

Measuring Mismatch in the U.S. Labor Market[†]

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 VERSION: NOVEMBER 2010 - THIS REVISION: MARCH 2011

PRELIMINARY - DO NOT CITE WITHOUT PERMISSION

Abstract

This paper measures mismatch in the U.S. labor market. Mismatch is defined as the distance between the observed allocation of unemployment across sectors and 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 are equated across sectors. We develop this condition into mismatch indexes that allow to quantify how much of recent rise in U.S. unemployment is associated to an increase in mismatch. We use two sources of cross-sectional data on vacancies, JOLTS and HWOL, together with unemployment data from the CPS for 2001-2010. We find that increased mismatch accounted for less than one percentage point of the rise in the unemployment rate from the start of the recession to 2010.

[†]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. 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. Receiving unemployment insurance (UI) benefits for a longer period might reduce the incentive of the unemployed to look for work. Similarly, it also increases their reservation wage, so that they may reject job offers that they would otherwise have accepted in the absence of these extended benefits.¹ A 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 eight 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. Given the decline in house prices that accompanied the recession, job applicants may be more reluctant to apply for and accept jobs that are not within commuting distance from their current residence and would require them to sell their homes. This phenomenon, which is generally referred to as “house lock,” appears consistent with recent data that showed that the rate of interstate migration in the U.S. has reached a postwar low. Additionally, recent work examining the link between house prices and mobility using data from 1985 to 2005 has found that mobility was lower for owners with negative equity in their homes (Ferreira, Gyourko, and Tracy, 2010), pointing to a potentially important negative effect of housing-related problems on the labor market. 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.² Lack of coincidence between unemployment and vacancies across labor markets is one

¹Various studies analyzed the effects of UI extensions on the unemployment rate. Estimates typically attribute around one percentage point of the rise in the unemployment rate to the UI extensions. This is due to fewer moves into employment, but also fewer people dropping out of the labor force. See Valletta and Kuang (2010) and Fujita (2011) for a detailed discussion.

²This observation has been emphasized before by Davis, Faberman, and Haltiwanger (2010), Elsby, Hobijn, and Şahin

of the candidate explanations for this shift.³

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 economic slump.⁴ In this paper we develop a simple framework to conceptualize the notion of mismatch unemployment and construct some intuitive mismatch indexes. We then use disaggregated data on vacancies and unemployment 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. 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, we examine whether, given the distribution of vacancies observed in the economy, it would be feasible to reallocate unemployed workers across markets in a way that reduces 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 aggregate job finding rate which, in turn, translates into additional unemployment. The difference in unemployment between the observed allocation and the efficient allocation provides an estimate of *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 (e.g., Lagos, 2006; Hsieh and Klenow, 2009; Restuccia and Rogerson, 2008): quantifying misallocation entails measuring how much the observed allocation deviates from a first-best benchmark.⁵

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

(2010), Hall (2010), and others.

³For 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.

⁴For an overview of this debate, see Roubini Global Economics at <http://www.roubini.com/>.

⁵In our case, the benchmark is a constrained first best, because the planner still faces the within-market frictional matching.

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 focus on two specific indices. The first, \mathcal{M}^u , is similar to traditional measures of the extent of misallocation that have been used to measure structural imbalance in the economy. It measures the fraction of unemployed workers searching in the wrong labor market, where “wrong” is defined relative to the optimal allocation of workers across markets. This index, however, cannot be used to compute a counterfactual measure of unemployment in the absence of mismatch because it does not provide any information on how the job-finding process changes across the two environments. Workers searching in the wrong labor market can still find jobs, albeit at a slower rate. At the same time, even in the optimal allocation, unemployed workers still face the frictions embodied in the within-market matching functions. Thus, to compute how equilibrium unemployment would change in the absence of mismatch, one needs to understand what happens to the job-finding rate in this case. The second index we develop, \mathcal{M}^h , does this by measuring the fraction of hires that are lost because of the misallocation. Since the presence of mismatch results in a loss of hires, it lowers the average job-finding rate for a given level of unemployment and vacancies. One can then make the appropriate correction for the job-finding rate and compute counterfactual equilibrium unemployment in the absence of mismatch. It is important to note that the effect of mismatch on the unemployment rate tends to be higher during recessions. When separations are high, the pool of unemployed is large, so the effect of the reduction in job finding induced by mismatch is amplified.

