Labor Market Heterogeneities, Matching Efficiency, and the Cyclical Behavior of the Job Finding Rate*

Regis Barnichon
CREI and Universitat Pompeu Fabra

Andrew Figura
Federal Reserve Board

December 2011
(First version: October 2010)

Abstract

The matching function implies that the job finding rate depends only on labor market tightness. We estimate such a matching function and find that the regression residual, or matching efficiency, varies substantially over the business cycle. We argue that labor market heterogeneities are not fully captured by the matching function. Using CPS microdata over 1976-2010 and new, highly disaggregated, data on local labor market tightness, we show that matching efficiency captures two factors, which, in addition to labor market tightness, affect the job finding rate: (i) composition of the unemployment pool, and (ii) dispersion in labor market conditions, the fact that tight labor markets coexist with slack ones.

JEL classifications: J6, E24, E32

Keywords: Matching Function, Heterogeneity, Composition, Dispersion.

*We thank Bruce Fallick, Shigeru Fujita, Jordi Gali, Bart Hobijn, Rob Valletta, Thijs van Rens, William Wascher and numerous seminar participants. We thank Peter Chen for excellent research assistance. The views expressed here do not necessarily reflect those of the Federal Reserve Board or the Federal Reserve System. Any errors are our own.

1 Introduction

The search and matching model (Mortensen and Pissarides, 1994) has become the canonical framework to introduce equilibrium unemployment in macroeconomic models. One of its building blocks is the aggregate matching function that relates the flow of new hires to the stocks of vacancies and unemployment. Like the aggregate production function, the matching function is a convenient device that "partially captures a complex reality [...] with workers looking for the right job and firms looking for the right worker" (Blanchard and Diamond, 1989).

The existence of an aggregate matching function implies that the aggregate job finding rate depends only on one aggregate variable: labor market tightness –the vacancy-unemployment ratio–. We estimate a matching function tying the job finding rate to labor market tightness, and we find that the regression residual, or matching efficiency, displays non-trivial cyclical fluctuations.[1] Matching efficiency declines in the aftermaths of recessions and increases in the later stages of expansions or during recessions, and generates fluctuations in unemployment of about 1 percentage point over 1967-2006. Moreover, in the 2008-2009 recession, an exceptional decline in matching efficiency added an estimated $1\frac{1}{2}$ percentage points to the unemployment rate.

Although the matching function is meant to capture "a trading technology between heterogeneous agents" (Pissarides 2000, p.4), this paper argues that labor market heterogeneities, in particular heterogeneities across individuals and labor markets, are not fully captured by the matching function but are key to understand fluctuations in the job finding rate (and thus unemployment).

Under fairly general conditions, we theoretically show how heterogeneity can generate movements in matching efficiency, and we emphasize the role of two factors: (i) composition of the unemployment pool, and (ii) dispersion in labor market conditions, the fact that tight labor markets coexist with slack ones. First, if composition changes, and a group with a lower than average job finding probability (such as workers on permanent layoff) becomes over-represented among the unemployed, the average job finding probability will decline more than what a matching function would imply. Second, changes in the location and nature (e.g., skill requirements) of new jobs can lead to the misallocation of jobs and workers across labor markets and generate dispersion in labor market conditions as tight labor markets coexist with slack labor markets.[2] Because of the concavity of the matching function, an increase in disper-

[1] In essence, matching efficiency is akin to a Solow residual; a parameter that adjusts to capture any hiring behavior that cannot be explained by the observed aggregate levels of unemployment and vacancy posting.

[2] Note that this is not the only possible mechanism. As argued by Abraham and Katz (1986), different

sion in labor market conditions will lower matching efficiency.[3] Moreover, because the effect of higher dispersion on matching efficiency may be exaggerated if workers can find a job outside of their local labor market (Abraham, 1991), we introduce the concept of "permeability" between labor market segments. With higher permeability, workers are more likely to cross local labor market barriers and find a job in a different labor market segment, and dispersion has a weaker effect on matching efficiency.

We then use matched CPS micro data on unemployment-employment transitions over four decades to estimate a model of job finding probability and to empirically relate aggregate matching efficiency to composition. In addition, because the extent of dispersion can only be properly assessed by observing labor market segments at a high level of disaggregation, we present a unique new dataset on labor market tightness by occupation and geographic location covering a total of 564 segments, a 55-fold improvement over publicly available data such as the Job Openings and Labor Turnover Survey (JOLTS). Further, because 564 segments may still be well below the true number of segments in the US, we propose a method using UK data to scale up our measure of dispersion over 564 segments to a larger number of labor market segments.

Over 1976-2006, changes in composition are responsible for the cyclical movements in matching efficiency, and the two key individual characteristics are reason of unemployment and unemployment duration, which likely capture unobserved heterogeneity across individuals.[4] Matching efficiency decreases in the early stages of recoveries and increases in the late stages of expansions because the fraction of long-term unemployed and the fraction of permanent job losers in the unemployment pool lags the business cycle.

Over 2006-2010, composition explains only 45 percent of a dramatic decline in matching efficiency. Instead, the decline in composition-adjusted matching efficiency –the level of matching efficiency holding composition constant– is highly (negatively) correlated with an increase in dispersion in labor market conditions. Quantitatively, our measure of dispersion over 564 segments explain about 40 percent of the decline in composition-adjusted matching efficiency since 2006, and at the higher level of disaggregation allowed with UK data, the increase in dispersion would explain all of the decrease in matching efficiency.

We build on a large literature studying the matching function (e.g., Pissarides 1986, Blanchard and Diamond 1989, Bleakley and Fuhrer 1997, and the review of Petrongolo and Pis-

cyclical sensitivities to aggregate shocks across labor markets can also generate dispersion.

[3] The effect of labor misallocation on matching efficiency in the context of the matching function is similar to the effect of capital misallocation on aggregate TFP in the context of the production function and emphasized in recent studies (e.g., Hsieh and Klenow, 2009).

[4] In particular, we show that duration dependence is weaker in recessions than in expansions, which strongly suggests that duration dependence captures unobserved heterogeneity across individuals, rather than hysteresis or scarring effects of unemployment.

sarides 2001).[5] While that literature focuses on aggregate labor market tightness as the main explanatory variable of the aggregate job finding rate, we emphasize the importance of heterogeneities and the fact that the aggregate job finding probability is an average of probabilities across heterogeneous workers working in different segments of the labor market. In this context, our paper closely relates to the literature on the heterogeneity hypothesis raised by Darby, Haltiwanger and Plant (1985), but rejected by Baker (1992) and Shimer (2007). The heterogeneity hypothesis states that the average job finding rate increases during recessions because workers who are unemployed during recessions are different from workers who are unemployed during expansions. Our paper revisits the heterogeneity hypothesis by considering a broader spectrum of heterogeneities and by merging the concept of an aggregate matching function with the existence of heterogeneities across individuals, jobs and labor markets.[6] Recently, Davis, Faberman and Haltiwanger (2010) study establishment level data from the JOLTS over 2001-2006 and argue that variations in firms' recruiting intensity generates movements in the vacancy yield, and thus matching efficiency. Although we cannot measure recruiting intensity over our time sample, the fact that our framework without varying recruiting intensities can successfully capture job finding probability movements over 1976-2006 suggests that labor market tightness can proxy for varying intensities. Finally, the literature on mismatch has typically relied on a variety of dispersion measures (Padoa Schioppa, 1991, Layard, Nickell and Jackman, 2005) to capture the extent of misallocation of jobs and workers, and absent a unifying framework, there was no consensus on the most appropriate measure. This paper contributes to that literature by providing a dispersion measure, the variance of labor market tightness across labor market segments, that can be analytically related to matching efficiency and to the equilibrium unemployment rate. In this context, two recent contributions are related to this paper. Herz and van Rens (2011) propose an approach to study the sources and the cyclicality of mismatch over a long time sample. Sahin, Song, Topa, and Violante (2010) construct mismatch indices based on a theoretical framework of mismatch. An important benefit of Sahin et al.'s approach is the possibility to rigorously define mismatch, and interestingly, our dispersion measure and their mismatch measure are closely related.[7] Our analyses are also

[5] See also the recent work from Sedlacek (2011) who studies the cyclicality of matching efficiency using a state-space model with unobserved vacancy posting.

[6] Baker (1992) tested the heterogeneity hypothesis by evaluating the contribution of demographics and reason for unemployment (but not of unemployment duration) to the cyclicality of the job finding rate. While finds evidence that an aggregate variable (such as labor market tightness) drives cyclical fluctuations in the job finding rate, he does find, consistent with this paper, that reason for unemployment explains close to 10% of the cyclicality of the job finding rate. Shimer (2007) also evaluates, and rejects, the heterogeneity hypothesis but he does not study the contribution of reason for unemployment and unemployment duration, the two most important contributors to the composition effect.

[7] In their framework, mismatch is defined as the distance between the observed allocation of unemployed workers across sectors and the optimal allocation that solves a planner's problem

complementary in the empirical dimension. Both studies rely on online help-wanted ads, but while Sahin et al. probe the extent of mismatch at a deep level of disaggregation for occupations, we probe dispersion at a deep geographic level by considering highly disaggregated geographic segments.

The next section estimates how an aggregate matching function fares at capturing movements in the job finding rate and highlights the existence of cyclical fluctuations in the regression residual, or matching efficiency. Section 3 presents an empirical framework to study how labor market heterogeneities affect matching efficiency. Section 4 uses micro data to estimate that framework and study how far composition can go in explaining movements in matching efficiency over 1976-2010, and Section 5 presents a new dataset with labor market tightness data over 564 segments over 2006-2010 and evaluates the effect of the measured dispersion on matching efficiency. Section 6 concludes.

2 Matching efficiency, the Solow residual of the matching function

The matching function relates the flow of new hires to the stocks of vacancies and unemployment. Like the production function, the matching function is a convenient device that partially captures a complex reality with workers looking for the right job and firms looking for the right worker. In a continuous time framework, the flow of hires can be modeled with a standard Cobb-Douglas matching function with constant returns to scale, and we can write

$$m_t = m_{0t} U_t^{\sigma} V_t^{1-\sigma} \tag{1}$$

with m_t, the number of new hires at instant t, U_t the number of unemployed, V_t the number of vacancies, and m_{0t} a (possibly time-varying) parameter that we will refer to as "aggregate matching efficiency".[8]

Since the job finding rate jf_t is the ratio of new hires to the stock of unemployed, we have $jf_t = \frac{m_t}{U_t}$ so that $jf_t = m_{0t}\theta_t^{1-\sigma}$with $\theta = \frac{v}{u}$ the aggregate labor market tightness, $u = U/LF$, $v = V/LF$ and LF the labor force. To identify m_{0t}, we regress

$$\ln jf_t = (1-\sigma)\ln\theta_t + E_T(\ln m_{0t}) + \mu_t \tag{2}$$

with $E_T(.)$ denoting the average over the estimation period so that $E_T(\ln m_{0t})$ denotes the

[8] The Cobb-Douglas matching function is used in almost all macroeconomic models with search and search and matching frictions (e.g., Pissarides, 2001). Allowing for non constant returns to scale or using a more general CES matching function $m_t = m_{0t}\left[\sigma U_t^{\rho} + (1-\sigma)V_t^{\rho}\right]^{1/\rho}$ gives very similar results.

intercept of the regression. Deviations of aggregate matching efficiency from its average level are then given by

$$\mu_t = \ln m_{0t} - E_T \ln m_{0t}. \tag{3}$$

We measure the job finding rate jf_t from unemployment-employment transitions from the Current Population Survey (CPS) over 1976-2010 and from the worker flows data tabulated by Joe Ritter for the period 1968-1975. We use the composite help-wanted index presented in Barnichon (2010) as a proxy for vacancy posting.[9] We use non-detrended quarterly data and estimate (2) over 1968-2007. Table 1 presents the results. The elasticity is estimated at 0.67. Using lagged values of v_t and u_t as instruments gives similar results, and the elasticity is little changed at 0.66.

Figure 1 plots the empirical job finding rate, its fitted value, and the regression residual, i.e., μ_t, the movements in aggregate matching efficiency.

While aggregate labor market tightness does a good job at capturing movements in the aggregate job finding rate, a testimony of the success of the matching function, the residual, or aggregate matching efficiency, displays a clear cyclical pattern. In units of unemployment rate, matching efficiency movements generates fluctuations in unemployment of about 1 percentage point over 1967-2006, a non-trivial contribution (Figure 2).[10] Matching efficiency typically lags the business cycle, increasing in the later stages of expansions, peaking in the late stages of recessions or the early stages of recoveries, and declining thereafter. In addition, in the 2008-2009 recession, the decline in matching efficiency occurred earlier than in previous recessions and was a lot more pronounced. In the fourth quarter of 2010, the residual reached an all time low of four standard-deviations.[11] In units of unemployment, the decline added about 1.5ppt to the unemployment rate (Barnichon and Figura, 2010). The period preceding the 2008-2009 recession also appears peculiar, because, compared to the previous two recessions, the increase in matching efficiency prior to the recession was a lot more muted.

[9] A measurement issue is that vacancies are not only filled from the unemployed pool (U) but also from the employment pool (E) and individuals outside the labor force (NLF). As a robustness check, we proceeded as in Blanchard and Diamond (1989) and estimated a regression over 1994-2007 of the sum of U-E flows, NLF-E flows and E-E flows (Fallick and Fleischman, 2004) on vacancies and the number of unemployed and individuals outside the labor force willing to work. The behavior of m_{0t} was broadly unchanged.

[10] See Barnichon and Figura (2010) for the method (based on the concept of steady-state unemployment) used to convert movements in matching efficiency into units of unemployment rate.

[11] Elsby, Hobijn and Sahin (2010) report a similar finding using the unemployment outflow rate, and Davis, Faberman and Haltiwanger (2010) also report a dramatic decline in the vacancy yield using JOLTS data.

