Mismatch in the Labor Market: Evidence from the U.K. and the U.S.[†]

Ayşegül Şahin
Federal Reserve Bank of New York

Joseph Song
Federal Reserve Bank of New York

Giorgio Topa
Federal Reserve Bank of New York and IZA

Giovanni L. Violante
New York University, CEPR and NBER

FIRST DRAFT: NOVEMBER 2010 - VERY PRELIMINARY - DO NOT CITE WITHOUT PERMISSION

Abstract

This paper provides a simple approach to measure mismatch in the labor market. We measure mismatch by comparing the observed allocation of unemployment and vacancies across sectors to the optimal allocation chosen by a planner who can freely move labor across sectors. We show that, in a rich dynamic stochastic economic environment, the planner's optimal allocation is dictated by a "generalized Jackman-Roper (JR) condition" where (productive and matching) efficiency-weighted vacancy-unemployment ratios should be equated across sectors. We then use this condition to develop mismatch indexes that allow us to quantify how much of the recent rise in unemployment is associated to an increase in mismatch. We apply our analysis to the U.K. and the U.S. labor markets.

[†]The opinions expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of New York or the Federal Reserve System.

1 Introduction

The unemployment rate in the U.S. rose from 4.7% in December 2007 to 10.1% in October 2009, and subsequently has been fairly stable at around 9.6% through most of 2010. In the U.K., unemployment increased by 2.5% points during the recession and leveled off at around 7.7% since then. This persistently high unemployment, in spite of the recovery in economic activity, has sparked a vibrant debate among policymakers. The main point of contention is the nature of this persistent rise. One view is that unemployment is high because aggregate labor demand is still low, and therefore reducing unemployment may require even more fiscal and monetary stimulus. A second view is that unemployment is high because of the extension of unemployment benefits. Here the policy trade-off is clear: supporting the living standards of jobless households may interfere with their incentives to quickly find new employment.

The third view –which is the focus of our study– is that unemployment is still high because of a more severe mismatch between vacant jobs and unemployed workers, i.e., the skills and locations of idle labor are poorly matched with the skill and geographical characteristics of unfilled job openings. Under this scenario, fiscal or monetary stimulus would be less effective to speed up recovery in the labor market.

This latter view is quite popular in the U.S. because several factors seem to suggest that the mismatch component of unemployment could now be significantly larger. First, half of the 8 million jobs lost in the recession belonged to construction and manufacturing, whereas a large chunk of the newly created jobs are in health care and education. Such a skill gap between job losers and job openings may hamper employment growth. Second, conditions in the housing market may slow down geographical mobility. The housing bust left almost a quarter of mortgage borrowers with negative equity. These households would have to foreclose to move to a different location and may therefore choose to keep their house at the cost of staying in low labor demand areas. Third, the U.S. Beveridge curve (i.e., the empirical relation between aggregate unemployment and aggregate vacancies) displays a marked rightward movement indicating that the current level of aggregate unemployment is higher than what it has been in the past for similar levels of aggregate vacancies.[1] Lack of coincidence between unemployment and vacancies across labor markets is one of the candidate explanations for this shift.[2]

Although there has been much debate on mismatch in policy circles, there has been no systematic and rigorous analysis of this issue in the context of the last recession.[3] In this paper we develop a

[1]This observation has been emphasized before by Davis, Faberman, and Haltiwanger (2010), Elsby, Hobijn, and Şahin (2010) and many others in the context of a shift in the Beveridge curve.

[2]For example, Phelps (2008), Elsby, Hobijn, and Şahin (2010), and Kocherlakota (2010) have argued that reallocation following the 2007-2009 recession might lead to a mismatch in skill-mix that might have resulted in a slower adjustment of the labor market than in previous recessions.

[3]For an overview of this debate, see Roubini Global Economics at http://www.roubini.com/.

simple framework to conceptualize the notion of mismatch unemployment and construct some intuitive mismatch indexes. We then use U.K. and U.S. data to quantify how much of the recent rise in unemployment is due to this channel and to identify what dimension of heterogeneity (occupation, industry, geographical location) is mostly responsible for mismatch dynamics.

To formalize the notion of mismatch, it is useful to envision the economy as comprising a large number of distinct labor markets (or sectors), segmented by industry, skill, occupation, geography, or a combination of these attributes. Consider an allocation of vacancies and unemployed workers across all these labor markets. Each labor market is frictional, i.e., the hiring process within a labor market is governed by a matching function. To assess the existence of mismatch, i.e., misallocation of unemployed workers across labor markets, we ask the following question: given the distribution of vacancies, is it feasible to reallocate unemployed workers across markets and reduce the aggregate unemployment rate?

Answering this question requires comparing the actual allocation of unemployed workers across sectors to an ideal allocation. The ideal allocation that we choose as our benchmark of comparison is the allocation which would be selected by *a planner who can freely move unemployed workers across sectors*. Since the only friction faced by this planner is the within-market matching function, unemployment arising in the efficient allocation is purely *frictional*. The differential distribution of unemployment between the observed equilibrium allocation and the ideal allocation induces a lower job finding rate which, in turn, translates into additional unemployment. This additional unemployment should be interpreted as *mismatch unemployment*.

This formalization of mismatch unemployment follows from the insight of Jackman and Roper (1987). It is, in essence, the same approach used in the large literature on misallocation and productivity (Hsieh and Klenow, 2009; Restuccia and Rogerson, 2008): quantifying misallocation entails measuring how much the observed allocation deviates from a first-best benchmark.[4]

The idea of mismatch (or structural) unemployment became popular in the 1980s when economists were struggling to understand why unemployment kept rising steadily in many European countries. The conjecture was that the oil shocks of the 1970s and the concurrent shift from manufacturing to services induced structural transformations in the labor market that permanently modified the skill and geographical map of labor demand. From the scattered data available at the time, there was also some evidence of shifts in the Beveridge curve for some countries. Padoa-Schioppa (1991) contains a number of empirical studies on mismatch and concludes that it was not an important explanation of the dynamics of European unemployment in the 1980s.[5] Similarly, the importance of the "sectoral

[4] In our case, the benchmark is a constrained first best, because the planner still faces the within-market frictional matching.

[5] Since then, it has become clear that explanations of European unemployment based on the interaction between technological changes in the environment and rigid labor market policies are more successful quantitatively (e.g., Ljungqvist and Sargent, 1998; Hornstein, Krusell and Violante, 2007).

shift hypothesis" developed by Lilien (1982) was much diminuished when Abraham and Katz (1984) pointed out that Lilien's empirical measure of dispersion of employment growth across industries could be correlated with aggregate unemployment rate even in the absence of sectoral shifts. More recently, Shimer (2007) developed a dynamic model of mismatch where workers and jobs are randomly assigned to labor markets and showed it is consistent with the aggregate empirical Beveridge curve. Alvarez and Shimer (2010), Birchenall (2010) and Carrillo-Tudela and Visscher (2010) have developed dynamic equilibrium models with mobility decisions across labor markets where unemployed workers, in equilibrium, may be misallocated.

We begin our analysis by laying out a dynamic stochastic economy with several sources of heterogeneity across sectors and show that the planner's optimal allocation of unemployed workers across sectors follows a "generalized Jackman-Roper (JR) condition" where (productive and matching) efficiency-weighted vacancy-unemployment ratios should be equated across sectors. The key feature of this optimality condition is that it is static, and hence it can be easily manipulated to construct simple mismatch indexes to use in the empirical analysis. We construct two indexes: the first, $\mathcal{M}_t^u$, measures the fraction of unemployed searching in the wrong sector at date t. The second, $\mathcal{M}_t^h$, measures the fraction of hires at date t that are lost because of such misallocation in job search. This second index allows us to construct a counterfactual unemployment rate purged of its mismatch component.

It is worth noting that our indexes capture an "ideal" notion of total mismatch as misallocation relative to an optimal unemployment distribution in the absence of *any* frictions across markets. Such frictions may include any moving or retraining costs that an unemployed worker may incur when she searches in a different sector than her original one, as well as any other distortions originating for instance from incomplete insurance, imperfect information, wage rigidities, or various government policies. Therefore, our approach yields a measurement device to compare actual unemployment to an ideal benchmark. We do not provide here a model of mismatch that analyzes its sources and delivers mismatch as an equilibrium outcome; as a consequence, we cannot say whether observed mismatch is efficient or not. We discuss the nature of our approach in more detail in Section 2.5.

We apply our analysis to the U.K. and the U.S. labor market. Constructing our indexes requires detailed information on vacancies and unemployment counts by "labor market". For the U.K., we make use of the administrative data collected by local employment agencies. The vacancy stocks and flows come from Jobcentre Plus Vacancy Statistics and the unemployment counts are from Jobseeker's Allowance Claimant Counts and are available, starting in 2005 at a monthly basis, for 2-, 3-, 4-digit occupation codes and for different travel-to-work areas (TTWAs). For the U.S., we use unemployment data from the Current Population Survey (CPS). Vacancy data are obtained from the Job Openings and Labor Turnover Survey (JOLTS) and from the Help Wanted OnLine (HWOL) series collected by the Conference Board which we have recently acquired and are currently exploring in detail.

Our main findings are as follows. In the U.K. labor market, there is no evidence of a worsening in geographical mismatch. At the occupational level, instead, mismatch rose sharply during the recession, but then quickly fell towards a value slightly higher than its pre recession level. Overall, at most 1% of the 2.5% increase in U.K. unemployment is attributable to mismatch. However, we find evidence that mismatch may be much more important for high-skill occupations.

In the U.S. labor market, mismatch at the sectoral level increased during the recession and started to come down in 2010. The increase in mismatch affected labor market outcomes during the period 2007-2010: if sectoral mismatch had stayed at its 2006 level, the number of hires would have been higher over this period. Our calculations show that the cumulative number of *lost hires* starting from the beginning of the recession added up to 2.6 million hires. However, one should keep in mind that, for the U.S., our mismatch measures only capture misallocation of workers across 14 broad industrial sectors. It is possible that these mismatch measures may not capture a significant portion of mismatch if it occurs *within* these broad sectors. Skill mismatch could be better captured by looking at mismatch between vacancies and job seekers at the occupational level. To further investigate these issues in detail, we have recently acquired Help Wanted OnLine data from the Conference Board on job vacancies by MSA, state, 6-digit occupation and education classifications.

The rest of the paper is organized as follows. Section 2 presents the theoretical framework and the mismatch indexes. Section 3 performs the empirical analysis on the U.K. labor market, and section 4 on the U.S. labor market. Section 5 concludes.

2 The framework

In this section, we generalize the insight of Jackman and Roper (1987) on how to measure mismatch unemployment (which they call "structural" unemployment). The generalization is twofold. First, we allow for a dynamic and stochastic economic environment, while their set up was static. Second, we allow for heterogeneity across sectors in a number of dimensions.[6]

Time is discrete. The economy is populated by a measure one of agents. There are I distinct labor markets (sectors) indexed by i. Labor markets are frictional: new matches, or hires, (h_i) between unemployed workers (u_i) and vacancies (v_i) in market i are determined by the matching function $\phi_i m\,(u_i, v_i)$, with m strictly increasing and strictly concave in both arguments and homogeneous of degree one in (u_i, v_i). The scalar ϕ_i measures matching efficiency (i.e., the level of fundamental frictions) in sector i. Existing matches produce z units of output, but new matches produce only a fraction $\gamma < 1$ of existing matches. This is a stylized way to capture training costs for hires of unemployed workers (regardless of the sector in which they are hired). Matches are destroyed

[6] In their model, there is no deep source of heterogeneity across sectors, even though they assume a non-degenerate distribution of vacancies across sectors. In other words, the Jackman and Roper model is not a fully specified economic environment in the tradition of modern macroeconomics.

exogenously at rate δ. To post new vacancies in sector i, firms pay the per-period cost κ_i.

The evolution of z and δ follows the joint Markov chain $\Gamma_{z,\delta}(z', \delta'; z, \delta)$. The evolution of κ_i and ϕ_i follows, respectively, $\Gamma_\kappa(\kappa'; \kappa, z', \delta')$ and $\Gamma_\phi(\phi'; \phi, z', \delta')$. We allow for a correlation between the sector specific variables (κ_i, ϕ_i) and the aggregate states, but conditional on these aggregate states, the realizations of (κ_i, ϕ_i) are independent across sectors.

Within each period, events unfold as follows. At the beginning of the period, the aggregate shocks (z, δ), matching efficiencies $\phi = \{\phi_1, ...\phi_I\}$ and vacancy costs $\kappa = \{\kappa_1, ...\kappa_I\}$ are observed. At this stage, the distribution of active matches $\mathbf{e} = \{e_1, ...e_I\}$, and vacancies $\mathbf{v} = \{v_1, ...v_I\}$ across markets are given. Since $u = \sum_{i=1}^{I}(1 - e_i)$, the number of unemployed workers is also given. Next, the unemployed workers direct their job search towards a labor market. Once the unemployed workers are allocated, the matching process takes place and $h_i = \phi_i m(u_i, v_i)$ new hires are made in each market. At this point, production takes place in the $e_i + h_i$ matches. Next, a fraction δ of matches is destroyed exogenously and s_i workers employed in sector i choose to quit into unemployment. Finally, at the end of the period, vacancies $\mathbf{v}'$ for next period are created.

We begin from studying an environment where heterogeneity across sectors is driven by matching efficiency and costs of vacancy creation. In Section 2.4, we add sector-specific productivity fluctuations to the model.