Our indexes capture an “ideal” notion of total mismatch defined as misallocation relative to an optimal unemployment distribution in the absence of *any* frictions across markets. Such frictions may include 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 2.2.

We apply our analysis to the U.S. labor market and construct measures of mismatch across seventeen major industry sectors and four Census regions using vacancy data from the JOLTS and unemployment data from the CPS for January 2000 to November 2010.⁶ We find that mismatch at the sectoral level increased during the recession and started to come down in 2010; an indication of a cyclical pattern for mismatch. Our calculations show that sectoral mismatch accounted for at most 0.7 percentage points of the increase in the unemployment rate from the start of the recession to

⁶In Şahin, Song, Topa and Violante (2011), we also apply our methodology to the U.K labor market.

2010.⁷ We also calculate geographic mismatch measures and find little role for geographic mismatch in explaining the increase in the unemployment rate. This finding is consistent with other recent work that investigated the house-lock mechanism using different methods.⁸

Our paper relates to an older literature that popularized the idea of mismatch (or structural) unemployment 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.⁹ Similarly, the importance of the “sectoral shift hypothesis” developed by Lilien (1982) was much diminished 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, Barnichon and Figura (2011) have contributed to reviving this literature by showing that the variance of labor market tightness across sectors, suggestive of mismatch between unemployment and vacancies, can be analytically related to aggregate matching efficiency.

Shimer (2007a) and Mortensen (2009) developed dynamic models 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 proposed dynamic equilibrium models with mobility decisions across labor markets where unemployed workers, in equilibrium, may be misallocated.

The rest of the paper is organized as follows. Section 2 presents the theoretical framework. Section 3 derives the mismatch indexes and explains how we run our counterfactuals. Section 4 describes the data and Section 5 performs the empirical analysis on the U.S. labor market. Section 6 concludes.

⁷One should also keep in mind that, for the U.S., our mismatch measures only capture misallocation of workers across 17 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.

⁸See, for example, Molloy, Smith, and Wozniak (2010) and Schulhofer-Wohl (2010).

⁹Since 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; Mortensen and Pissarides, 1999; Hornstein, Krusell and Violante, 2007).

2 Theoretical 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: 1) we allow for a dynamic and stochastic economic environment, while their set up was static; and 2) we allow for heterogeneity across sectors in a number of dimensions.¹⁰

Time is discrete. The economy comprises of a large number I of distinct labor markets (sectors) indexed by i . New production opportunities, corresponding to job vacancies (v_i) arise exogenously across sectors. The economy is populated by a measure one of risk-neutral individuals. Individuals choose to participate to the labor force. If they do, they can be either employed in sector i (e_i) or unemployed and searching in sector i (u_i). Therefore, the aggregate labor force is $\ell = \sum_{i=1}^I (e_i + u_i) \leq 1$. We normalize to zero utility from non participation and let ξ denote the disutility of search for the unemployed.

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 $\Phi \cdot \phi_i \cdot m(u_i, v_i)$, with m strictly increasing and strictly concave in both arguments and homogeneous of degree one in (u_i, v_i) . The term $\Phi \cdot \phi_i$ measures matching efficiency (i.e., the level of fundamental frictions) in sector i , with Φ denoting the aggregate component and ϕ_i the idiosyncratic component. Existing matches in sector i produce $Z \cdot z_i$ units of output, where Z is common across sectors. However, new matches produce only a fraction $\gamma < 1$ of output compared to existing matches –a stylized way to capture training costs for hires of unemployed workers (regardless of the sector in which they are hired). Matches are destroyed exogenously at rate δ , common across sectors.

Aggregate shocks Z , δ and Φ follow the joint Markov chain $\Gamma_{Z,\delta,\Phi}(Z', \delta', \Phi'; Z, \delta, \Phi)$ and the vector of vacancies $\mathbf{v} = \{v_i\}$ follows $\Gamma_{\mathbf{v}}(\mathbf{v}'; \mathbf{v}, Z', \delta', \Phi')$. The notation shows that we allow for autocorrelation in $\{Z, \delta, \Phi, \mathbf{v}\}$ and for correlation between vacancies and the aggregate shocks. The idiosyncratic sector-specific vectors of matching and productive efficiency $\phi = \{\phi_i\}$ and $\mathbf{z} = \{z_i\}$ follow, respectively, the Markov matrices $\Gamma_{\phi}(\phi'; \phi)$ and $\Gamma_{\mathbf{z}}(\mathbf{z}'; \mathbf{z})$. We assume that these idiosyncratic components of matching efficiency and productivity are uncorrelated across sectors, even though they can be correlated over time.