3 A framework to study the effect of heterogeneities on matching efficiency

To show how heterogeneities can generate movements in aggregate matching efficiency, this section generalizes the simple job finding model (2) by taking into account individual characteristics as well as local labor market characteristics. Under fairly general assumptions, we theoretically identify two factors that affect aggregate matching efficiency: the composition of the unemployment pool, and the amount of dispersion in labor market conditions, suggestive of mismatch in the labor market. The former arises if the characteristics of job seekers and job openings change throughout the cycle, making job finding more or less likely, while the latter is caused by the concavity of the matching function and arises if tight labor markets coexist with slack labor markets.

3.1 The determinants of an individual's job finding probability

Denote $JF_{ij,t}$ the job finding probability of an individual of type $j \in [1, J]$ in labor market segment $i \in [1, I]$.[12] Individual type j is defined by a vector X_{jt} of K characteristics $\left\{x_{jt}^k\right\}$. Labor market segment i has a matching technology and is characterized by its labor market tightness (or vacancy-unemployment ratio) $\theta_{it} = \frac{v_{it}}{u_{it}}$, and matching efficiency m_i.[13]A labor market segment can be defined by its geographic location, industry group or occupation group. The labor market segment i of individual type j can then be thought of as the labor market in which individual j is most likely to look for work and to find a job. Typically, this will be proxied by his location and past employment history.

An individual's job finding probability will depend on his characteristics, the labor market tightness θ_{it} in his segment and the characteristics of that segment. Because an unemployed worker may look for work outside of his labor market segment, his job finding probability can also depend on the aggregate labor market tightness θ_t. Another determinant of the job finding probability is the search intensity of worker type j. In particular, search intensity may decline when unemployment benefits become more generous, as has been the case in past recessions. We thus include search intensity in the vector of individual characteristics X_{jt}.

[12]The job finding rate jf and the the job finding probability JF can be related from $jf = -\ln(1 - JF)$.

[13]We do not explicitly consider the case where firms can vary their recruiting intensity at a given level of vacancies posted, which would lead to a time varying m_{it}. We do so for two reasons. First, this would be difficult to implement empirically as measuring firms' recruiting intensity is notoriously difficult (see however Davis, Faberman and Haltiwanger, 2010 for very interesting work in this direction). Second, as we report below, our framework without varying recruiting intensities can successfully capture job finding probability movements over 1976-2006. This suggests that aggregate labor market tightness (or sectoral tightness) can proxy for varying intensities and that our framework provides a good reduced-form approximation of a model of job finding probability with varying intensities.

Thus, we postulate that the job finding probability of individual type j in labor market i can be written

$$JF_{ij,t} = JF(X_{jt}, m_i, \theta_{it}, \theta_t) \tag{4}$$

so that the average job finding probability is given by

$$JF_t = \sum_{i,j} \frac{U_{ij,t}}{U_t} JF_{ij,t}. \tag{5}$$

To highlight the effects of composition and dispersion on the average job finding probability, we take a second-order Taylor expansion of JF_t with respect to θ_{it} around θ_t and (X_{jt}, m_i) around $\bar{X} \equiv \frac{1}{TJ} \sum_{t,j} X_{jt}$ and $m_0 \equiv \frac{1}{I} \sum_{i} m_i$ with $X_{jt} = [1, x^1_{jt}, ..., x^K_{jt}]$ and x^k_{jt} the kth observable characteristics of individual j, and we get[14]

$$JF_t = \overline{JF}_t(\theta_t) + \sum_k JF_t^{U,k} + JF_t^m - MM_t\left(\frac{\theta_{it}}{\theta_t}\right) + \eta_t. \tag{6}$$

The first term in (6), $\overline{JF}_t(\theta_t) = JF_{ij,t}(\bar{X}, \theta_t, \theta_t)$, is the average job finding rate absent worker and job heterogeneity and absent dispersion.

Composition: Because the aggregate job finding probability is the average of individual job finding probabilities, changes in the composition of the unemployment pool can affect JF_t.

The second term in (6), $\sum_k JF_t^{U,k}$, captures the composition effect coming from the composition of the unemployment pool with $JF_t^{U,k}$, the contribution of a given characteristic k to the average job finding probability:[15]

$$JF_t^{U,k} = \sum_j \frac{U_{j,t}}{U_t} \left.\frac{\partial JF}{\partial x^k_{jt}}\right|_{\theta_t, \bar{X}} \left(x^k_{jt} - \bar{x}^k\right). \tag{7}$$

This composition effect arises because of worker heterogeneity. If the share $\frac{U_{j,t}}{U_t}$ of a demographic group (e.g. job losers) with a lower than average job finding probability (i.e., $\left.\frac{\partial JF}{\partial x^k_{jt}}\right|_{\theta_t, \bar{X}} \left(x^k_{jt} - \bar{x}^k\right) < 0$) increases in recessions, then the *average* job finding probability will decline without any change in individuals' job finding probabilities.

[14] The cross-order term between $X_{jt} - \bar{X}$ and $\frac{\theta_{it}}{\theta_t} - 1$ is omitted and only described in the Appendix. This is done for clarity of exposition as that term is empirically very small at the level of disaggregation permitted by our data on labor market segments.

[15] We omitted the second-order term for clarity of exposition, but incorporated the (small) second-order term in all our calculations.

The third term in (6), JF_t^m, captures the effect of changes in the distribution of the unemployed across segments with different average matching efficiency:

$$JF_t^m = \sum_i \frac{U_{i,t}}{U_t} \left.\frac{\partial JF}{\partial m_i}\right|_{\theta_t,\bar{X}} (m_i - m_o)\,. \tag{8}$$

For instance, if a higher fraction of the unemployed becomes concentrated in a segment with higher matching efficiency, the average job finding probability will increase even if the aggregate numbers of vacancy and unemployed remain constant.

Dispersion: The third term in (6) captures the effect of dispersion in labor market conditions on the average job finding probability with

$$\begin{aligned} MM_t\left(\frac{\theta_{it}}{\theta_t}\right) &= MM_0\left(\theta_t\right)\sum_i \frac{U_{i,t}}{U_t}\left(\frac{\theta_{it}}{\theta_t}-1\right)^2 \\ &= MM_0\left(\theta_t\right) Var\left(\frac{\theta_{it}}{\theta_t}\right) \end{aligned} \tag{9}$$

and $MM_0\left(\theta_t\right) = -\frac{1}{2}\theta_t^2 \left.\frac{\partial^2 JF_{ij,t}}{\partial \theta_{it}^2}\right|_{\theta_t,\bar{X}}$.[16] Dispersion in labor market tightness across segments will negatively affect the *average* job finding probability if the individual job finding probability is a concave function of θ_{it} (as would be the case with a matching function). With $\left.\frac{\partial^2 JF_{ij,t}}{\partial \theta_{it}^2}\right|_{\theta_t,\bar{X}} < 0$, an increase in dispersion across labor market segments decreases the average job finding probability. For example, if some segments (such as health care) display a relatively tight labor market and some segments (such as manufacturing) display a slack labor market, the average job finding probability will be lower than in an economy where labor market tightness is identical across segments.

3.2 Postulating a functional form for $JF_{ij,t}$

To bring our framework to the data, we need to posit a functional form for the job finding probability $JF_{ij,t}$. We adopt a logistic functional form

$$\ln \frac{JF_{ij,t}}{1-JF_{ij,t}} = \beta X_{jt} + \ln \frac{1-e^{m_i \theta_{it}^{(1-\sigma)\omega}\theta_t^{(1-\sigma)(1-\omega)}}}{e^{m_i \theta_{it}^{(1-\sigma)\omega}\theta_t^{(1-\sigma)(1-\omega)}}} + \eta_{ij,t} \tag{10}$$

with $\theta_{it} = \frac{v_{it}}{u_{it}}$, $X_{jt} = [1, x_{jt}^1, .., x_{jt}^K]$ and $\omega \in [0,1]$.[17]

[16] The term corresponding to θ_{it}, $\sum_{i,j} \frac{U_{j,t}}{U_t} \left.\frac{\partial JF}{\partial \theta_{it}}\right|_{\theta_t,\bar{X}} (\theta_{it} - \theta_t)$, is nil because $\theta_t = \sum_{i,j} \frac{U_{ij,t}}{U_t}\theta_{it}$.

[17] In the estimation, we will demean the $\left\{x_{jt}^k\right\}$ so that we can assume $\overline{x}^k = 0$.

This specification has a number of advantages:

First, a logistic functional form is consistent with the fact that the job finding probability falls between 0 and 1.

Second, in the absence of worker and segment heterogeneity ($X_{jt}^U = \bar{X}$, $m_i = m_0$) and labor market dispersion ($\theta_{it} = \theta_t$), (10) reduces to $JF_{ij,t} = 1 - e^{-m_0\theta_t^{1-\sigma}}$, the reduced-form aggregate specification (2) with $m_{0t} = m_0$.

Third, to relate θ_{it} and θ_t to the job finding probability of individual type j, we assume that the job finding probability is a geometric average of local labor market tightness θ_{it} and the aggregate labor market tightness θ_t. Put differently, we allow for the possibility that a worker crosses barriers between labor market segments, so that his job finding probability is not solely a function of tightness in his local labor market. To get some intuition, consider the simpler case without heterogeneity. The job finding rate of a worker in segment i becomes

$$jf_{it} = m_i\theta_{it}^{(1-\sigma)\omega}\theta_t^{(1-\sigma)(1-\omega)}$$

i.e., a weighted (geometric) average of the segment labor market tightness and the aggregate labor market tightness.[18] The weight $\omega \in [0,1]$ captures the impermeability of the local labor market. When $\omega = 1$, as has been typically imposed in the mismatch literature (Layard, Nickell and Jackman, 2005), labor market segments are impossible to cross, and aggregate labor market tightness has no impact on the local job finding rate. In contrast, if there are no barriers between labor markets, $\omega = 0$, a worker's job finding rate only depends on the aggregate labor market tightness.

3.3 A decomposition of aggregate matching efficiency

Thanks to our decomposition (6), we can now link movements in aggregate matching efficiency to the composition of the unemployment pool and the amount of dispersion in labor market conditions. After some manipulation of (6) left for the Appendix, $\ln m_{0t} - \mathrm{E}_T \ln m_{0t}$, the deviations of aggregate matching efficiency from its average value, can be written

$$\ln m_{0t} - E_T \ln m_{0t} = \frac{e^{m_0\theta_t^{1-\sigma}}}{m_0\theta_t^{1-\sigma}}\left(\sum_k JF_t^{U,k} + JF_t^m\right) - \Delta mm_t + \zeta_t. \quad (11)$$

with

$$mm_t = \frac{1}{2}\omega(1-\sigma)\left[(1-\omega(1-\sigma)\left(1 - m_0\theta_t^{1-\sigma}\right)\right] Var\left(\frac{\theta_{it}}{\theta_t}\right) \quad (12)$$

and $\Delta mm_t = mm_t - \mathrm{E}_T mm_t$ with $\mathrm{E}_T(.)$ denoting the average over the estimation period.

[18] See also Burda and Profit (1996) and Burgess and Profit (1998) for related specifications.

Aggregate matching efficiency m_{0t} is a function of the distribution of individual characteristics and labor market segments' tightness. Movements in aggregate matching efficiency can be decomposed into an unemployment composition effect, the first two terms on the right-hand side of (11), a dispersion effect, the third term, and an unexplained component (that includes the approximation error), the last term.

Expression (12) describes the effect of dispersion on aggregate matching efficiency. Three observations are worth noting. First, the effect of dispersion is roughly proportional to the variance of relative labor market tightness, so that one can readily estimate the effect of dispersion on matching efficiency by looking at the dispersion in labor market conditions.[19] The literature on mismatch has used various measures to quantify the effect of misallocation on the unemployment rate. For example, some use $\sum_i \left|\frac{U_i}{U} - \frac{V_i}{V}\right|$ (e.g., Jackman and Roper 1987, Franz 1991, Brunello 1991), others unemployment rate dispersion measures $\sum_i u_i^2$ or $\sum_i \left(\frac{u_i}{u}\right)^2$ (e.g., Jackman, Layard and Savouri (1991), Attanasio and Padoa Schioppa (1991)), others $\sum_i \left(\frac{U_i}{U}\frac{V_i}{V}\right)^{1/2}$ (Bean and Pissarides, 1991), and others $\sum_i \frac{E_i}{E}\left(\frac{U_i/E_i}{U/E} - \frac{V_i/E_i}{V/E}\right)^2$ (Layard, Nickell and Jackman, 1991) with E_i the number of employed workers in segment i and E the total number of employed workers. Some measures were constructed using employment or labor force weights (but surprisingly, rarely unemployment weights), but other measures did not weight observations. While all these measures capture the extent of dispersion across labor markets, absent a unifying framework, there was no consensus on the most appropriate measure. The measure we propose has an important advantage over these other measures: It can be directly related to aggregate matching efficiency and thus to the equilibrium unemployment rate (Barnichon and Figura, 2010).[20]

Second, while the mismatch literature imposes tight labor market segment boundaries, our framework allows for some permeability between labor market segments. The effect of dispersion on the average job finding rate and matching efficiency depends on ω. For the range of plausible values for σ and $m_0\theta_t^{1-\sigma}$, the effect of dispersion increases with barriers between labor market segments (i.e. when ω increases) and is strongest when labor market segments' barriers are infinite. Conversely, with higher permeability, workers are more likely to cross local labor market barriers and find a job in a different labor market, and this possibility weakens the effect of dispersion on aggregate matching efficiency.

Finally, because average dispersion is positive ($Var\left(\frac{\theta_{it}}{\theta_t}\right) \geq 0$, $\forall t$), the effect of dispersion on aggregate matching efficiency movements μ_t is not given by mm_t, the level of dispersion,

[19] The coefficient of proportionality does depend on θ_t but its effect is small.

[20] See also Sahin, Song, Topa, and Violante (2011) who develop mismatch indices based on a theoretical framework of mismatch.

but by Δmm_t, the *deviations* of dispersion from its average level.