2.1 Planner's solution

Recall that we are interested in characterizing how a planner would choose allocations under free mobility of workers across sectors (i.e., occupation, location, industry). The efficient allocation at any given date is the solution of the following planner's problem that we write in recursive form:

$$\begin{aligned}
V(\mathbf{e}, \mathbf{v}; \kappa, \phi, z, \delta) &= \max_{\{u_i, s_i, v_i'\}} \sum_{i=1}^{I} (ze_i + \gamma z h_i - \kappa_i v_i') + \beta \mathbb{E}\left[V(\mathbf{e}', \mathbf{v}'; \kappa', \phi', z', \delta')\right] \\
s.t. &: \\
1 - \sum_{i=1}^{I} e_i &\geq \sum_{i=1}^{I} u_i \qquad (1) \\
h_i &= \phi_i m(u_i, v_i) \qquad (2) \\
e_i' &= (1-\delta)(e_i + h_i) - s_i \qquad (3) \\
&\Gamma_{z,\delta}(z, \delta'; z, \delta),\ \Gamma_\kappa(\kappa'; \kappa, z', \delta'), \Gamma_\phi(\phi'; \phi, z', \delta') \qquad (4)
\end{aligned}$$

The first constraint (1) states that the planner has $1 - \sum_{i=1}^{I} e_i$ unemployed workers available to allocate across sectors. The distribution of unemployment is not a state variable since all the unemployed are the same from the planner's viewpoint.Equation (2) states that, once the allocation $\{u_i\}$ is chosen, the frictional matching process in each market yields $\phi_i m(u_i, v_i)$ new matches which add to the existing

e_i active matches. Production takes place with output $z\left(e_i+\gamma h_i\right)$ in each market. Equation (3) describes (exogenous and endogenous) separations and the determination of next period distribution of active matches $\{e_i'\}$. The last line (4) in the problem collects all the exogenous Markov processes the planner takes as given.

It is easy to see that this is a concave problem where first-order conditions are sufficient for optimality. The choice of how many unemployed workers u_i to allocate in the i market yields the first-order condition

$$\gamma z\phi_i m_u\left(\frac{v_i}{u_i}\right)+\beta\mathbb{E}\left[V_{e_i}\left(\mathbf{e}',\mathbf{v}';\kappa',\phi',z',\delta'\right)\right](1-\delta)\,\phi_i m_u\left(\frac{v_i}{u_i}\right)=\mu, \tag{5}$$

where μ is the multiplier on the constraint (1) and where we used the linear homogeneity of m. Condition (5) states that, at the optimum, the marginal value of an unemployed worker must be equalized across markets. This value includes the current period output γz conditional on being matched to a vacancy plus the discounted expected value of being in an active match next period conditional on the match not being destroyed. Using the Envelope condition, we obtain:

$$V_{e_i}\left(\mathbf{e},\mathbf{v};\kappa,\phi,z,\delta\right)=z-\mu+\beta(1-\delta)\mathbb{E}\left[V_{e_i}\left(\mathbf{e}',\mathbf{v}';\kappa',\phi',z',\delta'\right)\right]. \tag{6}$$

This equation implies that $V_{e_i}=\bar{V}_e$ independent of i. Using (6) into (5) and collecting terms, we arrive at

$$\phi_i m_u\left(\frac{v_i}{u_i}\right)=\frac{\mu}{\gamma z+\beta\left(1-\delta\right)\mathbb{E}\left[\bar{V}_e{}'\right]}$$

which implies that the left hand side (LHS) of this last equation is equalized across markets, or

$$\phi_1 m_u\left(\frac{v_1}{u_1}\right)=\ldots=\phi_i m_u\left(\frac{v_i}{u_i}\right)=\ldots=\phi_I m_u\left(\frac{v_I}{u_I}\right). \tag{7}$$

The higher the matching efficiency parameter, the more unemployed workers the planner wants searching in market i. This is our key optimality condition. It states that, no matter what the distribution of vacancies is and no matter what the level of unemployment is, unemployed workers should be distributed in a certain way across markets. Condition (7) is a "generalized JR optimality condition" for a dynamic stochastic economy with heterogeneity across sectors.

It is easy to verify that the planner's two other decisions –separations and vacancies– have no bearing on the optimality condition (7) because of the Envelope theorem.For completeness, we report them below. The decision of how many workers to move from sector i employment into unemployment is:

$$\mathbb{E}\left[V_{e_i}\left(\mathbf{e}',\mathbf{v}';\kappa',\phi',z',\delta'\right)\right]\begin{cases} >0 & \rightarrow s_i=0\\ =0 & \rightarrow s_i\in\left(0,(1-\delta)\left(e_i+h_i\right)\right)\\ <0 & \rightarrow s_i=(1-\delta)\left(e_i+h_i\right)\end{cases}$$

Which inequality holds in this condition is immaterial for (7). Finally, the vacancy creation decision is summarized by the first-order condition $\kappa_i=\mathbb{E}\left[V_{v_i'}\left(\mathbf{e}',\mathbf{v}';\kappa',\phi',z',\delta'\right)\right]$ which, also, does not affect (7).

2.2 Mismatch indexes with no heterogeneity in ϕ across markets

With no heterogeneity in ϕ, the strict concavity of m implies that the planner wants to equate the vacancy-unemployment ratio across labor markets. Recall that, at every date t, before the planner chooses its allocation, the aggregate number of vacancies v_t and unemployed u_t is given. Let the aggregate market tightness be θ_t. Optimality requires $u_{it}^* = (1/\theta_t)v_{it}$. The number of mismatched unemployed workers (i.e., unemployed searching in the wrong sector) is therefore

$$u_t^M = \frac{1}{2}\sum_{i=1}^{I}|u_{it} - u_{it}^*| = \frac{1}{2}\sum_{i=1}^{I}|\frac{u_{it}}{u_t} - 1/\theta_t\frac{v_{it}}{u_t}|u_t = \frac{1}{2}\sum_{i=1}^{I}|\frac{u_{it}}{u_t} - \frac{v_{it}}{v_t}|u_t$$

and, as a share of total unemployment at date t, is equal to

$$\mathcal{M}_t^u = \frac{u_t^M}{u_t} = \frac{1}{2}\sum_{i=1}^{I}|\frac{u_{it}}{u_t} - \frac{v_{it}}{v_t}|. \tag{8}$$

Note that $\mathcal{M}_t^u \in [0,1]$ and in this sense it is an index. As explained, $\mathcal{M}_t^u = 0$ when the shares of unemployment and vacancies are the same in every sector. When, instead, all unemployed workers are in markets with zero vacancies and all vacancies in markets with zero unemployed, $\mathcal{M}_t^u = 1$.

It is important to note that $\mathcal{M}_t^u$ does not measure the extent to which unemployment would be reduced if we could eliminate mismatch. Even if workers searched in the wrong sector, they would find jobs at some (slower) rate. Conversely, even if unemployed workers were moved according to the optimal allocation u_{it}^*, they would still face the frictions arising from the within-market matching functions. The $\mathcal{M}_t^u$ index is therefore an upper bound on the fraction of unemployment due to mismatch.

A more precise calculation demands computing how many additional hires would be generated by switching to the optimal allocation of unemployed workers across sectors. To make progress in addressing this issue, we must state an additional assumption, well supported by the data as we show below: the individual-market matching function $m(u_i, v_i)$ is Cobb-Douglas, i.e.,

$$h_{it} = \phi v_{it}^{\alpha} u_{it}^{1-\alpha}.$$

Summing across market, with some simple algebra, we get an expression for the aggregate numbers of hires:

$$h_t = \phi v_t^{\alpha} u_t^{1-\alpha} \cdot \left[\sum_{i=1}^{I}\left(\frac{v_{it}}{v_t}\right)^{\alpha}\left(\frac{u_{it}}{u_t}\right)^{1-\alpha}\right]. \tag{9}$$

The first term denotes the highest number of new hires that can be achieved under the optimal allocation where market tightness is equated (to the aggregate value) across sectors. Therefore, we can define an alternative mismatch index as:

$$\mathcal{M}_t^h = 1 - \sum_{i=1}^{I}\left(\frac{v_{it}}{v_t}\right)^{1-\alpha}\left(\frac{u_{it}}{u_t}\right)^{\alpha}. \tag{10}$$

The index $\mathcal{M}_t^h \in [0,1]$ measures precisely what fraction of hires is lost because of misallocation. To express it as a fraction of the observed hires, we would have to compute $\mathcal{M}_t^h / \left(1-\mathcal{M}_t^h\right)$. Since the aggregate matching function becomes

$$h_t = \left(1-\mathcal{M}_t^h\right) \cdot \phi \cdot v_t^{\alpha} u_t^{1-\alpha},$$

the index $\mathcal{M}_t^h$ captures the shift in the aggregate matching function due to a change in mismatch.

This second index allows us to compute the counterfactual frictional unemployment rate that would arise in absence of mismatch, i.e., when all unemployed workers search in the right sector. Let $f_t = h_t/u_t$ be the job finding rate and, with a slight abuse of notation, s_t be the separation rate. In steady state, at date $t=0$, actual unemployment is $u_0 = s_0/\left(f_0+s_0\right)$ and the unemployment rate purged of the mismatch component is $u_0^* = s_0/\left(f_0^*+s_0\right)$ where $f_0^* = f_0/\left(1-\mathcal{M}_0^h\right)$. Given an estimated sequence of $\left\{\mathcal{M}_t^h\right\}$, and given the observed sequence of separation and job finding rates $\{s_t, f_t\}$ one can therefore compute the counterfactual frictional unemployment rate $u_{t+1}^* = s_t + \left(1-s_t-f_t^*\right) u_t^*$.

2.3 Mismatch indexes with heterogeneous matching efficiencies

Suppose now that individual labor markets differ in their frictional parameter ϕ_i and assume Cobb-Douglas matching function within markets, i.e., $h_{it} = \phi_i v_{it}^{\alpha} u_{it}^{1-\alpha}$. From equation (7), rearranging the optimality condition dictating how to allocate unemployed workers between market 1 and market i, we arrive at:

$$\frac{v_{1t}}{u_{1t}^*} = \left(\frac{\phi_i}{\phi_1}\right)^{\frac{1}{\alpha}} \cdot \frac{v_{it}}{u_{it}^*}.$$

Summing across i's

$$\begin{aligned} \sum_{i=1}^{I} u_{it}^* &= u_t = \left(\frac{u_{1t}^*}{v_{1t}}\right) \cdot \sum_{i=1}^{I} \left(\frac{\phi_i}{\phi_1}\right)^{\frac{1}{\alpha}} v_{it} \\ &= \left(\frac{1}{\phi_1}\right)^{\frac{1}{\alpha}} \left(\frac{u_{1t}^*}{v_{1t}}\right) \cdot \sum_{i=1}^{I} \phi_i^{\frac{1}{\alpha}} v_{it}. \end{aligned}$$

Let $v_{\phi t} \equiv \sum_{i=1}^{I} \phi_i^{\frac{1}{\alpha}} v_{it}$. Then re-expressing the above relationship for a generic market i (instead of market 1) and rearranging yields

$$u_{it}^* = \phi_i^{\frac{1}{\alpha}} \cdot \left(\frac{v_{it}}{v_{\phi t}}\right) \cdot u_t. \tag{11}$$

Recall that

$$u_t^M = \frac{1}{2} \sum_{i=1}^{I} |u_{it} - u_{it}^*|.$$

Substituting the expression for u_{it}^* from (11) and rearranging we arrive at:

$$u_t^M = \frac{1}{2}\sum_{i=1}^{I} |\frac{u_{it}}{u_t} - \phi_i^{\frac{1}{\alpha}}\left(\frac{v_{it}}{v_{\phi t}}\right)|u_t$$

which, after some simple manipulations yields the mismatch index

$$\mathcal{M}_{\phi t}^u = \frac{1}{2}\sum_{i=1}^{I} |\frac{u_{it}}{u_t} - \left(\frac{\phi_i}{\Phi_t}\right)^{\frac{1}{\alpha}} \cdot \frac{v_{it}}{v_t}| \tag{12}$$

where

$$\Phi_t = \left[\sum_{i=1}^{I} \phi_i^{\frac{1}{\alpha}}\left(\frac{v_{it}}{v_t}\right)\right]^{\alpha} \tag{13}$$

is a CES aggregator of the market-level matching efficiencies weighted by their vacancy share. The index in (12) is similar to the index (8) derived for the homogeneous markets case, except for the adjustment term in brackets which equals 1 when there is no heterogeneity in ϕ_i. The interpretation of this index is exactly the same as before: it measures the fraction of unemployed who are misallocated relative to the first best.

Let's now turn to the other alternative index measuring shifts in the aggregate matching function. Assuming that the local matching function is Cobb-Douglas with heterogeneous matching efficiencies and summing across sectors, we get the following expression for the optimal aggregate number of hires

$$h_t^* = v_t^{\alpha} u_t^{1-\alpha}\left[\sum_{i=1}^{I} \phi_i\left(\frac{v_{it}}{v_t}\right)^{\alpha}\left(\frac{u_{it}^*}{u_t}\right)^{1-\alpha}\right]. \tag{14}$$

Substituting the optimality condition (11) in equation (14), we arrive at:

$$h_t^* = v_t^{\alpha} u_t^{1-\alpha}\Phi_t$$

where Φ_t is defined in equation (13).