Within each period, events unfold as follows. At the beginning of the period, the aggregate shocks (Z, δ, Φ) , vacancies \mathbf{v} , matching efficiencies ϕ , and sector specific productivities \mathbf{z} are observed. At this stage, the distribution of active matches $\mathbf{e} = \{e_1, \dots, e_I\}$ across markets and the total number of unemployed workers u is also given. Next, the unemployed workers choose to direct their job search towards a specific labor market. Once the unemployed workers are allocated, the matching process

¹⁰In 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.

takes place and $h_i = \Phi\phi_i m(u_i, v_i)$ new hires are made in each market. Production takes place in the $e_i + h_i$ matches. Next, a fraction δ of matches is destroyed exogenously in each market and a number σ_i of workers separates from sector i , determining next period employment distribution $\{e'_i\}$. Finally, labor force decisions for next period are taken. Given ℓ' and $\{e'_i\}$, also the stock of unemployed workers u' for next period is determined.

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:

$$\begin{aligned}
V(u, \mathbf{e}; \phi, \mathbf{z}, \mathbf{v}, Z, \delta, \Phi) &= \max_{\{u_i, \sigma_i, \ell'\}} \sum_{i=1}^I Z z_i (e_i + \gamma h_i) - \xi u + \beta \mathbb{E}[V(u', \mathbf{e}'; \phi', \mathbf{z}', \mathbf{v}', Z', \delta', \Phi')] \\
\text{s.t.} &: \\
\sum_{i=1}^I u_i &\leq u & (1) \\
h_i &= \Phi\phi_i m(u_i, v_i) & (2) \\
e'_i &= (1 - \delta)(e_i + h_i) - \sigma_i & (3) \\
u' &= \ell' - \sum_{i=1}^I e'_i & (4) \\
u_i &\in [0, u], \ell' \in [0, 1], \sigma_i \in [0, (1 - \delta)(e_i + h_i)], & (5) \\
&\Gamma_{Z, \delta, \Phi}(Z', \delta', \Phi'; Z, \delta, \Phi), \Gamma_{\mathbf{v}}(\mathbf{v}'; \mathbf{v}, Z', \delta', \Phi'), \Gamma_{\phi}(\phi'; \phi), \Gamma_{\mathbf{z}}(\mathbf{z}'; \mathbf{z}) & (6)
\end{aligned}$$

The per period net output for the planner equals to production $Z z_i (e_i + \gamma h_i)$ in each market i minus the search costs. The first constraint (1) states that the planner has u unemployed workers available to allocate across sectors. Equation (2) states that, once the allocation $\{u_i\}$ is chosen, the frictional matching process in each market yields $\Phi\phi_i m(u_i, v_i)$ new hires which add to the existing e_i active matches. Equation (3) describes (exogenous and endogenous) separations and the determination of next period distribution of active matches $\{e'_i\}$. Equation (4) describes the law of motion of the stock of unemployment. The last line (6) in the problem collects all the exogenous Markov processes the planner takes as given. The planner chooses how to allocate $\{u_i\}$ across sectors, chooses how many employed workers to separate from their productive matches at the end of the period $\{\sigma_i\}$, and the size of the labor force next period ℓ' .

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

$$\gamma Z z_i \Phi \phi_i m_u \left(\frac{v_i}{u_i} \right) + \beta \mathbb{E} [-V_u(u', \mathbf{e}'; \phi', \mathbf{z}', \mathbf{v}', Z', \delta', \Phi') + V_{e_i}(u', \mathbf{e}'; \phi', \mathbf{z}', \mathbf{v}', Z', \delta', \Phi')] (1 - \delta) \Phi \phi_i m_u \left(\frac{v_i}{u_i} \right) = \mu, \quad (7)$$

where μ is the multiplier on constraint (1). The Envelope conditions with respect to the states u and e_i yield:

$$V_u(u, \mathbf{e}; \phi, \mathbf{z}, \mathbf{v}, Z, \delta, \Phi) = \mu - \xi \quad (8)$$

$$V_{e_i}(u, \mathbf{e}; \phi, \mathbf{z}, \mathbf{v}, Z, \delta, \Phi) = Z z_i + \beta(1 - \delta) \mathbb{E} [V_{e_i}(u', \mathbf{e}'; \phi', \mathbf{z}', \mathbf{v}', Z', \delta', \Phi')]. \quad (9)$$