4 The determinants of matching efficiency over 1976-2007

4.1 Estimation

We use matched monthly data from the Current Population Survey (CPS) covering January 1976 to December 2007 to estimate the Unemployment-Employment (UE) transition probability for an individual j in labor market segment i. We restrict the estimation to pre-2008 data so that any changes in matching efficiency post 2007 do not affect our coefficient estimates. In 1994, a major redesign of the CPS survey was implemented and introduced breaks in many important variables, such as reason for unemployment and duration of unemployment.[21] To control for these breaks, we estimate separate coefficients for the pre and post redesign periods. Our whole sample contains about 1.2 million observations.

In this section, we present the characteristics that influence the job finding probability, discuss our method for measuring labor market tightness by segment, and present results.

1. Measuring composition The CPS data provides information allowing us to control for changes in the characteristics of the unemployed. We use three main types of information to capture both observed and unobserved heterogeneity: demographic, reason for unemployment and duration of unemployment. We also include a set of industry dummies to control for different levels of matching efficiency across industry groups, and a set of monthly dummies to control for seasonality in job finding probabilities.[22]

Demographic information includes the age and sex of the unemployed individual. We model the effect of age on the job finding probability using a quadratic in age.

The CPS distinguishes between 5 main reasons for unemployment: permanent layoff, temporary layoff, new labor force entrant, reentering the labor force, and quit job. We use dummy variables for each reason. Reason for unemployment likely captures unobserved heterogeneity across individuals.

The CPS records the duration (in weeks) of individuals' current spells of unemployment. Prior research (e.g., Kaitz 1970, Machin and Manning 1999, Shimer 2008) suggests that the job finding probability declines with duration, and we include unemployment duration as an

[21] See, for example, Polivka and Miller (1998).

[22] We also experimented with education level and race/ethnicity but found that these characteristics play little role in the cyclicality of matching efficiency, consistent with the findings of Baker (1992). We thus omitted these characteristics for clarity of exposition. The industry dummies are agriculture/forestry, mining, construction, manufacturing, trade, transportation and utilities, information, financial activities, professional and business services, hospitality and leisure, health and education, other services, public administration.

explanatory variable. Duration dependence can arise through two channels: scarring or unobserved heterogeneity. First, prolonged unemployment may lower individuals' skills relative to other job seekers, making them less desirable to employers (e.g., Pissarides, 1992), or it may reduce their contacts in job–finding networks, making it harder to find employment. Henceforth, we describe this effect as scarring. Second, prolonged unemployment may signal that the individuals have unobserved characteristics that make it difficult to find employment.

In addition, because we also study the effect of duration in the time series dimension, we interact an individual's duration with average duration, and we postulate that

$$JF_{ijt} = ... + dur_{jt}(\beta^{dur} + \gamma_t^{dur} dur_t)$$

with dur_{jt} the unemployment duration of individual type j and dur_t the average unemployment duration across all types. This specification allows the slope of the relationship between job finding and duration to change as aggregate conditions change. When labor demand is low, it takes unemployed individuals longer to find jobs, and duration rises for everyone. Thus, the signal about job finding prospects from an individual's unemployment duration may be weaker when average duration is higher. To allow the signal from an individual's duration to change as aggregate conditions change, we interact individual duration with average duration.

Finally, while we do not observe workers' search effort and its effect on matching efficiency, the extension of unemployment benefits during recessions could lead workers eligible for unemployment insurance (UI) benefits to lower their search effort (or increase their reservation wage) and reduce matching efficiency. To identify the effect of extended and emergency unemployment benefits (EEB) on job finding probabilities, we extend a strategy used by Kuang and Valletta (2010) who note that job losers are predominantly eligible for UI benefits while job leavers and new labor force entrants are not. We identify the effect of EEB on the job finding probability by interacting a job loser dummy with an EEB dummy δ_{it}^{EEB} defined as

$$\begin{cases} \delta_{it}^{EEB} = 1 \text{ if \$ spent on EEB in state } i > 0 \\ \delta_{it}^{EEB} = 0 \text{ otherwise} \end{cases}.$$

2. Measuring labor market tightness by segment To properly assess the effect of dispersion in θ_{it} on aggregate matching efficiency, it is important to reach a good level of disaggregation as dispersion increases with the number of observed segments.

In Section 5, we present a new dataset with a unique level of disaggregation; with vacancy data by geography and occupation for a total of 564 segments.[23] However, since this dataset

[23] In contrast, the two public data sources with vacancy posting data, the JOLTS and the print Help-Wanted Indexes (HWI) from the Conference Board do not allow for a high level of disaggregation. The JOLTS measure

only starts in 2006, in this section, we use the ratio of unemployment rates $\frac{u_{it}}{u_t}$ to proxy for the ratio of labor market tightness $\frac{\theta_{it}}{\theta_t}$ over 1976-2010. Regional and industry data on vacancy posting from the JOLTS over 2000-2010 show that vacancies and unemployment rates are highly negatively correlated across regions or industries, and that $\left(\frac{u_{it}}{u_t}\right)^{\alpha}$ with $\alpha \simeq 1.3$ is a good proxy for $\frac{\theta_{it}}{\theta_t}$.[24] We thus use the CPS micro data to estimate the unemployment rate and labor market tightness of each segment. Since new entrants to the labor force cannot be easily classified in a particular industry, we use the average unemployment rate in their state of residence. While the CPS sample is large (about 60,000 households), we nonetheless face some limitations regarding the degree of disaggregation we can achieve. We define 150 segments based on the cross product of 50 states and three broad industry groups.[25]

Accordingly, we estimate the slightly modified form of equation (10)

$$\ln \frac{JF_{ij,t}}{1 - JF_{ij,t}} = \beta X_{jt} + \ln \frac{1 - e^{m_i \left(\frac{u_{it}}{u_t}\right)^{\gamma} \theta_t^{1-\sigma}}}{e^{m_i \left(\frac{u_{it}}{u_t}\right)^{\gamma} \theta_t^{1-\sigma}}} + \varepsilon_{ijt} \tag{13}$$

by maximum likelihood with $\gamma = \alpha\,(1-\sigma)\,\omega$.

4.2 Results

4.2.1 Coefficient estimates

Table 2 reports our coefficient estimates. The coefficient on the aggregate vacancy-unemployment ratio is highly significant but is lower than the coefficient estimated in Section 2 using only aggregate labor market tightness. This suggests that characteristics and/or dispersion are on average procyclical, and that failing to control for those parameters biases estimates of the aggregate matching function elasticity upward.

The impermeability coefficient of labor market segments is significantly smaller than one ($\omega = \gamma/(\alpha(1-\sigma)) = .2/(1.3 * .28) \simeq 0.55$). While barriers between labor market segments appear to be non trivial, they are also not insurmountable. As a result, the job finding

of job openings can be disaggregated into 10 industry groups, but the survey's sample size is too small to allow for a disaggregation by regions and industries. The print HWI can be disaggregated by regions (the nine US Census divisions), but not by industry. In the online Appendix, we nonetheless present a measure of industry dispersion using JOLTS data and a measure of geography dispersion using print HWI data.

[24] Formally, we used JOLTS and Conference Board vacancy data by region or industry to regress $\ln \frac{\theta_{it}}{\theta_t} = \alpha \ln \frac{u_{it}}{u_t} + \beta$. The regression results are shown in Table A1 in the online Appendix.

[25] Defining segments by occupation rather than industry does not qualitatively change our results. Our three broad industry groups roughly correspond to goods producing, business/health care/educational services, and other services. The industry groups are (1) manual workers-agriculture, mining, construction, manufacturing, transportation and public utilities, (2) professional workers-finance, professional and business services, health care and education, information and public sector, and (3) service workers-retail and wholesale trade, leisure and hospitality services, other services.

probability is a function of the labor market tightness within a segment (with elasticity $\omega(1-\sigma)=0.20$) and a function of aggregate labor market tightness (with elasticity $(1-\sigma)(1-\omega)=0.13$).

Turning to individual characteristics, JF is decreasing in unemployment duration, consistent with previous findings. The coefficient on unemployment duration implies that having a spell of unemployment lasting 6 months is associated with a decrease in an individual's job finding probability of about 1-$1\frac{1}{2}$ percentage point (Figure 3). In addition, we find that the slope flattens in downturns, a result, which, as far as we know, had not been documented previously, and which strongly suggests that duration dependence captures unobserved heterogeneity rather than scarring.[26] Indeed, while we show in the Appendix that a flattening of the slope is precisely what one would expect in a model with unobserved heterogeneity, it is much harder to explain in a model with hysteresis and loss of skills.[27]

The estimates of the effect of the reason for unemployment on the job finding probability are relative to that of a job leaver. The estimates reveal that it is particularly difficult for permanent job losers and new entrants to the labor force to find employment. Unsurprisingly, workers on temporary layoff have an easier time becoming reemployed.[28]

Turning to search effort and the effect of EEB, the increases in the maximum length of eligibility for unemployment insurance in recessions has a small but significant effect on job finding probabilities.

Turning to demographics, the coefficient on the male dummy indicates that males are more likely to find jobs than females. However, a comparison of the pre and post redesign coefficients shows that this relative advantage has lessened over time. The estimated coefficients on the age variables indicate that the probability of job finding initially increases and then decreases with age.[29]

[26] Shimer (2008) compares the job finding rate-duration relation in booms and in slumps and finds no significant difference in the slope of the relation. However, his different conclusion could be due to the fact that he does not explicitly control for cyclical fluctuations (only splitting his sample in two bins; high or low average job finding rate) and that he does not control for the range of workers and labor markets heterogeneities that we control for.

[27] With unobserved ability, the observed job finding probability decreases more slowly with duration during recesions because the high ability individuals exit the unemployment pool more slowly. In contrast, with hysteresis and skill depreciation, one would need to explain how unemployed individuals lose their skills faster in expansions.

[28] As expected, the CPS redesign, by restricting temporary layoffs to individuals expecting to be recalled within 6 months, increased the difference in exit hazards for permanently and temporarily laid off workers.

[29] In the pre redesign period, the age with the highest job finding probability is around 30. In the post redesign period, it is close to 17.

4.2.2 The effect of composition on the job finding probability

Next, we use our decomposition (6) to estimate the effect of individual characteristics and labor market dispersion on the average job finding probability over 1976-2009. Figure 4 graphs $\left\{JF_t^k\right\}$, the contributions of individual characteristics –reason for unemployment, unemployment duration, demographics, search effort and EEB–, JF_t^m capturing the effect of changes in the distribution of the unemployed across segments with different average matching efficiency, and MM_t, the contribution of dispersion in labor market tightness, to JF_t.

Figure 4 shows that, aside from aggregate labor market tightness, the key components of JF_t are the two individual characteristics reason for unemployment and unemployment duration. The component of JF_t driven by reason for unemployment declines slowly in the early phase of expansions and starts increasing only after a few years of expansion. This happens because the fraction of permanent job losers in the unemployment pool (which increases during recessions) is a persistent variable. Permanent job losers have a lower job finding rate than average (Table 2), and many years of expansion are necessary to bring their share back to pre-recession levels.[30] A similar reasoning holds with unemployment duration. The contribution of duration is to increase the job finding rate in the late stages of expansion and to decrease it the aftermaths of recessions. By definition, unemployment duration is an inertial variable, and average unemployment duration lags the cycle. As a result, the component of JF_t driven by duration also lags the cycle; peaking at the end of expansions and bottoming out a few years into the recovery.

Demographics generate a downward trend in the average job finding probability over the sample period, as the labor force ages and women's share of the labor force increases. Consistent with Baker (1992), demographic characteristics have little influence on the cyclical behavior of the job finding probability.

Lower search effort in response to more generous unemployment benefits reduces the job finding probability. The effect is most noticeable in the last two recessions, but it remains small, lowering the job finding probability by about $\frac{1}{4}$ percentage point in the 2008-2009 recession.[31]

[30] The share of permanent job losers over 1967-2010, as well as average unemployment duration, are plotted in Figure 1 in the online Appendix. Note also that reason for unemployment tends to lift the job finding rate in recessions. This pattern owes to an increasing share of temporary job losers during recessions (especially before 1985). At the onset of recessions, bursts of temporary layoffs lift the job finding rate because job losers on temporary layoffs have a higher job finding probability than average (Table 2). As shown in Figure 1 in the online Appendix, the fraction of unemployed on temporary layoffs increases sharply in recessions. This was especially the case in the 70s, and probably explains the sharp increases in matching efficiency in the 1970 and 1974 recessions (Figure 1) for which we do not have micro data. Moreover, because temporary job losers exit unemployment quickly, the share of workers on temporary layoffs reverts quickly to its mean (in contrast with the share of permanent job losers) and the contribution of temporary layoffs is concentrated around recessions.

[31] Our results do not necessarily contradict Moffitt (1985), Katz and Meyer (1990), Meyer (1990) and Kuang and Valletta (2010) who find a larger effect of extending benefits on unemployment duration. The reason is

Changes in the distribution of the unemployed across industries (JF_t^m) only have marginal effects on the job finding probability, because we found that average matching efficiency varies little across industries. Finally, the effect of dispersion, given by (9), is also very small because the cross-sectional variance of relative labor market tightness is too small, at least for the segments we observe, for dispersion to have a noticeable effect on aggregate matching efficiency.

4.2.3 Movements in aggregate matching efficiency

Using decomposition (11), the upper panel of Figure 5 presents movements in aggregate matching efficiency, in a similar fashion to Figure 1, but allowing for a richer specification than the reduced-form approach (2). The blue line is analogous to the residual of Figure 1 and shows $\ln m_{0t} - \mathrm{E}_T \ln m_{0t}$, the total movements in aggregate matching efficiency, given by (11). The green line plots the movements in matching efficiency predicted by composition $\frac{e^{m_0 \theta_t^{1-\sigma}}}{m_0 \theta_t^{1-\sigma}} \sum_k JF_t^{U,k}$. The lower panel plots the difference between the two other lines, i.e., the changes in aggregate matching efficiency that cannot be accounted for by composition.