Similarly, we can define the total number of observed new matches as

$$h_t = v_t^{\alpha} u_t^{1-\alpha}\left[\sum_{i=1}^{I} \phi_i\left(\frac{v_{it}}{v_t}\right)^{\alpha}\left(\frac{u_{it}}{u_t}\right)^{1-\alpha}\right]$$

and the counterpart of (10) in the heterogeneous markets case becomes

$$\mathcal{M}_{\phi t}^h = 1 - \sum_{i=1}^{I}\left(\frac{\phi_i}{\Phi_t}\right)\left(\frac{v_{it}}{v_t}\right)^{\alpha}\left(\frac{u_{it}}{u_t}\right)^{1-\alpha}. \tag{15}$$

Finally, one can rewrite the aggregate matching function as:

$$h_t = \left(1 - \mathcal{M}_{\phi t}^h\right) \cdot \Phi_t \cdot v_t^{\alpha} u_t^{1-\alpha} \tag{16}$$

which highlights that a shift in the aggregate matching function can have two separate sources: 1) a change in misallocation of idle labor across markets which affects $\mathcal{M}^h_{\phi t}$; 2) shifts in the demand for labor across sectors that, through a composition effect in the vacancy distribution, change the average value of matching efficiency Φ_t.

2.4 The economy with heterogeneous productivity

We now generalize our environment to allow for sectoral level stochastic productivity $\mathbf{y} = \{y_1, ..., y_I\}$. Let $\Gamma_y(y', y)$ be the law of motion of the sector-specific productivity shock and assume it is either degenerate along its diagonal (time-invariant heterogeneity) or a martingale.[7] The production function in each market is linear in labor and zy_i is output per worker in market i. We also let agents decide whether to participate in the labor force. Let $\ell = \sum_{i=1}^{I}(e_i + u_i) \leq 1$ be the aggregate labor force. The decision to participate takes place at the end of the period, just before vacancy creation. Finally, assume that unemployed workers suffer (linear) disutility b.[8]

The new planner problem can be written as:

$$V(u, \mathbf{e}, \mathbf{v}; \kappa, \phi, \mathbf{y}, z, \delta) = \max_{\{u_i, s_i, \ell', v_i'\}} \sum_{i=1}^{I} (zy_i e_i + \gamma z y_i h_i - \kappa_i v_i') - bu + \beta \mathbb{E}\left[V(u', \mathbf{e}', \mathbf{v}'; \kappa', \phi', \mathbf{y}', z', \delta')\right]$$

$$\begin{aligned}
s.t. &\quad : \\
u &\geq \sum_{i=1}^{I} u_i \\
h_i &= \phi_i m(u_i, v_i) \\
e_i' &= (1-\delta)(e_i + h_i) - s_i \\
u' &= \ell' - \sum_{i=1}^{I} e_i' \qquad (17) \\
&\Gamma_{z,\delta}(z', \delta'; z, \delta), \ \Gamma_\kappa(\kappa'; \kappa, z', \delta'), \Gamma_\phi(\phi'; \phi, z', \delta'), \Gamma_y(y', y)
\end{aligned}$$

The difference with the problem in Section 2.1 is that, because the size of the labor force is endogenous, we must keep track of the total number of unemployed workers besides the distribution of employment. The additional constraint (17) is precisely the law of motion of this additional state.

The choice of how many unemployed workers to allocate in market i yields the first-order condition

$$\gamma z y_i \phi_i m_u\left(\frac{v_i}{u_i}\right) + \beta \mathbb{E}\left[-V_u(u', \mathbf{e}', \mathbf{v}'; \kappa', \phi', \mathbf{y}', z', \delta') + V_{e_i}(\mathbf{e}', \mathbf{v}'; \kappa', \phi', \mathbf{y}', z', \delta')\right](1-\delta)\phi_i m_u\left(\frac{v_i}{u_i}\right) = \mu, \tag{18}$$

[7]Note also that Γ_y does not depend on z' or δ'.

[8]We normalize output of a nonparticipant to zero, and hence one can interpret b as the utility (or production) of a nonemployed net of the disutility of search.

Using the Envelope conditions, we obtain:

$$V_u(u,\mathbf{e},\mathbf{v};\kappa,\phi,\mathbf{y},z,\delta) = \mu - b \tag{19}$$

$$V_{e_i}(u,\mathbf{e},\mathbf{v};\kappa,\phi,\mathbf{y},z,\delta) = zy_i + \beta(1-\delta)\mathbb{E}\left[V_{e_i}(u',\mathbf{e}',\mathbf{v}';\kappa',\phi',\mathbf{y}',z',\delta')\right]. \tag{20}$$

According to the first condition, the marginal value of an unemployed worker equals to its shadow price for the planner net of the disutility b. Consider now the decision on the labor force size next period ℓ' which states that

$$\mathbb{E}\left[V_u(u',\mathbf{e}',\mathbf{v}';\kappa',\mathbf{y}',\phi',z',\delta')\right] = 0, \tag{21}$$

i.e. the marginal expected value of moving a nonparticipant into job search should be equal to its value as nonparticipant, which is normalized to zero.[9] Combining (21) with (19), we note that the planner will choose the size of the labor force so that the expected shadow value of an unemployed worker $\mathbb{E}[\mu']$ equals its disutility b.[10]

Using (21) into (18), the optimality condition for the allocation of unemployed workers across sectors becomes

$$\gamma z y_i \phi_i m_u\left(\frac{v_i}{u_i}\right) + \beta(1-\delta)\mathbb{E}\left[\bar{V}_e{}' y_i'\right]\phi_i m_u\left(\frac{v_i}{u_i}\right) = \mu. \tag{22}$$

Because of the assumptions made on $\Gamma_y(y',y)$, equation (22) becomes

$$y_i\phi_i m_u\left(\frac{v_i}{u_i}\right) = \frac{\mu}{\gamma z + \beta(1-\delta)\mathbb{E}\left[\bar{V}_e{}'\right]}$$

which implies that the left hand side of this last equation is equalized across markets, yielding the generalized JR condition:

$$y_1\phi_1 m_u\left(\frac{v_1}{u_1}\right) = \ldots = y_i\phi_i m_u\left(\frac{v_i}{u_i}\right) = \ldots = y_I\phi_I m_u\left(\frac{v_I}{u_I}\right). \tag{23}$$

As before, the separation decision and the vacancy creation decision do not alter condition (23).

2.4.1 Mismatch indexes with heterogeneous productivities

It is useful to define "overall market efficiency" as the product $x_i \equiv y_i\phi_i$ of productive and matching efficiency of sector i. The optimality condition dictating how to allocate unemployed workers between market 1 and market i is:

$$\frac{v_{1t}}{u_{1t}^*} = \left(\frac{x_i}{x_1}\right)^{\frac{1}{\alpha}} \cdot \frac{v_{it}}{u_{it}^*} \tag{24}$$

[9] We are assuming an interior solution, i.e. we implicitly assume the population is large enough to move workers in and out of the labor force to achieve equalization between $\mathbb{E}(\mu')$ and b.

[10] It is clear that our result is robust to allowing b to be stochastic and correlated with (z,δ).

and the new mismatch index becomes

$$\mathcal{M}^u_{xt} = \frac{1}{2}\sum_{i=1}^{I} \left| \frac{u_{it}}{u_t} - \left(\frac{x_i}{X_t}\right)^{\frac{1}{\alpha}} \cdot \frac{v_{it}}{v_t} \right| \tag{25}$$

where

$$X_t = \left[\sum_{i=1}^{I} x_i^{\frac{1}{\alpha}} \left(\frac{v_{it}}{v_t}\right)\right]^{\alpha} \tag{26}$$

is a CES aggregator of the market-level overall efficiencies weighted by their vacancy share.

Turning to our alternative index, define the highest number of hires that can be obtained by optimally allocating the available unemployed workers as

$$h^*_t = v_t^{\alpha} u_t^{1-\alpha} \left[\sum_{i=1}^{I} \phi_i \left(\frac{v_{it}}{v_t}\right)^{\alpha} \left(\frac{u^*_{it}}{u_t}\right)^{1-\alpha}\right]. \tag{27}$$

Substituting the optimality condition (24) in equation (27), we arrive at:

$$y^*_t = v_t^{\alpha} u_t^{1-\alpha} \Phi^y_t$$

where

$$\Phi^y_t = \frac{\sum_{i=1}^{I} \left(\phi_i^{\frac{1}{\alpha}} \cdot y_i^{\frac{1-\alpha}{\alpha}}\right) \left(\frac{v_{it}}{v_t}\right)}{\left[\sum_{i=1}^{I} (\phi_i y_i)^{\frac{1}{\alpha}} \left(\frac{v_{it}}{v_t}\right)\right]^{1-\alpha}}.$$

Since total new hires are

$$h_t = v_t^{\alpha} u_t^{1-\alpha} \left[\sum_{i=1}^{I} \phi_i \left(\frac{v_{it}}{v_t}\right)^{\alpha} \left(\frac{u_{it}}{u_t}\right)^{1-\alpha}\right],$$

we obtain the counterpart of (15)

$$\mathcal{M}^h_{xt} = 1 - \sum_{i=1}^{I} \left(\frac{\phi_i}{\Phi^y_t}\right) \left(\frac{v_{it}}{v_t}\right)^{\alpha} \left(\frac{u_{it}}{u_t}\right)^{1-\alpha} \tag{28}$$

which measures the fraction of output from new hires lost because of mismatch unemployment at date t. As explained in Section 2.2, this index can be used to construct a counterfactual rate of frictional unemployment in absence of mismatch.

In the paper, we also use the notation $\mathcal{M}^u_{yt}$ and $\mathcal{M}^h_{yt}$ to denote indexes for an economy where there is productivity heterogeneity but all markets have the same matching efficiency ϕ.

2.5 Taking stock

Our mismatch indexes measure the fraction of unemployed workers misallocated (and the consequent fraction of hires lost) relative to a first best benchmark. This allows us to quantify the portion of observed unemployment due to mismatch in the labor market between idle labor and vacant jobs in two ways. The first way (through $\mathcal{M}_t^u$) is an upper bound but, in its simplest form, it does not require any functional form assumption on the sectoral matching function. The second (through $\mathcal{M}_t^h$) yields a more precise estimate but hinges on the unit elasticity specification for the sectoral matching function.

Our methodology is based on comparing the actual (u_i, v_i) distribution to the optimal distribution in an environment with costless labor mobility. In this sense, as noted in the Introduction, it should be viewed as a measurement device that delivers an "ideal" notion of total mismatch, as misallocation relative to an optimal unemployment distribution in the absence of *any* frictions across markets. In particular, we have abstracted from any moving or retraining costs that an unemployed worker may incur when she searches in a different sector than her original one, as well as any other sources of frictions across markets.

A partial list of these other sources of possible frictions includes incomplete insurance, imperfect information, wage rigidities, or government policy distortions. Under incomplete insurance, workers may choose not to switch occupation because of the temporary earnings loss associated to their productivity in the new occupation ($\gamma < 1$ in our set up).[11] In the presence of imperfect information, workers may be reluctant to move because they do not know where the vacant jobs are or what their prospects are in the new location, new occupation or new industry. With wage rigidities, workers may choose not to move because wages remain relatively high (low) in the declining (expanding) sectors. Finally, generous unemployment benefits or distortions in housing policies may prevent mobility, or sector specific taxes/transfers may be a source of misallocation.

We leave to future work a model of mismatch that analyzes the contributions of these various costs and distortions to equilibrium mismatch. Such a model would allows us to determine which component of mismatch, if any, can be considered as "efficient", and to study the effects of various shocks and policies on mismatch. A partial step in the direction of decomposing mismatch into components attributable to various causes could be to pursue a variation of our approach, which would let the planner explicitly take into account moving or retraining costs, while doing away with other possible frictions that may limit mobility across markets. We have not yet solved for the optimal allocation under this alternative scenario. It is reasonable to conjecture that we would obtain a lower level of mismatch (and hence our current calculations should be interpreted as *upper bounds for the level of mismatch unemployment*), but one would not expect that the dynamics of the index over time

[11] In our model of Section 2.3, earnings are equally uncertain in all occupations, but as explained we can allow for different dispersion in different occupations. With incomplete markets, mobility towards occupations with higher (lower) uncertainty will be lower (higher) than in the first best, ceteris paribus.

would be much affected.

In sum, our mismatch measures are based on a comparison between the planner's optimal allocation of unemployment and vacancies, where all potential sources of cross-market frictions do not play a role, and the actual allocation. As such, they can be interpreted as measures of "total" mismatch. An additional advantage of our current approach is that it allows us to construct some simple and informative mismatch indexes which can be easily computed from repeated cross-sectional data on the distribution of unemployment and vacancies across labor markets, without the need to solve complex dynamic equilibrium models featuring all the forces discussed above.

3 Mismatch in the U.K. labor market

We begin this section by describing the data. Next we analyze the issue of specification of the matching function at the sectoral level. Finally, we present the analysis of mismatch in the U.K. labor market.

3.1 Data Description

Our analysis requires detailed information on vacancies and unemployment. In particular, for each labor market we consider, we need monthly vacancy and unemployment statistics. We make use of the administrative data collected by local employment agencies that are available through Nomis.[12] The vacancy stocks and flows come from Jobcentre Plus Vacancy Statistics and the unemployment counts are from Jobseeker's Allowance Claimant Counts.[13] Both the vacancy and unemployment stocks and flows are available starting in 2005 on a monthly basis. The administrative data have the advantage of being available at a regular basis and at a disaggregated level which is ideal for analysis of mismatch. The only drawback of the data is its coverage. Not all vacancies are reported to the Jobcentres and not all unemployed qualify or choose to collect jobseekers' allowance. Thus employers and workers who do not use Jobcentres as one of their search channels are not captured by the administrative data. In particular, low-skill occupations are likely to be over-represented in the vacancy data.

