Macroeconomic Dynamics, **19**, 2015, 643–668. Printed in the United States of America.
doi:10.1017/S1365100513000527

AN INCENTIVE THEORY OF MATCHING

ALESSIO BROWN
Kiel Institute for the World Economy

CHRISTIAN MERKL
Friedrich-Alexander-University Erlangen-Nürnberg
and
Kiel Institute for the World Economy

DENNIS SNOWER
Kiel Institute for the World Economy
and
Christian-Albrechts-University Kiel

This paper presents a theory of the labor market matching process in terms of incentive-based, two-sided search among heterogeneous agents. The matching process is decomposed into its two component stages: the contact stage, in which job searchers make contact with employers, and the selection stage, in which they decide whether to match. We construct a theoretical model explaining two-sided selection through microeconomic incentives. Firms face adjustment costs in responding to heterogeneous variations in the characteristics of workers and jobs. Matches and separations are described through firms' job offer and firing decisions and workers' job acceptance and quit decisions. Our calibrated model for the United States can account for important empirical regularities, such as the large volatilities of labor market variables, that the conventional matching model cannot.

Keywords: Matching, Incentives, Adjustment Costs, Unemployment, Employment, Quits, Firing, Job Offers, Job Acceptance

1. INTRODUCTION

This paper aims to explain labor market matching explicitly in terms of the microeconomic decisions of firms and job searchers. In particular, the matching process

We would like to thank participants at the Euro Area Business Cycle Network conference in Amsterdam, the EES conference "Macroeconomic fluctuations and the labor market" (CREI, Barcelona), the NBER Summer Institute, the annual meeting of the Verein für Socialpolitik, the annual congress of the EEA, the MidWest Macro Meeting, and seminars at the Université Catholique Louvain-la-Neuve, the Federal Reserve Bank of Philadelphia, the Federal Reserve Bank of Richmond, the Kiel Institute, and the Institute for International Economic Studies (Stockholm) for comments. We are indebted to Dale Mortensen and Pieter Gautier for discussing previous versions of this paper and to Larry Christiano, Britta Kohlbrecher, Wolfgang Lechthaler, Chris Reicher, and Thijs van Rens for very thorough comments. We thank Anja Bauer, Britta Gehrke, and Christoph Hiemer for excellent research assistance. Christian Merkl thanks the Fritz Thyssen Foundation for support during his visit at the NBER. Address correspondence to: Christian Merkl, Friedrich-Alexander-Universität Erlangen-Nürnberg, Lange Gasse 20, 90403 Nürnberg, Germany; e-mail: christian.merkl@fau.de.

 1365-1005/13

is decomposed into (a) firms' incentives to make job offers and to fire and (b) workers' incentives to accept job offers and to quit. Thus we call this analysis an "incentive theory of matching," because the matching process is motivated by the incentives of agents on both sides of the market.

In addition to providing a microfoundation for the matching process, this paper contributes to the vibrant, ongoing debate on the driving sources of labor market dynamics [see, for example, the new dynamic labor market models of Christiano et al. (2010) and Gali (2011),[1] which make interesting advances in fitting the U.S. data]. These new models do not only help to better understand the transmission of aggregate shocks; they also have far-reaching normative implications. This aspect is highlighted in Faia (2008, 2009), showing that in contrast to a neoclassical labor market, monetary policy may face severe trade-offs in a labor market with search and matching.

The mainstream literature on labor market search and matching views the number of unemployed job searchers and vacancies as inputs into a matching process, whose outcome is the number of hired workers. The matching function, meant as a summary description of this matching process, may represent all kinds of heterogeneities and frictions.[2] The matching function is also assumed to be stable. This paper attempts to look into this black box and improve our understanding of the matching process in terms of the choices made by firms and job searchers. The paper also seeks to do the same for the breakup of employment relationships. By decomposing the matching and separation processes into the maximization problems of heterogeneous firms and heterogeneous workers, our analysis makes it possible to examine how business cycle variations and labor market policies influence the incentives of firms and workers to find one another and separate from one another, and thereby may well affect the frequency of matches per unit of time.

In our new framework, we distinguish between two component stages of the matching process: (i) the contact stage, in which job searchers make contact with employers who have vacancies, and (ii) the selection stage, in which both potential employers and job searchers gain some information about one another and decide whether to match. In the selection stage, we decompose hires into firms' job offer decisions and workers' job acceptance decisions. Furthermore, we decompose separations into firms' firing decisions and workers' quit decisions.

Although some contributions to the search and matching literature take these two stages into account as separate decision-making processes, the full implications of this distinction for labor market dynamics have thus far not been worked out. An interesting exception is Silva and Toledo (2009), which has a contact and selection margin. However, the role of ex post labor turnover costs is very different in their model and ours. Whereas our model uses labor turnover costs to drive a wedge between selection and firing rate, theirs uses these costs mainly to reduce the surplus in a search and matching model and thereby amplify the reaction to aggregate shocks (see Section 5.3 for a detailed discussion with analytical derivations). To the best of our knowledge, this paper is the first to show that a pure selection model can generate amplification effects and the Beveridge curve at the same time.

We address these issues through a theoretical model of two-sided selection among heterogeneous firms and workers and calibraton of this model for the U.S. economy. Our theoretical model has the following distinctive features: (a) The matching process is decomposed into contact and selection stages, and the selection stage is viewed as the outcome of two-sided search among heterogeneous agents. (b) Selection and separation are viewed as analogous phenomena, viz., the making and the breaking of employment relationships. Whereas selection is analyzed in terms of decision makers' incentives to offer and accept jobs, the separation process is analyzed in terms of incentives to fire and quit. (c) The match-specific shocks that give rise to selection and deselection are not just productivity perturbations, but shocks to both firms' costs and workers' disutility of work.[3] (d) The making and breaking of matches in our model are influenced by hiring and firing costs,[4] which drive a wedge between the job-finding and the retention rate, so that the proportion of contacts that lead to new hires is less than the proportion of incumbent workers that are retained.

The paper is organized as follows. In Section 2 we present a dynamic model of two-sided selection in terms of optimizing decisions of firms and workers. This model is calibrated in Section 3. Section 4 presents the numerical results. We show that the calibrated model can account for important empirical regularities. First, our model generates labor market volatilities that are close to what can be found in the empirical data, specifically for the unemployment rate, the job finding rate, and the separation rate. This is remarkable, as we do not rely on any form of real wage rigidity. The standard calibration of the conventional matching model (with exogenous or endogenous separations) is unable to generate these high volatilities of labor market variables [see Shimer (2005)]. The "standard" calibration of the model excludes rigid wages and small surplus calibrations. Although the rigid wage version of the search and matching model can also generate higher volatilities [Hall (2005)], it implies the counterfactual prediction that wages are acyclical. Thus we do not make this assumption here. We also do not rely on Hagedorn and Manvoskii's (2008) small-surplus calibration, in which the average unemployed worker is basically indifferent between working and not working. Second, our model generates a strong negative correlation between vacancies and unemployment (i.e., the Beveridge curve correlation). The standard calibrations of the matching model, with endogenous job destruction [see Krause and Lubik (2007)], have trouble accounting for this stylized fact. The search and matching model with exogenous job destruction actually has a strong Beveridge curve [see Shimer (2005)]. However, there is an intensive debate in the literature as to whether separations are exogenous or not [see, for example, Hall (2006) and Fujita and Ramey (2009) for opposing views]. Separations are endogenous in our analysis.

Section 5 examines a highly simplified analytical version of our model in order to explore the intuition underlying our results. The intuitive reason that our model is more successful than the conventional matching model at replicating the preceding stylized facts is that macroeconomic shocks are propagated differently. In the conventional matching models, the employment effect of a change in aggregate productivity depends on the change in new hires generated by the matching function, and this matching function exhibits diminishing returns (i.e., a declining marginal product of matches with respect to unemployment and vacancies). In our incentive model, the adjustments are made on a different margin. Because both firms and workers in our model face heterogeneous match-specific shocks, a change in aggregate productivity affects the range of match-specific shocks over which firms are willing to make job offers and workers are willing to accept these offers. Because aggregate productivity shocks are persistent, they can have a substantial leverage effect on the expected present value of profit generated by newly hired workers and incumbent workers, and thereby a strong effect on the hiring and firing thresholds. The persistence of aggregate productivity shocks, via its association with the persistence of wages, also affects the expected present value of workers' utility, thereby influencing the job acceptance and quitting thresholds. In short, whereas an aggregate productivity shock affects employment via the matching function in conventional matching models, it affects employment in our model via the mass of the distribution of match-specific shocks at which job-offer decisions and job-acceptance decisions, as well as firing and quitting decisions, are made. This explains why our incentive model is more successful than the conventional matching model in generating the observed high volatilities of the unemployment rate and the job finding rate. The other stylized facts can be understood intuitively along the same lines.

