemployed that receive job offers and increasing the endogenous separation rate of already employed low skill workers. Low skill workers become a larger fraction of the unemployed pool, since the probability of separation is always higher than for high skill workers. Also, after a positive aggregate shock (even *i.i.d.*) the average duration of unemployment increases, as the low skill workers lose jobs faster and have a harder time finding new employment since they are more likely to be screened out during the interview process.

3.2 The labor and goods markets

3.2.1 The wholesale sector

Wholesale firms post vacancies, interview and screen applicants, make hiring and retention decisions, and produce a homogenous output. There are N_t matched workers and firms that produce in period t, and $U_t = 1 - N_t$ unmatched workers. We normalize the productivity of high-skill workers to equal 1; low skill workers have individual productivity $a_{i,t} < 1$. Let h_t^h denote hours worked by high-skill workers and let $h_{i,t}^l$ be hours worked by low-skill worker i. All type h workers will work the same hours since they have the same productivity, but the hours of low-skill workers will depend on their idiosyncratic productivity realizations. Output of wholesale goods is obtained by aggregating over the output produced by employed high-skill workers and the output produced by employed low-skill workers with productivity levels greater than $\bar{a}_t$:

$$\begin{aligned} Q_t &= z_t^l N_t^l \frac{\int_{\bar{a}_t}^1 a_{i,t} h_{i,t}^l dF(a_i)}{1-F(\overline{a}_t)} + z_t^h h_t^h N_t^h \\ &= \left[z_t^l \xi_t \frac{\int_{\bar{a}_t}^1 a_{i,t} h_{i,t}^l dF(a_i)}{1-F(\overline{a}_t)} + z_t^h (1-\xi_t) h_t^h \right] N_t \end{aligned} \tag{3}$$

where z_t^j is aggregate productivity for workers of skill level $j = [l, h]$ and $F(a)$ is the c.d.f. of the idiosyncratic productivity shocks. Since $F(\bar{a})$ is the probability $a_{i,t} \leq \bar{a}_t$, $F(\bar{a}) = \rho_t^n$ is also the endogenous separation and screening rate. We assume the productivity of a match depends on a common productivity disturbance z_t, with the productivity z_t^l of l workers equal to z_t, and the productivity of h workers equal to $z_t^h = z^h z_t$. The constant z^h is used to parameterize the relative average productivity of l and h workers.

The homogenous output of wholesale firms is sold to retail firms in a competitive goods market. The price of the wholesale good is P_t^w; the aggregate price index for retail goods is P_t. We define $\mu_t = P_t/P_t^w$ as the retail-price markup.

Expressed in terms of final retail goods, the current surplus of a firm-worker match involving a high-skill worker is

$$s_t^h = \left(\frac{z_t h_t^h}{\mu_t}\right) - \frac{v(h_t^h)}{\lambda_t} - w_t^{u,h} + q_t^h, \tag{4}$$

where h_t^h is chosen optimally to maximize the match surplus, $v(h_t^h)$ is the disutility of hours worked, λ_t is the marginal utility of consumption, $w_t^{u,h}$ is an unmatched high-skilled worker's opportunity utility, and q_t^h is the value of a match with a high-skill worker that continues into $t+1$. Since all type h workers have the same productivity, they will all work the same number of hours and generate the same surplus. Thus, we do not need to index h^h or s^h by i.

The surplus of a match involving a low-skill worker is

$$s_{i,t}^l = \left(\frac{a_{i,t} z_t h_{i,t}^l}{\mu_t}\right) - \frac{v(h_{i,t}^l)}{\lambda_t} - w_t^{u,l} + q_t^l, \tag{5}$$

This differs from the expression for high-skill worker/firm matches because of the idiosyncratic productivity disturbance and the non-degenerate distribution of hours worked among low-skill workers. As is common in the literature on unemployment, we assume complete consumption risk sharing, so λ_t is the same for all workers.

Because the idiosyncratic productivity shocks are assumed to be serially uncorrelated, q_t^j depends on the skill-type of the worker in a match but is the same for all matches of the same skill-type. Let $f(a_i)$ be the density function for $a_{i,t}$. The continuation values are therefore given by

$$q_t^h = \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t} \right) \left[(1-\rho^x) s_{t+1}^h + w_{t+1}^{u,h} \right]. \tag{6}$$

and

$$\begin{aligned} q_t^l &= \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t} \right) \left[(1-\rho^x)(1-\rho_{t+1}^n) E_t(s_{i,t+1}^l | a_{i,t} > \overline{a}_{i,t}) + w_{t+1}^{u,l} \right] \\ &= \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t} \right) \left[(1-\rho^x) \int_{\overline{a}_{t+1}}^{1} s_{i,t+1}^l f(a_i) da_i + w_{t+1}^{u,l} \right], \end{aligned} \tag{7}$$

To determine $w_t^{u,j}$, we assume that that w^j is the value of time spent unemployed (home production or an unemployment benefit) and that wages are determined by Nash bargaining with the worker receiving a constant share η of the match surplus. Then the value of unemployment is equal to w^j plus the expected probability of being employed and receiving the surplus share ηs_{t+1}^j plus the expected value of remaining unemployed. For a high-skilled worker this is

$$w_t^{u,h} = w^h + \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t} \right) \left(k_{t+1}^w \eta s_{t+1}^h + w_{t+1}^{u,h} \right), \tag{8}$$

while for a low-skilled worker it is

$$\begin{aligned} w_t^{u,l} &= w^l + \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t} \right) \left[k_{t+1}^w \eta (1-\rho_{t+1}^n) E_t(s_{i,t+1}^l | a_{i,t} > \overline{a}_{i,t}) + w_{t+1}^{u,l} \right] \\ &= w^l + \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t} \right) \left[k_{t+1}^w \eta \int_{\overline{a}_t+1}^{1} s_{i,t+1}^l f(a) da + w_{t+1}^{u,l} \right]. \end{aligned} \tag{9}$$

If a low-skilled worker's productivity is too low, the surplus will be negative, leading to endogenous separation (or screening in the case of an interviewed job seeker). From

(5), the cutoff value of worker productivity at which the surplus produced by a low-skill worker equals zero is

$$\bar{a}_t = \frac{\mu_t \left(w_t^{u,l} + \frac{v(\hat{h}_t^l)}{\lambda_t} - q_t^l \right)}{z_t \hat{h}_t^l},$$

where $\hat{h}_t^l$ maximizes the surplus and satisfies

$$v_h(\hat{h}_t^l) \equiv \frac{\partial v(\hat{h}_t^l)}{\partial \hat{h}_t^l} = \left(\frac{\bar{a}_t z_t}{\mu_t} \right) \lambda_t.$$

That is, hours $\hat{h}_t^l$ maximizes the joint surplus in a match with a low-skill worker of productivity $\bar{a}_t$. Matches of low-skill workers separate endogenously if $a_{i,t} < \bar{a}_t$. As claimed previously, $\bar{a}_t$ is the same for all firm considering the retention or hire of a low–skill worker. The probability of endogenous separation for a low-skilled worker/firm match is

$$\rho_t^n = F(\bar{a}_t).$$

This is also the probability a low-skill worker who receives an interview is not hired.