Up until 2006, composition does a good job explaining the cyclical movements in aggregate matching efficiency. This is especially true after 1994, where the fit is excellent. Prior to 1994, the ability of composition to account for matching efficiency movements is not as good, but this is not surprising given the lower quality of the data prior 1994. The 1994 CPS redesign particularly improved the measurement of the two key variables responsible for the effect of composition on matching efficiency, reason for unemployment and unemployment duration.[32]

Note that the behavior of unemployment duration and reason of unemployment shown in Figure 4 provides insights on the behavior of matching efficiency. For instance, matching efficiency entered the 2008-2009 recession at an unusually low level compared to the previous two recessions, because both duration and the fraction of permanent job losers were not back to their pre-2001 level when the 2008-2009 recession started.[33]

Since 2007, a large fraction of the decline in matching efficiency has been due to unobserved factors. Initially, the deterioration in late 2007 and 2008 owed almost entirely to unobservable

that unemployment duration is determined by the unemployment-employment (UE) transition rate *and* by the unemployment-nonparticipation (UN) transition rate. In on-going work, we estimate a logistic regression for the UN transition probability after controlling for characteristics of the unemployed. We find a much larger effect of EEB on UN transitions as job losers' probability of remaining in the labor force increased significantly more than the UN probability of other unemployed. See also Fujita (2010) who argues that EEB significantly lowered male workers' job finding probability in the 2008-2009 recession.

[32] The CPS redesign, by restricting temporary layoffs to individuals expecting to be recalled within 6 months, made the distinction a lot sharper (as seen from the evolution of the coefficient on temporary layoffs pre and post-1994 in Table 2), which certainly improved our measure of composition.

[33] In addition, the (unusual) decline in the share of temporary layoffs during the recession explains why matching efficiency declined during the recession, in contrast with previous experiences. See Figure 1 in the online Appendix.

factors, as observable components were relatively constant. Thereafter, observable factors, especially unemployment duration and reason for unemployment, began to contribute to the deterioration, and this contribution has grown steadily, while the unexplained component has been relatively constant. As a result, as of 2010Q4, composition accounts for about 45 percent of the 0.32 log points decline in aggregate matching efficiency since 2007, and the unexplained decline in matching efficiency amounts to about 0.17 log points.[34]

5 Dispersion and matching efficiency over 2006-2010

While we found little effect of dispersion on matching efficiency so far, our empirical exercise has relied on an imperfect proxy of local labor market tightness (the unemployment rate by state and industry).

In this section, we relax this assumption and present a new dataset with online Help-Wanted ads from the Conference Board that allows us to directly observe labor market tightness by geography and occupation. While the dataset only starts in 2006, it allows us to probe the effect of dispersion on matching efficiency at a unique level of disaggregation (564 segments), a 55-fold improvement over publicly available data such as the Job Openings and Labor Turnover Survey (JOLTS). This feature of the data is key, since the effect of dispersion arises out of the concavity of the matching function, and can thus only be observed at a sufficiently high level of disaggregation. Nonetheless, since the level of disaggregation allowed by our dataset may still be too coarse to capture the full extent of dispersion, we present a method using UK data to infer the extent of dispersion at an even finer level of disaggregation.

5.1 A simpler empirical framework

To clarify the presentation, we abstract from worker heterogeneity and instead study the effect of dispersion on *composition-adjusted* matching efficiency, i.e., the component of matching efficiency *not* explained by composition (the lower panel of Figure 5). As discussed in Section

[34] In the online Appendix, we test whether there has been a break in the coefficients capturing the effect of composition, in particular in the coefficient on permanent job losers and in the coefficient on unemployment duration. The main change is in the elasticity of the aggregate matching function, suggesting that the deterioration in matching efficiency is not related to the observed characteristics of the unemployed. We also test whether the unexplained decline in matching efficiency was concentrated in certain states (perhaps caused by a house-lock which lowered workers mobility and thus matching efficiency in states with particularly distressed housing markets) or in certain industries. The decline in matching efficiency appears relatively widespread across industries, and the geographic distribution of match efficiency does not suggest real estate, or any other factor, as a single cause.

3.2, without worker heterogeneity, the job finding rate in segment i is given by

$$jf_{it} = m_i \theta_{it}^{\omega(1-\sigma)} \theta_t^{(1-\omega)(1-\sigma)}. \tag{14}$$

5.2 A new dataset with job openings data by geography and occupation

Since November 2006, the Conference Board has published the *number* of help-wanted online ads by state *and* occupation, as well as the number of ads by metropolitan statistical areas (MSA) *and* occupation. To the best of our knowledge, these datasets are the first ones to contain information on vacancy posting in the US by occupation *and* geographic location at a high level of disaggregation.

Moreover, we expand the coverage of each dataset by combining the state-level information with the MSA-level information to produce series of vacancy posting by occupation and geographic areas across the US. Indeed, while help-wanted ads by state cover the whole US population (unlike ads by MSA), probing the data only at the state level may lead us to underestimate the extent of geographic dispersion because U.S. states are far from being uniformly populated.[35] To address that issue, we combine state-level vacancy information with MSA-level vacancy information to create smaller geographic cells. With 50 states and 52 MSAs, we get a total of 94 geographic areas.[36]

The Conference Board reports online ads for six occupation groups, allowing us to observe vacancies across 94 geographic units and 6 occupation groups.[37] After combining the vacancy series with the number of unemployed by geographic area and occupation estimated from the CPS, we can survey the extent of dispersion over $94 * 6 = 564$ labor market segments during the 2008-2009 recession.

[35] For instance, California hosts 17% of all the unemployment, and treating California as one geographic unit could lead us to understate the extent of dispersion.

[36] More specifically, we proceed as follows: when a state comprises n MSAs, we decompose the number of ads for that state into $n + 1$ geographic areas. The additional area is the difference between total ads in the state and the total number of ads across the MSAs from that state. We obtain less than 50+52 areas because some MSAs span different states (such as New York City). When this is the case, we group these states together. The additional area is then the difference between total ads in those states and the total number of ads in the MSAs of those states. A list of the geographic areas is presented in Table 5. For instance, California is split in 7 cells, and the largest cell, NewYork City hosts less than 5% of the unemployed.

[37] These groups are Management & Business, Professional & Related, Services, Sales & Office, Construction & Maintenance, Production & Transportation. They correspond to the SOC high level aggregations, except Management & Business/Financial and Professional & Related which split the high-level aggregation group "Management, Business, Science and Arts" into two subgroups.

5.3 Accounting for vacancy heterogeneity

Vacancy data across segments may not be directly comparable because some segments can have a higher share of informal hiring than others. For example, it is likely that a lot of hiring in construction occurs without the formal posting of a vacancy. While the aggregate vacancy-unemployment ratio from Conference Board data averaged 1.1 in 2006, the vacancy-unemployment ratio averaged 0.5 in construction and maintenance but about 4 in management and business/financial. Similarly, because a broad industry group may contain industries with different levels of informal hiring, the levels of job openings may not be comparable across regions with different industry specializations. For example, labor market tightness in services is on average three times higher in Denver than in New York. Similarly, rural areas and urban areas need not display the same fraction of informal hiring.

Because of such differences in informal hiring, the matching function constant m_i may differ across segments as some sectors (such as construction) produce more matches than others (such as business/financial) for a given number of job openings and job searchers. To address this issue, we treat informal hiring as a measurement error in vacancy posting. Formally, we do not measure V_{it}, the effective number of job openings that includes formal and informal job openings, but $\widetilde{V}_{it}$, the number of published job openings, and we have

$$V_{it} = \alpha_i \widetilde{V}_{it}$$

with α_i the inverse of the share of formal hiring. As a result, $\theta_{it} = \alpha_i \tilde{\theta}_{it}$ and from $\theta_t = \sum_i \frac{U_{it}}{U_t}\theta_{it}$, we get

$$\theta_t = \alpha_{0t}\tilde{\theta}_t \text{ and } \alpha_{0t} \equiv \sum_i \frac{V_{it}}{V_t}\alpha_i.$$

After controlling for the different fractions of informal hiring, the matching function constants m_is are identical across segment and equal m_0.[38] The job finding rate in segment i can then be written

$$jf_{it} = m_0 \left(\alpha_i \tilde{\theta}_{it}\right)^{\omega(1-\sigma)} \left(\alpha_{0t}\tilde{\theta}_t\right)^{(1-\omega)(1-\sigma)}. \tag{15}$$

[38] We thus assume that, apart from measurement error, there are no differences in matching efficiency across occupation and geographic area so that $m_i = m_0$, $\forall i$ (an assumption we implicitly made in Section 3 with our functional form (10)). With different matching efficiency levels, (14) becomes $jf_{it} = m_i \alpha_i^{\omega(1-\sigma)} \tilde{\theta}_{it}^{\omega(1-\sigma)} \alpha_{0t}^{(1-\omega)(1-\sigma)} \tilde{\theta}_t^{(1-\omega)(1-\sigma)}$. With α_i and m_i both affecting matching efficiency in an observationally equivalent manner, it is not possible to disentangle the two phenomena using information on the job finding rate and measured labor market tightness alone.

5.4 Estimating the fraction of informal hiring

With a little bit of algebra similar to the previous section, the effect of dispersion on aggregate matching efficiency movements is given by

$$\Delta mm_t \simeq \frac{1}{2}\omega(1-\sigma)(1-\omega(1-\sigma))\left[Var\left(\frac{\alpha_i}{\alpha_{0t}}\frac{\tilde{\theta}_{it}}{\tilde{\theta}_t}\right) - \mathrm{E}_T Var\left(\frac{\alpha_i}{\alpha_{0t}}\frac{\tilde{\theta}_{it}}{\tilde{\theta}_t}\right)\right]. \qquad (16)$$

To calculate the extent of dispersion, we thus need a measure labor market tightness by segment *as well as* an estimate of $\frac{\alpha_i}{\alpha_{0t}}$, the share of informal hiring in segment i relative to average informal hiring. To estimate $\frac{\alpha_i}{\alpha_{0t}}$ for each segment, we proceed as follows:

After taking the log of (15) for segment i and for an arbitrarily chosen segment 1, we estimate the regression

$$\ln jf_{it} - \ln jf_{1t} = \omega(1-\sigma)\ln\frac{\alpha_i}{\alpha_1} + \omega(1-\sigma)\left(\ln\tilde{\theta}_{it} - \ln\tilde{\theta}_{1t}\right) + \xi_{it} \qquad (17)$$

where we take the log-difference between two segments to remove the non-observable parameter α_{0t}.

Given the high level of disaggregation of the online HWI data, the sample size of the CPS is a limitation and the monthly measures of the job finding rate and unemployment over 564 segments are noisy. We thus take yearly averages of equation (17), and build a panel with 564 segments and 4 time periods (2006-2007, 2008, 2009 and 2010).

Table 3 presents the results of regression (17). With an estimated $\omega(1-\sigma) = 0.22$ and using $1-\sigma = 0.33$ from Table 1, we get a permeability coefficient $\omega = 0.65$.[39] Even at a relatively high level of disaggregation, workers' job finding rate depends to a large extent on the tightness of the local labor market. From the value of $\omega(1-\sigma)$ and the constant term in (17), we then get estimates of $\frac{\alpha_i}{\alpha_1}$. We can then obtain $\frac{\alpha_{0t}}{\alpha_1}$ from $\frac{\alpha_{0t}}{\alpha_1} = \sum_i \frac{V_{it}}{V_t}\frac{\alpha_i}{\alpha_1}$, and we get $\frac{\alpha_i}{\alpha_{0t}}$ from $\frac{\alpha_i}{\alpha_{0t}} = \frac{\alpha_i}{\alpha_1} / \frac{\alpha_{0t}}{\alpha_1}$.

5.5 A measure of dispersion over 564 segments

With measures of $\frac{\tilde{\theta}_{it}}{\tilde{\theta}_t}$ and estimates of informal hiring $\frac{\alpha_i}{\alpha_{0t}}$, we can now study the effect of dispersion on matching efficiency. Figure 6 plots $Var\left(\frac{\theta_{it}}{\theta_t}\right)$ with $\frac{\theta_{it}}{\theta_t} = \frac{\alpha_i}{\alpha_{0t}}\frac{\tilde{\theta}_{it}}{\tilde{\theta}_t}$, the dispersion in labor market tightness over 2006-2010, along with composition-adjusted matching efficiency over 2001-2010.

[39] Our previous estimate based on estimating (13) in which we used $\frac{u_{it}}{u_t}$ to proxy for $\frac{\theta_{it}}{\theta_t}$ was smaller at 0.55. This is not surprising given that our previous estimate likely suffered from an attenuation bias because of measurement error.

Dispersion by geography and occupation shows a clear increase in the 2008-2009 recession, and Figure 6 shows that the behavior of the dispersion lines up strikingly well with the behavior of composition-adjusted matching efficiency. In fact, the correlation between the two series is very high at 0.91. Over 2007-2008, the increase in dispersion coincides with the decline in match efficiency. In contrast, composition was flat during that period (Figure 4). In 2009, both dispersion and the negative of composition-adjusted matching efficiency peaked before declining slightly, and then rebounding in 2010.

This evidence suggests that dispersion may be responsible for the behavior of matching efficiency over 2006-2010 and its exceptionally low value. To quantify the effect of dispersion in units of lower matching efficiency, we use

$$\Delta mm_t \simeq \frac{1}{2}\omega(1-\sigma)(1-\omega(1-\sigma))\left[\Delta Var\left(\frac{\theta_{it}}{\theta_t}\right)\right]. \tag{18}$$

With the variance of $\frac{\theta_{it}}{\theta_t}$ increasing from about 0.95 in November 2006 to about 1.6 in December 2009, dispersion can explain about 38 percent (0.055 log-points out of the 0.17 unexplained log-decline in jf, cf. Figure 5).[40] Dispersion appears to be responsible for a large fraction of the unexplained decline in matching efficiency.