According to the first condition, the marginal value of an unemployed to the planner equals the shadow value of being available to search (μ) net of the disutility of search ξ . The second condition states the marginal value of an employed worker is its flow output plus its discounted continuation value, conditional on the match not being destroyed.

The decision of how many workers to separate from sector i employment into unemployment is:

$$\mathbb{E} [V_{e_i}(u', \mathbf{e}'; \phi', \mathbf{z}', \mathbf{v}', Z', \delta', \Phi') - V_u(u', \mathbf{e}'; \phi', \mathbf{z}', \mathbf{v}', Z', \delta', \Phi')] \begin{cases} > 0 & \rightarrow \sigma_i = 0 \\ = 0 & \rightarrow \sigma_i \in (0, (1 - \delta)(e_i + h_i)) \\ < 0 & \rightarrow \sigma_i = (1 - \delta)(e_i + h_i) \end{cases} \quad (10)$$

depending on whether at the optimum a corner or interior solution arises.

Consider now the decision on the labor force size next period ℓ' which states that

$$\mathbb{E} [V_u(u', \mathbf{e}'; \phi', \mathbf{z}', \mathbf{v}', Z', \delta', \Phi')] = 0, \quad (11)$$

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. Combining (11) with (8), 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 search disutility ξ .¹¹ Note that the first order condition (11) and the Envelope condition (9) imply that the optimality condition (10) holds with the “ $>$ ” inequality and hence, $\sigma_i = 0$. Intuitively, if the number of unemployed can be freely adjusted by moving individuals into (out of) unemployment out of (into) non participation, the planner will prefer to keep the employed workers producing.

Consider now the Envelope condition (9) and make an additional assumption about the stochastic process for z_i , i.e., $\mathbb{E}(z'_i) = \rho z_i$, or that z_i follows a *linear* first-order autoregressive process. We now conjecture that

$$V_{e_i}(u, \mathbf{e}; \phi, \mathbf{z}, \mathbf{v}, Z, \delta, \Phi) = z_i \Psi(Z, \delta, \Phi), \quad (12)$$

¹¹We 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 ξ . It is clear that our result is robust to allowing ξ to be stochastic and correlated with (z, δ, Φ) .

where $\Psi(Z, \delta, \Phi)$ is a function of Z , δ and Φ alone. Using this conjecture into (9), we arrive at

$$V_{e_i}(u, \mathbf{e}; \phi, \mathbf{z}, \mathbf{v}, Z, \delta, \Phi) = Zz_i + \beta(1 - \delta)\mathbb{E}[z'_i\Psi(Z', \delta', \Phi')] = Zz_i + \beta(1 - \delta)\rho z_i\mathbb{E}[\Psi(Z', \delta', \Phi')].$$

Let's verify the conjecture:

$$\begin{aligned} z_i\Psi(Z, \delta, \Phi) &= Zz_i + \beta(1 - \delta)\rho z_i\mathbb{E}[\Psi(Z', \delta', \Phi')] \\ \Psi(Z, \delta, \Phi) &= Z + \beta(1 - \delta)\rho\mathbb{E}[\Psi(Z', \delta', \Phi')] \end{aligned}$$

which confirms the conjecture, since $\mathbb{E}[\Psi(Z', \delta', \Phi')]$ is only a function of (Z, δ, Φ) because of the assumed Markov structure for $\Gamma_{Z, \delta, \Phi}$.