3.2.2 Hours

Hours maximize the joint surplus in a match. For a high-skill worker, this implies

$$\left(\frac{z_t^h z_t}{\mu_t} \right) \lambda_t = v_h(h_t^h)$$

Hours h_t^l maximizes the joint surplus in a match with a low-skill worker of productivity $a_{i,t}$:

$$\left(\frac{a_{i,t} z_t}{\mu_t} \right) \lambda_t = v_h(h^l)$$

3.2.3 Vacancies

Wholesale firms post vacancies after observing aggregate variables, so their decisions are conditional on $\bar{a}_t$. If κ is the cost of posting a vacancy, expressed in terms of final goods, the job posting condition is

$$k_t^f (1-\eta) \left[(1-\gamma_t) s_t^h + \gamma_t \int_{\bar{a}_t}^1 s_{i,t}^l f(a_i) da_i \right] = \kappa, \tag{10}$$

since with probability $(1-\gamma_t)$ the firm interviews (and hires) a high-skill worker and with probability γ_t it interviews a low-skilled worker. This condition can be expressed as

$$k_t^f(1-\eta)\left[s_t^h - \gamma_t\left(s_t^h - \int_{\bar{a}_t}^1 s_{i,t}^l f(a_i)da_i\right)\right] = \kappa.$$

Since the surplus from a high skill worker is greater than that from a low skill worker, a fall in the quality of the unemployment pool (a rise in γ_t) reduces the incentive to post vacancies.

Given the pool of job seekers S_t and the number of vacancies V_t posted by firms, the number of new interviews is determined by a standard matching function $m(S_t, V_t)$. This is taken to be Cobb-Douglas with constant returns to scale:[8]

$$m(S_t, V_t) = \psi S_t^\alpha V_t^{1-\alpha} = \psi\theta_t^{1-\alpha} S_t, \qquad 0 < \alpha < 1,\, \psi > 0, \tag{11}$$

where $\theta_t \equiv V_t/S_t$ is the standard measure of labor market tightness. Because of worker heterogeneity, the probabilities of being interviewed and being hired will differ by the worker's skill level. The probability an unemployed worker obtains an interview, k_t^w, is

$$k_t^w = \frac{m(S_t, V_t)}{S_t} = \psi\theta_t^{1-\alpha}. \tag{12}$$

This is the same for all job seekers. Similarly, the probability a firm with a posted vacancy finds an applicant, k_t^f, is

$$k_t^f = \frac{m(S_t, V_t)}{V_t} = \psi\theta_t^{-\alpha}. \tag{13}$$

Compared to the standard single-skill setup: k_t^w is the probability obtains an interview, and k_t^f is the probability an interview slot will not go unfilled. The job finding probability is identical to the interview rate for high-skill workers, while it is lower, and equal to

$$k_t^{w,l} = k_t^w(1-\rho_t^n) < k_t^w$$

for low-skill workers. The overall job finding probability can be defined as $\gamma_t k_t^{w,l} + (1-\gamma_t)k_t^w$. With heterogeneous worker skills, a job opening that would be filled and lead to production if a high-skill applicant is interviewed may go unfilled if a low-skill worker is

[8] Constant returns to scale is consistent with the empirical evidence when applied to new hires; see Petrongolo and Pissarides (2001).

interviewed.

3.3 Households

The representative household purchases consumption goods, holds bonds, and supplies labor. Since some workers will be matched while others will not be, and workers differ their productivity and hours worked, distributional issues arise. To avoid these issues, we follow the literature in assuming households pool consumption by viewing the household as consisting of a continuum of members of various skill levels, some of whom will be employed, others unemployed.[9] Households are also the owners of all firms in the economy.

Households maximize

$$\mathrm{E}_t \sum_{i=0}^{\infty} \beta^i \left[D_t \frac{(\mathcal{C}_{t+i} - \phi \mathcal{C}_{t+i-1})^{1-\sigma}}{1-\sigma} - v(h^h_{t+i})(1-\xi_{t+i})N_{t+i} - \xi_{t+i} N_{t+i} \int_{\bar{a}_t}^{1} v(h^l_{i,t+i}) f(a) da \right], \tag{14}$$

where $\sigma > 0$ is the coefficient of relative risk aversion, $\phi > 0$ is a measure of the degree of internal habit persistence in consumption, D_t is an aggregate preference shock. In models with sticky prices, output responds to demand shifts; if consumption is purely forward looking and there is no investment, consumption and output jump immediately in response to interest rate shocks. To match the hump shaped response of output seen in the data, habit persistence has become a standard component of new Keynesian models (Fuhrer 2000, Christiano, Eichenbaum, and Evans 2001). To incorporate habit persistence, preferences of the representative household are defined over $\mathcal{C}_t$ and $\mathcal{C}_{t-1}$, where $\mathcal{C}_t$ is the sum of a market purchased composite consumption good C_t and home produced consumption C^h_t. The latter is defined as $C^h_t = (1 - N^l_t)w^l + (1 - N^h_t)w^h$. Thus, we allow high-skill and low-skill workers to have different productivity in home production if $w^l \neq w^h$. measures nontradeable home production when unemployed.

Market consumption C_t is a Dixit-Stiglitz composite good consisting of the differentiated products produced by retail firms and is defined as

$$C_t = \left[\int_0^1 c_{kt}^{\frac{\theta-1}{\theta}} dk \right]^{\frac{\theta}{\theta-1}} \qquad \theta > 0.$$

[9] This assumption is common; see Merz (1995), Andolfatto (1996), den Haan, Ramey, and Watson (2000), Cooley and Quadrini (1999), and Hairault (2002).

Given prices p_{kt} for the final goods, this preference specification implies the household's demand for good j is

$$c_{kt} = \left(\frac{p_{kt}}{P_t}\right)^{-\theta} C_t, \tag{15}$$

where the aggregate retail price index P_t is defined as

$$P_t = \left[\int_0^1 p_{kt}^{1-\theta} dj\right]^{\frac{1}{1-\theta}}.$$

In (14),

$$v(h^h_{t+i})(1-\xi_{t+i})N_{t+i} - \xi_{t+i}N_{t+i}\int_{\bar{a}_t}^1 v(h^l_{i,t+i})f(a)da$$

is the disutility to the household of having N_t members working, where hours worked depends on type and the idiosyncratic productivity shocks. We assume $v(h_{t+i}) = \ell h_{t+i}^{1+\chi}/(1+\chi)$.

If i_t is the nominal rate of interest. the representative household's first order conditions imply the following must hold in equilibrium:

$$\lambda_t = \beta(1+i_t)\mathrm{E}_t\left(\frac{P_t}{P_{t+1}}\right)\lambda_{t+1}, \tag{16}$$

where λ_t denotes the total marginal utility of consumption at time t and is given by

$$\lambda_t \equiv (\mathcal{C}_t - \phi\mathcal{C}_{t-1})^{-\sigma} - \beta\phi\left(\mathrm{E}_t\mathcal{C}_{t+1} - h\mathcal{C}_t\right)^{-\sigma}. \tag{17}$$

3.4 Retail firms

Each retail firm purchases wholesale output which it then converts into a differentiated final good that is sold to households and wholesale firms. Retail firms maximize profits subject to a CRS technology for converting wholesale goods into final goods, the demand functions (15), and a restriction on the frequency with which they can adjust their price.

Retail firms adjust prices according to the Calvo updating model. Each period a firm can adjust its price with probability $1-\omega$. The real marginal cost for retail firms is the price of the wholesale good relative to the price of final output, P_t^w/P_t. This is just the inverse of the markup of retail over wholesale goods.