Section 6 summarizes extensions of our model and interesting future routes for research. Section 7 concludes.

2. A DYNAMIC, TWO-SIDED INCENTIVE MODEL

To set the stage, we begin by presenting a dynamic incentive model containing two-sided selection in the labor market. In the context of conventional calibrations, we will show that the incentive model fares better than the standard matching model in reproducing the volatilities of major labor market variables. The sequence of decisions may be summarized as follows. First, vacancies are posted. Second, unemployed workers make contact with firms. Third, the aggregate productivity shock and the idiosyncratic shocks are revealed. Fourth, the firms make their hiring and firing decisions and the households make their job acceptance and refusal decisions, based on the realization of the aggregate and idiosyncratic shocks and anticipating the bargaining results. Fifth, the wage is determined. The assumption that employment decisions are made before wage decisions parallels what is assumed in traditional search and matching models [for example, in Pissarides (2000, Chap. 1), vacancies are posted first, some workers are matched, and then wages are determined]. This assumption also permits us to distinguish between quitting and firing decisions. In contrast, if wages are determined prior to the employment decisions and are the outcome of bargaining between each employee and her employer, then wage formation takes place only when there is a positive bargaining surplus to be shared, so that firing and quitting do not occur after wage setting. We now proceed to consider these decisions in reverse order.

2.1. Wage Determination

In endogenizing the real wage, our aim is to formulate a model that (i) is simple and tractable, (ii) enables us to distinguish between job offer decisions and job acceptance decisions in the hiring process, and between firing decisions and quit decisions in the separation process, and (iii) contains a wage bargaining process that is comparable to the one in the conventional matching models and is able to reproduce the stylized fact that wages are as volatile as productivity. This comparability allows us to infer that the predictions of our model differ from those of the conventional matching models on account of the differences between our two-sided selection process and the matching function, rather than on account of differences in wage determination.

The distinction between quitting and firing is important. Employers and employees generally know whether a worker has been fired or left of her own accord. Various aspects of labor law (such as firing "with cause") depend on this distinction. It is well known that quitting is quantitatively at least similarly important for separations as firing [e.g., Hall (2006)].

It is important to emphasize at the outset that our main quantitative results—accounting for the observed labor market volatilities—are not driven by our particular timing assumptions and model of wage formation. If we allow wages to be determined before employment, either as individualistic bargaining or as union bargaining, we still obtain broadly similar results. This suggests that our results depend primarily on our labor market selection mechanism rather than on our particular specification of wage determination.[5]

Because the wage is set after the employment decisions, the hiring and firing costs, as well as the match-specific random shocks, are already sunk.[6] Thus, all workers obtain the same wage, which, for simplicity, is assumed proportional to productivity a:

$$w_t = \gamma a_t, \tag{1}$$

where γ $(0 > \gamma > 1)$ is a constant. The average aggregate productivity of each worker is a, a positive constant subject to random aggregate productivity shocks. This wage equation may be interpreted as the outcome of Nash bargaining[7] between each employer and employee. Given the timing of economic decisions, wages are privately efficient because the idiosyncratic shock realizations are already sunk when the wage formation takes place.

Choosing this simple wage equation has three advantages. First, it ensures that our results are not generated by any kind of wage rigidity (because the wage moves proportionally with productivity). This is important because it is well

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press

known that rigid wages imply that labor market shocks have larger amplification effects and thereby generate greater labor market volatilities [e.g. Hall (2005)]. Second, as noted, it enables us to separate the decisions of workers and firms, thereby distinguishing especially firms' firing from workers' quitting decisions. Third, our simple wage equation ensures analytical tractability and allows us to compare the steady state elasticities of our model with the traditional search and matching model (under the same wage formation) and explain how these models differ in their amplification effects.

2.2. The Firm's Behavior

We assume that the profit generated by a particular worker at a particular job is subject to a match-specific random shock ε_t in period t, which is meant to capture idiosyncratic variations in workers' suitability for the available jobs.[8] For example, workers in a particular skill group and sector may exhibit heterogeneous profitabilities due to random variations in their state of health, levels of concentration, and mobility costs, or to random variations in firms' operating costs, screening, training, and monitoring costs, and so on. The random shock ε_t is i.i.d. across workers, with a stable probability density function $G_\varepsilon(\varepsilon_t)$, known to the firm. Let the corresponding cumulative distribution be $J_\varepsilon(\varepsilon_t)$; specifically, the cumulative distribution at the point ν is $J_\varepsilon(\nu) = \int_{-\infty}^{\nu} G_\varepsilon(\varepsilon_t) d\varepsilon_t$. In each period of analysis a new value of ε_t is realized for each worker. The unemployment benefits b, the hiring cost h, and the firing cost f are all constant. The hiring cost includes the administrative costs, screening costs, retraining costs, and relocation costs, as well as the basic instruction, mentoring, and on-the-job training costs that are required to integrate the worker into the firm's workforce.

The firing decision. The firm maximizes the present value of its expected profit, with a time discount factor δ. The expected present value of profit generated by an incumbent employee, after the random cost term ε_t is observed, is[9]

$$\pi_t^{\mathrm{I}}(\varepsilon_t) = (a_t - w_t - \varepsilon_t) + \delta E_t \left[(1 - \sigma_{t+1}) \pi_{t+1}^{\mathrm{I}} - \phi_{t+1} f \right], \qquad (2)$$

where δ is the time discount factor, a_t is the incumbent employee's productivity, w_t is the real wage, the superscript "I" stands for the incumbent employee who has been retained, σ_{t+1} is the separation rate, f is the firing cost per worker, assumed constant and paid with the firing probability ϕ_{t+1}, and $E(\pi_{t+1}^{\mathrm{I}})$ denotes the future expected average profit of an incumbent:

$$E_t\left(\pi_{t+1}^{\mathrm{I}}\right) = E_t\left\{a_{t+1} - w_{t+1} - \left[\varepsilon_{t+1} | \left(\varepsilon_{t+1} < \nu_{t+1}^{\mathrm{I}}\right)\right] + \delta\left[(1 - \sigma_{t+2})\pi_{t+2}^{\mathrm{I}} - \phi_{t+2} f\right]\right\}. \qquad (3)$$

$E_t(\varepsilon_{t+1} | (\varepsilon_{t+1} < \nu_{t+1}^{\mathrm{I}}))$ is the expectation of the random term ε_{t+1}, conditional on this random cost being sufficiently small to permit retention of the incumbent

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press

employee. We define the *incumbent employee's retention incentive* as

$$\nu_t^{\mathrm{I}} = a_t - w_t + \delta E_t \left[(1 - \sigma_{t+1}) \pi_{t+1}^{\mathrm{I}} - \phi_{t+1} f \right] + f; \tag{4}$$

i.e., the firm's retention incentive (its payoff from retaining a worker) is the difference between the gross expected profit from retaining the employed worker $\{a_t - w_t + \delta E_t[(1-\sigma_{t+1})\pi_{t+1}^{\mathrm{I}} - \phi_{t+1} f]\}$[10] and the expected profit from firing her $(-f)$.

An incumbent worker is fired in period t when the realized value of the random cost ε_t is greater than the incumbent worker's employment incentive: $\varepsilon_t > \nu_t^{\mathrm{I}}$. Because the cumulative distribution of ε_t is $J_\varepsilon(\nu_t^{\mathrm{I}})$, the employed worker's firing rate is

$$\phi_t = 1 - J_\varepsilon \left(\nu_t^{\mathrm{I}} \right). \tag{5}$$

The job offer decision. The expected present value of profit generated by an entrant $\pi_t^{\mathrm{E}}(\varepsilon_t)$, given that a contact has been made and the random cost ε_t has been observed, is

$$\pi_t^{\mathrm{E}}(\varepsilon_t) = a_t - w_t - \varepsilon_t - h + \delta E_t \left[(1 - \sigma_{t+1}) \pi_{t+1}^{\mathrm{I}} - \phi_{t+1} f \right], \tag{6}$$

where h is the constant hiring cost and the superscript "E" stands for "entrant."

We define the firm's expected *job offer incentive* ν_t^{E} as the difference between the gross expected profit from a hired worker $\{a_t - w_t - h + \delta E_t[(1-\sigma_{t+1})\pi_{t+1}^{\mathrm{I}} - \phi_{t+1} f]\}$ and the profit from not hiring him (i.e., zero):

$$\nu_t^{\mathrm{E}} = a_t - w_t - h + \delta E_t \left[(1 - \sigma_{t+1}) \pi_{t+1}^{\mathrm{I}} - \phi_{t+1} f \right]. \tag{7}$$

A job is offered when $\nu_t^{\mathrm{E}} > \varepsilon_t$. Thus, given that the cumulative distribution of ε_t is $J_\varepsilon(\nu_t^{\mathrm{E}})$, the job offer rate is

$$\eta_t = J_\varepsilon \left(\nu_t^{\mathrm{E}} \right). \tag{8}$$

Note that because of the hiring and firing costs, the retention incentive exceeds the job offer incentive ($\nu_t^{\mathrm{I}} > \nu_t^{\mathrm{E}}$) and thus the retention rate exceeds the job offer rate ($1 - \phi_t > \eta_t$).