Using (12) into (7), the optimality condition for the allocation of unemployed workers across sectors becomes

$$\gamma Z z_i \Phi \phi_i m_u \left(\frac{v_i}{u_i} \right) + \beta (1 - \delta) \rho \mathbb{E}[\Psi(Z', \delta', \Phi')] z_i \Phi \phi_i m_u \left(\frac{v_i}{u_i} \right) = \mu, \quad (13)$$

and rearranging:

$$z_i \phi_i m_u \left(\frac{v_i}{u_i} \right) = \frac{\mu}{\gamma Z \Phi + \beta (1 - \delta) \Phi \rho \mathbb{E}[\Psi(Z', \delta', \Phi')]},$$

where the right hand side is a magnitude independent of i . We conclude that the left hand side of this last equation is equalized across markets, yielding:

$$z_1 \phi_1 m_u \left(\frac{v_1}{u_1^*} \right) = \dots = z_i \phi_i m_u \left(\frac{v_i}{u_i^*} \right) = \dots = z_I \phi_I m_u \left(\frac{v_I}{u_I^*} \right), \quad (14)$$

where we have used the “*” to denote the optimal allocation. This is our key optimality condition for the allocation of unemployed workers across labor markets. It states that the higher the matching and productive efficiency parameters in market i , the more unemployed workers the planner wants searching in that market. Condition (14) is the “generalized Jackman-Roper optimality condition” for a dynamic stochastic economy with heterogeneity across sectors.

2.2 Comparison between actual and optimal allocation: what do we measure?

Our approach to quantify the mismatch component of unemployment at date t is based on comparing the actual (equilibrium) distribution $\{u_{it}\}$ observed directly from the data to the optimal (planner's) distribution $\{u_{it}^*\}$ implied by (14), for an (exogenously given) distribution of vacancies $\{v_{it}\}$ across sectors of the economy. This approach is at the heart of the misallocation literature (Hsieh and Klenow, 2009).

In equilibrium, there are a number of sources of misallocation that may induce $\{u_{it}\}$ to deviate from $\{u_{it}^*\}$ including imperfect information, wage rigidities, government policies, and moving/retraining costs. Under imperfect information, workers may be reluctant to move because they do

not know where the vacancies are or what their prospects might be in the new location, occupation or industry. In the presence of wage rigidities, workers may choose not to move because wages deviate from productivity remaining relatively high (low) in the declining (expanding) sectors. An array of government interventions (e.g., generous unemployment benefits, housing and mortgage related policies, sector specific taxes/transfers) may hamper mobility and be a source of misallocation. Moving or retraining costs associated to working in a new location, industry or occupation can also reduce mobility.

By following our approach, one does not need to model explicitly any of the sources of misallocation since the distribution $\{u_{it}\}$ comes straight from the data and the distribution $\{u_{it}^*\}$ is the solution to a planner problem with free mobility of labor across markets. The crucial advantage is that optimality can be fully characterized analytically and boils down to the intuitive static condition (14). This condition can be easily manipulated into mismatch indexes –measuring the distance between the actual and optimal allocation– that can be estimated using micro data. In the context of the recent US experience, these indexes can answer the question of whether the observed rise in unemployment is due to increased mismatch.

The transparency of our approach must be traded off with two drawbacks. First, some of the impediments to labor mobility, in particular moving and retraining costs, would be part of the physical environment in a constrained planner’s problem and will likely lead to a lower measured mismatch. Therefore, our approach should be thought of as a measurement device that (for a given level of disaggregation) delivers an *upper bound for the level of mismatch unemployment*, $u_{it} - u_{it}^*$.

Second, our methodology offers a measurement tool for mismatch unemployment, but does not get at the questions of why unemployed workers are misallocated or whether mismatch is “constrained efficient”. Answering these questions would require solving an equilibrium model incorporating all the potential sources of limited labor mobility across sectors.¹² Within our approach, we can still learn something useful by looking at various measures of mismatch for various definitions of sectors (occupation, industry, location) and for various groups of workers (young vs old, skilled vs unskilled, etc.).

3 Mismatch indexes and counterfactual analysis

We now show how to derive, from the optimality condition (14), indexes measuring the size of the mismatch component of unemployment. To fix ideas, we begin with the case where there is no heterogeneity in ϕ and z across markets, and then we move to the case with heterogeneity. Finally, we describe how to use these indexes to construct counterfactual experiments that show how much of

¹²For example, if the key sources of limited mobility are moving costs, one would conclude that mismatch is largely constrained efficient. If, instead, the main sources are informational frictions, wage rigidities or government policies, one would conclude that it is not.

the recent rise in US unemployment is due to mismatch.