A retail firm k that can adjust its price in period t chooses $P_t(k)$ to maximize

$$\sum_{s=0}^{\infty} (\omega\beta)^s \, \mathrm{E}_t \left[\left(\frac{\lambda_{t+s}}{\lambda_t} \right) \left(\frac{P_t(k) - P^w_{t+s}}{P_{t+s}} \right) Y_{t+s}(k) \right],$$

subject to

$$Y_{t+s}(k) = Y^d_{t+s}(k) = \left[\frac{P_t(k)}{P_{t+s}} \right]^{-\varepsilon} Y^d_{t+s}, \tag{18}$$

where Y^d_t is aggregate demand for the basket of final goods. The first order condition for those firms adjusting their price in period t is

$$P_t(k)\mathrm{E}_t \sum_{s=0}^{\infty} (\omega\beta)^s \left(\frac{\lambda_{t+s}}{\lambda_t} \right) \left[\frac{P_t(k)}{P_{t+s}} \right]^{1-\varepsilon} Y_{t+s} = \left(\frac{\varepsilon}{\varepsilon - 1} \right) \mathrm{E}_t \sum_{s=0}^{\infty} (\omega\beta)^s \left(\frac{\lambda_{t+s}}{\lambda_t} \right) \left(\frac{1}{\mu_{t+s}} \right) \left[\frac{P_t(k)}{P_{t+s}} \right]^{1-\varepsilon} Y_{t+s}.$$

The standard pricing equation obtains. These can be written as

$$(1 + \pi_t)^{1-\varepsilon} = \omega + (1 - \omega) \left[\frac{\tilde{G}_t}{\tilde{F}_t} (1 + \pi_t) \right]^{1-\varepsilon}, \tag{19}$$

where

$$\tilde{G}_t = \mu \lambda_t \mu_t^{-1} Y_t + \omega\beta \mathrm{E}_t \tilde{G}_{t+1} (1 + \pi_{t+1})^{\varepsilon}$$

$$\tilde{F}_t = \lambda_t Y_t + \omega\beta \mathrm{E}_t \tilde{F}_{t+1} (1 + \pi_{t+1})^{\varepsilon - 1}.$$

When linearized around a zero-inflation steady state yields a new Keynesian Phillips curve in which the retail price markup

$$\mu_t \equiv \frac{P_t}{P^w_t}$$

is the driving force for inflation. As in a standard Phillps curve, the elasticity of inflation with respect to real marginal costs will be $\delta \equiv (1 - \omega)(1 - \beta\omega)/\omega$.

3.5 Monetary policy

We assume that the monetary authority in this economy implements monetary policy through a simple Taylor-type instrument rule with inertia of the form

$$\ln(1 + i_t) = -\ln\beta + \chi_i \ln(1 + i_{t-1}) + (1 - \chi_i) \left[\phi_\pi \pi_t + \phi_y \left(\ln Y_t - \ln \overline{Y} \right) \right]. \tag{20}$$

As a baseline policy we assume $\phi_\pi = 1.5$, $\phi_y = 0$ and $\chi = 0.8$.

3.6 Market clearing

Goods market clearing requires that household consumption of market produced goods equals the output of the retail sector minus final goods purchased by wholesale firms to cover the costs of posting job vacancies Hence, goods market equilibrium takes the form

$$Y_t = C_t + \kappa V_t. \tag{21}$$

4 Results

4.1 Model Parameterization for US and EU

The baseline model is very parsimonious, and has a limited number of parameters. We further reduce the set of parameters by assuming the value of home production is independent of market skill levels, so $w^l = w^h$. The coefficient ℓ scaling the disutility of labor hours, the cost of vacancy posting κ, the productivity of the matching technology ψ, and the labor force share of low-skilled workers $\overline{\Lambda}$, to match the steady-state values for five data points, as described in table 1. The steady state aggregate separation rate is about half as large in our European calibration, labeled EU, relative to the US, and it is set according to available average separation data (Shimer 2005, Blanchard and Gali 2010). The steady state unemployment rate is the data point for the second quarter of 2009, for the European calibration which includes the 27 member states of the European Union, and the second quarter of 2007, for the US calibration. We distinguish among h and l-skill workers by using unemployment data by age. For the EU, the youth unemployment rate includes the labor force aged below 24, while in the US includes the labor force 16-19 years of age. The ratio between the l and h-skill unemployment rate is about 4 in the US, and only about 2.6 in the EU.[10]

[10] The calibration is only illustrative, to highlight the impact of the skill heterogeneity on the model dynamics. Using data averaged over a longer sample does not change qualitatively the result. In the US, the unemployment rate in the total, 16 to 19 year old and over-20 year old populations over the 1948-2010 sample averaged respectively 5.6%, 15.66% and 5%. For the 16 to 24 and over-24 year old population the respective average is equal to 11.6% and 4.4%. In the over-25 year old population, the average unemployment rate for the labor force participants with at least a college degree, with a high school degree and with no high school is respectively 2.6%, 5.3% and 9%.

Our model endogenously generates heterogeneous unemployment duration. While the short and long-term unemployed in our model can be drawn potentially from any level of skills, the long-term unemployed group is for the most part composed by low-skill workers. This is consistently with the empirical evidence in Villena-Roldan (2008), showing that the dependence of job-finding probability on unemployment duration can be nearly completely explained by skill heterogeneity across workers. Our parameterization implies a share of l workers in the pool of job seekers of about 20% for the *EU* calibration. In the second quarter of 2009, the EU-27 share of long-term unemployed was 32.3%, a value not very distant once we consider that the baseline model has no firing or training costs. The choice for other parameters common to both calibrations follow the recent literature on business cycle models with search unemployment and nominal rigidities.

Table 1: Parameterization			
		US	*EU*
Steady state aggregate separation rate	ρ_{ss}	7%	3.7%
Steady state unemployment rate	u_{ss}	4.6%	8.8%
Steady state unemployment rate - $l-skill$ labor	u^l_{ss}	16%	19.3%
Steady state unemployment rate - $h-skill$ labor	u^h_{ss}	4%	7.5%
Steady state average hours per worker	h^{av}_{ss}	0.33	0.25
Workers' share of surplus	η	0.4	0.3
Common Parameters			
Vacancy elasticity of matches	α	0.6	
Discount factor	β	0.99	
Inverse of labor hours supply elasticity	χ	2.5	
Relative risk aversion	σ	1	
Steady state inflation rate	π_{ss}	1	
Steady state vacancy filling rate	$k^{job,f}_{ss}$	0.07	
Vacancy posting cost share of output	$\frac{\kappa V_{ss}}{Y_{ss}}$	0.05	
AR(1) parameter for technology shock z_t	ρ_z	0.95	
Calvo pricing parameter values			
Price elasticity of retail goods demand	ε	6	
Average retail price duration (quarters)	$\frac{1}{1-\omega}$	3.33	
Steady state markup	μ	1	

4.2 Steady State

Low skill workers are over-represented in the pool of unemployed. Our parameters imply that the share of l workers $\overline{\Lambda}$ in the labor force L is 5.3% in the US and 10.4% in the EU. Because the separation rate of l workers is about twice as large as the overall separation rate, their share $\overline{\gamma}$ in the pool of job seekers is around 10% for the US calibration and 20% for the EU calibration. This result is key also to the dynamic behavior of the model, since it implies that when deciding whether to hire, a firm faces a 1 in 10 probability of interviewing a low skill worker in the US, but a 1 in 5 probability in the EU case. This affects the incentive of firms to post vacancies, given that the implied expected relative productivity of an h worker compared to a l worker is 1.47 in the US, 1.38 in the EU case.