2.3. The Worker's Behavior

The worker faces a discrete choice of whether or not to work. Her disutility of work effort at a given job is e_t, a random variable, which is i.i.d., with a stable probability density function $G_e(e_t)$, known to the worker. The corresponding cumulative distribution is $J_e(e_t)$. The random variable captures match-specific heterogeneities in the disagreeability of work, due to such factors as idiosyncratic reactions to particular workplaces or variations in the qualities of these workplaces. The worker's utility is linear in consumption and work effort. She consumes all her income.

The incumbent employed worker's expected present value of utility from working $\Omega_t^{\mathrm{N}}(e_t)$ for a given e is[11]

$$\Omega_t^{\mathrm{N}}(e_t) = w_t - e_t + \delta E_t\left[(1-\sigma_{t+1})\,\Omega_{t+1}^{\mathrm{N}} + \sigma_{t+1}\Omega_{t+1}^{\mathrm{U}}\right], \tag{9}$$

where $E_t(\Omega_{t+1}^{\mathrm{N}})$ is the expected present value of utility of the following period (before the realized value of the shock e_{t+1} is known), the superscript "N" stands for "employed," and "U" for "unemployed":[12]

$$E_t\left(\Omega_{t+1}^{\mathrm{N}}\right) = E_t\left\{w_{t+1} - (e_{t+1}|e_{t+1} < \iota_{t+1}) + \delta\left[(1-\sigma_{t+2})\,\Omega_{t+2}^{\mathrm{N}} + \sigma_{t+2}\Omega_{t+2}^{\mathrm{U}}\right]\right\}. \tag{10}$$

The expected present value utility from unemployment is

$$\Omega_t^{\mathrm{U}} = b + \delta E_t\left[\mu_{t+1}\Omega_{t+1}^{\mathrm{N}} + (1-\mu_{t+1})\,\Omega_{t+1}^{\mathrm{U}}\right]. \tag{11}$$

An unemployed worker's expected "work incentive" ι_t is the expected gross difference[13] between these two utility streams:

$$\iota_t = \Omega_t^{\mathrm{N}}(e) - e - \Omega_t^{\mathrm{U}}; \tag{12}$$

$$\iota_t = w_t - b + \delta E_t\left[(1-\sigma_{t+1}-\mu_{t+1})\,\Omega_{t+1}^{\mathrm{N}} - (1-\sigma_{t+1}-\mu_{t+1})\,\Omega_{t+1}^{\mathrm{U}}\right]. \tag{13}$$

Thus the unemployed accepts a job offer when $e_t < \iota_t$. Consequently, the job acceptance rate is

$$\alpha_t = J_e(\iota_t). \tag{14}$$

Along the same lines, the incumbent worker decides to quit his job when the present value of becoming unemployed exceeds the present value of remaining employed $[\Omega_t^{\mathrm{N}}(e) < \Omega_t^{\mathrm{U}}]$, so that his expected work incentive is lower than the utility cost $e_t > \iota_t$. Thus the quitting rate[14]

$$\chi_t = 1 - J_e(\iota_t). \tag{15}$$

2.4. Employment

The change in employment is the difference between the number of hires and the number of fires. The number of hires depends on the contact probability, the job offer probability, and the job acceptance probability. We assume that contacts take place before firms make their job offer decision, after which workers make their job acceptance decision.

The contact and selection stages are distinct in practice. In the contact stage, the job searchers and potential employers have relatively little information about one another, so that workers and vacancies each appear relatively homogeneous (as assumed in conventional matching functions). At this stage, workers and firms are engaged in a process of "outreach," i.e., reaching out to people who were hitherto unknown. In the selection stage, the two parties exchange enough information

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press

about one another to permit them to decide whether to consummate the match. On the basis of this additional information, workers and vacancies appear more heterogenous. At this stage, workers and firms are in the process of assessing the "match suitability." The labor market frictions relevant to the contact stage are search costs; the frictions relevant to the selection stage are hiring costs for the firm and job acceptance costs for the worker.[15] The outcome of the contact stage is an interview; the outcome of the selection stage is a hire or a rejection. A job searcher who makes contact with a potential employer becomes an applicant; an applicant who is selected becomes an entrant into the firm's workforce.

The traditional search and matching literature did not pay much attention to the distinction between contact and selection.[16] Specifically, the matching function has been interpreted in two ways. In the first, traditional interpretation [e.g., Pissarides (2000, Chap. 1)], the matching function describes the outcome of both contact and selection, explaining how given job searchers and vacancies lead to new hires. We call this the "encompassing matching function," because it encompasses both the contact and selection stages.

In the second interpretation, the matching function covers only the contact stage, and thus we call it the "contact function." Matching models with endogenous separations—where workers and firms are first matched through a contact function and then decide whether to continue or to sever the contact in response to productivity perturbations—can be interpreted in this vein [see, for example, the models of Mortensen and Pissarides (1994) and den Haan et al. (2000)]. In these "productivity perturbation models," however, the distinction between contact and selection is incomplete for two reasons. First, when these models are calibrated, the calibration relates unemployment to new hires, not to contacts (such as interviews).[17] Second, these models invariably assume that the proportion of interviews that do not lead to hiring is equal to the proportion of currently employed workers who separate from their jobs. But, in practice, interviews fail far more frequently than existing employment relationships, and the job finding rate is much lower than the retention rate. In the United States, the average monthly job-finding rate is 0.45 and the average retention rate is around 0.97; see, e.g., Shimer (2005). This indicates that we need to distinguish between the breaking of contacts and the breaking of selection (i.e., deselection of employees).

In the context of our model, the number of unemployed workers who get jobs in period t is $\mu_t U_{t-1}$, where U_{t-1} is the number unemployed in the previous period. The number of employed people who separate from their jobs in period t is $\sigma_t N_{t-1}$, where N_{t-1} is the number employed in the previous period.

The change in employment is $\Delta N_t = N_t - N_{t-1} = \mu_t U_{t-1} - \sigma_t N_{t-1}$. Let L be the labor force (assumed constant), whereas $n_t = N_t/L$ and $u_t = U_t/L$ are rates of employment and unemployment, respectively, so that $u_t = 1 - n_t$.

Let C_t be the number of contacts made in period t and c_t be an unemployed worker's contact probability:

$$c_t = C_t/U_{t-1}. \tag{16}$$

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press

An unemployed worker gets a job when three conditions are fulfilled: (i) she makes contact with an employer, (ii) she receives a job offer, and (iii) she accepts that offer. Thus the match probability (μ_t) is the product of the contact probability (c_t), job offer probability (η_t), and acceptance probability (α_t):

$$\mu_t = c_t \eta_t \alpha_t. \tag{17}$$

An employee separates from her job when at least one of two conditions is satisfied: (i) she is fired or (ii) she quits. Thus the separation probability is

$$\sigma_t = \phi_t + \chi_t - \phi_t \chi_t, \tag{18}$$

and normalizing the labor force to unity, equilibrium employment n_t may be described by the associated employment dynamics equation,

$$n_t = \mu_t + (1 - \sigma_t - \mu_t)\, n_{t-1}, \tag{19}$$

where the degree of employment persistence—measured by the persistence parameter $\omega_t = (1 - \mu_t - \sigma_t)$—depends inversely on the matching rate and separation rate.

2.5. Contacts

For simplicity, we let the contact function be[18]

$$C_t = U_{t-1}. \tag{20}$$

Vacancies are posted before ε_t is realized. As in the conventional search and matching literature, we assume free entry of firms, so that the number of vacancies V_t^* is determined by a zero-profit condition.[19] Let V_t be the number of posted vacancies, κ be the cost of posting a vacancy, and U_{t-1} be the number unemployed in the previous period.

The probability that a vacancy is filled is $(U_{t-1}\mu_t)/V_t$, i.e., the probability of a contact times the probability that the contact leads to a match, divided by the number of vacancies. The expected profit per match is $\{a_t - w_t - h + \delta E_t[(1 - \sigma_{t+1})\pi^{\mathrm{I}}_{t+1} - \phi_{t+1} f] - E_t(\varepsilon_t | \varepsilon_t < \nu^{\mathrm{E}})\}$, where $E_t(\varepsilon_t | \varepsilon_t < \nu^{\mathrm{E}})$ is the expected value of the idiosyncratic productivity shock ε_t conditional on match formation.