5.6 Using UK data to infer the amount of dispersion at a higher level of disaggregation

Despite the striking correlation between labor market dispersion and composition-adjusted matching efficiency, dispersion cannot account for all of the decline in matching efficiency. One reason for this result may be that the level of disaggregation achieved by our dataset is not sufficient to probe the full extent of the increase in dispersion in the labor market. Indeed, because the effect of dispersion arises out of the concavity of the matching function, it is crucial to reach a good level of disaggregation. Our highest level of disaggregation covers *only* 564 labor market segments, probably a small amount compared to the true number of segments in the US.[41]

[40] Note that our calculation implicitly assumed that the estimated value of $Var\left(\frac{\theta_{it}}{\theta_t}\right)$ in November 2006 corresponds to $\mathrm{E}_T Var\left(\frac{\theta_{it}}{\theta_t}\right)$, the average dispersion level over 1976-2007. This seems plausible given that Figures 1 and 5 show that aggregate matching efficiency was at its average level in late 2006, suggesting that dispersion was at its average level.

[41] As Shimer (2007) noted, the Occupational Employment Statistics (OES) counts about 800 occupations, and there are 362 metropolitan statistical areas and 560 micropolitan statistical areas, so a total of about 740,000 labor market segments. Of course, 740,000 is an extreme example. At this level of disaggregation, the boundaries between segments are less clearly defined, and ω is likely close to zero. Moreover, a matching function is unlikely to capture the functioning of the labor market at this level of disaggregation.

In this subsection, we propose a method using UK data to scale up our measure of dispersion over 564 segments to a larger number of labor market segments. Our results suggest that, at the higher level of disaggregation allowed with UK data, dispersion would explain all of the decrease in matching efficiency.

We define a unit as the smallest labor market segment in which the concept of a matching function remains valid and the job finding rate is described by (14). We will refer to an elementary segment as a unit.

The effect of dispersion on matching efficiency is thus given by the dispersion in labor market conditions across such units. We cannot observe labor market tightness at the unit level, but we can observe the average value of θ_i over several units. Specifically, we observe

$$\bar{\theta}_j = \frac{1}{m} \sum_{j_k \in I_j} \theta_{j_k}$$

the average value of θ_i in segment j, consisting of $m \equiv \frac{\bar{N}}{N}$ units indexed by $I_j = \{j_1, .., j_m\} \subset [1, \bar{N}]$ with $\bar{N}$ the total number of units and N the number of observed (larger) segments. Thus, we cannot measure $Var\left(\frac{\theta_i}{\theta}\right)$, the dispersion over the $\bar{N}$ units, but we can measure $Var_n\left(\frac{\bar{\theta}_j}{\theta}\right)$, the variance in labor market tightness over N larger segments with $n \equiv \frac{N}{\bar{N}} \leq 1$. As we refine the level of disaggregation and $n \rightarrow 1$, $Var_n\left(\frac{\bar{\theta}_j}{\theta}\right)$ will converge to $Var\left(\frac{\theta_i}{\theta}\right)$. The goal of this section is to construct an estimator of $Var\left(\frac{\theta_i}{\theta}\right)$ from $Var_n\left(\frac{\bar{\theta}_j}{\theta}\right)$.

A simple theoretical framework left for the Appendix shows that we could expect a relation between $Var\left(\frac{\theta_i}{\theta}\right)$ and $Var_n\left(\frac{\bar{\theta}_j}{\theta}\right)$ of the form

$$Var\left(\frac{\theta_i}{\theta}\right) = \frac{Var_n\left(\frac{\bar{\theta}_j}{\theta}\right)}{f(n_{geo}, n_{occ})} \qquad \text{with } n_{geo} = \frac{N_{geo}}{\bar{N}_{geo}}, n_{occ} = \frac{N_{occ}}{\bar{N}_{occ}} \tag{19}$$

with $f(n_{geo}, n_{occ}) \leq 1$, $f'(.)>0$, $f \rightarrow 1$ when $(n_{geo}, n_{occ}) \rightarrow (1, 1)$.

To make further progress on $f(.)$ and (19), we turn to UK data. Unlike the US, the UK public employment office collects vacancies by occupation and geography at very different levels of disaggregation, from low levels of disaggregation to very high levels of disaggregation (as high as 80,000 segments). These numbers can in turn be matched to the number of job seekers' allowance claimants to construct measures of labor market tightness across various occupation and geographic segments.[42] With these data, we can then establish an empirical "scaling law" that captures how $Var_n\left(\frac{\bar{\theta}_j}{\theta}\right)$ converges towards $Var\left(\frac{\theta_i}{\theta}\right)$ when we raise the number of observations N and observe smaller labor market segments.

[42] See Sahin, Song, Topa, and Violante (2010) for a detailed study of mismatch in the UK.

The UK data by occupation are available at the one- to four- digits SOC levels, consisting of respectively 9, 25, 81 and 353 groups, and we use data by geographic region at three disaggregation levels; government office regions (11 segments), Job Center plus Districts (48 segments), and Travel to Work Areas (232 segments). Thanks to these different levels of disaggregation, we can probe how $Var_n\left(\frac{\bar{\theta}_j}{\theta}\right)$ varies as we increase the precision of observations from $N = N_{occ} * N_{geo} = 9 * 11 = 99$ segments to $353 * 232 = 81,896$ segments. To increase the sample size, we took averages of unemployment and vacancy data over the whole sample period July 2006-July 2010.[43]

Empirically, a power law

$$\ln Var_n(\frac{\bar{\theta}_j}{\theta}) = \ln a_0 + a_{geo} \ln n_{geo} + a_{occ} \ln n_{occ}$$

implying

$$f(n_{geo}, n_{occ}) = n_{geo}^{a_{geo}} n_{occ}^{a_{occ}} \text{ and } Var\left(\frac{\theta_i}{\theta}\right) = a_0$$

fits the UK data extremely well with $a_{occ} = 0.67$ and $a_{geo} = 0.13$ and an R^2 of 0.98 (Table 4).[44] To illustrate the empirical relation, Figure 7 plots the relationship between the total number of observed segments and $Var_n\left(\frac{\bar{\theta}_j}{\theta}\right)$ as we increase the number of occupation categories from 9 (comparable with the 6 occupations observed using Conference Board data) to 353, and holding the number of geographic areas constant at 48 (comparable with the 94 areas observed using Conference Board data).

Assuming that $f(.)$ is time invariant and is similar in the UK and in the US (i.e. that the scaling law parameters a_{occ} and a_{geo} are not country specific)[45] and given an estimate of the number of labor market units $\bar{N}_{geo}$ and $\bar{N}_{occ}$, we can use the UK scaling law to build an estimator of $Var\left(\frac{\theta_{it}}{\theta_t}\right)$:

$$\widehat{Var}_n\left(\frac{\theta_i}{\theta}\right) \equiv \frac{Var_n\left(\frac{\bar{\theta}_j}{\theta}\right)}{f(n_{geo}, n_{occ})}. \tag{20}$$

Assuming that there are 81 occupation units in the US and 232 geographic units, we get that $f(\frac{94}{232}, \frac{6}{81}) \simeq 1/6$, so that an increased in measured dispersion in online HWI from 0.95 to 1.6 between November 2006 and December 2010 for (N_{occ}=6, N_{geo}=94) translates into an increase from 5.5 to 9.5 when (N_{geo}=232, N_{occ}=81). At this high level of disaggregation, it is likely that permeability is a lot higher and thus ω a lot lower. For instance, using $\omega = 0.3$ in (18), the

[43] We do not use data prior to May 2006 because of a break in methodology.

[44] Estimating the scaling law using only yearly data (and leaving out the highest disaggregation levels $N_{occ} = 353$) gives similar a_1 and a_2 but different intercepts a_0, supporting our assumption that $f(.)$ is time invariant.

[45] In the Appendix, we show that one can apply the UK scaling law to US data if the average correlation in labor market conditions within an occupation group and/or a geographic area is similar in both countries.

increase in dispersion can explain all of the decline in matching efficiency (i.e., $\Delta mm_t \simeq$0.17 log points).[46]

6 Conclusion

An important assumption in search and matching models is the existence of an aggregate matching function with a constant matching technology. This paper shows that a substantial fraction of the cyclical variations in the job finding rate cannot be explained within the context of such an aggregate matching function. Movements in matching efficiency, i.e., movements in the aggregate job finding rate not captured by a matching function with constant matching technology, display non-trivial cyclical fluctuations. In units of unemployment rate, matching efficiency movements generate fluctuations in unemployment of about 1 percentage point over 1967-2006. Moreover, in the 2008-2009 recession, an exceptional decline in matching efficiency added an estimated $1\frac{1}{2}$ percentage points to the unemployment rate.

Although the matching function is meant to capture a trading technology between heterogeneous agents, we argue that heterogeneities across individuals and labor markets are not captured by the matching function but are key to understand fluctuations in the job finding rate (and thus unemployment). Under fairly general assumptions, we identify two factors, which, in addition to labor market tightness, affect the job finding rate: (i) composition of the unemployment pool, in particular the fraction of long-term unemployed and the fraction of job losers, and (ii) dispersion in labor market conditions, the fact that tight labor markets coexist with slack ones.

Using CPS micro data over 1976-2010, we find that until 2006, composition of the unemployment pool is responsible for most of the cyclical movements in matching efficiency. Since 2006 however, composition explains only 45 percent of a dramatic decline in matching efficiency. Instead, the decline in matching efficiency not explained by composition is very highly correlated with dispersion in labor market conditions. Quantitatively, our measure of dispersion over 564 segments explains about 40 percent of the decline in composition-adjusted matching efficiency since 2006. Using UK data to infer the increase in dispersion over a higher number of segments suggests that dispersion can explain all of the decline in matching efficiency. Nonetheless, our quantitative exercise does not rule out another mechanism that could have also contributed to the exceptionally low level of matching efficiency since 2006: As suggested by Davis, Faberman and Haltiwanger (2010), firms could have displayed an exceptionally low level of recruiting intensities since the 2008-2009 recession, which would have automatically

[46] If the average level of dispersion is given by the November 2006 reading of 5.5, this implies that, on average, dispersion in labor market conditions depresses the US job finding rate by about 26 percent.

lead to lower matching efficiency.

An implication of our results is that ignoring heterogeneity across workers and labor markets (as typically assumed in search and matching models, e.g., Pissarides, 2001) may yield an incomplete depiction of unemployment fluctuations.[47] As such, explicitly incorporating heterogeneity across agents and modeling mobility decisions across segmented labor markets are important research projects.[48] Moreover, while increased dispersion is suggestive of an increase in mismatch in the US economy since 2007, dispersion could also arise because of disproportionate responses to an aggregate shock. Better understanding how dispersion can arise in equilibrium is thus an important task for future research.

[47] While the Mortensen-Pissarides model allows for heterogeneity across matches, it assumes homogenous individuals and the existence of a single labor market.

[48] For recent work in this direction, see Alvarez and Shimer (2011), Birchenall (2011), Carrillo-Tudela and Visscher (2011), and Hertz and Van Rens (2011).

References

[1] Abraham, K. "Mismatch and Labour Mobility: some final remarks," in *Mismatch and Labor Mobility*, Cambridge, Cambridge University Press, 1991.

[2] Abraham, K. and L. Katz. "Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances?," *Journal of Political Economy*, Vol. 94, No. 3, pp. 507-522, 1986.

[3] Alvarez F. and R. Shimer "Search and Rest Unemployment," *Econometrica*, 79(1): 75-122, 2011.

[4] Attanasio, O. and F. Padoa Schioppa. "Regional Inequalities, migration and mismatch in Italy, 1690-86," in *Mismatch and Labor Mobility*, Cambridge, Cambridge University Press, 1991.

[5] Baker, M. "Unemployment Duration: Compositional Effects and Cyclical Variability," *American Economic Review,* 1992.

[6] Barnichon, R. "Building a composite Help-Wanted index," *Economics Letters*, 2010.

[7] Barnichon, R., M. Elsby, B. Hobijn, A. Sahin. "Which Industries are Shifting the Beveridge Curve? Working Paper, 2010.

[8] Barnichon, R. and A. Figura. "What Drives Movements in the Unemployment Rate? A Decomposition of the Beveridge Curve," FEDS Working Paper, 2010.

[9] Bean, C. and C. Pissarides. "Skill shortages and structural unemployment in Britain: a (mis)matching approach," in *Mismatch and Labor Mobility*, Cambridge, Cambridge University Press, 1991.

[10] Birchenall, J. "A Competitive Theory of Mismatch," Working Paper, 2011.

[11] Blanchard O. and P. Diamond. "The Beveridge Curve," *Brookings Paper on Economic Activity*, 1:1-60, 1989.

[12] Bleakley H. and J. Fuhrer. Shifts in the Beveridge Curve, Job Matching, and Labor Market Dynamics," *New England Economic Review*, 1997.

[13] Brunello, G. "Mismatch in Japan," in *Mismatch and Labor Mobility*, Cambridge, Cambridge University Press, 1991.

[14] Burda, M. and S. Profit. "Matching across space: Evidence on mobility in the Czech Republic," *Labour Economics*, vol. 3(3), 255-278, 1996.

[15] Burgess, S. and S. Profit. "Externalities in the matching of workers and firms in Britain," *Labour Economics*, vol. 8(3), 313-333, 2001.

[16] Carrillo-Tudela, C. and L. Visscher. "Unemployment and Endogenous Reallocation over the Business Cycle," Wrking Paper, 2011.

[17] Darby, Michael, John Haltiwanger, and Mark Plant. "Unemployment Rate Dynamics and Persistent Unemployment under Rational Expectations," American Economic Review, 75(4), 614–637, 1985.

[18] Davis, S., J. Faberman and J. Haltiwanger. "The Establishment-Level Behavior of Vacancies and Hiring," NBER Working Paper, 2010.

[19] Elsby, M., B. Hobijn and A. Sahin. "The Labor Market in the Great Recession," *Brookings Papers on Economic Activity*, 1-48, 2010.

[20] Farber, H "Job Loss and Unemployment in the 21st Century: the Great Recession in Labor Market Perspective," Industrial Relations Section , Princeton University Working Paper,560, 2010.