The different incentives faced in the EU and US case for firms and workers to form matches - including the composition effect of the labor force on incentives to post vacancies - result in a job turnover rate (the sum of job-creation and job destruction rate) which is close across parameterizations, 4.6% and 6.4% for the US and EU case. At the same time, the worker turnover rate (the sum of all hires and separations relative to the labor force) is roughly equal to the job turnover rate in the EU case, and three times as large as the turnover rate in the US case. This result matches empirical evidence in Burgess, Lane and Stevens (2000) for the US and available cross country evidence (see Pries and Rogerson, 2005). Intuitively, the EU case describes an economy where firms hire much more cautiously, and employees have longer tenure. At the same time, unemployment is overall higher: once a worker enters into the unemployed pool, it is much more difficult to find a new match. The US case describes instead an economy with plenty of worker reallocation, where workers enter and exit the unemployed pool much more frequently.

Finally, part of the difference in the unemployment rate across parameterizations also obtains because the value of home production is 34% higher in EU (the ratio of home consumption relative to the market consumption obtained by participating in the labor market is 0.38 for the US, 0.47 for the EU), and the disutility of work hours measured by ℓ is nearly three times as large in the EU calibration.

As it is, our framework is inadequate to explain differences in unemployment duration. The implied steady-state unemployment duration for low-skill workers is only 9% longer (12% longer) than for high-skill workers in the EU (US). It should not be surprising that the US calibration results in longer unemployment duration for l workers. When the share of low-skill workers in the labor force is smaller, as in the US case, a firm that interviews a low-skill worker has a greater incentive to screen and postpone filling the vacancy in hopes of finding a high-skill worker. The expected surplus of any future hire is higher - both because total separations are lower when there are more high-skill workers and because the unconditional expected productivity of an interviewee is higher, leading to a higher separation rate for l workers relative to h workers. Given our parameterization, the probability for an l worker of being screened out at the interview is twice as large in the US case (14.5%) compared to the EU case (7.6%).[11]

[11] Our parameterization also implies that in the EU economy - with a lower share of high skill workers in the labor force, and longer overall unemployment duration - the duration of unemployment of high

A clue for understanding why the model fails to account for large differences in unemployment duration comes from the screening-out rate: the unconditional probability an interviewee will not be hired. Workers fail to receive a job offer, conditional on being interviewed, with a probability around only 8.5% in both the EU and US case. Given that the only heterogeneity across workers in the model is attributed to a skill differential, combined with the low share of l-skill workers in the labor force, the model does not generate a strong incentive to screen out applicants. Different sunk costs across workers, such as training or firing cost, would provide a greater incentive for firms to screen more aggressively, affecting directly the duration of unemployment. Adding such costs is one area in which we plan to extend the model in future work.

To study the dynamics implied by the model, we compute the response to a persistent 1% fall in total factor productivity.

Unemployment rate Figure 3 shows the impact of the negative productivity shock on the aggregate unemployment rate, and on the unemployment rate for the two groups of workers. The plot is scaled in terms of percentage points of the overall labor force, and of the labor force for each group of workers. The impact on the overall unemployment rate is relatively small in the US case, a feature that is common to search models of the labor market with Nash bargaining. A vast array of mechanism to address this shortcoming has been proposed in the literature. In the EU case, the composition of the unemployment pool very significantly amplifies the impact of the shock on employment flows. We focus our attention on the implications for unemployment across the two subgroups of workers. First, in both the US and EU calibrations, the change in the unemployment rate for the low-skill workers is nearly an order of magnitude larger than for the high-skill workers. Second, in the EU case, unemployment among low-skill workers increases by five percentage points - about five times the increase observed in the US case. Table 2 shows that this behavior is consistent with the dynamics of unemployment rates over the period 1983-2007, for which youth unemployment data is available. Volatility of youth and long term unemployment is much higher in Euro area countries, though obviously the moments of the data reflect all business cycle shocks, rather than just TFP shocks. The volatility of the youth unemployment rate is 200% higher than that of the

skill workers is closer to the one of low skill workers, when compared with the US. That is, relative to low skill workers, high skill workers have a larger comparative advantage in leaving unemployment in the economy where their share is *larger*.

aggregate unemployment rate in the EU-27 data, and only 32% higher in the US data. In our model, low-skill workers experience higher volatility in both job-finding probability and unemployment duration over a business cycle driven by TFP shocks, relative to high-skill workers. Several mechanisms are at work in generating this result, and are discussed in a later section in detail.

Table 2: Unemployment rate, 1983-2007

		Average	Standard deviation
Euro area	Unemployment (% labor force)	10.11%	1.33
	Unemployment - youth (% labor force age 15-24)	22.16%	4.06
	Unemployment - long term (% total unemployment)	48.74%	4.11
France			
	Unemployment (% labor force)	9.98%	1.36
	Unemployment - youth (% labor force age 15-24)	22.32%	3.16
	Unemployment - long term (% total unemployment)	40.47%	3.14
US			
	Unemployment (% labor force)	5.84%	1.28
	Unemployment - youth (% labor force age 15-24)	12.03%	1.69
	Unemployment - long term (% total unemployment)	9.25%	2.40

Note: Annual data. Source: World Development Indicators (2009).

Timing Relative to search models with homogeneous worker-skills, our framework generates considerable delay in the response of employment to productivity shocks. The peak response in overall unemployment happens after 7 quarters in the EU case, and 4 quarters in the US case. The lag is even more pronounced for h workers. This response depends on the combination of a change in productivity and in the implied changes in the composition of both the unemployment pool and the stock of employed workers.

Unemployment pool composition Figure 4 shows the log-deviation of selected variables in response to the recessionary productivity shock. The difference in the response of output across the two parameterizations is less pronounced than the response of employment, since an important share of the output decline comes directly from the fall in aggregate productivity. The composition of employment shifts in favor of h workers, much more so in the EU case which sees a large increase in the separation rate experienced by (formerly) employed low-skill workers. The increase in the separation rate - driven entirely by the firing of low skill workers - raises the share of less productive workers in the unemployment pool by over 15% (versus only about 4% in the US case). This in turn increases the likelihood that any firm that posts a vacancy will end up interviewing a low-skill worker. As a consequence, the probability an interview actually results in a hire decreases as more interviewee will be screened out, lowering firms' incentives to post vacancies for any given level of separations. Thus, a negative productivity shock increases the inflow into unemployment and reduces the outflow into employment - worsening the unemployment effects of the recession. In summary, low skill workers are more vulnerable to recessions in the EU case, and the worsening of the average quality of the unemployment pool causes firms' behavior to further exacerbate the severity of the recessions for low skill workers.

Job and worker dynamics Our model provides a natural framework to generate different job turnover and worker turnover rates, because the chance of hiring a higher-skill worker creates incentives for firms to separate from low-skill workers without destroying an employment position. The incentives driving the relative size of job reallocations and worker reallocations change with the level of aggregate productivity. The impact of the recession on unemployment reflects radically different employment flows across the EU and US model economies. Figure 5 illustrates the job and worker dynamic behavior. In the EU, firms lose employment by shedding low skill workers, and destroying job positions. At the same time, job creation increases, to replace some of the separations. In the EU, worker and job turnover both increase substantially. In the US case, firms drastically reduce job creation, while retaining workers. As a consequence, job turnover falls while worker turnover increases slightly.

4.2.1 The Composition Effect

The difference in flows of high and low skill workers has an important impact on the composition of the unemployment pool. The change in unemployment pool composition affects hiring and firing in two ways: first, through a direct channel by changing the quantity of low skill workers (the direct composition effect), and second, by changing the incentive of firms and unemployed to form matches (the indirect incentive effect).