Thus the zero-profit condition for posting vacancies is

$$a_t - w_t - E_t\left[\varepsilon_t | \left(\varepsilon_t < \nu_t^{\mathrm{E}}\right)\right] - h - \frac{\kappa}{\frac{U_{t-1}}{V_t^*}\eta_t \alpha_t} + \delta E_t\left[(1 - \sigma_{t+1})\, \pi^{\mathrm{I}}_{t+1} - \phi_{t+1} f\right] = 0, \tag{21}$$

where the vacancy posting cost is κ.

2.6. The Labor Market Equilibrium

The labor market equilibrium is the solution of the system comprising the following equations:

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press

- *Incentives:* the incumbent worker retention incentive ν_t^{I} equation (4), the job offer incentive ν_t^{E} equation (7), and the work incentive ι_t equation (12).
- *Employment decisions:* the firing rate ϕ_t equation (5) and the job offer rate η_t equation (8).
- *Work decisions:* the job acceptance rate α_t equation (14) and the quitting rate χ_t equation (15).
- *Contacts and vacancies*: the contact function C_t equation (20) and the number of vacancies V_t^* equation (21).
- *Match and separation probabilities:* the match probability μ_t equation (17) and the separation probability σ_t equation (18).
- *Employment and wage:* the employment n_t equation (19) and the negotiated wage w_t equation (1).

3. CALIBRATION

We now calibrate our incentive model for the U.S. economy. The calibration is done on a monthly basis. The simulation results are aggregated to quarterly frequency to make them comparable to the empirical data.

Our monthly discount factor $\delta = \frac{1}{1+r}$ is consistent with an annual real interest rate of 4% . We normalize the average productivity (a) to 1. As in Hall (2005) and Shimer (2005), we set b by applying a replacement rate of $\beta = 40\%$ of the wage. The wage parameter γ is set to 0.5, the value commonly found in the literature on workers' bargaining power.

In this section, we use the simplifying assumptions that each unemployed worker makes one distinct contact in each period: $C_t = U_{t-1}$, implying a contact rate of unity: $c_t = 1$.[20] In the working paper version these assumptions are relaxed by assuming a Cobb–Douglas contact function. The vacancy posting costs κ of 0.19 are chosen to satisfy the zero-profit condition.

Dolfin (2006) shows that an average U.S. worker spends 203 hours in training activities during her first three months of employment, whereas other employees spend around 146 hours training her. In line with this evidence, we set hiring costs, h, to 130% of the monthly productivity.[21]

The literature does not provide reliable direct estimates of the magnitude of US firing costs. Thus, we assess these costs indirectly. For this purpose, note that Belot et al. (2007) provide index measures of employment protection for regular jobs in the United States and United Kingdom, and that Bentolila and Bertola (1990) provide estimates of the average magnitude of U.K. firing costs on a yearly basis.[22] Assuming that the index measures of employment protection are proportional to the estimates of the magnitude of firing costs, we multiply the magnitude of the U.K. firing costs by the ratio of the U.S. to the U.K. employment protection indices to derive a rough estimate of the magnitude of U.S. firing costs.Accordingly, the magnitude of monthly U.S. firing costs, relative to productivity, is 0.08. The same exercise based on other industrialized countries (France, Germany, and Italy),

TABLE 1. Steady state values for a contact function $C_t = U_{t-1}$

Variable	Description	Steady state value
u	Unemployment rate	0.056
μ	Match probability	0.450
c	Contact rate	1
η	Hiring/job offer rate	0.456
σ	Separation rate	0.0268
ϕ	Firing rate	0.0134
$\chi = 1 - \alpha$	Job quit rate	0.0136
θ	Market tightness	1

however, yields higher estimates of U.S. firing costs. Thus we choose a value of 0.1 for our baseline calibration.[23]

We assume that the random profitability term ε_t and the utility shock e_t have cumulative distributions given by logistic functions with scale factors s_ε and s_e and expected values $\bar{\varepsilon}$ and $\bar{e}$, respectively.

We set the expected value of the utility shock to $\bar{e} = 0.15$, which implies that the current-period value of unemployment is 70% of the value of employment. Although this is an intermediate value between Shimer's (2005) and Hagedorn and Manovskii's (2008) calibrations, in contrast to Hagedorn and Manovskii (2008), workers are consequently not indifferent between working and not working.[24]

After having set all our other parameter values, we set the remaining three free distributional parameters (the average operating costs $\bar{\varepsilon}$, the scale factor of the cumulative distribution of s_ε, and the scale factor of the cumulative distribution of s_e) to replicate three steady state labor market flow rates: the job acceptance rate α, the job offer rate η, and the firing rate ϕ (i.e., given our preceding parameter choices and the targeted flow rates, we constrain ourselves with respect to the remaining distributional parameters to be chosen). These steady state values are calibrated as follows. The match probability μ, which is the probability of a worker finding a new job within one period, is calibrated to 45%, as in Shimer (2005) and Hagedorn and Manovskii (2008). The unemployment rate u is calibrated to 5.6% [as in Shimer (2005)]. According to our employment dynamics equation (19), steady state unemployment is $u = \frac{\mu}{\mu+\sigma}$, which implies a separation rate of 2.68%. Based on Hall (2006), who shows that fires and quits have approximately the same share in separation, we assume firings to account for 50% of the separations, namely $\phi = 1.34\%$. Equation (18) then yields the quit rate of $\chi = 1.36\%$. Because α is equal to $1 - \chi$, the job acceptance rate is set at 98.64%. Recalling that $\mu = c\alpha\eta$ and that we have assumed $c = 1$, the implied job offer rate η is 45.6%.We determine the number of vacancies by targeting a market tightness equal to 1. See Table 1 for a summary of targeted steady state values.

The variance of the idiosyncratic cost shock is of significant importance for the aggregate dynamics, because the lower it is, the stronger are the reactions to

TABLE 2. Parameter values for a contact function $C_t = U_{t-1}$

Parameter	Description	Value
a	Productivity	1
δ	Discount factor	0.997
β	Replacement rate b/w	0.4
γ	Bargaining power	0.5
f	Firing cost	0.1
h	Hiring cost	1.3
κ	Vacancy posting cost	0.188
r	Discount factor	0.997
$\bar{e}$	Average value of leisure	0.15
$\bar{\varepsilon}$	Average operating costs	0.483
s_ε	Scale factor of the cumulative distribution of ε_t	0.313
s_e	Scale factor of the cumulative distribution of e_t	0.075
ρ_a	Autocorrelation of the aggregate productivity shock	0.975
ϖ_a	Standard error of the aggregate productivity shock	0.007

productivity in our model. This value could be approximated by the residuals of the variance of log male wages as calculated by Heathcote et al. (2010), which is on the average approximately 0.32 between 1970 and 2005. Our baseline calibration yields a variance (standard deviation) of the idiosyncratic productivity shock of $\sigma_\varepsilon^2 = 0.32$ ($\sigma_\varepsilon = 0.57$), which matches the value in Heathcote et al. (2010). This value is larger than values commonly used in the literature [e.g., den Haan et al. (2000) choose a standard deviation of $\sigma_\varepsilon = 0.1$, Krause and Lubik (2007) use $\sigma_\varepsilon = 0.12$, and Mortensen and Pissarides (1994) choose $\sigma_\varepsilon = 0.0375$]. Using this relatively more conservative value, we bias the dynamics against our model, and this ensures that our volatilities are not driven by an unrealistically small standard deviation of the idiosyncratic cost shock.

Finally, we normalize the autocorrelation (ρ_a) of the aggregate productivity shock and normalize the standard error so that we obtain the empirical values for the autocorrelation and the volatility of productivity in the model simulation that follows. Table 2 summarizes our calibrated parameter values.

4. RESULTS: LABOR MARKET VOLATILITIES AND CORRELATIONS

Costain and Reiter (2008) and Shimer (2005) show that the conventional calibration of the matching model is unable to replicate the volatility of the job-finding rate, the unemployment rate, and other labor market variables in response to productivity shocks. Table 3 shows the empirical volatilities for the United States from 1964 to 2009, HP filtered data with smoothing parameter 100,000.[25] Note that the empirical business cycle statistics in Shimer (2005) from 1951 to 2003 are very similar to those in Table 3.

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press

TABLE 3. Empirical labor market statistics

Volatilities	U	Match. r.	Sep. rate	Vac.	M. tight.	Prod.
Standard deviation	0.17	0.11	0.06	0.18	0.34	0.02
Relative to productivity	9.5	5.7	3.1	9.2	17.1	1
Quarterly autocorr.	0.92	0.88	0.71	0.94	0.93	0.88
Correlation (U, V)	−0.89					

Sources: Michaillat (2012) and own calculations.