[21] Franz, W. "Match and mismatch on the German labour market," in *Mismatch and Labor Mobility*, Cambridge, Cambridge University Press, 1991.

[22] Fujita, S. "Effects of the UI Benefit Extensions: Evidence from the Monthly CPS," Philadelphia Fed Working Paper 10-35, 2010.

[23] Herz, B. and T. van Rens. "Structural Unemployment," mimeo, 2011.

[24] Hsieh, C. and P. Klenow. "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124, 1403-1448, 2009.

[25] Jackman, R. and S. Roper. "Structural Unemployment," *Oxford Bulletin of Economics and Statistics*, 49(1), 9-36, 1987.

[26] Jackman, R., R. Layard, and S. Savouri. "Mismatch: a framework for thought," in *Mismatch and Labor Mobility*, Cambridge, Cambridge University Press, 1991.

[27] Layard R., S. Nickell, and R. Jackman. *Unemployment: Macroeconomic Performance and the Labour Market*, 2nd Edition, Oxford University Press, 2005.

[28] Kaplan, G. and S. Schulhofer-Wohl. "Interstate migration has fallen less than you think: Consequences of hot deck imputation in the Current Population Survey," Working Paper, 2010.

[29] Kaitz H. "Analyzing the Length of Spells of Unemployment," *Monthly Labor Review*, 93(11):10-20 , 1970.

[30] Katz, Lawrence, and Bruce Meyer. "The Impact of the Potential Duration of Unemployment Benefits on the Duration of Unemployment," *Journal of Public Economics*, 41 pp. 45-72, 1991.

[31] Kuang, K. and R. Valletta. "Extended Unemployment and UI Benefits," FRBSS Economic Letter, 2010.

[32] Machin, Stephen, and Alan Manning. The Causes and Consequences of Long Term Unemployment in Europe. In Handbook of Labor Economics Vol. 3, ed Orley Ashenfelter and David Card, 3085-3139.Elsevier: Amsterdam, 1999.

[33] Meyer, Bruce. "Unemployment Insurance and Unemployment Spells," *Econometrica*, 58:4, pp. 757-82, 1990.

[34] Moffitt, Robert. "Unemployment Insurance and the Distribution of Unemployment Spells," *Journal of Econometrics*, 28, pp. 85-101, 1985

[35] Molloy, R., C. Smith and A. Wozniac. "Internal Migration in the US: Updated Facts and Recent Trends," mimeo, University of Notre Dame, 2010.

[36] Padoa Schioppa, F. *Mismatch and Labor Mobility*, Cambridge, Cambridge University Press, 1991.

[37] Petrongolo, B. and C. Pissarides. "Looking into the black box: A survey of the matching function," *Journal of Economic Literature*, 39: 390-431, 2001.

[38] Pissarides, C. Loss of Skills during Unemployment and the Persistence of Employment Shocks," *The Quarterly Journal of Economics*, 107(4): 1371-91, 1992.

[39] Pissarides, C. *Equilibrium Unemployment Theory*, 2nd ed, MIT Press, 2000.

[40] Sahin A., J. Song, G. Topa, and G. Violante. "Mismatch in the Labor Market: Evidence from the US," mimeo, 2011.

[41] Sedlacek, P. "Match Efficiency and the Cyclical Behavior of Job Finding Rates," mimeo, 2011.

[42] Shimer, R. "Mismatch," *American Economic Review*, 97(4): 1074-1101, 2007.

[43] Shimer, R. "Reassessing the Ins and Outs of Unemployment," NBER Working Paper No. 13421, 2007.

[44] Shimer, R. "The Probability of Finding a Job," *American Economic Review (Papers and Proceedings)*, 98(2): 268-273, 2008.

Appendix

A second-order Taylor expansion of the job finding probability

Rewriting (10), the individual job finding probability is given by

$$JF_{ij,t} = \frac{\left(1 - e^{m_i\left(\frac{\theta_{it}}{\theta_t}\right)^{(1-\sigma)\omega}\theta_t^{1-\sigma}}\right)e^{\beta X_{jt}}}{e^{m_i\left(\frac{\theta_{it}}{\theta_t}\right)^{(1-\sigma)\omega}\theta_t^{1-\sigma}} + \left(1 - e^{m_i\left(\frac{\theta_{it}}{\theta_t}\right)^{(1-\sigma)\omega}\theta_t^{1-\sigma}}\right)e^{\beta X_{jt}}} + \eta_{it}.$$

Expanding with respect to X_{jt} around $\bar{X}$ and θ_{it} around θ_t to a second-order, $JF_t = \sum\limits_{i,j} \frac{U_{ij,t}}{U_t} JF_{ij,t}$ becomes

$$JF_t = \overline{JF}_t\left(\theta_t\right) + \sum_k JF_t^{U,k} + JF_t^m - MM_t\left(\frac{\theta_{it}}{\theta_t}\right) + f_t + \eta_t$$

with

$$\begin{aligned} JF_t^{U,k} \quad = \quad & e^{-m_0\theta_t^{1-\sigma}}\left(1 - e^{-m_0\theta_t^{1-\sigma}}\right)\sum_j \frac{U_{j,t}}{U_t}[\beta_k\left(x_{jt}^k - \bar{x}^k\right) - \frac{1}{2}\left(1 - 2e^{-m_0\theta_t^{1-\sigma}}\right)\beta_k^2\left(x_{jt}^k - \bar{x}^k\right)^2 \\ & + \frac{1}{2}\sum_{l\neq k}\left(1 - 2e^{-m_0\theta_t^{1-\sigma}}\right)\beta_k\left(x_{jt}^k - \bar{x}^k\right)\beta_l\left(x_{jt}^l - \bar{x}^l\right)] \end{aligned}$$

the (second-order) composition effect from characteristic k on the average job finding rate,

$$MM_t\left(\frac{\theta_{it}}{\theta_t}\right) = \frac{1}{2}\omega(1-\sigma)\left[(1 - \omega(1-\sigma)\left(1 - m_0\theta_t^{1-\sigma}\right)\right] m_0\theta_t^{1-\sigma}e^{-m_0\theta_t^{1-\sigma}}\sum_i \frac{U_{i,t}}{U_t}\left(\frac{\theta_{it}}{\theta_t} - 1\right)^2$$

the term capturing the effect of dispersion on the average job finding rate, and

$$f_t = \omega(1-\sigma)\sum_{i,j}\sum_k \frac{U_{ij,t}}{U_t} m_0\theta_t^{1-\sigma}\left(2m_0\theta_t^{1-\sigma} - 1\right)\left(\frac{\theta_{it}}{\theta_t} - 1\right)\beta_k\left(x_{jt}^k - \bar{x}^k\right).$$

the term capturing the interaction of composition and dispersion.[49]

[49] This effect comes from the concavity of the matching function, as above average workers would have a stronger positive impact on matching efficiency than below average workers if above average workers were located in looser labor markets. Interestingly, this also implies that matching efficiency is lower when workers with above average characteristics are concentrated in tighter labor markets.

Decomposing movements in aggregate matching efficiency

To establish a link between aggregate movements in matching efficiency and changes in composition and dispersion, we write the job finding rate jf_t as a function of the job finding probability JF_t, and use (6) to get

$$\begin{aligned} jf_t &= -\ln\left(1 - JF_t\right) \\ &= -\ln\left(1 - \left(\overline{JF}_t\left(\theta_t\right) + \sum_k JF_t^{U,k} + JF_t^m - MM_t\left(\frac{\theta_{it}}{\theta_t}\right) + \eta_t\right)\right) \\ &= -\ln\left(1 - \overline{JF}_t\left(\theta_t\right)\right) - \ln\left(1 - \frac{1}{1 - \overline{JF}_t\left(\theta_t\right)}\left(\sum_k JF_t^{U,k} + JF_t^m - MM_t\left(\frac{\theta_{it}}{\theta_t}\right) + \eta_t\right)\right) \\ &\simeq \overline{jf}\left(\theta_t\right) + \frac{1}{1 - \overline{JF}_t\left(\theta_t\right)}\left(\sum_k JF_t^{U,k} + JF_t^m - MM_t\left(\frac{\theta_{it}}{\theta_t}\right) + \eta_t\right) \end{aligned}$$

with $\overline{jf}\left(\theta_t\right) \equiv -\ln\left(1 - \overline{JF}_t\left(\theta_t\right)\right)$.

Using the functional form (10), we have $\overline{jf}\left(\theta_t\right) = m_0\theta_t^{1-\sigma}$, and taking the log of the previous expression gives us

$$\begin{aligned} \ln jf_t &\simeq \ln m_0 + (1-\sigma)\ln\theta_t \\ &\quad + \ln\left(1 + \frac{e^{m_0\theta_t^{1-\sigma}}}{m_0\theta_t^{1-\sigma}}\left(\sum_k JF_t^{U,k} + JF_t^m - MM_t\left(\frac{\theta_{it}}{\theta_t}\right) + \eta_t\right)\right) \\ &\simeq \ln m_0 + (1-\sigma)\ln\theta_t \\ &\quad \underbrace{+\frac{e^{m_0\theta_t^{1-\sigma}}}{m_0\theta_t^{1-\sigma}}\left(\sum_k JF_t^{U,k} + JF_t^m - MM_t\left(\frac{\theta_{it}}{\theta_t}\right) + \eta_t\right)}_{A_t} \end{aligned} \tag{21}$$

where for the last expression, we used the fact that $m_0\theta_t^{1-\sigma}e^{-m_0\theta_t^{1-\sigma}} \gg \sum_k JF_t^{U,k} + JF_t^m - MM_t\left(\frac{\theta_{it}}{\theta_t}\right) + \eta_t$. Expression (21) has the same form as our aggregate regression (2).

Thus, the deviations of aggregate matching efficiency from its average level can be written

$$\begin{aligned} \mu_t &= \ln m_{0t} - E_T \ln m_{0t} \\ &= \ln jf_t - (1-\sigma)\ln\theta_t - \mathrm{E}_T\left(\ln jf_t - (1-\sigma)\ln\theta_t\right) \\ &\simeq A_t - \mathrm{E}_T A_t \text{ using (21)} \\ &\simeq \frac{e^{m_0\theta_t^{1-\sigma}}}{m_0\theta_t^{1-\sigma}}\left(\sum_k JF_t^{U,k} + JF_t^m\right) - \Delta mm_t + \zeta_t \end{aligned}$$

with $mm_t = \frac{e^{m_0\theta_t^{1-\sigma}}}{m_0\theta_t^{1-\sigma}} MM_t\left(\frac{\theta_{it}}{\theta_t}\right)$ and $\Delta mm_t = mm_t - \mathrm{E}_T mm_t$.[50] Using the expression for $MM_t\left(\frac{\theta_{it}}{\theta_t}\right)$, the effect of dispersion on matching efficiency movements is then given by

$$\Delta mm_t \simeq \frac{1}{2}\omega(1-\sigma)\left[\left(1-\omega(1-\sigma)\left(1-m_0\theta_t^{1-\sigma}\right)\right]\left[Var\left(\frac{\theta_{it}}{\theta_t}\right) - \mathrm{E}_T Var\left(\frac{\theta_{it}}{\theta_t}\right)\right].$$

Duration dependence and cyclical fluctuations

In this section, we show that a simple model with unobserved heterogeneity can generate a downward sloping relation between the observed job finding rate and duration *as well as* a flattening of the slope in recessions.

For ease of presentation, we consider a continuous time framework. Assume that there exists two types of individuals with unobserved ability. Individuals type H have job finding rate f_H and Individuals type L have job finding rate f_L with $f_H > f_L$.

Assume that both types are equally represented at $T = 0$, the beginning of an unemployment spell. Then, the observed job finding rate f depends on unemployment duration T according to

$$f = \frac{U_H}{U} f_H + \frac{U_L}{U} f_L$$

with $\frac{U_H}{U} = \frac{U_{H0}e^{-f_H T}}{U_{H0}e^{-f_H T} + U_{L0}e^{-f_L T}} = \frac{1}{1+e^{(f_H-f_L)T}}$ the fraction of type H unemployed with duration T amongst the unemployed with duration T. Rewriting

$$f = \frac{f_H - f_L}{1 + e^{(f_H-f_L)T}} + f_L$$

which implies

$$\frac{\partial f}{\partial T} = \frac{-(f_H-f_L)^2}{\left(1+e^{(f_H-f_L)T}\right)^2} < 0.$$

This is the standard result that unobserved heterogeneity generates a job finding rate declining with unemployment duration as the good workers disappear from stock of unemployed.

To capture the effect of cyclical fluctuations in f_H and f_L on $\frac{\partial f}{\partial T}$, the slope of the job finding rate-unemployment duration relation, assume that $f_H = \alpha\overline{f_H}$ and $f_L = \alpha\overline{f_L}$ with $\alpha < 1$. α is a cyclical indicator being high in booms and low in slumps. Then, it is easy to show that

$$\frac{\partial^2 f}{\partial\alpha\partial T} > 0,$$

[50] For the last expression, we assumed that $\mathrm{E}_T \sum_k JF_t^k$ is small, which is empirically the case since the second-order approximation of JF_t^k is very small and since we demeaned the X_{jt} variables before estimating (10) so that the first-order term of JF_t^k is nil.

so that the slope of the job finding rate-duration relation flattens in recessions: when α declines and the economy moves from a boom to a recession, the slope $\frac{\partial f}{\partial T}$ becomes less negative. To see why, denote $\delta = \overline{f_H} - \overline{f_L}$, and note that $\frac{\partial}{\partial \alpha} \ln \left(\frac{-(f_H - f_L)^2}{\left(1+e^{(f_H - f_L)T}\right)^2} \right) = \frac{-2}{\alpha} - \frac{2\delta\alpha T e^{\alpha\delta T}}{1+e^{\alpha\delta T}} < 0$, which implies that $\frac{\partial^2 f}{\partial\alpha\partial T} > 0$ since $\frac{\partial}{\partial\alpha} \ln \frac{\partial f}{\partial T} = \frac{1}{\frac{\partial f}{\partial T}} \frac{\partial^2 f}{\partial\alpha\partial T}$.