The direct composition effect can be illustrated through the dynamic behavior of the job finding probability. The probability of finding a job for a l worker depends only on the interviewing rate k_t^w and on the endogenous separation rate ρ_t^n. Both will fall in a recession, so the job finding probability falls by more (and the unemployment duration increases by more) for an l worker than for an h worker. Thus, the unconditional probability that an unemployed worker enters into a match falls by more when the unemployed pool worsens. The top panel of figure 6 shows the behavior of the unconditional, l worker and h worker job finding probability. The unconditional probability falls in part because both k_t^w and $k_t^{w,l}$ fall, but also because the weight on $k_t^{w,l}$ increases in the overall average job finding rate. This effect is larger in the case of the EU. Note also that average unemployment duration reflects the composition effect.

The indirect effect of the change in the composition of the unemployment pool occurs through changes in the value of vacancies over the business cycle, a point made clear by Pries (2010). The presence of heterogenous skills among workers implies that a firm with a low-skill worker may terminate the match in hopes of finding a high-skill replacement. This leads to an increase in worker reallocation. Additionally, some low-skill applicants who are interviewed are not hired since the firm does not wish to forego the opportunity of finding a high-skill worker if the position is kept open. Both these margins are affected as the composition of the pool of job seekers changes. In a recession, the quality of the unemployment pool deteriorates, and this reduces the likelihood a firm will find a high-skill worker to hire. The composition effect then dampens the incentive to terminate existing low-skill matches and helps limit the decline in the inflow to unemployment. At the same time, by reducing the incentive to post vacancies, the composition effect acts to reduce the outflow from unemployment. In equilibrium, unemployment composition changes will impact employment flows, and the ratio of the duration of unemployment spells between high and low-skill workers.[12]

[12] The composition effect and incentive effect may work in opposite direction. Assume a marginal

Finally, screening has a negative externality on other firms. Since an individual firm hires with a higher probability an h worker it interviews compared to an l worker, it tends to deteriorate the average skill level of the pool of unemployed as it hires, making it less likely other firms will successfully fill vacancies.

4.2.2 The Impact of Screening

The bottom panel of figure 6 illustrates the impact of screening in a model with heterogenous worker skills. We define the screening rate as unconditional rate at which an interviewee is screened out. This rate is given by

$$scr_t = \gamma_t \rho_t^n = \gamma_t [1 - \Pr(s_{i,t}^l > 0)].$$

In a recession, the screening rate increases for three reasons. First, as in any search model of the labor market with endogenous separation, the separation rate ρ_t^n increases. The impulse response of the endogenous separation rate is shown in figure 6 as the screening rate net of the composition effect. Second, the likelihood that an interviewee is a low skill worker γ_t also increases. In the EU case, the composition effect accounts for around a third of the dynamics of the screening rate. Finally, the incentive effect may play a role in changing both ρ_t^n and the number of vacancies posted, since the probability of filling a position with a high skill worker drops, and the probability that a low skill interviewee results in a positive surplus decreases.

Since in a productivity-driven recession the share γ_t of l workers in the unemployment pool is positively correlated with k^w and with ρ_t^n, ceteris paribus, the skills heterogeneity will increase the volatility of unemployment relative to a model without screening.

5 Monetary Policy Shocks

In our baseline model, monetary policy has been represented by a simple policy rule in which the nominal interest rate was adjusted in response to inflation and to the lagged

increase in the share of h-workers in the labor force. The composition effect will drive down the unconditional job finding probability: there is less churning of workers since the share of employed workers who can separate endogenously is smaller. The incentive effect though may drive up the unconditional job finding probability, since the likelihood that an open vacancy will be filled with a high skill worker increases, leading possibly to a higher endogenous separation rate.

value of the policy rate. We now consider a monetary policy shock to this policy rule to see how a policy shock affects unemployment of workers of different skill levels.

Figure 7 shows the effect of a negative interest rate shock on the overall unemployment rate and the unemployment rates of the high and low skill workers. Results are shown for the US calibration (solid lines) and the EU calibration (dotted lines). Comparing this with figure 3 shows that productivity and policy shocks produce quite different dynamic responses in unemployment. For the EU, overall unemployment and the unemployment rates of both low-skill and high-skill workers are more persistent than for the US. Unemployment of high-skill workers is much less volatile than low-skill unemployment under either calibration. For the US calibration, however, the immediate impact of the policy shock on unemployment among high-skill workers is larger than in the EU case, but it is also much less persistent, consistent with the perception that labor flows adjust quickly in the US. From a policy perspective, figure 7 suggests that monetary policy has much large and long lasting effects on unemployment in the EU than in the US.

6 Conclusions and Extensions

We have developed a simple model of worker heterogeneity that incorporates endogenous separation. Heterogeneity causes the composition of the pool of unemployed workers to vary over the business cycle in ways that cannot occur in standard models with homogenous labor. A negative productivity shock reduces output and employment, but it also lowers the average quality of the unemployed, as low-skill workers experience greater unemployment. This compositional effect reduces the incentive for firms to post vacancies, as they are less likely to find a worker who is sufficiently productive to generate a positive surplus if hired.

As den Haan, et. al. (2000) had previously shown, endogenous separation can contribute to both the amplitude of employment responses to productivity shocks and the persistence generated by such shocks. We find that these effects are further strengthened by compositional affects that arise with heterogeneous workers. Moreover, the compositional effect has the potential to amplify the impact of productivity shocks on unemployment.

One simplifying assumption of the model was that the same critical productivity level determined whether existing employed low-skill workers would be retained and whether a low-skill job seeker would be hired. Hiring and/or firing costs would drive a wedge between the productivity level that determines if an existing worker is retained and the level sufficient to justify hiring a new low-skill worker. Introducing these costs would imply that for some productivity levels, a firm would be willing to retain an existing worker while simultaneously be unwilling to hire an identical job seeker.[13]

Despite the introduction of only two worker types, the model generates a rich set of implications for unemployment inflows and outflows. It provides a platform on which to investigate the role of labor market dynamics in affecting the transmission of monetary policy, the effects of macroeconomic fluctuations on unemployment flows in different countries or global regions characterized by different labor market structures, and to evaluate the implications of heterogeneity and endogenous separation on the design of optimal monetary policy.

References

[1] Benigno, P. and M. Woodford, "Linear-quadratic approximation of optimal policy problems", NBER working paper 12672, November 2006.

[2] Blanchard, O. J. and Jordi Galí, "A New Keynesian Model with Unemployment," *American Economic Journal: Macroeconomics,* (forthcoming), 2010.

[3] Blanchard, O. J. and Jordi Galí, "Real wage rigidity and the new Keynesian model," *Journal of Money, Credit and Banking*, 39 (1), 2007.

[4] Brown, A., Merkl, C. and Snower, D., "An incentive theory of matching", Universitat Kiel, 2009.

[5] Burgess, S., Lane, J., and Stevens, D., "Job flows, worker flows, and churning", *Journal of labor economics*, 18 (July): 473-500, 2000.

[6] Christoffel, K. and T. Linzert, "The role of real wage rigidity and labor market frictions for unemployment and inflation dynamics", ECB Woring Paper 556, 2005.

[7] Davis, Steven J., John C. Haltiwanger, and Scott Schuh. 1996. *Job Creation and Job Destruction*, Cambridge, MA: The MIT Press.

[13] See Lechthaler, Merkl, and Snower (2010).

[8] Elsby, M., B. Hobiijn, and A. Sahin, "Unemployment dynamics in the OECD," NBER Working Paper No. 14617, Dec. 2008.