Both unemployment and vacancy counts are available for 2-, 3-, 4-digit occupation codes and for different TTWAs (travel-to-work areas).[14] Throughout the analysis for the U.K. we focus on the

[12]https://www.nomisweb.co.uk/Default.asp

[13]Pissarides (1986), Layard and Nickell (1986), Jackman and Roper (1987) all used published vacancy statistics notified to the Employment Service run by the Department of Employment for their analysis of mismatch for 1960s and 1970s. The vacancy data used in these studies can be thought of as the predecessor of the Jobcentre vacancy data. More recently, Coles and Smith (1996) and Burgess and Profit (2001) both used the Jobcentre data to estimate matching functions for TTWAs for the UK between 1985-1995.

[14]TTWAs are defined by the Office for National Statistics as zones that are labor market areas. The fundamental criterion is that, of the resident economically active population, at least 75% actually work in the area, and that, of everyone working in the area, at least 75% live in the area. 243 TTWAs were defined in 2007 by using 2001 Census data.

following definitions of labor markets: 1) 2-digit level occupations; 2) 3-digit level occupations; 3) Travel To Work Areas (henceforth, TTWA's); 4) 2-digit level occupations and TTWA's. The first two definitions will enable us to study occupational mismatch; the third refers to geographic mismatch, and the last one defines a local labor market as a specific occupation in a given location.

We use *unfilled live vacancies* as our measure of stock of vacancies and the total claimant count as our measure of stock of unemployment. Both of these stocks are reported at the end of each month. As for the number of total matches formed in a given month we use *vacancy outflows* corresponding to the number of outflow of vacancies during a given month.[15]

We start our analysis from July 2006. This choice is motivated by a change in Jobcentre Plus's vacancy handling procedure which was introduced in May 2006. In particular, prior to May 2006, vacancies notified to Jobcentre Plus were followed up with the employer to ascertain whether (a) they should remain available to jobseekers, or (b) they should be closed or had been filled by clients referred by Jobcentre Plus. Starting from May 2006 vacancies notified to Jobcentre Plus have a fixed closure date. Vacancies are automatically withdrawn on the closure date unless the employer advises that a later closure date is required. Due to this change, there is a sharp decline in the number of live unfilled vacancies in May 2006.

Starting in March 2007, ONS added UK armed forces vacancies into the data under "Protective Service Occupations" (SOC = 33) and "Protective Service Officers" (SOC=117). This caused approximately a ten-fold increase in the number of vacancies in these occupations. Also, all the UK armed forces vacancies were allocated to the "Lincoln" Travel to Work Area. To resolve this issue, we have excluded these occupations and geographical area from our analysis. Lastly, there was an irreconcilable spike in vacancy outflows for "administrative occupations: government and related organizations" (SOC=411) in May 2009. We impute the May 2009 value by taking the average of April 2009 and June 2009. The aggregate 2-digit occupation code (SOC=41) was also imputed for May 2009 in the same way.

Figure 1 (left panel) shows the total number claimants together with unemployment measured by the Labor Force Survey. As expected, survey-based unemployment is higher than claimant count unemployment since not all unemployed workers collect Job Seekers Allowance. The level of claimant count unemployment is about two thirds of labor force unemployment. However, the two measures are highly correlated with a correlation of 0.98. In the right panel, we plot the Jobcentre Plus's vacancy measure against the Office of National Statistics' (ONS) economy-wide survey-based vacancy measure. Similar to unemployment, Jobcentre vacancy measure lies below the ONS measure. However, the two series are again highly correlated with a correlation coefficient of 0.90.

[15]Another possibility is to use claimant off-flows who reported finding work as a measure of matches. However, the percentage of off-flows with a "not known" or "failed to sign" destination has increased since the start of the series (representing 44% of total UK off-flows in July 2009) complicating the interpretation of the decline in unemployment outflows. This is because the completion levels of the forms filled in by Job Seekers's Allowance leavers have decreased.

The wage data that we use to calculate mismatch for low- and high-wage occupations come from the Annual Survey of Hours and Earnings (ASHE) which provides information about the levels, distribution and make-up of earnings and hours paid for employees within industries, occupations and regions.

3.1.1 Measurement error in vacancies

As discussed above, one of the challenges we face is that the Jobcentre Plus vacancy data may over-represent low-skilled occupations. Suppose that true vacancies (V_{it}) in market i are a factor μ_i^v of the observed vacancies (v_{it}), i.e., $V_{it} = v_{it}\mu_i^v$. Similarly, since hires are measured as vacancy outflows, $H_{it} = h_{it}\mu_i^v$, where H_{it} are the true hires. Since this problem appears to be less severe for claimant count data, we assume that measurement error in unemployment is constant across markets, i.e., $U_{it} = u_{it}\mu^u$. In other words, there is incomplete coverage in unemployment data, but no systematic differences across occupations or geographical areas.

For simplicity, consider the economy of Section 2.2 without productivity or matching efficiency heterogeneity. Recall that the mismatch index is

$$\begin{aligned} \mathcal{M}_t^u &= \frac{1}{2}\sum_{i=1}^{I} \left|\frac{U_{it}}{U_t} - \frac{V_{it}}{V_t}\right| \\ &= \frac{1}{2}\sum_{i=1}^{I} \left|\frac{u_{it}}{u_t} - \frac{v_{it}\mu_i^v}{\sum_{i=1}^{I} v_{it}\mu_i^v}\right| \end{aligned}$$

where the second line expresses the index in terms of observable variables. Rearranging, we obtain

$$\mathcal{M}_{\mu t}^u = \frac{1}{2}\sum_{i=1}^{I} \left|\frac{u_{it}}{u_t} - \frac{\mu_i^v}{\sum_{i=1}^{I}\left(\frac{v_{it}}{v_t}\right)\mu_i^v} \cdot \frac{v_{it}}{v_t}\right|, \tag{29}$$

which is the mismatch index corrected for measurement error in vacancies.

Is it possible to identify measurement error in vacancies μ_i^v in each sector? With Cobb-Douglas specification, the true sectoral matching function is $H_{it} = \phi_t V_{it}^{\alpha} U_{it}^{1-\alpha}$. Substituting observed variables measured with error in place of true ones, we arrive at

$$\begin{aligned} h_{it}\mu_i^v &= \phi_t \left(v_{it}\mu_i^v\right)^{\alpha} \left(u_{it}\mu^u\right)^{1-\alpha} \\ h_{it} &= \phi_t \cdot \left(\mu^u\right)^{1-\alpha} \cdot \left(\frac{1}{\mu_i^v}\right)^{1-\alpha} v_{it}^{\alpha} u_{it}^{1-\alpha} \end{aligned}$$

Therefore, in a panel regression of log hires on log vacancies and log unemployment augmented with time dummies and fixed sector-specific effect, the estimated sector fixed effect is $(\mu_i^v)^{\alpha-1}$. Given an estimate of α, one can therefore obtain an estimate of μ_i^v. For example, sectors where vacancies are especially underreported (i.e., $\mu_i^v >> 1$) will look like sectors with lower matching efficiency because hires are reduced proportionately but vacancies are reduced only at rate $\alpha < 1$. The estimates of μ_i^v can be used to correct the mismatch index as shown in equation (29). [16] In the next version of the paper, we will report some calculations based on this adjustment.

3.2 Matching function specification

We start by showing that a matching function with unit elasticity is a very reasonable representation of the hiring process at the sectoral level. For the 2-digit occupation definition of sectors (24 occupations) and the period July 2006-August 2010, we estimate the parameters of the following CES matching function via minimum distance:[17]

$$\ln\left(\frac{h_{it}}{u_{it}}\right) = \ln\phi_i + \frac{1}{\sigma}\ln\left[\alpha\left(\frac{v_{it}}{u_{it}}\right)^{\sigma} + (1-\alpha)\right]. \tag{30}$$

Recall that $\sigma \in (-\infty, 1)$ with $\sigma = 0$ being the Cobb-Douglas case.[18] We find that $\hat{\sigma} = 0.35$ $(S.E. = 0.06)$ implying an elasticity around 1.54, hence only slightly larger than the Cobb-Douglas benchmark. Figure 2 plots the iso-matching curves for the CES and the Cobb-Douglas specification over the empirical range of vacancies and unemployment, demonstrating the closeness of the two specifications. In light of this finding, and given the analytical convenience of the unit elasticity benchmark, we restrict σ to be zero.

In Table 1 we report the estimation results of panel regressions for a Cobb-Douglas matching function of the form

$$\ln\left(\frac{h_{it}}{u_{it}}\right) = \ln\phi_i + \alpha\ln\left(\frac{v_{it}}{u_t}\right), \tag{31}$$

where we fix the vacancy share α to be constant across markets and over time. We run a separate regression for each definition of sector and we report results both for the model where ϕ_i is allowed to vary across sectors and for the model where it is restricted to be the same. The estimates for α, the elasticity of hires with respect to vacancies, range from 0.57 to 0.80 depending on the local labor market definition and the restriction on ϕ. We also run an aggregate level regression whose estimate of α is well within that range. To maintain comparability across the U.S. and the U.K. results, we choose the value $\alpha = 0.67$ which appears to be a plausible elasticity for U.S. data as well.

Figure 3 illustrates the estimated heterogeneity of frictional parameters ϕ_i across markets. Recall that ϕ_i measures the fundamental matching efficiency at the sectoral level. Higher matching efficiency

[16] This derivation is under the assumption of no heterogeneity in productivity or matching efficiency across markets. If there is such heterogeneity, then the estimates of μ_i^v will pick up partly measurement error, partly heterogeneity in these efficiency parameters.

[17] Note that to be consistent with the timing of the measurement of flows and stocks, we use the unemployment and vacancy stocks at the beginning of the month (which are given by the stocks in month t-1) and the vacancy flows during the month (which are given by flows in month t) in all regressions throughout the paper.

[18] The estimation is performed by simulated annealing to ensure what we obtain is a local minimum. Results are very robust to the weighting matrix used.

may reflect a variety of factors from matching being intrinsically easier in certain jobs because skill requirements are easier to satisfy to differential use of informal hiring methods. See Davis, Faberman and Haltiwanger (2010), for a discussion on the sources of heterogeneity in vacancy yields.

Overall, we do not uncover a large heterogeneity in ϕ_i. Secretarial (administrative), customer service, and public sector occupations have the largest ϕ_i (lowest matching friction), while arts, leisure (sports), agricultural, and science and technology professional occupations are those with the smallest ϕ_i. One interpretation of these differences is that general skill labor markets have the highest ϕ_i and specialized skill labor markets the lowest ϕ_i. Recall that the distribution of ϕ_i is a key input of our mismatch indexes $\mathcal{M}_{\phi}^{u}$ and $\mathcal{M}_{\phi}^{h}$.

3.3 Mismatch indexes

In Figure 4, we report the evolution over time of the $\mathcal{M}_t^u$ and $\mathcal{M}_{\phi t}^u$ indexes for our four labor market definitions. Several observations are in order. First, in the pre-recession period, $\mathcal{M}_t^u$ varies between 0.2 and 0.4 depending on the labor market definition. The mismatch index is higher at the finer level of occupational disaggregation, and is lowest for our geographic definition of labor markets. Recall that $\mathcal{M}_t^u$ ought to be interpreted as the fraction of total unemployment that is misallocated relative to the first best. When we correct our mismatch indexes for heterogeneity in matching efficiency, the implied $\mathcal{M}_{\phi t}^u$ indexes are 10-20 percent higher, on average, suggesting that the fraction of total unemployment due to mismatch is larger once one appropriately takes into account the heterogeneity of matching functions across labor markets.

Second, all our measures of mismatch by occupation –both $\mathcal{M}_t^u$ and $\mathcal{M}_{\phi t}^u$– exhibit a marked rise in mismatch during the recession. By contrast, geographic mismatch (by TTWA) does not exhibit a strongly cyclical behavior and, if anything, it declines slightly over the time period under consideration.[19] We therefore conclude that occupation is the only serious candidate as a source of mismatch in the U.K. labor market.

Third, most of the observed increase in mismatch during the latest recession is relatively short lived. For instance, at the 3-digit occupation level, $\mathcal{M}_t^u$ varies between roughly 0.35 and 0.4 between summer 2006 and the Fall of 2008, rises to a peak of about 0.5 during the course of 2009, and falls back to about 0.4 from the Fall of 2009 onwards. A similar pattern occurs at the 2-digit occupation level. For our finest definition of mismatch (2-digit occupations by TTWAs), the rise in the mismatch index lasts longer. We note, though, that after the end of the recession the mismatch indexes have plateaued at levels slightly higher than their pre-recession values. This pattern is qualitatively similar for our ϕ-corrected index $\mathcal{M}_{\phi t}^u$.

Finally, for the two-digit occupation classification, we also compute $\mathcal{M}_{yt}^u$ with sector specific pro-

[19] In particular, the secular decline of geographical mismatch is largely attributable to the downward trend in the differential between unemployment and vacancy shares in the London TTWA.

ductivity measured through hourly wages. The level of the index is slightly higher, and its dynamics are similar to the other two indexes $\mathcal{M}_t^u$ and $\mathcal{M}_{\phi t}^u$, but $\mathcal{M}_{yt}^u$ displays a more pronounced and persistent increase.