The empirical volatilities are far greater than the corresponding volatilities in response to productivity shocks, as generated by the simulation of the conventional matching model [in its standard calibration, as calculated by Shimer (2005); see Table 4].

To compare our model with the conventional matching theory and with the empirical labor market volatilities, we used our baseline calibration (with robustness checks in the working paper version) to simulate our model for 200 quarters (i.e., 600 months). We repeated this exercise 1,000 times, and we report the average of the macroeconomic volatilities (HP filtered simulated data with smoothing parameter 100,000) in Table 5.

The differences between our model and the conventional matching model are striking. Our model can generate the high macroeconomic volatilities found in the data. Our results are all the more remarkable, as we have to resort neither to Hall's (2005) real wage rigidity assumption nor to Hagedorn and Manovskii's (2008) small surplus calibration (see wage volatility in Table 5).[26]

Specifically, the more rigid the wage in the conventional matching model [Hall (2005)], the greater the share of productivity variations that is captured by the firm and thus the greater the volatility of vacancies. However, there is evidence against the rigid-wage hypothesis from both the microeconometric and the macro perspective. Haefke et al. (2008) infer that wages for newly created jobs (i.e., those modeled in the matching model) are completely flexible. Hornstein et al. (2005) show that wages are roughly as volatile as the labor productivity on a macroeconomic level. See also Carneiro et al. (2012), Martins et al. (2012a, 2012b), and Stüber (2012) for further evidence that wages seem to be quite volatile over the business cycle. The debate on wage rigidity is certainly not resolved yet. However,

TABLE 4. Labor market statistics generated by the search and matching model from Shimer (2005)

Volatilities	U. rate	Match. r.	Sep. rate	Vac.	M. tight.	Prod.
Standard deviation	0.01	0.01	—	0.03	0.04	0.02
Relative to productivity	0.5	0.5	—	1.4	1.8	1
Quarterly autocorr.	0.94	0.88	—	0.84	0.88	0.88
Correlation (U, V)	−0.93					

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press

TABLE 5. Labor market statistics generated by the incentive theory of matching

Volatilities	U. rate	Match. r.	Sep. rate	Vac.	M. tight.	Wage	Prod.
Standard deviation	0.14	0.17	0.01	0.10	0.27	0.02	0.02
Relative to productivity	7.0	8.4	0.5	4.9	13.4	1	1
Quarterly autocorr.	0.91	0.87	0.88	0.66	0.86	0.88	0.88
Correlations (U, V)	−0.81						

at this stage, it remains to be emphasized that our model generates high labor market volatilities, even though our wages are as volatile as productivity. Our analytical analysis in Section 4 and our numerical robustness analyses in Brown et al. (2012) show that this result is not driven primarily by our specific assumptions concerning wage determination and the timing of wage versus employment decisions.

Hagedorn and Manovskii (2008) choose a small-surplus calibration to resolve the volatility puzzle of the matching model. Under this calibration, aggregate profits are only a very small share of the overall production in the steady state, so that a positive productivity shock sharply increases the relative profits. This gives a large incentive to firms to post more vacancies (because of the free entry condition). Consequently, all labor market variables become volatile. This type of calibration has several shortcomings. Besides the unrealistically low profit share, the utility value of unemployment is extremely high and workers' bargaining power is very low in the calibration. Therefore workers are almost indifferent between working and not working (i.e., unemployment and business cycle fluctuations do not create large welfare costs). We do not need to rely on any of these mechanisms in our calibration. As noted, we assume that workers' bargaining power is 50% . Furthermore, the average worker's disutility of labor and unemployment benefits make up only 70% of the current wage. As a consequence, the average worker is nowhere near indifferent between unemployment and employment. (But because workers are heterogeneous, the marginal worker with $e_t = \iota_t$ is of course indifferent between working and not working.)

Our model has an additional advantage compared to the conventional matching framework. Although a matching model with endogenous job destruction and flexible wages has trouble generating a negative correlation between vacancies and unemployment (i.e., the dynamic Beveridge curve), in our simulation, we obtain a strong negative correlation between vacancies and unemployment (namely, −0.81).[27]

In order to evaluate the performance of our model with regard to quitting, we use JOLTS data from 2000 and 2010 and calculate the correlation between quits and output (HP-filtered with the smoothing parameter 100,000). The correlation in the data is 0.91; i.e., quits are strongly procyclical. Our simulated model generates a correlation close to 1; i.e., the results are very close to those found in the data.

The reason that the distinction between quitting and being fired is so important in our model is that quitting depends on household variables, whereas being fired depends on firm variables. When this distinction ceases to be made, there are no

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press

involuntary separations. This is clearly at odds with the empirical evidence, where quitting is quantitatively at least similarly important for separations as firing, and legal procedures, where quitting and firing are commonly distinguished from one another.

From a macroeconomic perspective, quitting is strongly procyclical, and firing is strongly countercyclical, both in the data and in our model. Thus, quitting and firing roughly neutralize themselves in terms of their cyclicality. When we set the job acceptance and quitting decisions exogenously, the volatility of the separation rate relative to productivity increases to 24; i.e., our endogenous household decisions prevent firings from becoming too cyclical. It is well known from search and matching models [e.g., Krause and Lubik (2007)] that excessively volatile separations lead to a collapse of the Beveridge curve. The same happens in our framework: with exogenous household decisions, the negative business cycle correlation between vacancies and unemployment disappears.[28]

To sum up, endogenous quitting is important for the aggregate dynamics of our model in two respects. First, it prevents an excessive separation volatility. Second, it maintains the Beveridge curve in the presence of endogenous separations.

5. INTUITIVE EXPLANATION OF THE RESULTS

In this section, we seek to gain an intuitive understanding of why our model behaves differently from the conventional search and matching models. For this purpose, we make some simplifying assumptions that enable us to explore analytically the mechanism whereby productivity shocks are amplified in a model. For better comparability to the standard search and matching model, we focus on the firm side of our two-sided model, although the intuition is analogous for the household side. In particular, let us assume that aggregate productivity is constant, that the job separation rate is exogenously given, and that households always accept job offers.

5.1. The Incentive Model

Under the simplifying assumption from before, the job-finding rate is

$$\eta = J_{\varepsilon}(\nu^{\mathrm{E}}) = J_{\varepsilon}\left[\frac{a\,(1-\gamma)}{1-\delta\,(1-\phi)} - h\right]. \quad (22)$$

The corresponding elasticity of the job-finding rate with respect to productivity is

$$\frac{\partial \ln \eta}{\partial \ln a} = J'_{\varepsilon}\frac{a\,(1-\gamma)}{[1-\delta\,(1-\phi)]\,\eta}, \quad (23)$$

where J'_{ε} is the first derivative of the cumulative distribution.

If wages are determined before the employment decision is made, the realization of the operating cost shock ε is relevant to the wage bargaining. Let us suppose

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press

that the wage distribution for entrants in the incentive model is given by $w^{\mathrm{E}}(\varepsilon) = \gamma(a - \varepsilon)$, and that incumbents get the uniform wage $w^{\mathrm{I}} = \gamma a$ (under exogenous separations, they are homogeneous; see Appendix A.1 for details). Then the job-finding rate becomes

$$\eta = J_\varepsilon(\nu^{\mathrm{E}}) = J_\varepsilon \left[\frac{\frac{1-\gamma}{\gamma} a}{1 - \delta(1-\phi)} - h \right], \qquad (\mathbf{24})$$

and the elasticity of the job-finding rate with respect to productivity becomes

$$\frac{\partial \ln \eta}{\partial \ln a} = J'_\varepsilon \frac{a \frac{1-\gamma}{\gamma}}{[1 - \delta(1-\phi)]\, \eta}. \qquad (\mathbf{25})$$

Thus, the elasticity of the job-finding rate with respect to productivity is higher under the assumption of a wage distribution than under the assumption of a uniform wage, ceteris paribus. The intuition for this result is straightforward. In an upswing (characterized by a rise in aggregate productivity), firms have an incentive to hire workers associated with higher operating costs. In an economy where wages depend on workers' idiosyncratic productivity, the marginal worker then earns a lower wage than in an economy with the uniform wage. Consequently, firms have greater incentives to hire in the former economy. This shows that our assumption of a uniform wage in the dynamic incentive model biases our results against amplification of productivity shocks.