[9] Esteban-Pratel, J. and Faraglia, E., "Monetary shocks in a model with loss of skills", University of Tokio, 2009.

[10] Gertler, M. and Antonella Trigari, "Unemployment Fluctuations with Staggered Nash Wage Bargaining", *Journal of Polictical Economy,* (forthcoming), 2009.

[11] Gertler, M., Sala, L. and Antonella Trigari, "An Estimated Monetary DSGE Model with Unemployment and Staggered Nominal Wage Bargaining", 2007.

[12] Hall, R. E., "Lost jobs," *Brookings Papers on Economic Activity*, 1995:1, 221- 256.

[13] Hall, R. E., "Employment fluctuations with equilibrium wage stickiness", *American Economic Review* 2005, 95: 50-71.

[14] Hosios, Arthur J, "On the Efficiency of matching and Related Models of Search and Unemployment," *Review of Economic Studies*, 57(2), 279-298, 1990.

[15] Hornstein, A., Krusell, P. and Violante, G., "Frictional wage dispersion in search models: a quantitative assessment", Federal Reserve Bank of Richmond Working Paper, 2007.

[16] Lechthaler, W., C. Merkl, and D. Snower, "Monetary persistence and the labor market: a new perspective," *Journal of Economic Dynamics and Control* 34 (2010): 968-983.

[17] Levin, A., Onatski, A., Williams, N. and John Williams, "Monetary Policy under Uncertainty in Micro-Founded Macroeconomic Models "in Gertler, Mark and Kenneth Rogoff, eds., NBER Macroeconomics Annual 2005, MIT Press: Cambridge MA, 2006, 229-287

[18] Machin, S. and Manning, A., "The causes and consequences of long-term unemployment in Europe", in Ashenfelter, O. and Card, D., eds., Handbook of Labor Economics, Volume 3, 1999.

[19] Mortensen, D. T., "Job search, the duration of unemployment, and the Phillips curve," *American Economic Review* 60(5), Dec. 1970, 847-862.

[20] Mortensen, D. T. and C. A. Pissarides, “Job creation and job destruction in the theory of unemployment,” *Review of Economic Studies*, 61 (3), July 1994, 397-416.

[21] Nagypal, E., "Learning-by-Doing Versus Learning About Match Quality: Can We Tell Them Apart?", *Review of Economic Studies*, 74 (2), pp. 537-566, 2007.

[22] Pries, M., "Worker heterogeneity and labor market volatility in matching models", *Review of Economic Dynamics,* (forthcoming), 2010.

[23] Pries, M. and Rogerson, R., "Hiring policies, labor market institutions, and labor market flows", *Journal of Political Economy,* 113(4), 2005.

[24] Ravenna, F. and C. E. Walsh, “Vacancies, Unemployment, and the Phillips Curve,” *European Economic Review*, forthcoming, 2008a.

[25] Ravenna, F. and C. E. Walsh, “Welfare-based optimal monetary policy with unemployment and sticky prices: A linear-quadratic framework,” June 2008b.

[26] Ravenna, F. and C. E. Walsh, The welfare consequences of monetary policy and the role of the labor market: a tax interpretation,” June 2009.

[27] Simon, H., "International differences in wage inequality: a new glance with European matched employer-employee data", MPRA Working Paper 7932, 2008.

[28] Shimer, Robert, "The Consequences of Rigid Wages in Search Models", *Journal of the European Economic Association* (Papers and Proceedings), 2: 469-479, 2004

[29] Shimer, Robert, “The Cyclical Behavior of Equilibrium Unemployment and Vacancies,” *American Economic Review*, 2005, 95(1), 25-49.

[30] Tasci, M., "On the job search and labor market reallocation", *Federal Reserve Bank of Cleveland Working paper* 7-25, 2007.

[31] Thomas, Carlos, “Search and matching frictions and optimal monetary policy,” February 2008, forthcoming in the *Journal of Monetary Economics.*

[32] Villena-Roldan, B., "Aggregate implications of employer search and recruiting selection", University of Rochester, 2008.

[33] Walsh, C. E., "Labor market search and monetary shocks," in *Elements of Dynamic Macroeconomic Analysis*, S. Altuĝ, J. Chadha, and C. Nolan (eds.), Cambridge: Cambridge University Press, 2003, 451-486.

[34] Walsh, C. E., "Labor market search, sticky prices, and interest rate policies," *Review of Economic Dynamics*, 8(4), Oct. 2005, 829-849.

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

7 Appendix

7.1 Equilibrium conditions: Definitions and market clearing

$$\theta_t = \frac{V_t}{S_t}$$

$$k_t^w = \psi \theta_t^{1-\alpha}$$

$$k_t^f = \psi \theta_t^{-\alpha}$$

$$\rho_t = \rho^x + (1-\rho^x)\xi_{t-1}\rho_t^n.$$

$$\rho_t^n = F(\bar{a}_t),$$

$$\xi_t = (1-\rho_t^n)\left[\frac{\xi_{t-1}(1-\rho^x)N_{t-1} + \gamma_t k_t^w S_t}{N_t}\right].$$

$$\gamma_t = \frac{\gamma_{t-1}S_{t-1}\left[1-(1-\rho_{t-1}^n)k_{t-1}^w\right] + \rho^x(1-\rho_{t-1}^n)\xi_{t-1}N_{t-1} - \rho_{t-1}^n\xi_{t-1}N_{t-1}}{S_t}$$

$$S_t = 1 + (1-\rho^x)\,N_{t-1}$$

$$U_t = 1 - N_t.$$

$$H_t = (1-\gamma_t\rho_t^n)\,k_t^w S_t.$$

$$N_t = \left(1-\xi_{t-1}\rho_t^n\right)(1-\rho^x)\,N_{t-1} + H_t$$

$$Q_t = z_t N_t^l \frac{\int_{\overline{a}_t}^1 a_{i,t} h_{i,t}^l dF(a_i)}{1 - F(\overline{a}_t)} + z_t^h z_t h_t^h N_t^h$$

$$\mathcal{C}_t = C_t + (1 - N_t) b$$

$$Y_t = C_t + \kappa V_t$$

$$Q_t = Y_t f_t$$

7.2 Equilibrium conditions: Behavioral

7.2.1 Households

$$\lambda_t = \beta(1 + i_t) \mathrm{E}_t \left(\frac{P_t}{P_{t+1}} \right) \lambda_{t+1}$$

$$\lambda_t \equiv (\mathcal{C}_t - \phi \mathcal{C}_{t-1})^{-\sigma} - \beta^h f \left(\mathrm{E}_t \mathcal{C}_{t+1} - h \mathcal{C}_t \right)^{-\sigma}$$

7.2.2 Low-skill workers

$$v_h(\hat{h}_t^l) = \left(\frac{\bar{a}_t z_t}{\mu_t} \right) \lambda_t.$$

$$\bar{a}_t = \frac{\mu_t \left(w_t^{u,l} + \frac{v(\hat{h}_t^l)}{\lambda_t} - q_t^l \right)}{z_t \hat{h}_t^l},$$

$$v_h(h_{i,t}^l) = \left(\frac{a_{i,t} z_t}{\mu_t} \right) \lambda_t \text{ for } a_{i,t} > \bar{a}_t$$

$$s_{i,t}^l = \left(\frac{a_{i,t} z_t h_{i,t}^l}{\mu_t} \right) - \frac{v(h_{i,t}^l)}{\lambda_t} - w_t^{u,l} + q_t^l$$

$$q_t^l = \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t} \right) \left[\int_{\bar{a}_{t+1}}^1 (1 - \rho^x) s_{i,t+1}^l f(a_i) da_i + w_{t+1}^{u,l} \right]$$

$$w_t^{u,l} = w^l + \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t} \right) \left\{ k_{t+1}^w \eta \int_{\bar{a}_t+1}^1 s_{i,t+1}^l f(a) da + w_{t+1}^{u,l} \right\}$$