3.1 Mismatch indexes with no heterogeneity across markets

The \mathcal{M}_t^u index. We start by computing an index measuring the fraction of unemployed workers searching in the “wrong” sector at a date t . Recall that, at the beginning of period t , the distribution of vacancies $\{v_{it}\}$ and the number of unemployed u_t are given for the planner. The planner only chooses how to allocate unemployed workers across sectors. With no heterogeneity in ϕ and z , the strict concavity of m and equation (14) imply that the planner wants to equate the vacancy-unemployment ratio across labor markets, i.e., $u_{it}^* = (1/\theta_t)v_{it}$ where $v_t/u_t \equiv \theta_t$ is the aggregate market tightness. The number of unemployed workers misallocated in their job search, compared to the planner’s allocation, is therefore

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

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

$$\mathcal{M}_t^u = \frac{u_t^M}{u_t} = \frac{1}{2} \sum_{i=1}^I \left| \frac{u_{it}}{u_t} - \frac{v_{it}}{v_t} \right|. \quad (15)$$

It is easy to see that $\mathcal{M}_t^u \in [0, 1]$ and therefore it is an index. $\mathcal{M}_t^u = 0$ 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 answer the question of how much unemployment would be reduced if we could eliminate mismatch. Even if workers search in the wrong sector, they would find jobs at some (slower) rate. Addressing such question requires computing how many additional hires would be generated by switching to the optimal allocation of unemployed workers across sectors.

The \mathcal{M}_t^h index. 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.,

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

Summing across market, the aggregate numbers of hires can be expressed as:

$$h_t = \Phi_t v_t^\alpha u_t^{1-\alpha} \cdot \left[\sum_{i=1}^I \left(\frac{v_{it}}{v_t} \right)^\alpha \left(\frac{u_{it}}{u_t} \right)^{1-\alpha} \right]. \quad (16)$$

The first term in (16) denotes the highest number of new hires that can be achieved under the optimal allocation where market tightness is equated (to its aggregate value) across sectors. Therefore, we can

define an alternative mismatch index as:

$$\mathcal{M}_t^h = 1 - \frac{h_t}{h_t^*} = 1 - \sum_{i=1}^I \left(\frac{v_{it}}{v_t} \right)^\alpha \left(\frac{u_{it}}{u_t} \right)^{1-\alpha}. \quad (17)$$

The index \mathcal{M}_t^h measures precisely what fraction of hires is lost because of misallocation.¹³ It is easy to see that $\mathcal{M}_t^h \leq 1$. To show that $\mathcal{M}_t^h \geq 0$, note that

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

where the \leq sign follows from Hölder's inequality.

Properties of mismatch indexes. Both indexes \mathcal{M}_t^h and \mathcal{M}_t^u are invariant to pure aggregate shocks that shift the number of vacancies and unemployed up or down, but leave the vacancy and unemployment shares across markets unchanged.

Moreover, both indexes are non-decreasing in the level of disaggregation (i.e., the number of sectors). To see this, consider an economy where the aggregate labor market is described by two dimensions indexed by (i, j) , e.g., I regions \times J occupations. The mismatch index \mathcal{M}^u is

$$\mathcal{M}_{IJ}^u = \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J \left| \frac{v_{ij}}{v} - \frac{u_{ij}}{u} \right|.$$

Now, suppose we can only measure mismatch among the I regions, each containing J occupations. This coarser index is

$$\mathcal{M}_I^u = \frac{1}{2} \sum_{i=1}^I \left| \frac{v_i}{v} - \frac{u_i}{u} \right| = \frac{1}{2} \sum_{i=1}^I \left| \sum_{j=1}^J \left(\frac{v_{ij}}{v} - \frac{u_{ij}}{u} \right) \right| < \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J \left| \frac{v_{ij}}{v} - \frac{u_{ij}}{u} \right| = \mathcal{M}_{IJ}^u.$$