To better understand these movements in the $\mathcal{M}_t^u$ mismatch index, we report in Figure 5 the eight 2-digit occupations with the largest average $(u_{it}/u_t - v_{it}/v_t)$ over the period.[20] This is the difference between the share of unemployed and the share of vacancies in a given occupation, and describes the contribution of that occupation to the mismatch index. The occupations that contributed the most to the observed rise in mismatch during the recession are "Elementary trades, plant and storage", "Skilled construction and building trades", "Transport and mobile machine drivers" and, with the opposite sign, "Caring and Personal Services".

Figure 6 reports our alternative mismatch index $\mathcal{M}_t^h$ representing the fraction of total potential hires that are lost because of mismatch. The qualitative pattern is similar to that for $\mathcal{M}_t^u$, with a strong cyclical component and a slight long-run rise in mismatch during the course of 2009 for the occupation-based definitions of labor markets, particularly so for the 3-digit occupations. Geography, once again, plays no role.

Based on equation (10), one can use the $\mathcal{M}_t^h$ index to interpret the shifts of the Beveridge curve. The left panel of Figure 7 plots and $\ln\left(1 - \mathcal{M}_t^h\right)$ and $\ln \chi_t \equiv \ln\left(\frac{h_t}{u_t}\right) - \alpha \ln\left(\frac{v_t}{u_t}\right)$ computed from aggregate data and smoothed using a 12-month moving average (12MMA). This latter term measures the shift in the aggregate job finding rate unexplained by changes in aggregate market tightness. The right panel plots the results of the same exercise for the specification of the matching function that allows for heterogeneity in ϕ_i. While mismatch does not explain the rapid improvement in efficiency in the first half of the sample, it is able to explain about half of the decline in the aggregate job finding rate.

In Figure 8, we summarize our findings. The left panel plots the observed LFS unemployment rate (u_t) and the fraction due to mismatch according to our $\mathcal{M}_t^u$ index, based on 3-digit occupations sectors: aggregate unemployment rate increased from 5.3% to 8% from July 2006 to August 2010 and, over the same period, mismatch unemployment rate (the fraction of the labor force searching for jobs in the wrong sector) rose from 2% to almost 4% and then fell to 3% at the end of the sample period. Because of the upper bound nature of $\mathcal{M}_t^u$, according to this calculation, at most half of the rise in U.K. unemployment during the recession can be attributed to more severe mismatch. The right panel of Figure 8 translates the change in $\mathcal{M}^h$ into lost hires. In particular, it shows how many more hires would have been generated if all workers were in the right sector.

[20] These eight sector account for 80 percent of the level of the index $\mathcal{M}_t^u$ and the bulk of its dynamics

3.3.1 Low wage vs high wage sectors

In our benchmark, we took the view that planner can freely move unemployed workers across all sectors. At the other end of the spectrum, one could assume that mobility is costless only between sectors of similar skill levels, but it is infinitely costly between skill levels. Then, the economy would feature segregated labor markets and a different planner problem would apply to each skill level.

As a first step, we explore this idea by studying mismatch separately for high- vs. low-productivity occupations, using wages as a proxy for productivity. We compute median and mean hourly and weekly gross wages for our 2-digit occupational categories over our sample period. We then divide the twenty four 2-digit occupations into high and low-wage occupations using the median across these occupations as a threshold.

Figure 9 plots $\mathcal{M}_t^u$ separately for these two groups. For the high wage group, we find a more substantial increase and, interestingly, a more persistent one. While mismatch for the low wage occupations goes back to its pre recession level, in the high wage ones, it is still almost twice as large.[21] Figure 10 and 11 report the contributions of specific occupations to mismatch for each wage group, plotting $(u_{it}/u_t - v_{it}/v_t)$ over time. In the high wage group, "Skilled Construction and Building Trades" and "Health and Social Welfare Associate Professionals" are causing the spike during the recession, but for opposite reasons. In the low wage group, the "Caring Personal Service" and "Elementary Administrative and Services" occupations are driving the temporary spike, but they quickly return back to their pre-recession levels of unemployment-vacancy share differential.[22]

4 Mismatch in the U.S. Labor Market

The unemployment rate has been persistently high in the U.S. at around 9.6% through most of 2010 despite the ongoing economic recovery. Concurrently, the U.S. Beveridge curve (i.e., the empirical relation between aggregate unemployment and aggregate vacancies) has displayed a marked rightward movement implying a decline in match efficiency. One possible reason for a persistent reduction in match efficiency is a mismatch between the skills and the skill requirements of job openings. It is of first order importance to understand to which extent labor market mismatch is affecting the speed of recovery in the labor market. Developing measures similar to the ones we computed for the U.K. labor market would be extremely useful in addressing the concerns about the persistence of unemployment. Yet, as we have discussed in Section 3, the data requirements are quite demanding for constructing

[21] In the next version of the paper, we will report a calculation of the fraction of the rise in unemployed due to mismatch for both skill groups. It is likely that we'll conclude that mismatch is extremely important for high-skilled workers unemployment dynamics.

[22] "Health and Social Welfare Associate Professionals" are nurses, doctors, therapists and social welfare workers. "Caring Personal Service" are assistant nurses, dental nurses, orderlies, ambulance drivers (excluding paramedics), child care and animal care providers. The latter group is less skilled than the former, perhaps making it easier to fill vacancies.

mismatch measures. Administrative vacancy data similar to the Jobcentre Plus vacancies have never been available in the U.S. making it hard to assess the degree of mismatch in the labor market. The introduction of Job Openings and Labor Turnover Survey (JOLTS) starting in 2000 was an important step in that direction since JOLTS started providing survey-based measures of vacancies and hires at a monthly frequency for major industry groups.

In this section we use the JOLTS vacancy data and combine it with unemployment measures from the Current Population Survey (CPS) to calculate our measures of mismatch for the U.S. economy for January 2000 to July 2010 period. The JOLTS sample size allows us to disaggregate vacancies at the industry level to 14 groups.[23] Similarly we calculate unemployment counts from the CPS for the same industry classifications.[24]

As discussed in the previous section, the computation of the indexes with heterogeneity requires estimating market-specific match efficiency parameters. In particular, we estimate ϕ_i to vary across markets and estimate

$$\ln\left(\frac{m_{it}}{u_{it}}\right) = \ln\phi_i + \alpha\ln\left(\frac{v_{it}}{u_{it}}\right)$$

by using vacancies from the JOLTS and unemployment from the CPS. We use hires from the JOLTS as our measure of matches.[25] The estimate for α is 0.64.[26] Note that this estimate is consistent with our choice of $\alpha = 0.67$ throughout the paper. Our estimation also provides us industry-specific estimates of match efficiency (ϕ_i). These are reported in Table 1 in the Appendix. Industry-specific match efficiency estimates (ϕ_i) vary between 0.67 to 1.7. Among the industries, education, health, and information stand out as low-efficiency sectors while construction stands out as a high efficiency sector. High efficiency might be an outcome of different hiring practices in different industries as well as underreported vacancies as discussed in Davis, Faberman, and Haltiwanger (2010).

We now move on to computing our mismatch measures for the U.S. economy. We first start with computing $\mathcal{M}^u$ and $\mathcal{M}^u_\phi$ by using data for 14 industry groups. The time-series of these two indexes are shown on the left panel of Figure 14. $\mathcal{M}^u$ has risen from about 0.23 in 2005-2007 to about 0.33 in early 2009, and has since declined to about 0.3. $\mathcal{M}^u_\phi$ has been lower then $\mathcal{M}^u$, however the rise in this index is more pronounced. The bottom line is the fraction of unemployed workers that are misallocated increased during the recession and then started to come down.

What is the reason underlying this change in mismatch? Recall that $\mathcal{M}^u$ increases when the sum of $|u_i/u - v_i/v|$ increases. In other words when the difference between an industry's unemployment share and vacancy share increases. Figure 15 plots $(u_i/u - v_i/v)$ for 14 industries starting from 2001.

[23]We exclude government in our analysis and focus on the private sector.

[24]Note that industry affiliations are not available for all unemployed workers in the CPS. From 2000-2010, on average about 24% of unemployed do not have industry information. Some of these workers have never worked before and some are self-employed.

[25]An alternative is to use the unemployment outflow rate or the unemployment to employment transition rate. We do not pursue this approach here since JOLTS provides a direct measure of industry-specific hires.

[26]t-statistics is 60.00.

$(u_i/u - v_i/v)$ is normalized to zero for each industry to be able to isolate the contribution of each industry to the increase in mismatch. Industries that contributed to the increase in mismatch were construction, durable goods manufacturing, health, and education. For construction and durable goods manufacturing, the increase in mismatch came from an increase $(u_i/u-v_i/v)$ while for health and education mismatch increased because $(u_i/u-v_i/v)$ declined. Other industries are mostly concentrated close to zero line, contributing marginally to the increase in mismatch.

We next calculate $\mathcal{M}^h$ which has a different interpretation from $\mathcal{M}^u$. $\mathcal{M}^h$ measures the fraction of matches lost because of misallocation in the labor market. The right panel of Figure 14 shows that this measure averaged around 0.035 before the recession started, increased to 0.075 in 2009 and declined to 0.06 since then. The heterogenous index $\mathcal{M}^h_\phi$ implies a lower level of mismatch than $\mathcal{M}^h$, but the pattern is very similar.

Figure 16 summarizes our findings. The left panel plots the observed unemployment rate (u_t) and the fraction due to mismatch according to our $\mathcal{M}^u_t$ index. Unemployment rate increased from 4.7% to 10.1% from December 2007 to October 2009 and mismatch unemployment rate (the fraction of the labor force searching for jobs in the wrong sector) rose from 1.2% to almost 3.3%. Mismatch unemployment starting falling in 2010 declining to 2.8% by July 2010. The right panel of Figure 16 translates the change in $\mathcal{M}^h$ into lost hires. In particular, it shows how many more hires would have been generated if all workers were in the right sector. Our calculation shows that hires were increasingly lower due to the increase in mismatch throughout 2007-2010. The cumulative number of *lost hires* starting from the beginning of the recession is 2.6 million hires.

As we have discussed in Section 2, $\mathcal{M}^h$ captures the shift in the aggregate matching function due to a change in mismatch. In Figure 17, we plot $\ln(1-\mathcal{M}^h)$ and the shift in the matching function which is $\ln\left(\frac{H_t}{v_t}\right) - \alpha \ln\left(\frac{v_t}{u_t}\right)$. The figure shows that during the recession, there was a decline in match efficiency, causing the observed hires to be lower than implied by the matching function. The left panel of Figure 17 shows the role of mismatch in this shift of the matching function by using the index $\ln(1-\mathcal{M}^h)$ and the right panel repeats the same exercise for $\ln(1-\mathcal{M}^h_\phi)$. The figure shows that some of the decline in match efficiency could be attributed to the increase in mismatch. However, the magnitude of this component is not very large. At least with this measure of mismatch, mismatch is not the main reason underlying the decline in the observed shift in the matching function.

5 Conclusions

This paper collects work in progress where we are attempting to formalize and measure the notion of mismatch unemployment. This concept has recently become central to the macro policy debate. We have started from Jackman-Roper (1986) insight and we have obtained a “generalized J-R conditions” in a variety of dynamic stochastic economy for the optimal allocation of unemployed workers across sectors. These conditions are easily manipulated into indexes that permit to quantify how much of the recent rise in unemployment is associated to more severe mismatch. We have applied our analysis to the U.K. and the U.S. labor markets.

For the U.K. we find that mismatch has worsened across occupations but not geographical areas, and that at most it can account for 40% of the recession-driven rise in unemployment. Our findings indicate that imbalances between vacancies and unemployed workers may be much more important for skilled (high wage) occupations.

In the U.S. labor market, we find that mismatch at the industry level increased during the recession and started to come down in 2010 as the sectoral composition of the job-openings started to normalize. The increase in mismatch seemed to have contributed to the rise in the unemployment rate. However, mismatch itself can not explain the big decline in the match efficiency. It is important to emphasize that our mismatch measures only capture the misallocation of workers across 14 broad industrial sectors. It is possible that these mismatch measures do not capture a significant portion of mismatch if it occurs within these broad sectors. In the next draft of the paper, we will report results based on the recently acquired Help Wanted OnLine (HWOL) series collected by the Conference Board.

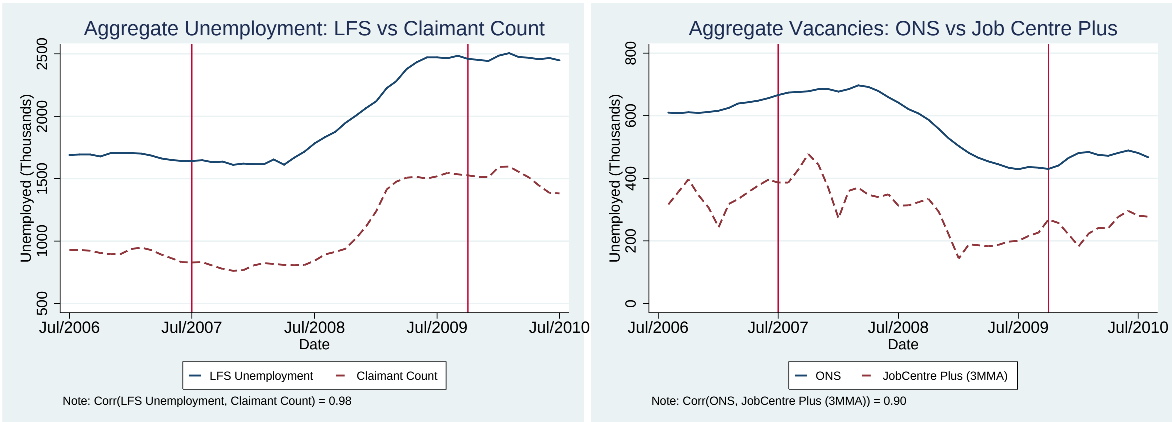

Figure 1: Unemployment and Vacancies in the U.K.