7.2.3 High-skill workers

$$v_h(h_t^h) = \left(\frac{z_t}{\mu_t}\right)\lambda_t.$$

$$s_t^h = \left(\frac{z_t h_t^h}{\mu_t}\right) - \frac{v(h_t^h)}{\lambda_t} - w_t^{u,h} + q_t^h$$

$$q_t^h = \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t}\right)\left[(1-\rho^x)s_{t+1}^h + w_{t+1}^{u,h}\right]$$

$$w_t^{u,h} = w^h + \beta \mathrm{E}_t \left(\frac{\lambda_{t+1}}{\lambda_t}\right)\left\{k_{t+1}^w \eta s_{t+1}^h + w_{t+1}^{u,h}\right\}$$

7.2.4 Job-posting condition

$$k_t^f(1-\eta)\left[\gamma_t \int_{\bar{a}_t}^1 s_{i,t}^l f(a_i) da_i + (1-\gamma_t)s_t^h\right] = \kappa.$$

7.2.5 Job destruction and creation rates

$$jd_t = \left[\rho^x + \rho_t^n (1-\rho^x)\, \xi_{t-1} - \rho^x (H_t/V_t)\right] N_{t-1}$$

$$jc_t = k_t^f (H_t/I_t)\, V_t - \rho^x N_{t-1} k_t^f (H_t/I_t)$$

7.2.6 Retail firms

$$[(1+\pi_t)]^{1-\varepsilon} = \omega + (1-\omega)\left[\frac{\tilde{G}_t}{\tilde{F}_t}(1+\pi_t)\right]^{1-\varepsilon},$$

where

$$\tilde{G}_t = \mu \lambda_t \mu_t^{-1} Y_t + \omega\beta \tilde{G}_{t+1}(1+\pi_{t+1})^{\varepsilon}$$

$$\tilde{F}_t = \lambda_t Y_t + \omega\beta \tilde{F}_{t+1}(1+\pi_{t+1})^{\varepsilon-1}$$

$$1+\pi_t = \frac{P_t}{P_{t-1}}$$

$$f_t \equiv \int_0^1 \left[\frac{P_t(z)}{P_t}\right]^{-\varepsilon} dz$$

7.2.7 Policy

$$\ln(1 + i_t) = -\ln\beta + \chi_i \ln(1 + i_{t-1}) + (1 - \chi_i)\left[\phi_\pi \pi_t + \phi_y \left(\ln Y_t - \ln \overline{Y}\right)\right]$$

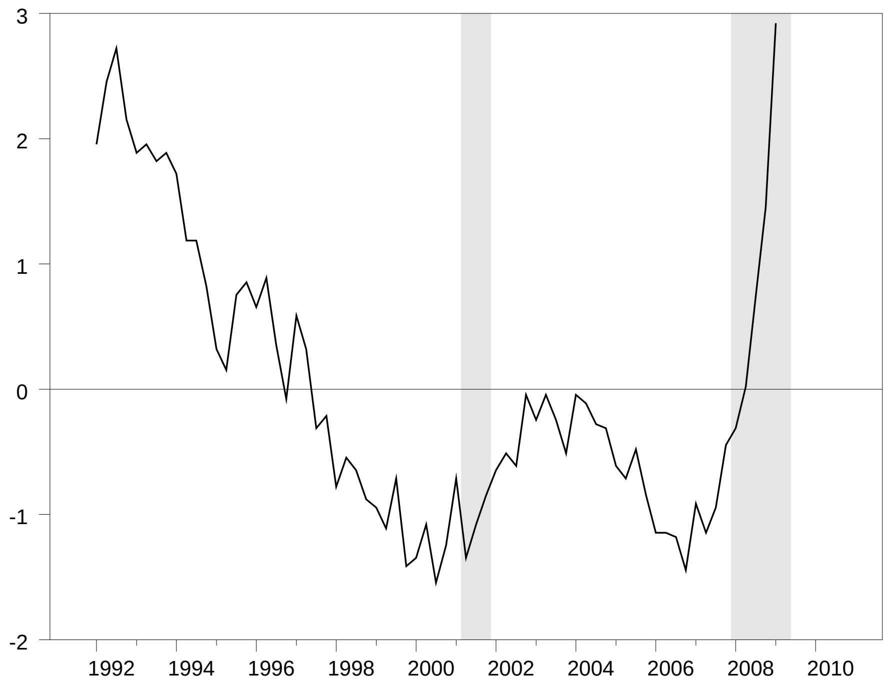

Figure 1: Unemployment rate for those with less than a high school diploma minus the unemployment rate for those with a college degree (both demeaned). Shaded regions are NBER business cycle recessions.

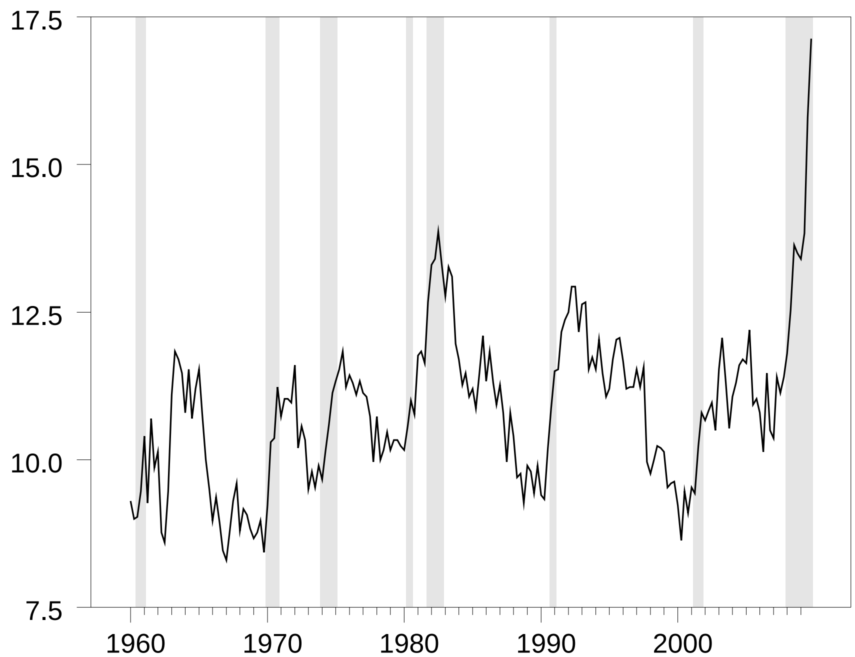

Figure 2: Unemployment rate for those 16-19 years of age minus the aggregate unemployment rate. Shaded regions are NBER business cycle recessions.