To get an idea of the amplification effects in the baseline scenario, let us assume a uniform distribution for the match-specific random cost shock with lower and upper support $-p$ and p, i.e., $U[-p, p]$. In this case, we can give a closed-form expression for the job-finding rate and for its elasticity. For the uniform wage bargaining, we obtain

$$\eta = J_\varepsilon(\nu^{\mathrm{E}}) = \left\{ \frac{a(1-\gamma)}{[1 - \delta(1-\phi)]} - h + p \right\} \Big/ 2p. \qquad (\mathbf{26})$$

Thus, the elasticity of the job-finding rate with respect to productivity is

$$\frac{\partial \ln \eta}{\partial \ln a} = \frac{a(1-\gamma)}{2[1 - \delta(1-\phi)]\, \eta p}. \qquad (\mathbf{27})$$

Given our simplifying assumptions, let us find the elasticity of the quarterly job-finding rate with respect to productivity. We use standard values of the parameters:[29] $a = 1$, $\gamma = 0.5$, $\delta = 0.99$, $\phi = 0.1$, $\eta = 0.83$. Consequently we obtain the elasticity $\partial \ln \eta / \partial \ln a = 2.76/p$. This means that there is amplification (i.e., $\partial \ln \eta / \partial \ln a > 1$), as long as $p < 2.76$. Given that the mean of productivity is normalized to 1, an upper and lower bound of the operating costs of 2.76 seems extremely large, and thus amplification can be expected to occur.

5.2. The Search and Matching Model

To illustrate the differences in the transmission mechanism between our model and the search and matching model, we impose the same simplifying assumptions (exogenous separations, wages are proportional to productivity, and no aggregate uncertainty) on a standard search and matching model. To nest Silva and Toledo's (2009) results and to illustrate their key mechanism, we assume that there are ex post hiring costs for entrants h. (Note that these ex post hiring costs are the same as the hiring costs in our model. Because the search and matching literature often calls the vacancy posting costs divided by the worker finding rate hiring costs, Silva and Toleda denote h as ex post, namely, costs that accrue after the match has taken place.[30])

Under these assumptions, we obtain the following elasticity of the job-finding rate with respect to productivity (see A.2 for some intermediate steps):

$$\frac{\partial \ln \eta}{\partial \ln a} = \frac{1-\lambda}{\lambda} \frac{\frac{a(1-\gamma)}{\kappa[1-\delta(1-\phi)]}}{\frac{a(1-\gamma)}{\kappa[1-\delta(1-\phi)]} - h}, \tag{28}$$

where λ is the elasticity of the matching function with respect to unemployment and κ are vacancy posting costs [see also Haefke et al. (2008)].

Let us start with the case of zero ex post hiring costs (i.e., $h = 0$). In this case, $\frac{\partial \ln \eta}{\partial \ln a} = \frac{1-\lambda}{\lambda}$. In the standard search and matching model (with nonrigid wages), the parameter values of the Cobb–Douglas matching function determine the elasticity of the job-finding rate with respect to productivity changes. Note that Petrongolo and Pissarides (2001) show that sensible estimations for the matching elasticity λ are in the range between 0.5 and 0.8. According to equation (28), this implies an elasticity of the job-finding rate with respect to productivity lower than 1 in the search and matching model. Thus, search and matching models with nonrigid wages and a plausible parametrization for the matching function do not have any amplification mechanism.

Silva and Toledo (2009) have proposed postmatch labor turnover costs to reduce the amplification problem in search and matching models. Equation (28) shows that hiring costs have the potential to generate very strong amplification effects. If $h \to \frac{a(1-\gamma)}{\kappa[1-\delta(1-\phi)]}$, the denominator converges to zero and the elasticity of the job-finding rate with respect to productivity converges to infinity.

5.3. The Role of Labor Turnover Cost in the Two Models

The roles of labor turnover costs in our "Incentive Theory" and in the search and matching model are very different. Intuitively, hiring costs are needed to reduce the surplus in search and matching models and thereby generate an amplifying effect. This is not the case in our model, which is easy to see by taking the first derivative of the job-finding rate with respect to productivity (i.e., not the log-deviations as before):

$$\frac{\partial \eta}{\partial a} = J'_{\varepsilon} \frac{(1-\gamma)}{[1-\delta(1-\phi)]}. \tag{29}$$

Equation (29) shows that hiring costs do not have any direct amplifying effect in the "Incentive Theory."[31] In the case of a uniform distribution, $\partial\eta/\partial a$ is completely independent of the level of hiring costs.

The search and matching model yields the following result for the first derivative of the job-finding rate with respect to productivity:

$$\frac{\partial \eta}{\partial a} = \frac{1-\lambda}{\lambda} \left\{ \frac{a(1-\gamma)}{\kappa[1-\delta(1-\phi)]} - h \right\}^{\frac{1-2\lambda}{\lambda}} \left\{ \frac{(1-\gamma)}{\kappa[1-\delta(1-\phi)]} \right\}. \tag{30}$$

For the conventionally used parameter values $0.5 < \lambda < 0.8$, $\partial^2\eta/\partial a \partial h > 0$; i.e., hiring costs amplify the model's reaction to productivity shocks. Silva and Toledo (2009, p. 77) provide an intuition for this phenomenon: Hiring costs reduce the surplus of a new match. "Therefore, a given productivity shock has a relatively larger impact on the value of new position and, in consequence, on job creation and market tightness."

5.4. Comparison

The underlying intuition for these striking differences is that the job-creation mechanisms are quite different in these two models. Because the agents in the incentive model face heterogeneous match-specific shocks, a change in aggregate productivity affects the range of match-specific shocks over which firms make their decisions. As can be seen in equation (23), changes in aggregate productivity can have a substantial leverage effect on the expected present value of profit generated by newly hired workers, and thus a strong effect on the hiring threshold. In particular, a rise in aggregate productivity increases the range of shocks ε over which the firm is willing to hire and reduces the range of shocks ε over which it has an incentive to fire. At the same time, it increases the range of shocks e over which the household is willing to accept jobs and reduces the range of shocks e over which the household has an incentive to quit. These channels determine not only the size of the impact effect of the shock, but also the strength of the propagation mechanism. In short, whereas in the conventional matching models an aggregate productivity change affects employment via the matching function and the free entry condition for vacancies, in our model this productivity change affects employment via the mass of the distribution of match-specific shocks at which employment decisions are made. This explains why the labor market volatilities for sensible parameter values are so different.

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press

6. ROBUSTNESS CHECKS AND OUTLOOK

It can be shown that our model economy continues to generate strong amplification effects, along with labor market volatilities that are close to those in the data, under different wage determination schemes. In Brown et al. (2012) this is shown for individualistic bargaining (in which each employee bargains individually with the employer) and union bargaining (where workers are paid the wage that maximizes the welfare of the median worker).

Brown et al. (2012) also relax the simplifying assumption that the number of contacts is equal to the number of unemployed workers, in two ways. First, we assume that an exogenous fraction (smaller than one) of unemployed workers gets in contact with firms. We show that as the contact probability falls, the elasticity of vacancies with respect to productivity declines, because vacancies are less likely to lead to hires. Consequently, the Beveridge curve relation weakens. Second, we assume a standard endogenous Cobb–Douglas contact function. The Beveridge curve then is somewhat stronger with endogenous contacts (because vacancies become allocationally relevant, firms have greater incentives to post vacancies).

We show that as the elasticity of contacts with respect to vacancies is increased, the model's performance in tracking the aggregate labor market data does not improve. In particular, unemployment becomes less volatile and in some cases, depending on the parameterization of the contact function, the model generates a counterfactual strong positive correlation between the matching rate and the separation rate. We view this as preliminary evidence that the responsiveness of matches to vacancies is not essential for explaining the observed labor market volatilities. It is the selection margins in hiring and separations that play the central role in this respect. These margins have been underemphasized in the literature thus far.

There are various interesting empirical exercises that may enhance our understanding of the transmission of aggregate shocks via the labor market. Further microeconomic data sets on quitting and firing, as well as interviews, would certainly be helpful. In addition, our model mechanism may be integrated into larger scale models and may be estimated using standard macroeconometric tools [e.g., Christiano et al. (2005) and Smets and Wouters (2007)].

7. CONCLUSION

This paper has presented a theory of two-sided labor market matching that distinguishes sharply between contacts and selection. The selection takes place in the presence of frictions, heterogeneous jobs, and heterogeneous workers. Our empirical results suggest that selection has a particularly important role to play in accounting for the observed labor market volatilities.

Our theory replaces the traditional encompassing matching function with a contact function combined with optimizing, incentive-based, two-sided selection decisions. The basic idea that motivates our incentive theory of matching is that the matching and separation probabilities can be understood in terms of job offer, job acceptance, firing, and quitting probabilities, which may be derived from the optimizing decisions of firms and workers. These optimizing decisions—in the presence of heterogeneous workers and jobs, as well as costs of adjustment—explain why some job-seeking workers remain unemployed and some vacant jobs remain unfilled. We have shown that, even on the basis of our radically simplifying assumptions, our calibrated incentive model can account for various important empirical regularities that have eluded the conventional matching models. In particular, our model comes close to generating the empirically observed volatilities of the unemployment rate, vacancies, the job finding rate and the separation rate. Furthermore, our model can also account for the observed strong negative correlation between vacancies and unemployment.