A theoretical link between actual labor market dispersion and observed dispersion

Assume that the $\{\theta_i\}$, $i \in [1, \bar{N}]$, are identically distributed across $\bar{N}$ elementary labor market segments of equal size (i.e., with the same number of unemployed). For clarity of exposition, we will refer to the elementary segments as units. Denote h the distribution of θ_i across units so that $\theta_i \sim h(\theta, \sigma_\theta^2)$ (mean θ and variance σ_θ^2). We cannot observe the value of θ_i for the $\bar{N}$ units, but we have $N < \bar{N}$ observation of the average value of θ_i across several units. Specifically, for N segments that consist of $m \equiv \frac{\bar{N}}{N}$ units, we observe $\bar{\theta}_j = \frac{N}{\bar{N}} \sum\limits_{j_k \in I_j} \theta_{j_k}$, the average value of the θ_is over segment j indexed by $I_j = \{j_1, .., j_m\} \subset [1, \bar{N}]$, a segment consisting of $m = \frac{\bar{N}}{N}$ observations of θ_i. To estimate the true amount of dispersion in the labor market, we want to recover σ_θ^2 from the observed variance across larger segments, i.e., $Var(\bar{\theta}_j)$.

If the observed segments were random samples of the θ_is (i.e., if the θ_is were independently distributed in each observed segment), we would have a linear relation linking the true dispersion to the observed variance

$$Var_n(\bar{\theta}_j) = \sigma_\theta^2 n. \tag{22}$$

which converges to σ_θ^2 when $n \equiv \frac{N}{\bar{N}} \to 1$.

In fact, however, the segments that we observe are not random samples of the θ_is. Instead, because the segments that we observe correspond to an occupation group or/and a geographic location, the labor market units inside those segments are likely to be correlated. Denote ρ the average correlation between labor market units within a segment. Specifically, $\rho = \frac{1}{n-1} \sum_{m \neq n} corr(\theta_{j_m}, \theta_{j_n})$, and, for simplicity, we assume that the average correlation ρ is the same for all observed segments. ρ is likely to increase as we refine our definition of a segment. For example, the average correlation between labor markets across the large US Census region "West" is certainly lower than the average correlation between labor markets across the city of Los Angeles. Denoting N_{geo} the number of observed geographic locations and N_{occ} the number of observed occupations, we have $N{=}N_{geo}N_{occ}$ and $n_{geo}{=}\frac{N_{geo}}{\bar{N}_{geo}}$ and $n_{occ}{=}\frac{N_{occ}}{\bar{N}_{occ}}$, and we assume that $\rho = \rho(n_{geo}, n_{occ})$ with $\rho_1' > 0$, $\rho_2' > 0$ and $\rho(n_{geo}, n_{occ}) \to 1$ when $n = n_{geo}n_{occ} \to 1$. A little bit of algebra gives us

$$Var_n(\bar{\theta}_j) = \sigma_\theta^2 \cdot f(n_{geo}, n_{occ}) \tag{23}$$

with

$$f(n_{geo}, n_{occ}) = n\,(1 + \rho(n_{geo}, n_{occ})\,(1 - n))$$

which also converges to σ_θ^2 when $N \to \bar{N}$. With $\rho(n_{geo}, n_{occ}) \neq 0$, this generalization of (22) is not linear. Instead, because $\rho(n_{geo}, n_{occ}) \to 1$ when $n \to 1$, one can show that, as n converges to 1, $\frac{\partial^2 Var(\bar{\theta}_j)}{\partial n^2} < 0$ and the curve flattens out, in line with the UK evidence that we document in the main text.

Thus, we can build an estimator of σ_θ^2

$$\widehat{Var}_n\left(\theta_i\right) \equiv \frac{Var_n\left(\bar{\theta}_j\right)}{f(n_{geo}, n_{occ})}. \tag{24}$$

This theoretical framework is useful to clarify what kind of assumptions are necessary to use the UK scaling law and the estimator $\widehat{Var}_n$ with US data. Note that $f(n_{geo}, n_{occ})$ is independent of σ_θ^2. Thus, $f(n_{geo}, n_{occ})$ will be identical in the US and the UK, if both countries have identical $\rho(n_{geo}, n_{occ})$, i.e., if, within a segment of size $\frac{\bar{N}}{N_{geo}N_{occ}}$, the *average* correlation in labor market tightness between units of that segment is the same for both countries. For example, within the "West" region of each country, the *average* correlation between two neighboring geographic units must be the same in the US and the UK. Or, within the occupation group "Construction", the *average* correlation between subcategories of construction must the same. Assuming as a first approximation that the average correlations across occupation and geographic are time invariant and similar in the US and the UK, we can apply the UK scaling law to US data.

Finally, we do not observe $Var(\bar{\theta}_j)$ but the sample variance $\frac{1}{N}\sum_{j=1}^{N}\left(\bar{\theta}_j - \theta\right)^2$. As a result, we can only use (24) if N is large enough to ensure $\frac{1}{N}\sum_{j=1}^{N}\left(\bar{\theta}_j - \theta\right)^2 \simeq Var(\bar{\theta}_j)$. For low values of N (as would be the case with JOLTS data with only 10 industry groups and 1 area), the sample variance may not be a good approximation of the actual variance, and the scaling law could give misguided results.

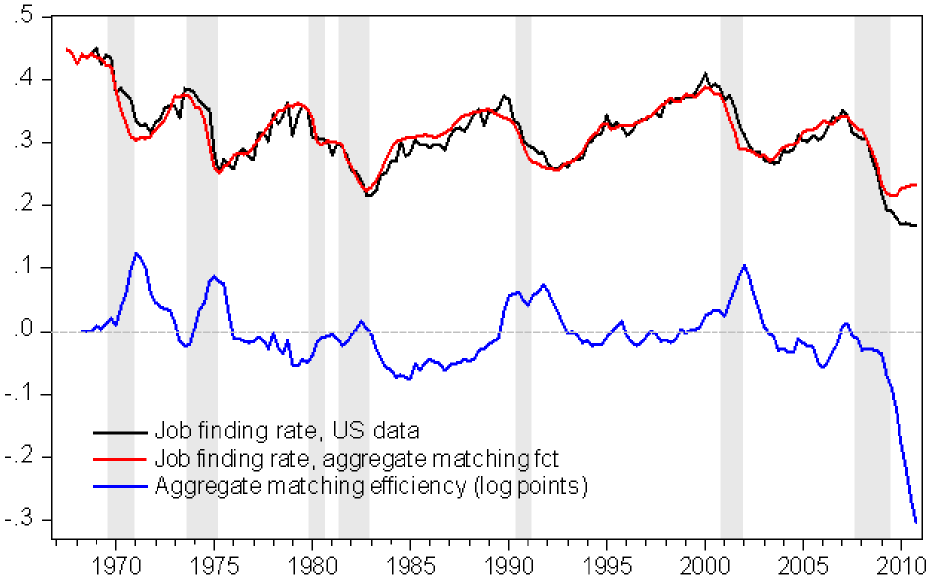

Figure 1: Empirical job finding rate, job finding rate predicted by an aggregate matching function and (log) aggregate matching efficiency, the (log) difference between the empirical and the predicted job finding rate, 1968-2010. For aggregate matching efficiency, the plotted series is the 4-quarter moving average. Grey bars indicate NBER recession dates.

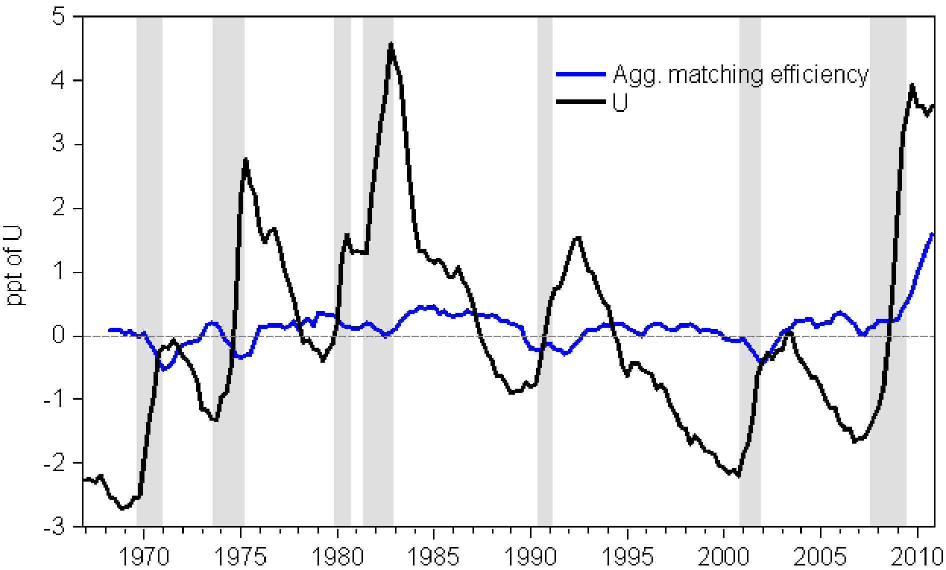

Figure 2: Contribution of movements in aggregate matching efficiency to movements in the unemployment rate, 1968-2010. See Barnichon and Figura (2010) for the methodology. The plotted series are 4-quarter moving averages. Grey bars indicate NBER recession dates.

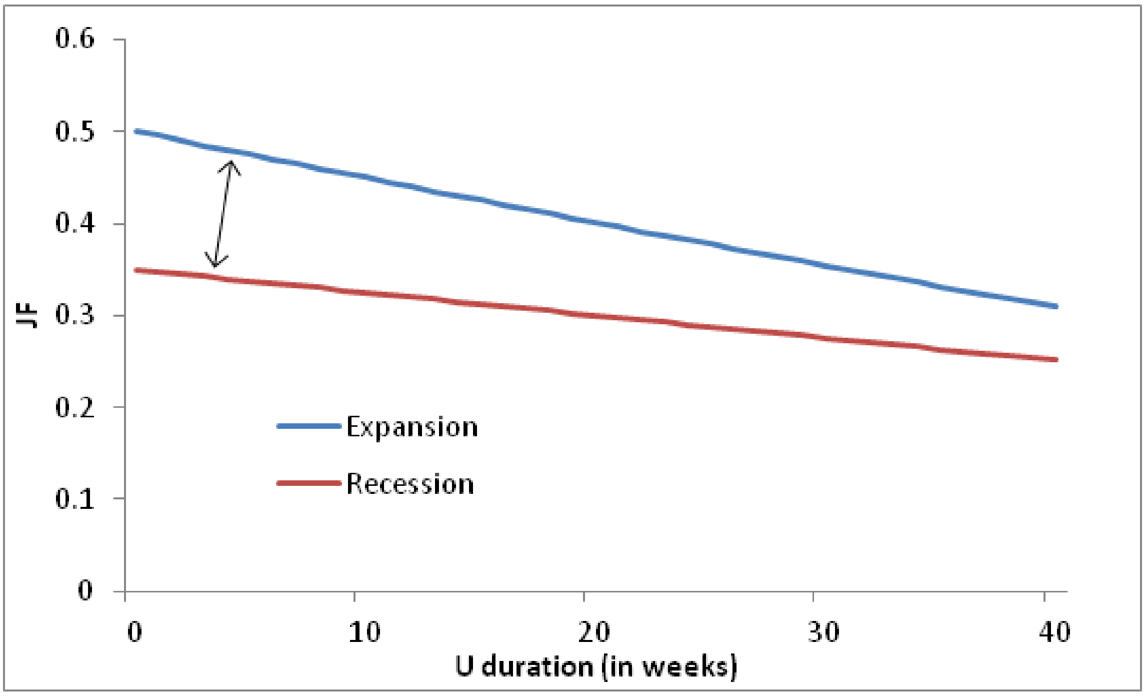

Figure 3: The estimated relation between job finding probability and unemployment duration. The slope in an expansion corresponds to an average duration of 10 weeks, and the slope in a recession corresponds to an average duration of 20 weeks.

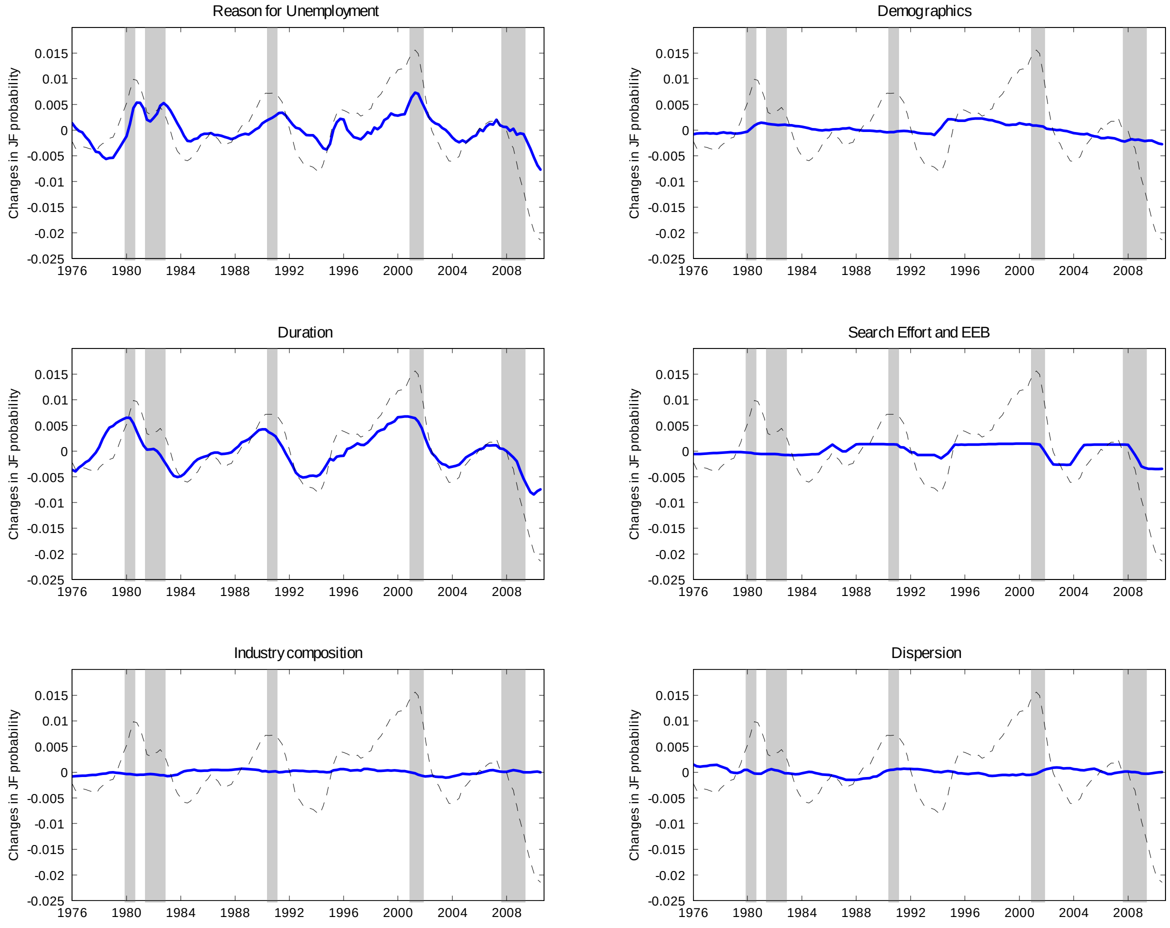

Figure 4: Decomposition of the effect of composition and dispersion on a worker's job finding probability into reason for unemployment, demographics, unemployment duration, search effort (due to emergency and extended unemployment benefits), industry composition, and dispersion. The dashed line represents the total effect of composition/dispersion. The results were obtained from a regression estimated over 1976-2007. All series are 4-quarter moving averages. Grey bars indicate NBER recession dates.