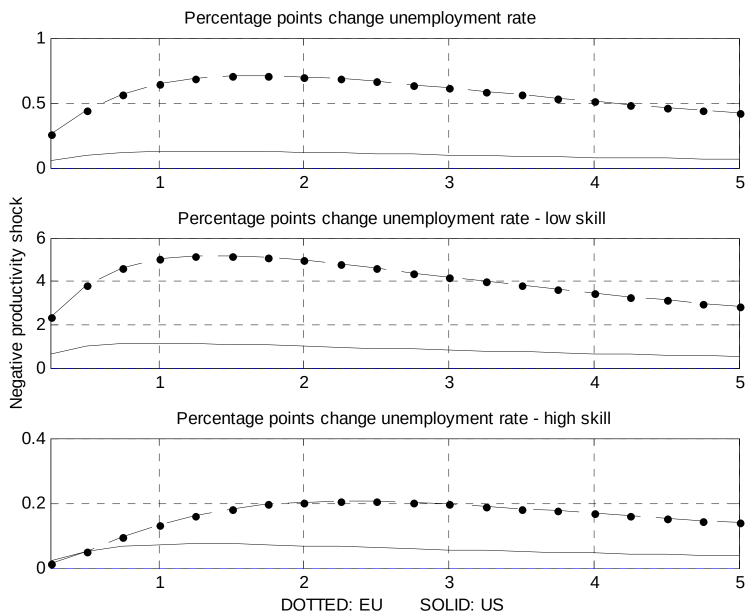

Figure 3: Response to a negative productivity shock: unemployment

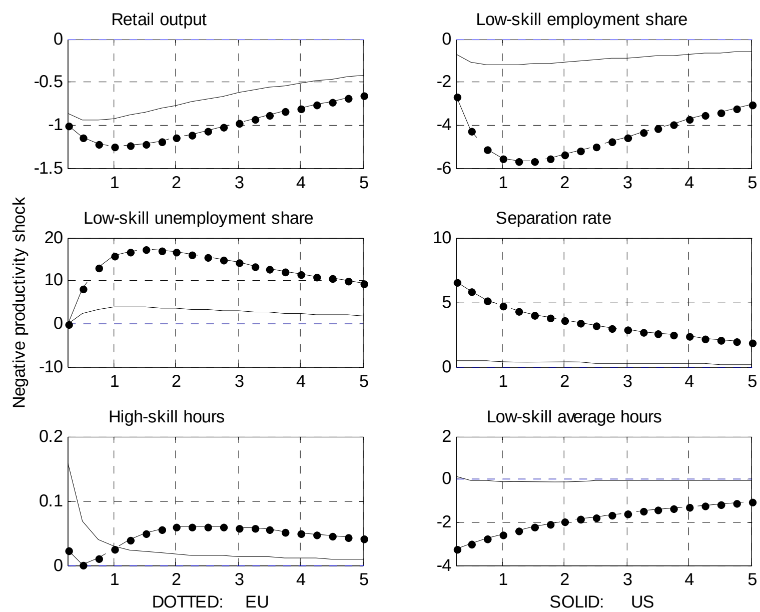

Figure 4: Response to a negative productivity shock: output, employment and unemployment shares, and hours

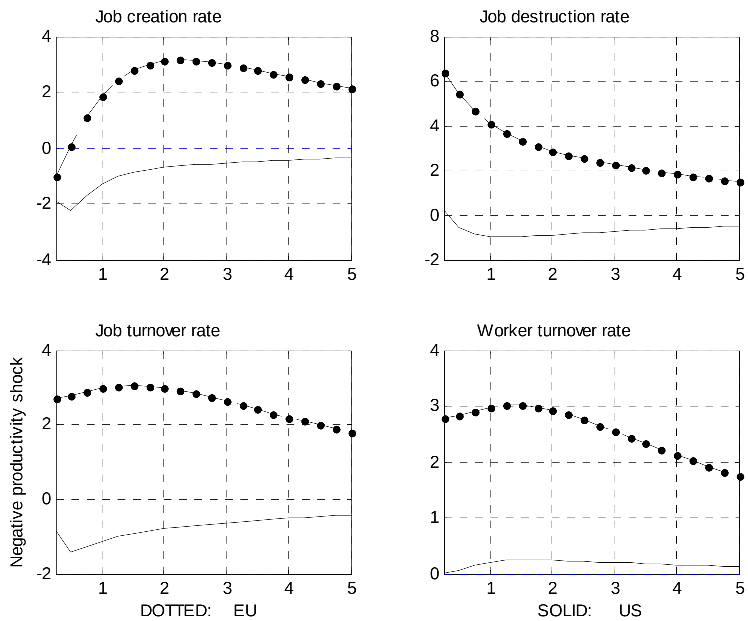

Figure 5: Response to a negative productivity shock: Job creation and destruction, job and worker turnover rates

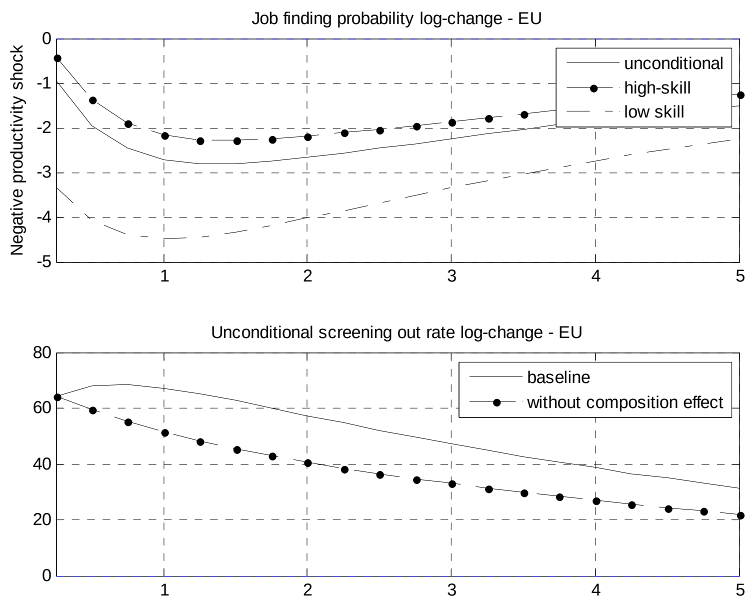

Figure 6: Response to a negative productivity shock: Job finding and screening rates

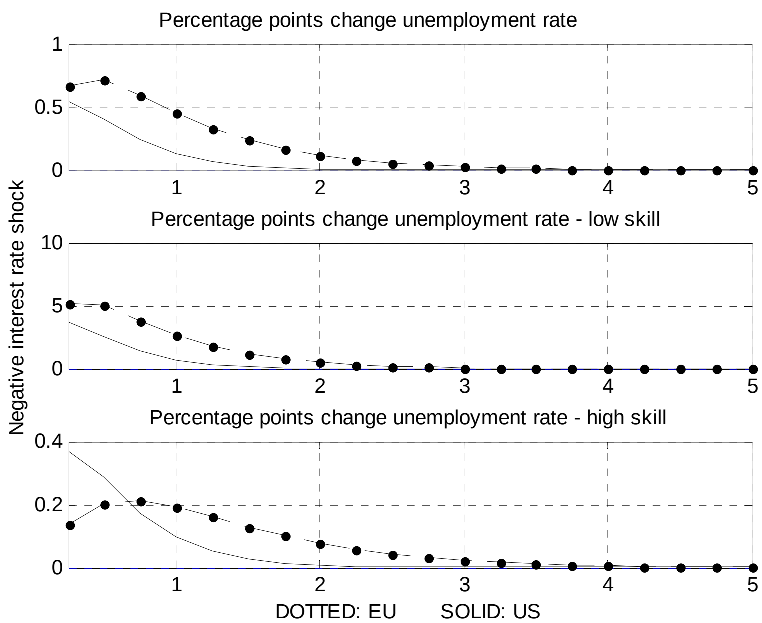

Figure 7: Response to a contractionary monetary policy shock: unemployment