NOTES

1. See Gali et al. (2011) for an estimated version. For recent contributions to how different labor market frictions help explain various aspects of the data and interact with macroeconomic policies see, for example, Lechthaler et al. (2010) and Lechthaler and Snower (2013) or Khalifa (2012). For the interaction of labor market frictions with optimal labor market policies see, e.g., Moyen and Stähler (in press).

2. Pissarides (2000, pp. 3–4) claims that the matching function summarizes "heterogeneities, frictions and information imperfections" and represents "the implications of the costly trading process without the need to make the heterogeneities and the other features that give rise to it explicit." For a discussion of the interpretation of the matching function as a contact function, see Section 2.4.

3. Thus the matching rate in our model is not the same as the job offer rate (as in conventional search and matching models), but depends on both the firms' job offer rate and the workers' job acceptance rate. Danthine (2005) also develops a two-sided search model with shocks to firms' and workers' productivities, but applies it to different questions than we do.

4. The hiring costs are not to be confused with vacancy posting costs, because the vacancy posting costs are incurred before the contact is made, whereas the hiring costs are incurred after the contact. See also Silva and Toledo (2009), who introduce post match labor turnover costs into a search and matching model, or Pissarides (2009), who adds fixed job creation costs.

5. The results for a different timing (individualistic bargaining and union bargaining) are given in the working paper version [Brown et al. (2012)].

6. This is the same assumption as in Pissarides (2009); see p. 1364 and the corresponding footnote 30.

7. In particular, this interpretation involves assuming that fallback profit is zero (leaving future values unaffected) and that the fallback wage is either zero or an exogenous constant that is proportional to the negotiated wage. The bargaining game is in line with Hall and Milgrom (2008) and Binmore et al. (1986).

8. Because each worker draws from the same distribution of random shocks, ε_{it}, we omit the subscript i for notational simplicity.

9. $\pi_t^{\mathrm{I}}(\varepsilon_t)$ is a function of the actual realization of the cost shock, whereas $E_t(\pi_{t+1}^{\mathrm{I}})$ is a function of the expected future average cost shock $E_t(\varepsilon_{t+1}|(\varepsilon_{t+1} < \nu_{t+1}^I))$.

10. This "gross" profit is the expected profit generated by retaining an incumbent employee or by hiring an unemployed worker, without taking the match-specific shock ε_t into account.

11. $\Omega_t^{\mathrm{N}}(e_t)$ is a function of the actual realization of the disutility shock, whereas $E_t(\Omega_{t+1}^{\mathrm{N}})$ is a function of the expected future average utility shock $E_t(e_{t+1}|e_{t+1} < \iota_{t+1})$.

12. Observe that on the firm's side, we distinguish between entrants (E) and incumbent workers (I), whereas on the workers' side, we distinguish between employed (N) and unemployed (U) workers. The rationale for these two distinctions is that the firm can hire two types of workers (entrants and incumbents), whereas the worker can be in two states (employment and unemployment).

13. "Gross" means that the disutility shock e_t is not taken into account.

14. Note that, for simplicity, we have assumed that the job acceptance rate is identical to the job staying rate ($\alpha_t = 1 - \chi_t$). When unemployed workers face costs of adjusting to employment (e.g., buying a car to get to work, or psychic costs of changing one's daily routine) or when employed workers face costs of adjusting to unemployment (e.g., building networks of friends with potential job contacts, psychic costs of adjusting to joblessness), then the job acceptance rate will fall short of the job retention rate. Specifically, for example, the unemployed worker's job acceptance incentive could be expressed as $\iota^{\mathrm{U}} = \Omega_t^{\mathrm{N}}(e) - e - \Omega_t^{\mathrm{U}} - \xi^{\mathrm{U}}$, where ξ^{U} is the cost of adjusting to employment, and the incumbent worker's job retention incentive could be expressed as $\iota_t^{\mathrm{N}} = \Omega_t^{\mathrm{N}}(e) - e - \Omega_t^{\mathrm{U}} + \xi^{\mathrm{N}}$, where ξ^{N} is the cost of adjusting to unemployment. Then the job acceptance rate becomes $\alpha_t = C_e(\iota_t^{\mathrm{U}})$ and the job retention rate becomes $C_e(\iota_t^{\mathrm{N}})$, so that the quitting rate becomes $\chi_t = 1 - C_e(\iota_t^{\mathrm{N}})$.

15. For analytical simplicity and calibration tractability, the latter costs are not considered in the model presented here.

16. As noted, however, see Silva and Toledo (2009). Another recent contribution distinguishing these two stages is Menzio and Shi (2011). A contact function brings workers and vacancies together. Based on the idiosyncratic productivity of their match, the firm makes a hiring decision. There are at least two major differences from our model. First, we have a two-sided decision-process. Second, in the corresponding version of their model where the quality of the match is observed before the match is created (i.e., matches are inspection good), the amplification effects of their model are negligible.

17. Calibrating with respect to contacts would require vast new data sets on formal and informal meetings between searching workers and searching employers, and these data sets are not currently available.

18. In the working paper version of this paper [Brown et al. (2012)], we also use a Cobb–Douglas contact function of the Mortensen and Pissarides (1994) type, $C_t = \zeta U_{t-1}^{\lambda} V_t^{1-\lambda}$, where λ is the contact elasticity and ζ the contact efficiency . However, because this does not improve business cycle results, we stick to the simple contact function in our baseline scenario.

19. Note that we do not have to specify the number of firms, as they face constant returns (there is only an ex post heterogeneity, once there are particular worker–firm pairs).

20. Under this simplified assumption, vacancies play no allocative role in our model: the number of vacancies has no effect on employment and unemployment. The reason is that vacancies do not influence the number of contacts made by a given number of unemployed job searchers, because we have assumed that the number of contacts is equal to the number of job searchers. In this context, vacancies are simply an attention-seeking device: The greater the number of vacancies that a firm posts for a given job, the greater is the number of job applicants it attracts relative to other firms. The greater the aggregate number of vacancies, the lower is the probability that they will be filled by a given number of job searchers, and in the labor market equilibrium, the aggregate number of vacancies has no effect on aggregate employment.

21. We take only the direct training costs into account. We divide 203 by 8, thereby obtaining 25.4 working days. Assuming 20 working days per month, this yields 1.3 months.

22. We take averages over the time periods provided by these authors.

23. We provide a robustness analysis for other values in the working paper version of this paper. Specifically, we provide simulation results for firing costs calculated with respect to the United Kingdom, $f = 0.08$, and as an upper bound we choose $f = 0.2$. The respective amplification results are qualitatively the same, but vary in magnitude. Thus our main conclusions are not driven by the value of the firing cost parameter.

24. Note that our simulation results are robust with respect to reasonable variations in this parameter.

25. The business cycle statistics for unemployment, vacancies, market tightness, and labor productivity are taken from Michaillat (2012). The time series for the matching rate and the separation

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press

rate are constructed based on the method by Shimer (2012). The structural break in the short-term unemployment rate in the CPS data is corrected with a level shift in 1994 as proposed by Shimer (2012). All data are from 1964:Q1 to 2009:Q2. Strong amplification effects can also be found for other countries [e.g., Gartner et al. (2012) for Germany].

26. In contrast to Sveen and Weinke (2008), who use a New Keynesian model with search and matching frictions, we also do not resort to additional aggregate demand shocks under sticky prices, which generate larger employment fluctuations in the data. See also Lechthaler et al. (2010) for the amplification effects of aggregate supply and demand shocks. Thus, we generate comparability to Hall (2005) and Shimer (2005) and show that our model has a stronger endogenous amplification mechanism.

27. We calculate the dynamic Beveridge curve as the correlation between contemporaneous unemployment and vacancies. If we used the lagged unemployment, we would obtain a correlation of -0.72; i.e., there remains a strong negative relationship. Many empirical studies use the contemporaneous correlation. Therefore, we do the same to be in line with the data. The reality might lie somewhere in between.

28. Note that our model generates a negative correlation between unemployment and vacancies with exogenous separations.

29. To be consistent with the previous calibration, we calculate a quarterly job-finding rate $\eta^q = (1 - \eta^m)^3$, where q denotes quarterly and m monthly values for the job-finding rate.

30. Note that Silva and Toledo's (2009) setup is far more complex. It contains, for example, endogenous separations and firing costs. They show simulation results, whereas we show analytical results.

31. Obviously, hiring costs ceteris paribus depress the job-finding rate η, as in the search and matching model, and thus increase the log-deviation. However, in contrast to the search and matching model, for a given calibration target η and for a uniform distribution, hiring costs do not generate any extra amplification effects.

REFERENCES

Belot, Michèle, Jan Boone, and Jan C. van Ours (2007) Welfare-improving employment protection. *Economica* 74, 381–396.