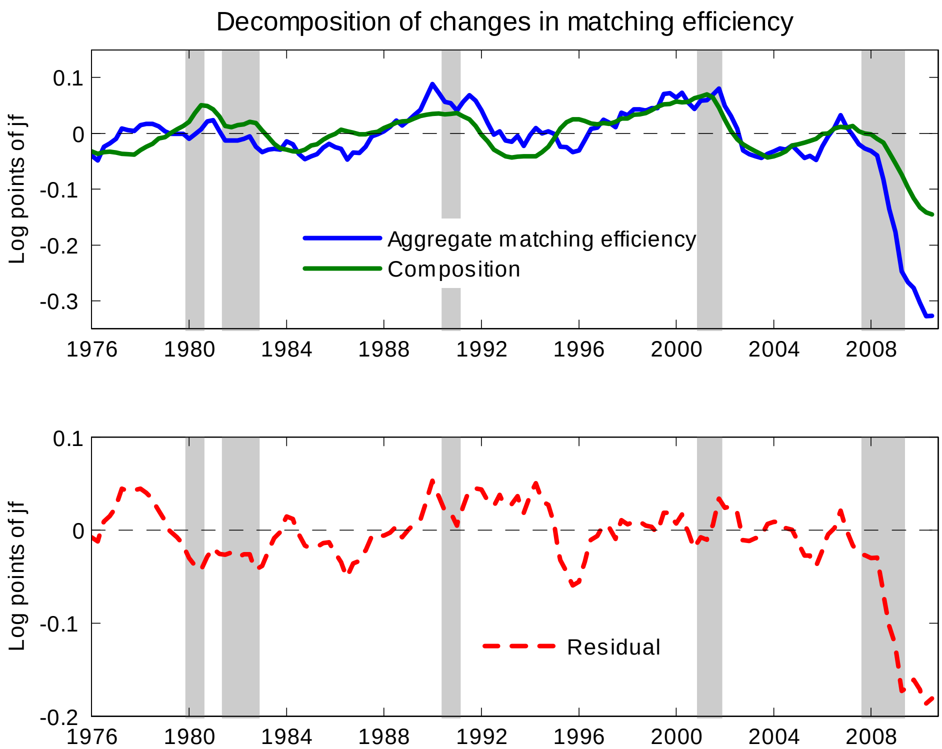

Figure 5: Upper panel: the contribution of changes in composition to movements in aggregate matching efficiency. Lower panel: unexplained movements in aggregate matching efficiency (the residual from the upper panel). The results were obtained from a regression estimated over 1976-2007. All series are 4-quarter moving averages. Grey bars indicate NBER recession dates.

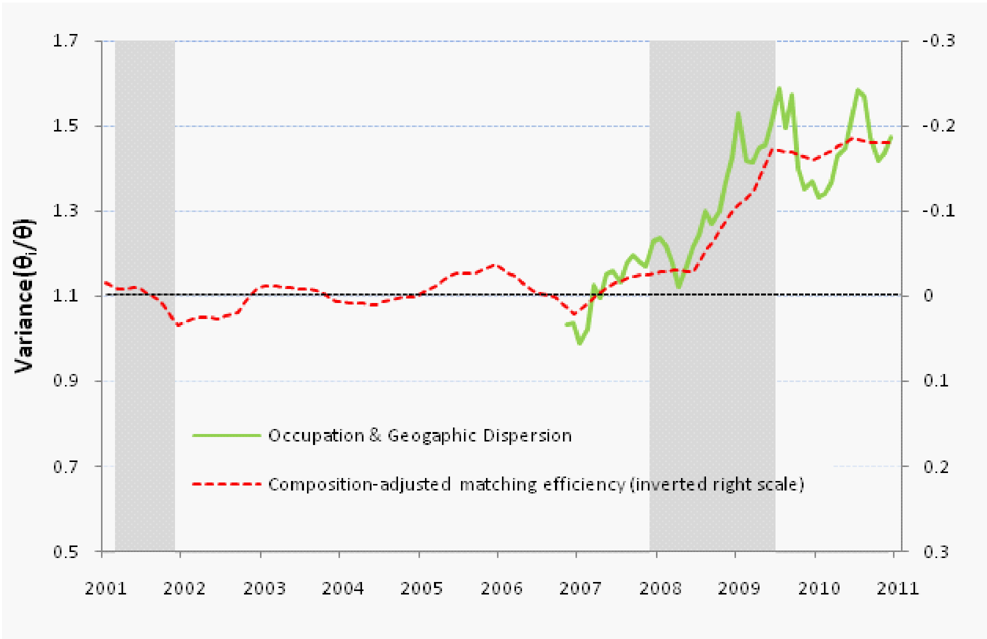

Figure 6: Mismatch and matching efficiency. Left scale: dispersion in labor market tightness across 564 occupation/region segments. Right scale: Composition-adjusted matching efficiency (y-axis in reverse order). 2006-2010. Grey bars indicate NBER recession dates.

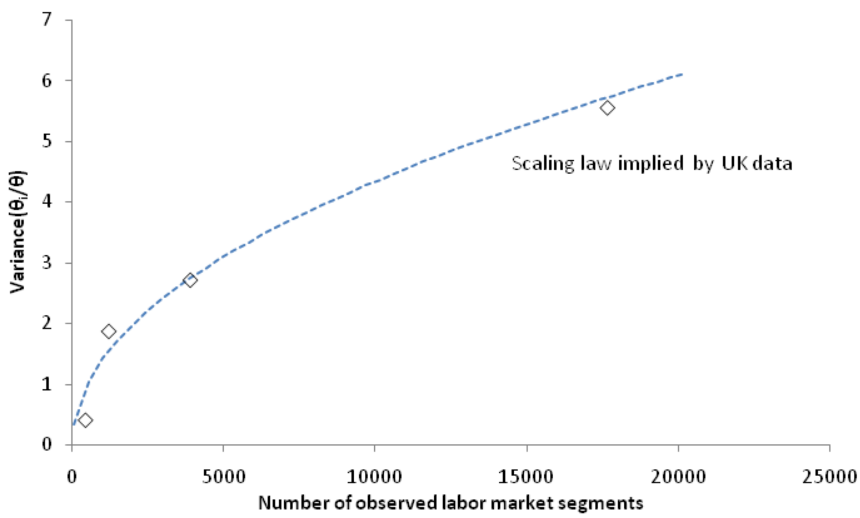

Figure 7: Relationship between $Var(\frac{\theta_{it}}{\theta})$ and the number of observed labor market segments in the UK, keeping the number of geographic units fixed (48) but increasing the number of observed occupations (9, 25, 81, 353). Labor market tightness measures constructed from jobseekers allowance claimants and vacancy posting data from Jobcentre Plus.

Table 1: Estimating a Cobb-Douglas matching function

Dependent variable:	λ^{UE}	λ^{UE}
Sample (quarterly frequency)	1968-2007	1968-2007
Regression	(1)	(2)
Estimation	OLS	GMM
1-σ	0.33***	0.34***
	(0.01)	(0.01)
R^2	0.85	--

Note: Standard-errors are reported in parentheses. In equation (2), we use 3 lags of v and u as instruments. We allow for first-order serial correlation in the residual.

Table 2 Estimated Coefficients for Job Finding probability regression, 1976-2007

Explanatory Variable	**Pre-redesign 1976-1993**		**Post-redesign 1994-2007**
Matching Function parameter			
Aggregate elasticity: 1-σ		0.28**	
		(0.01)	
Local elasticity: γ		0.20**	
		(0.01)	
Other parameters			
Age	0.003**		0.0004
	(0.001)		(0.001)
Age squared	-0.0002**		-0.0002**
	(0.0000)		(0.0000)
Male dummy	0.18**		0.09**
	(0.01)		(0.01)
Permanent layoff dummy	-0.26**		-0.27**
	(0.01)		(0.01)
Temporary layoff dummy	0.39**		0.68**
	(0.001)		(0.01)
Reentrant dummy	-0.27**		-0.33**
	(0.01)		(0.01)
New Entrant dummy	-0.63**		-0.94**
	(0.01)		(0.02)
Unemployment duration	-0.02**		-0.02**
	(0.00)		(0.00)
Duration interacted with average duration	0.0005**		0.0004**
	(0.0001)		(0.0001)
EEB variable interacted with layoff dummy (permanent layoffs only after 1994)	-0.02**		-0.05**
	(0.01)		(0.01)
Pseudo R^2		0.0455	

Note. Explanatory variables also include monthly dummies and 13 industry dummies. All variables, except age after 1993, are significant at conventional levels. Standard errors are in parentheses. ** indicates significance at the 95% confidence interval.

Table 3 Estimating the fraction of informal hiring and permeability, 2006-2010

Dependent Variables	$\ln\left(\frac{jf_{it}}{jf_{1t}}\right)$	$\ln\left(\frac{jf_{it}}{jf_{1t}}\right)$
$\omega(1-\sigma)$	0.22**	0.23**
	(0.02)	(0.02)
Number of observations	2246	2246
Adj. R^2	0.30	0.30

Note. In the first column, the regression includes 6 occupation dummies and 93 geographic dummies. In the second column, the regression includes 563 geographic/occupation dummies. Standard errors are in parentheses. ** indicates significance at the 95% confidence interval.

Table 4: Estimating a functional form for the UK scaling law

Dependent variable:	$\text{var}_n\left(\frac{\bar{\theta}_{jt}}{\theta_t}\right)$
a_0	5.08
	(0.50)
a_{occ}	0.67
	(0.03)
a_{geo}	0.13
	(0.04)
R^2	0.98

Note: The sample includes 12 observations, with N_{geo}=11, 48, 232 and N_{occ}=9, 25, 81, 353.

Table 5: List of geographic areas

AK	**FL**	**LA**	**WA**	**SD**
Other, AK	Jacksonville, FL	New Orleans, LA	Other, WA	Other, SD
AL	Miami, FL	Other, LA	Seattle-Tacoma, WA	**TN**
Birmingham, AL	Orlando, FL	**MA/RI**	**WI**	Memphis, TN
Other, AL	Other, FL	Boston, MA	Milwaukee, WI	Nashville, TN
AR	Tampa, FL	Other, MA/RI	Other, WI	**TD**
Other, AR	**GA**	Providence, RI	**WV**	Other, TD
AZ	Atlanta, GA	**MD**	Other, WV	**TX**
Other, AZ	Other, GA	Baltimore, MD	**WY**	Austin, TX
Phoenix, AZ	**HI**	Other, MD	Other, WY	Dallas, TX
Tucson, AZ	Honolulu, HI	**ME**	**NH**	Houston, TX
CA	Other, HI	Other, ME	Other, NH	Other, TX
Los Angeles, CA	**IA**	**MI**	**NM**	San Antonio, TX
Other, CA	Other, IA	Detroit, MI	Other, NM	**UT**
Riverside, CA	**ID**	Other, MI	**NV**	Other, UT
Sacramento, CA	Other, ID	**MN**	Las Vegas, NV	Salt Lake City, UT
San Diego, CA	**IN/IL/KS/MO**	Minneapolis-St. Paul, MN	Other, NV	**VA**
San Francisco, CA	Chicago, IL	Other, MN	**OK**	Other, VA
San Jose, CA	Indianapolis, IN	**MS**	Oklahoma City, OK	Richmond, VA
CO	Kansas City, MO	Other, MS	Other, OK	Virginia Beach, VA
Denver, CO	Other, IN/IL/KS/MO	**MT**	**OR**	**VT**
Other, CO	St. Louis, MO	Other, MT	Other, OR	Other, VT
CT	**KY/OH**	**NC/SC**	Portland, OR	
Hartford, CT	Cincinnati, OH	Charlotte, NC	**PA/NJ/NY**	
Other, CT	Cleveland, OH	Other, NC/SC	Buffalo, NY	
DC	Columbus, OH	**ND**	New York, NY	
Washington, DC	Louisville, KY	Other, ND	Other, PA/NJ/NY	
DE	Other, KY/OH	**NE**	Philadelphia, PA	
Other, DE		Other, NE	Pittsburgh, PA	
			Rochester, NY	