Table 1: α Across UK Markets 2006:07-2010:08

Labor Market	ϕ Fixed	ϕ Varying
Aggregate	0.72	
	(18.58)	
2 Digit Occupation	0.78	0.74
	(109.81)	(90.39)
3 Digit Occupation	0.80	0.72
	(162.87)	(112.92)
Travel to Work Area	0.74	0.68
	(199.38)	(154.77)
TWA x 2 Digit Occ.	0.67	0.57
	(471.26)	(305.95)

Note: t-statistics in parenthesis

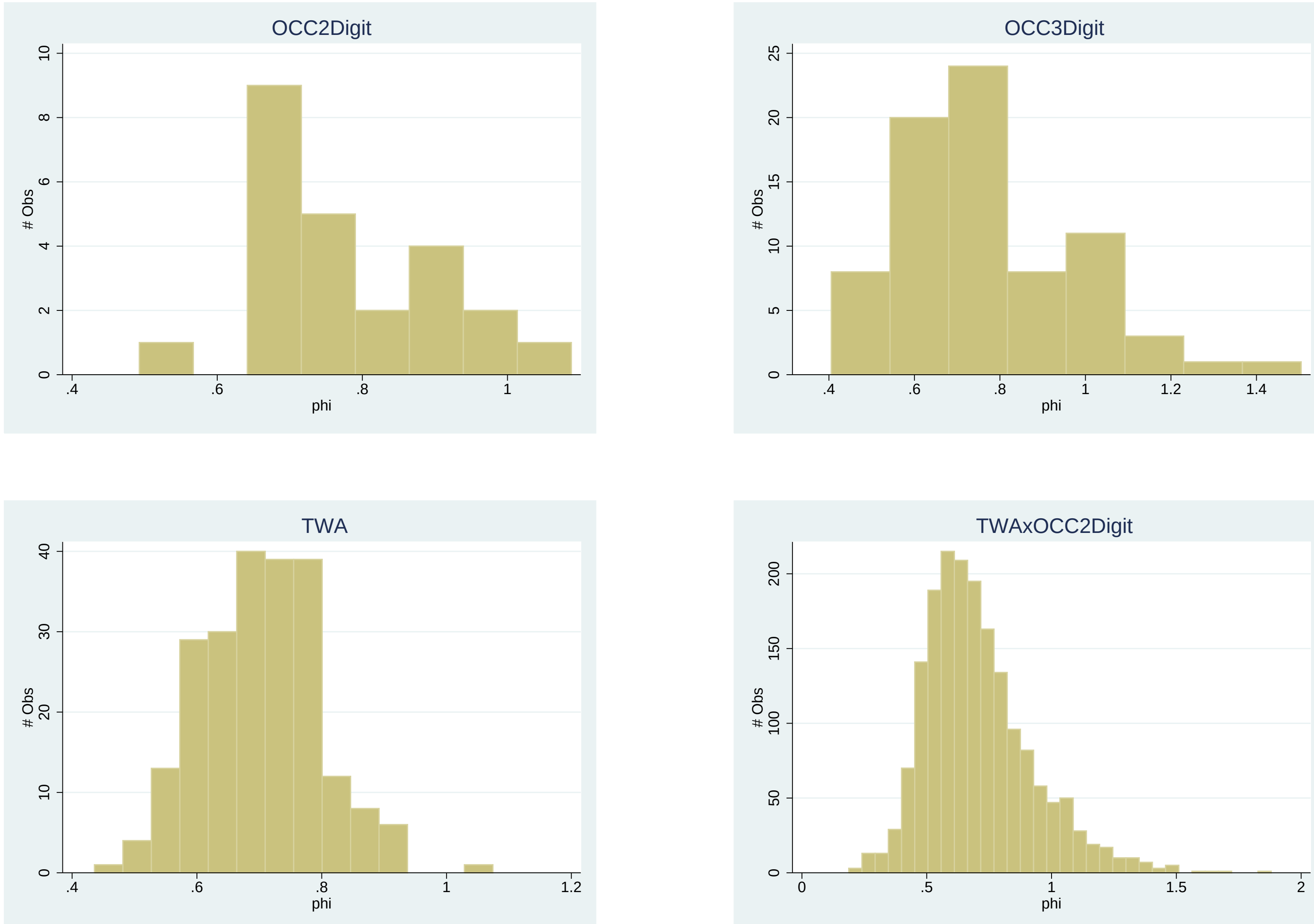

Figure 2: Heterogeneous ϕ Across Markets

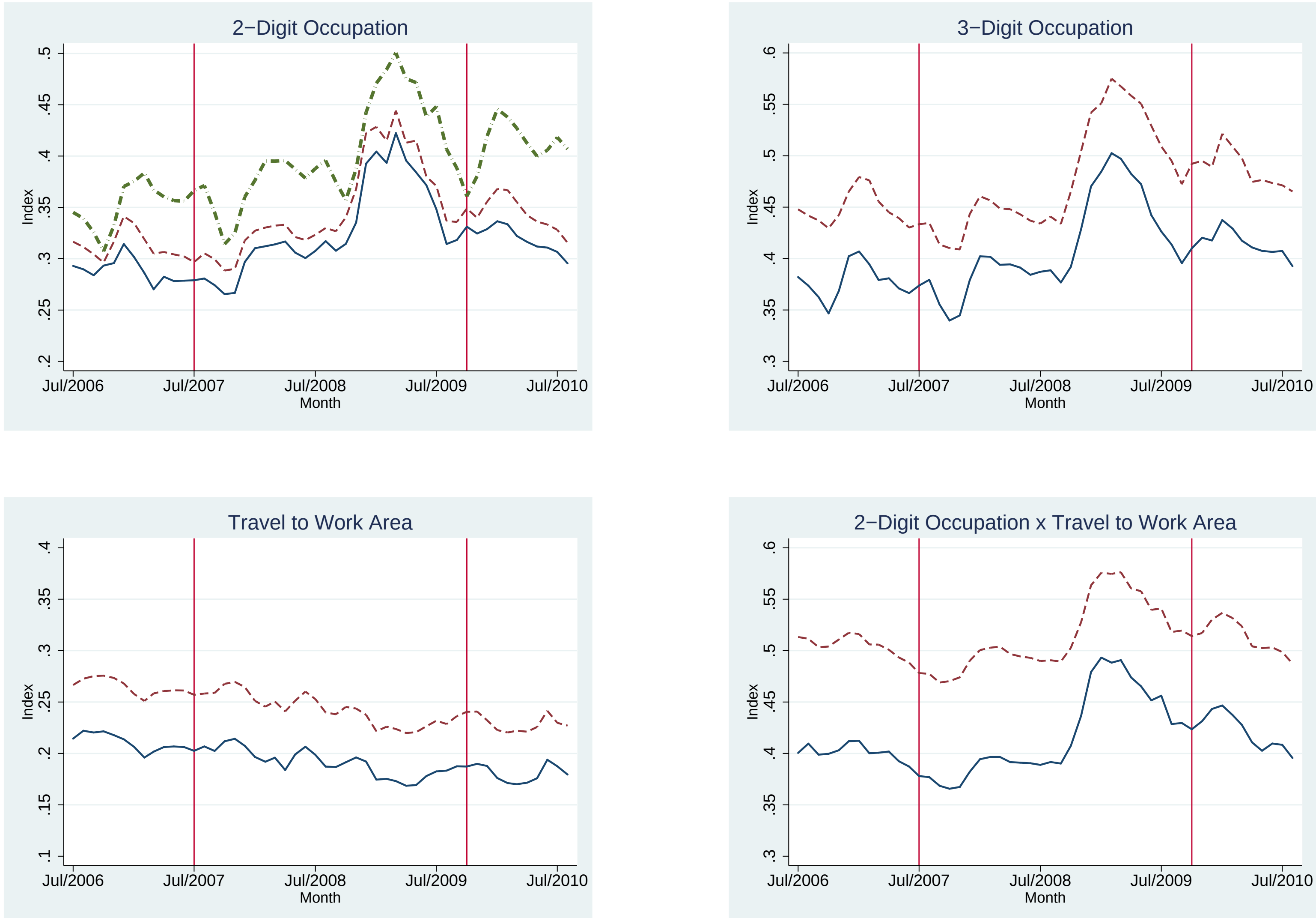

Figure 3: Mismatch $\mathcal{M}^u$ (Solid) $\mathcal{M}^u_\phi$ (Dashed) $\mathcal{M}^u_y$ (Dashed-Dot); UK Recession denoted by the vertical red lines.

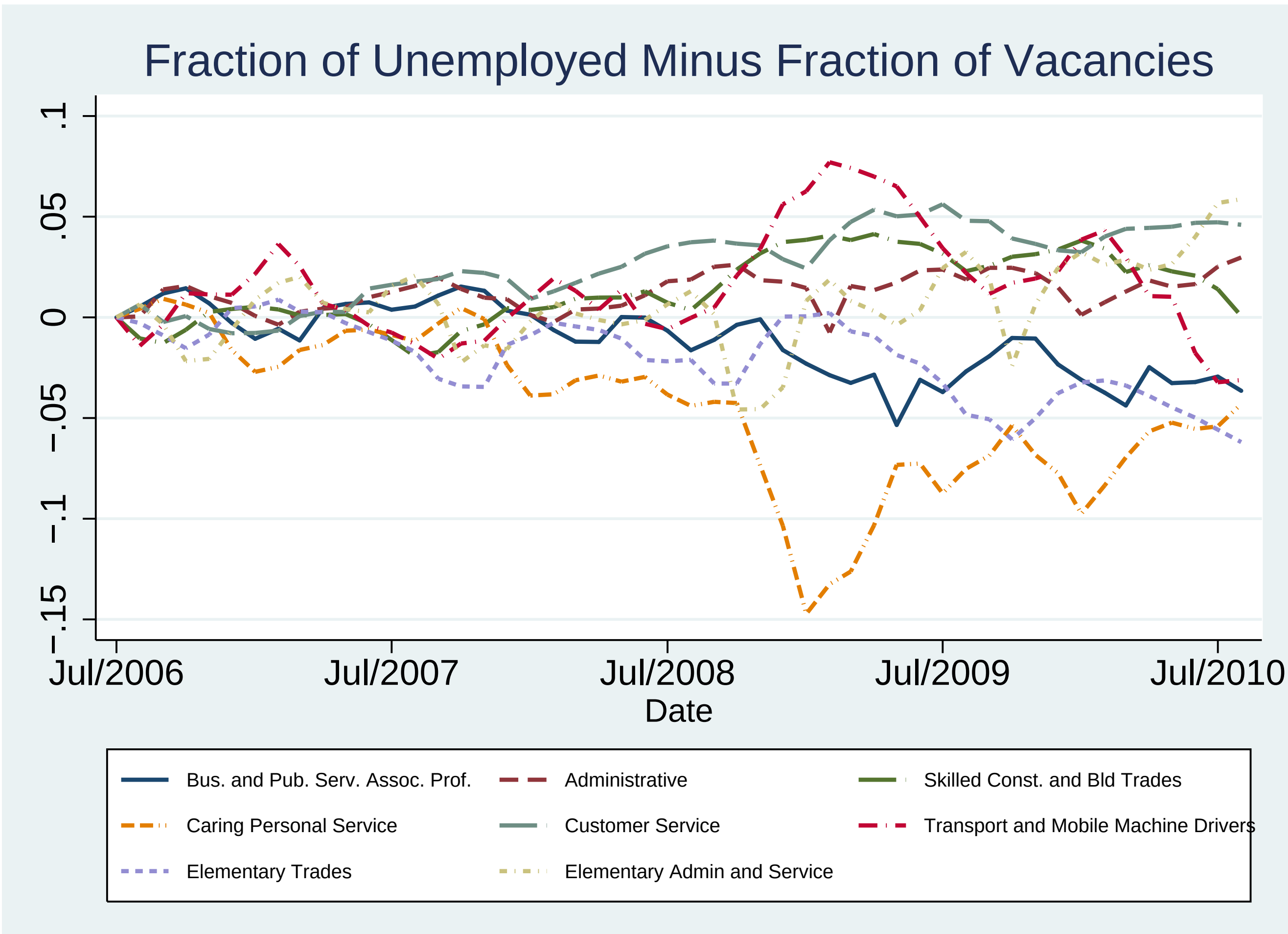

Figure 4: Fraction of Unemployed Minus Fraction of Vacancies (Top Eight)