Bentolila, Samuel and Guiseppe Bertola (1990) Firing costs and labour demand: How bad is eurosclerosis? *Review of Economic Studies* 57(3), 381–402.

Binmore, Ken, Ariel Rubinstein, and Asher Wolinsky (1986) The Nash bargaining solution in economic modelling. *RAND Journal of Economics* 17(2), 176–188.

Brown, Alessio J.G., Christian Merkl, and Dennis J. Snower (2012) An Incentive Theory of Matching. Working paper 1512, Kiel Institute for the World Economy (first version 2009; updated 2012).

Carneiro, Anabela, Paulo Guimarães, and Pedro Portugal (2012) Real wages and the business cycle: Accounting for worker, firm and job-title heterogeneity. *American Economic Journal: Macroeconomics* 4(2), 133–152.

Christiano, Lawrence, Martin Eichenbaum, and Charles Evans (2005) Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* 113(1), 1–45.

Christiano, Lawrence, Mathias Trabandt, and Karl Walentin (2010) Involuntary Unemployment and the Business Cycle. NBER working paper 15801.

Costain, James S. and Michael Reiter (2008) Business cycles, unemployment insurance, and the calibration of matching models. *Journal of Economic Dynamics and Control* 32(4), 1120–1155.

Danthine, Samuel (2005) Two-Sided Search, Heterogeneous Skills and Labor Market Performance. IZA working paper 1572.

Den Haan, Wouter J., Garey Ramey, and Joel Watson (2000) Job destruction and propagation of shocks. *American Economic Review* 90(3), 482–498.

Dolfin, Sarah (2006) An examination of firms' employment costs. *Applied Economics* 38(8), 861–878.

Faia, Ester (2008) Optimal monetary policy rules with labor market frictions. *Journal of Economic Dynamics and Control* 32(5), 1600–1621.

Faia, Ester (2009) Ramsey monetary policy with labor market frictions. *Journal of Monetary Economics* 56(4), 570–581.

Fujita, Shigeru and Garey Ramey (2009) The cyclicality of separation and job finding rates. *International Economic Review* 50(2), 415–430.

Gali, Jordi (2011) The return of the wage Phillips curve. *Journal of the European Economic Association* 9(3), 436–461.

Gali, Jordi, Frank Smets, and Rafael Wouters (2011) Unemployment in an Estimated New Keynesian Model. NBER discussion paper 17084.

Gartner, Hermann, Christian Merkl, and Thomas Rothe (2012) Sclerosis and large volatilities. Two sides of the same coin. *Economics Letters* 117(1), 196–109.

Haefke, Christian, Markus Sonntag, and Thijs van Rens (2008) Wage Rigidity and Job Creation. IZA working paper 3714.

Hagedorn, Marcus and Iourii Manovskii (2008) The cyclical behavior of equilibrium unemployment and vacancies revisited. *American Economic Review* 98(4), 1692–1706.

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

Hall, Robert E. (2006) Job loss, job finding and unemployment in the US economy over the past fifty years. *NBER Maroeconomics Annual* 20, 101–137.

Hall, Robert E. and Robert Milgrom (2008) The limited influence of unemployment on the wage bargain. *American Economic Review* 98(4), 1653–1674.

Heathcote, Jonathan, Perri Fabrizio, and Giovanni L. Violante (2010) Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006. *Review of Economic Dynamics* 13, 15–51.

Hornstein, Andreas, Per Krusell, and Giovanni L. Violante (2005) Unemployment and vacancy fluctuations in the matching model: Inspecting the mechanism. *Federal Reserve Bank of Richmond Economic Quarterly* 91(3), 19–51.

Khalifa, Sherif (2012) Job competition, crowding out, and unemployment fluctuations. *Macroeconomic Dynamics* 16, 1–34.

Krause, Michael U. and Thomas A. Lubik (2007) The (ir)relevance of real wage rigidity in the New Keynesian model with search frictions. *Journal of Monetary Economics* 54(3), 706–727.

Lechthaler, Wolfgang, Christian Merkl, and Dennis Snower (2010) Monetary persistence and the labor market: A new perspective. *Journal of Economic Dynamics and Control* 34(5), 968–983.

Lechthaler, Wolfgang and Dennis Snower (2013) Quadratic labor adjustment costs, business cycle dynamics and optimal monetary policy. *Macroeconomic Dynamics* 17, 464–475.

Martins, Pedro, Andy Snell, Heiko Stueber, and Jonathan Thomas (2012a) Minu, Startu and All That: Pitfalls in Estimating the Sensitivity of a Worker's Wage to Aggregate Unemployment. Mimeo, University of Edinburgh (version March 15, 2012).

Martins, Pedro, Gary Solon, and Jonathan Thomas (2012b) Measuring what employers do about entry wages over the business cycle: A new approach. *American Economic Journal: Macroeconomics* 4(4), 36–55.

Menzio, Guido and Shi Shou-Yong (2011) Efficient search on the job and the business cycle. *Journal of Political Economy* 119(3), 468–510.

Michaillat, Pascal (2012) Do matching frictions explain unemployment? Not in bad times. *American Economic Review* 102(4), 1721–1750.

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

Moyen, Stéphane and Nikolai Stähler (in press) Unemployment insurance and the business cycle: Should benefit entitlement duration react to the cycle? *Macroeconomic Dynamics*, First View Article. http://dx.doi.org/10.1017/S1365100512000478.

Petrongolo, Barbara and Christopher A. Pissarides (2001) Looking into the black box: A survey of the matching function. *Journal of Economic Literature* 39(2), 390–431.
Pissarides, Christopher A. (2000) *Equilibrium Unemployment Theory*. Cambridge, MA: MIT Press.
Pissarides, Christopher A. (2009) The unemployment volatility puzzle: Is wage stickiness the answer? *Econometrica* 77(5), 1339–1369.
Shimer, Robert (2005) The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review* 95(1), 25–49.
Shimer, Robert (2012) Reassessing the ins and outs of unemployment. *Review of Economic Dynamics* 15(2), 127–48.
Silva, Jose I. and Manuel Toledo (2009) Labor turnover costs and the cyclical behavior of vacancies and unemployment. *Macroeconomic Dynamics* 13(S1), 76–96.
Smets, Frank and Rafael Wouters (2007) Shocks and frictions in the US business cycle: A Bayesian DSGE approach. *American Economic Review* 97(3), 586–606.
Stüber, Heiko (2012) Are Real Entry Wages Rigid over the Business Cycle? Empirical Evidence for Germany from 1977 to 2009. IAB discussion paper 6/2012.
Sveen, Tommy and Lutz Weinke (2008) New Keynesian perspectives on labor market dynamics. *Journal of Monetary Economics* 55(5), 921–930.

APPENDIX: DERIVATION OF STEADY STATE ELASTICITIES

A.1. SIMPLIFIED INCENTIVE THEORY OF MATCHING

To make analytical statements, we use the deterministic version of our model (i.e., without aggregate uncertainty) and we assume that the job acceptance rate is 1 (i.e., the quit rate is 0). Thus, we combine equations (28) and (29) and calculate the job-finding rate,

$$\eta = J_\varepsilon \left[\frac{a(1-\gamma)}{1-\delta(1-\phi)} - h \right]. \tag{A.1}$$

Thus, we derive the elasticity of the job-finding rate with respect to productivity:

$$\frac{\partial \ln \eta}{\partial \ln a} = \frac{\partial \ln \eta}{\partial a} \frac{\partial a}{\partial \ln a} = J'_\varepsilon \frac{a(1-\gamma)}{[1-\delta(1-\phi)]\,\eta}. \tag{A.2}$$

A.2. SEARCH AND MATCHING MODEL

Let us assume a standard Cobb–Douglas matching function,

$$C_t = U_{t-1}^{\lambda} V_t^{1-\lambda}. \tag{A.3}$$

The job-finding rate in the deterministic version of the matching model with ex post

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press

hiring costs, h, for entrants is

$$\eta = \left\{ \frac{a - w}{\kappa \left[1 - \delta \left(1 - \phi \right) \right]} - h \right\}^{\frac{1-\lambda}{\lambda}}, \tag{A.4}$$

where κ are vacancy posting costs.

We use the same assumption for wages as before and calculate the steady state elasticity,

$$\frac{\partial \ln \eta}{\partial \ln a} = \frac{1 - \lambda}{\lambda} \frac{\frac{a \left(1 - \gamma \right)}{\kappa \left[1 - \delta \left(1 - \phi \right) \right]}}{\frac{a \left(1 - \gamma \right)}{\kappa \left[1 - \delta \left(1 - \phi \right) \right]} - h}. \tag{A.5}$$

https://doi.org/10.1017/S1365100513000527 Published online by Cambridge University Press