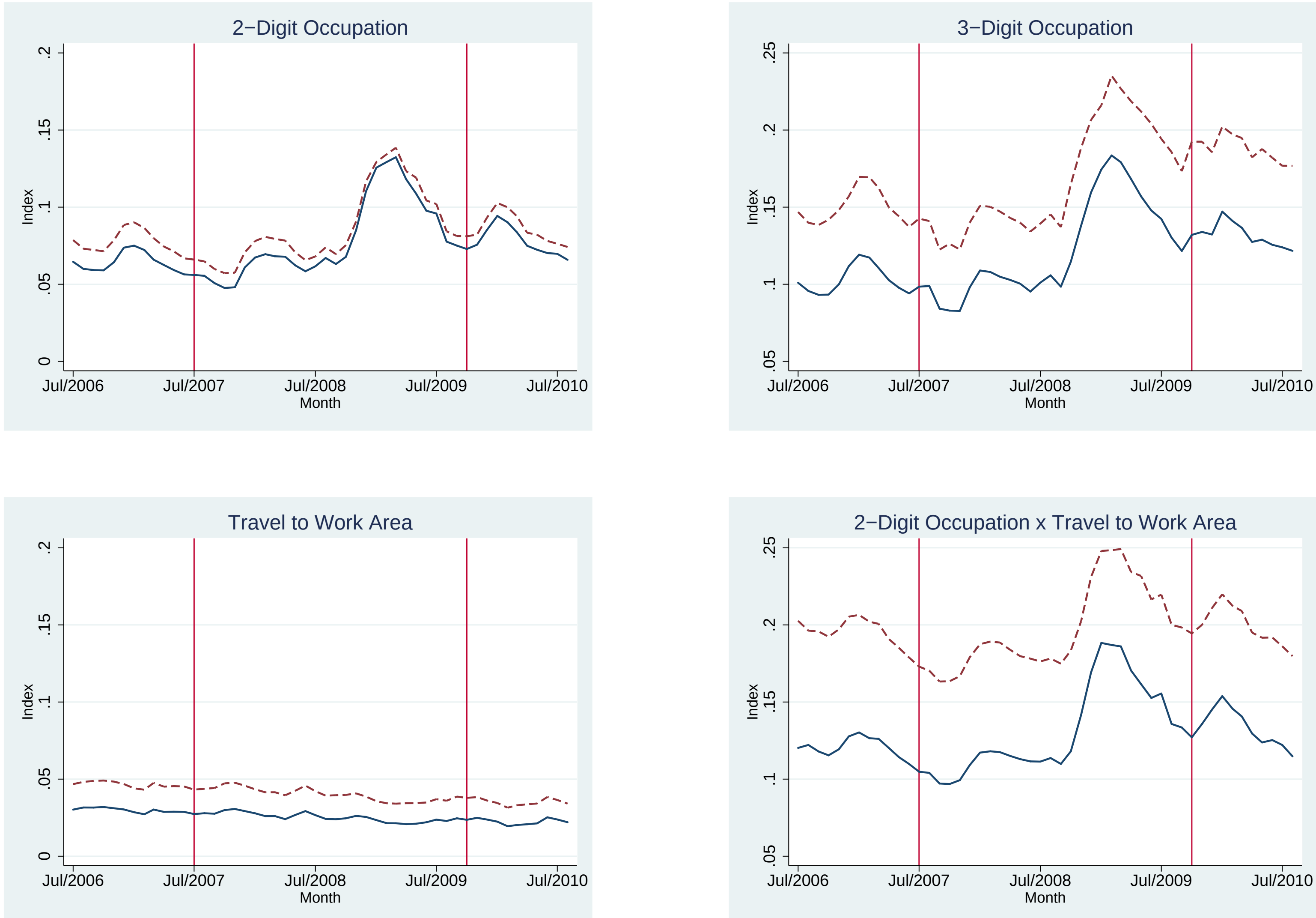

Figure 5: Mismatch $\mathcal{M}^h$ (Solid) $\mathcal{M}^h_\phi$ (Dashed); UK Recession denoted by the vertical red lines.

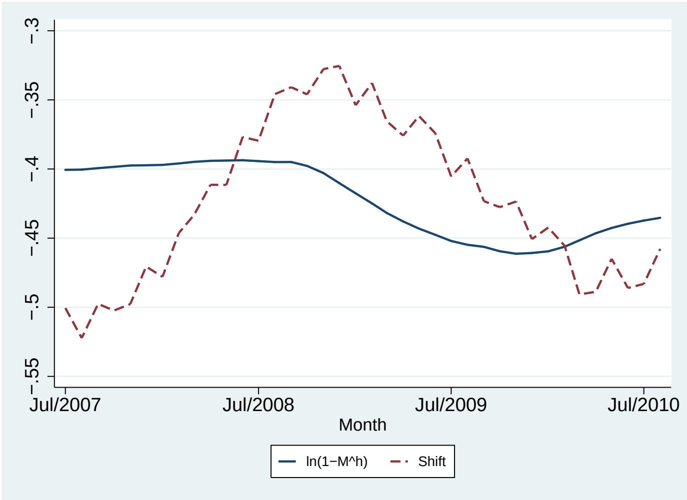

Figure 6: Insert Caption Here...this is the ln(1-Mismatch) v. ln(chi)

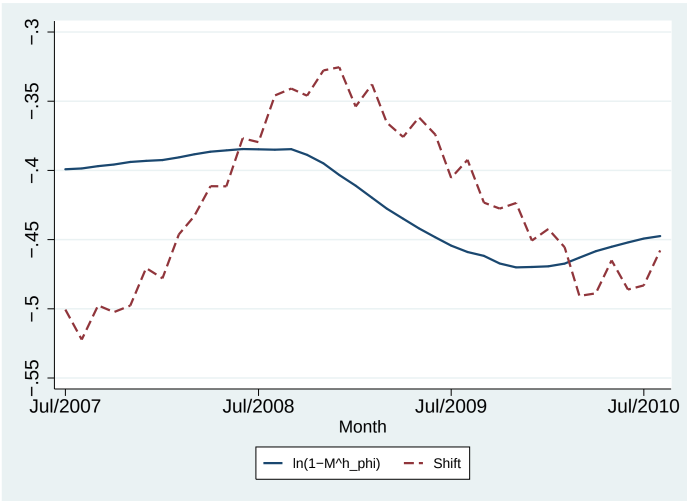

Figure 7: Insert Caption Here...this is the heterogeneous ln(1-Mismatch) v. ln(chi)

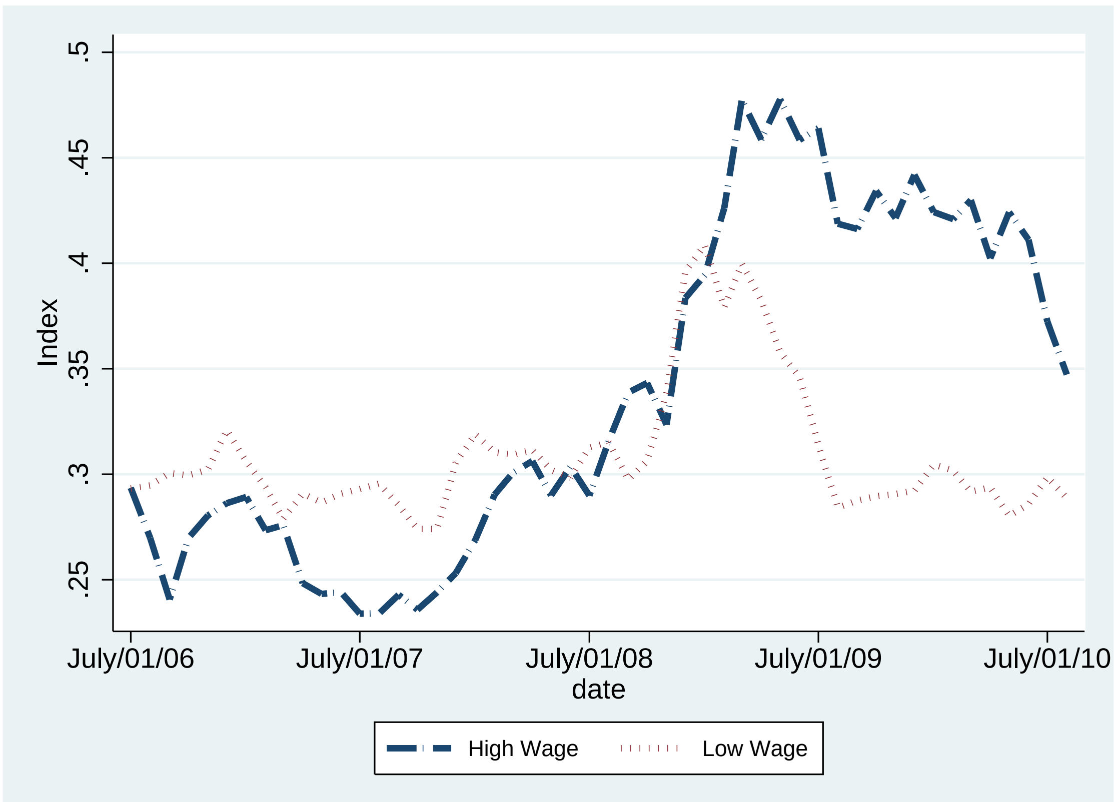

Figure 8: UK Mismatch Index : $\mathcal{M}^u$ By Wage

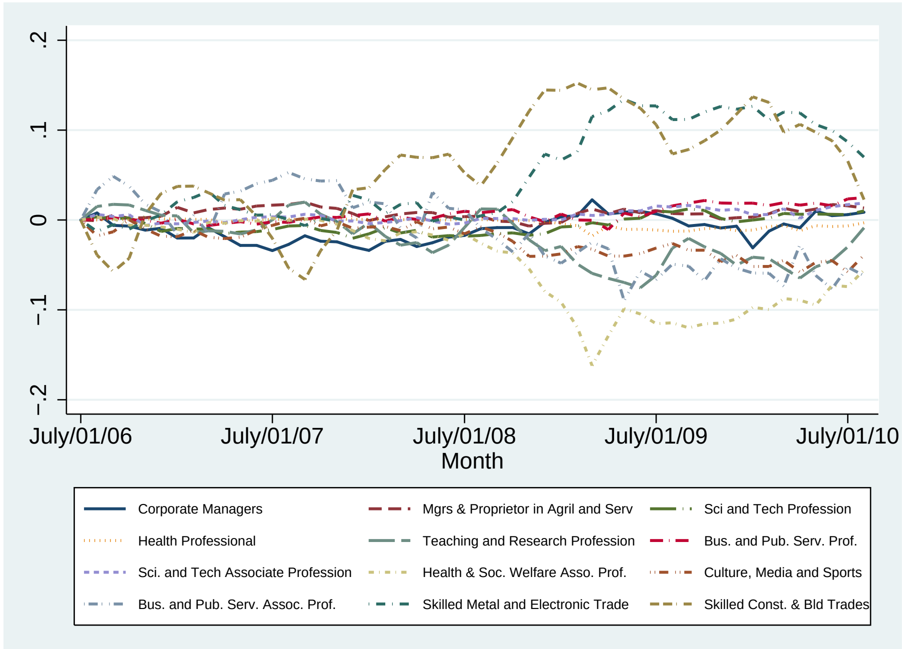

Figure 9: Fraction of Unemployed Minus Fraction of Vacancies (High Wage Group)

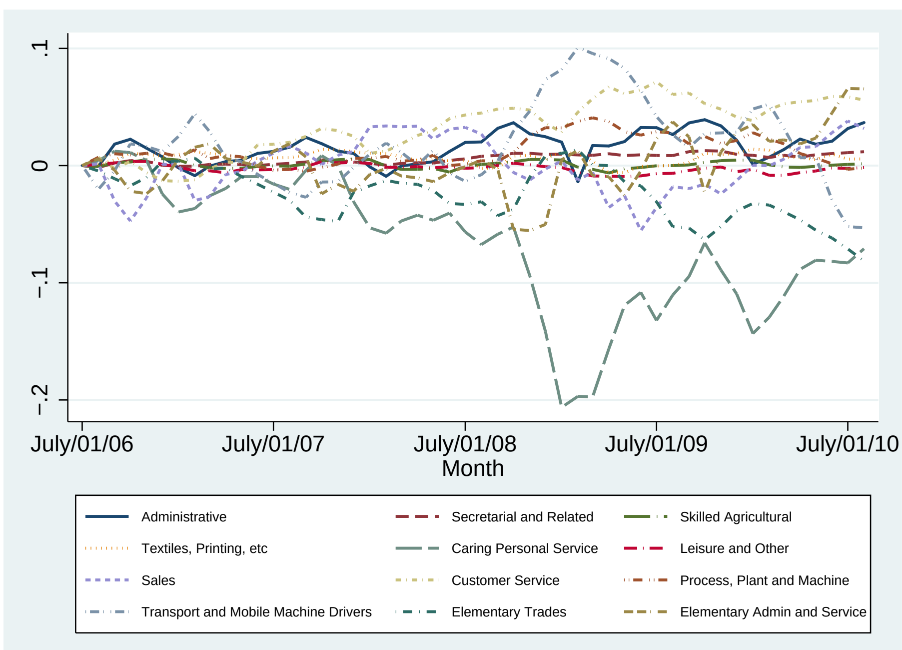

Figure 10: Fraction of Unemployed Minus Fraction of Vacancies (Low Wage Group)

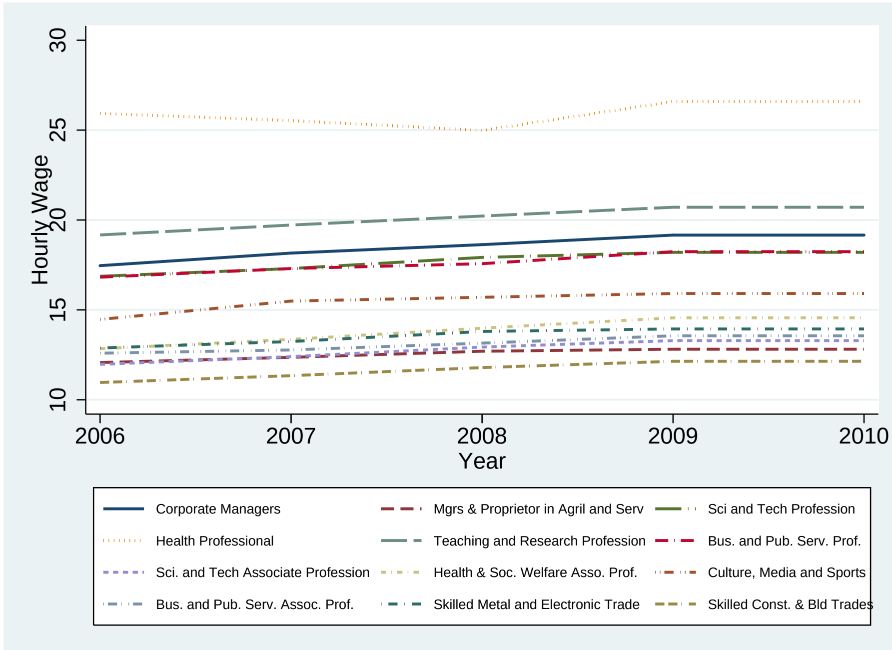

Figure 11: Median Hourly Wage (High Wage Group)

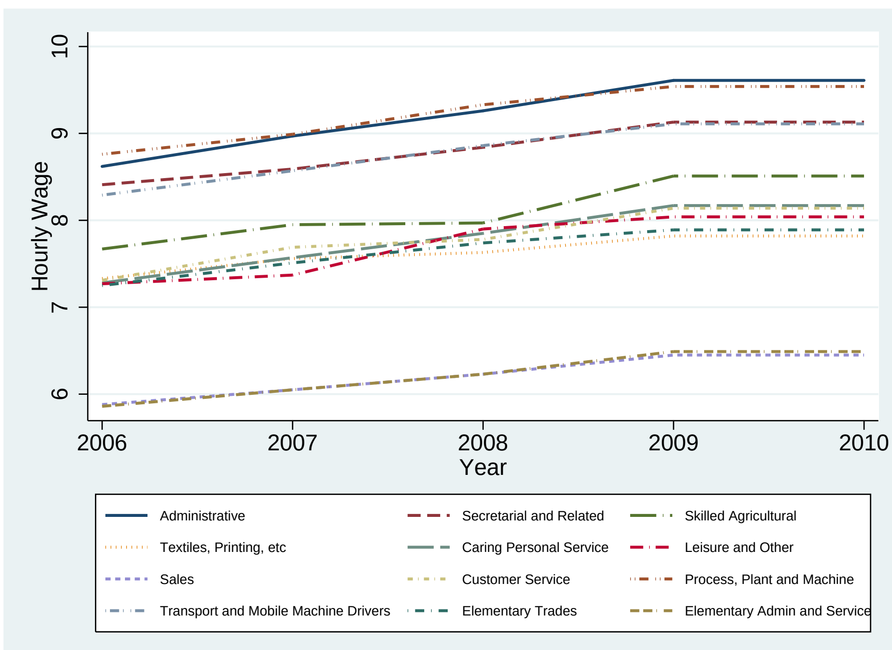

Figure 12: Median Hourly Wage (Low Wage Group)

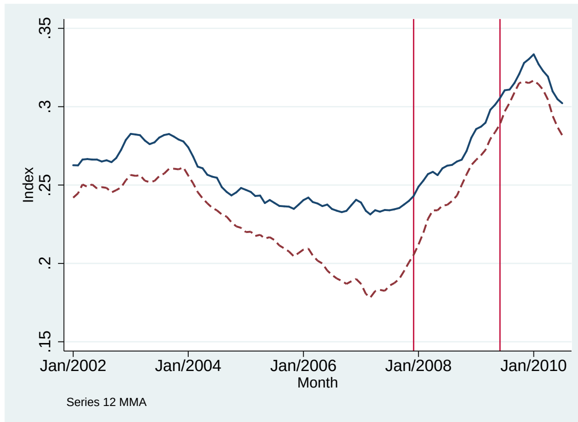

Figure 13: Mismatch indexes: $\mathcal{M}^u$ (solid line) and $\mathcal{M}^u_\phi$ (dashed line). All series are 12-month moving averages.

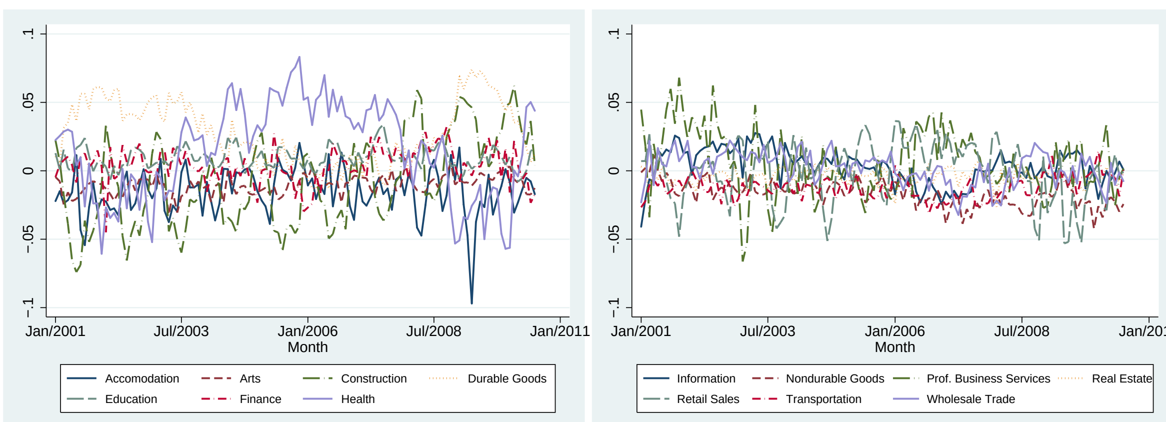

Figure 14: $u_i/u - v_i/v$ for 14 industries. Note that $u_i/u - v_i/v$ is normalized to zero at the beginning of the sample period.

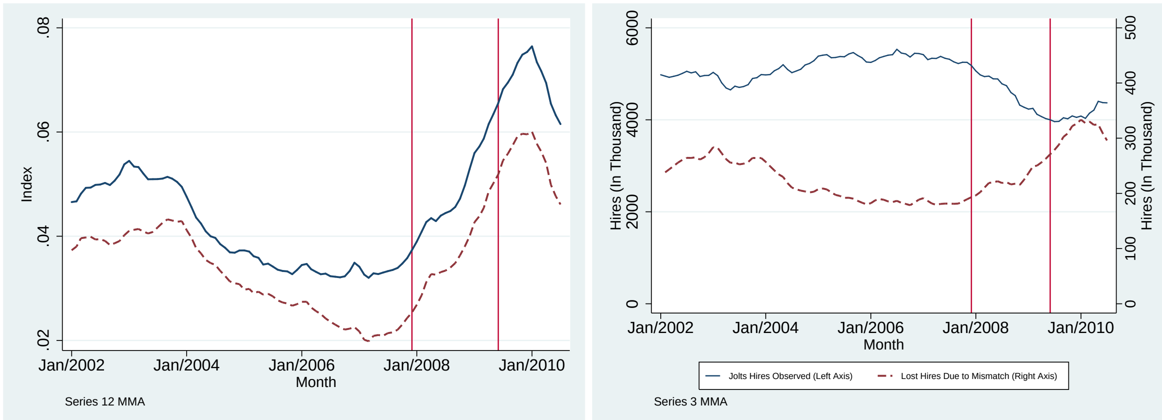

Figure 15: Left Panel: Mismatch indexes $\mathcal{M}^h$ (solid line) and $\mathcal{M}^h_\phi$ (dashed line). Right Panel: Hires (solid line, seasonally adjusted) and hires lost due to mismatch (dashed line, 12-month moving average).

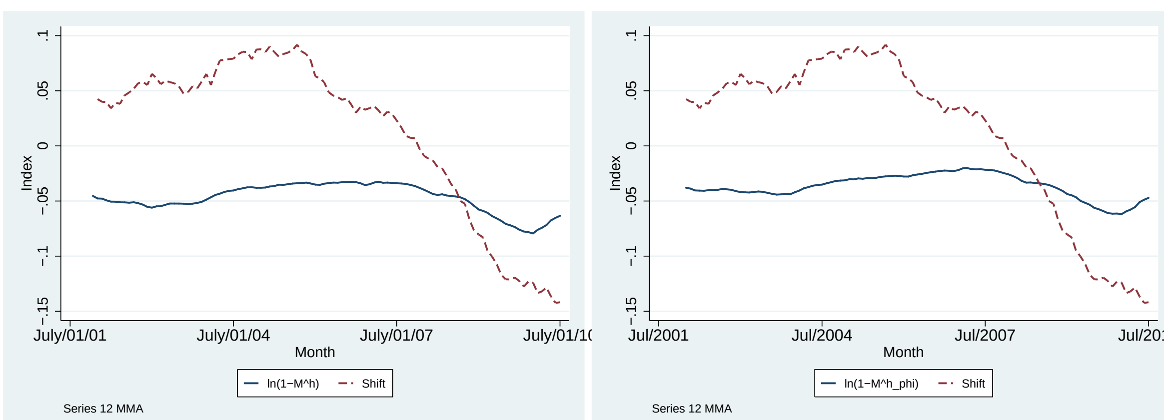

Figure 16: Left Panel: $\ln(1 - \mathcal{M}^h)$ and the shift in the matching function; Right Panel: $\ln(1 - \mathcal{M}^h_\phi)$ and the shift in the matching function. (All series are 12-month moving averages.)

References

[1] Abraham, K. and L.F. Katz (1986). “Cyclical Unemployment: Sectoral Shifts or Aggregate Disturbances? *Journal of Political Economy*, 94, 507-522.

[2] Alvarez F. and R. Shimer (2010). “Search and Rest Unemployment,” forthcoming, *Econometrica*.

[3] Birchenall, J. A. (2010). “A Competitive Theory of Mismatch,” mimeo, UCSB.

[4] Burgess, S. and S. Profit (2001). “Externalities in the Matching of Workers and Firms in Britain,” *Labour Economics*, 8, 313-333.

[5] Carrillo-Tudela and L. Visscher (2010). “Search and Endogenous Reallocation over the Business Cycle,” mimeo.

[6] Coles M. G. and E. Smith (1996). “Cross-Section Estimation of the Matching Function: Evidence from England and Wales,” *Economica* 63, 589-597.

[7] Davis, S. J., Faberman, R. J., Haltiwanger, J. C. (2010). “The Establishment-Level Behavior of Vacancies and Hiring,” NBER Working Paper 16265.

[8] Elsby, M., B. Hobijn, and A. Şahin (2010). “The Labor Market in the Great Recession,” *Brookings Papers on Economic Activity*, Spring 2010, 1-48.

[9] Hsieh, C. and P. Klenow (2009). “Misallocation and Manufacturing TFP in China and India,” *Quarterly Journal of Economics*, 124, 1403-1448.

[10] Hornstein, A., P. Krusell, and G. Violante (2007). “Technology-Policy Interaction in Frictional Labor Markets,” *Review of Economic Studies*, 74, 1089-1124.

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

[12] Kocherlakota, N. (2010). “Inside the FOMC,” Speech at Marquette, Michigan on August 17, 2010.

[13] Layard, R. and S. Nickell (1986). “Unemployment in Britain,” *Economica*, 53, 121-169.

[14] Lilien, D.M. (1982). “Sectoral Shifts and Cyclical Unemployment,” *Journal of Political Economy*, 90, 777-793.

[15] Ljungqvist, L. and T. J. Sargent (1998). “The European Unemployment Dilemma,” *Journal of Political Economy* 106, 514-550.

[16] Mortensen, D. T and C. A. Pissarides (1999). “Unemployment Responses to ’Skill-Biased’ Technology Shocks: The Role of Labour Market Policy,” *Economic Journal*, 109, 242-65.

[17] Padoa-Schioppa F. (1991). Mismatch and Labour Mobility. Centre for Economic Policy Research.

[18] Pissarides, C. A. (1986). “Unemployment and Vacancies in Britain,” *Economic Policy*, 1, 499-559.

[19] Phelps, E. (2008). “U.S. Monetary Policy and the Prospective Structural Slump,” Speech to the 7th Annual BIS Conference on Monetary Policy, Luzern, Switzerland.

[20] Restuccia, D. and R. Rogerson (2008). “Policy Distortions and Aggregate Productivity with Heterogeneous Plants,” *Review of Economic Dynamics*, 11, 707-720.

[21] Shimer R. (2007). “Mismatch,” *American Economic Review*, 97, 1074-1101.

6 Appendix

Table A1: Estimated ϕ for U.S. Industries

Labor Market	Industry	ϕ
Sectors	Information	0.66
	Finance	0.69
	Health	0.77
	Durable Goods	0.78
	Education	0.82
	Nondurable Goods	0.90
	Transportation	1.04
	Wholesale	1.06
	Retail Sale	1.20
	Professional Business Services	1.21
	Accommodations	1.38
	Arts	1.58
	Construction	1.68