Turning to the \mathcal{M}^h index,

$$\begin{aligned} 1 - \mathcal{M}_I^h &= \frac{1}{v^\alpha u^{1-\alpha}} \sum_{i=1}^I (v_i)^\alpha (u_i)^{1-\alpha} \\ &= \frac{1}{v^\alpha u^{1-\alpha}} \sum_{i=1}^I \left(\sum_{j=1}^J v_{ij} \right)^\alpha \left(\sum_{j=1}^J u_{ij} \right)^{1-\alpha} \\ &= \frac{1}{v^\alpha u^{1-\alpha}} \sum_{i=1}^I \left(\sum_{j=1}^J \tilde{v}_{ij}^{\frac{1}{\alpha}} \right)^\alpha \left(\sum_{j=1}^J \tilde{u}_{ij}^{\frac{1}{1-\alpha}} \right)^{1-\alpha} \\ &> \frac{1}{v^\alpha u^{1-\alpha}} \sum_{i=1}^I \sum_{j=1}^J \tilde{v}_{ij} \tilde{u}_{ij} = \frac{1}{v^\alpha u^{1-\alpha}} \sum_{i=1}^I \sum_{j=1}^J v_{ij}^\alpha \cdot u_{ij}^{1-\alpha} = 1 - \mathcal{M}_{IJ}^h \end{aligned}$$

where the third line defines $\tilde{v}_{ij} \equiv v_{ij}^\alpha$ and $\tilde{u}_{ij} \equiv u_{ij}^{1-\alpha}$, and the last line uses Hölder's inequality.

¹³To express it as a fraction of the observed hires, we would have to compute $\mathcal{M}_t^h / (1 - \mathcal{M}_t^h)$.

3.2 Mismatch indexes with heterogeneous matching efficiencies

The $\mathcal{M}_{\phi t}^u$ index. Suppose now that individual labor markets differ in their frictional parameter ϕ_i and assume Cobb-Douglas matching functions within markets, i.e., $h_{it} = \Phi_t \phi_i v_{it}^\alpha u_{it}^{1-\alpha}$. From equation (14), 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

$$u_{it}^* = \phi_i^{\frac{1}{\alpha}} \cdot \left(\frac{v_{it}}{v_{\phi t}} \right) \cdot u_t. \quad (18)$$

Recall that the share of unemployed workers searching in the wrong sector is $u_t^M = \frac{1}{2} \sum_{i=1}^I |u_{it} - u_{it}^*|$.

Substituting the expression for u_{it}^* from (18) into the definition of u_t^M gives:

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

which, after some simple manipulations, yields the mismatch index

$$\mathcal{M}_{\phi t}^u = \frac{u_t^M}{u_t} = \frac{1}{2} \sum_{i=1}^I \left| \frac{u_{it}}{u_t} - \left(\frac{\phi_i}{\bar{\phi}_t} \right)^{\frac{1}{\alpha}} \cdot \frac{v_{it}}{v_t} \right| \quad (19)$$

where

$$\bar{\phi}_t = \left[\sum_{i=1}^I \phi_i^{\frac{1}{\alpha}} \left(\frac{v_{it}}{v_t} \right) \right]^\alpha \quad (20)$$

is a CES aggregator of the market-level matching efficiencies weighted by their vacancy share. The index in (19) is similar to the index (15) derived for the homogeneous markets case, except for the adjustment term in brackets which equals 1 when there is no heterogeneity in ϕ_i . This term corrects the index for the fact that the planner may want to allocate a share of unemployed workers larger than the vacancy share in market i when its matching efficiency ϕ_i is higher than the average $\bar{\phi}_t$.

The $\mathcal{M}_{\phi t}^h$ index. The optimal aggregate number of hires is

$$h_t^* = \Phi_t v_t^\alpha u_t^{1-\alpha} \left[\sum_{i=1}^I \phi_i \left(\frac{v_{it}}{v_t} \right)^\alpha \left(\frac{u_{it}^*}{u_t} \right)^{1-\alpha} \right]. \quad (21)$$

Substituting the optimality condition (18) in equation (21), the total number of optimal new hires is $h_t^* = \Phi_t \bar{\phi}_t v_t^\alpha u_t^{1-\alpha}$, where $\bar{\phi}_t$ is defined in equation (20). Similarly, we can define the total number of observed new hires as

$$h_t = \Phi_t v_t^\alpha u_t^{1-\alpha} \left[\sum_{i=1}^I \phi_i \left(\frac{v_{it}}{v_t} \right)^\alpha \left(\frac{u_{it}}{u_t} \right)^{1-\alpha} \right], \quad (22)$$

and hence the counterpart of (17) in the heterogeneous markets case becomes

$$\mathcal{M}_{\phi t}^h = 1 - \frac{h_t}{h_t^*} = 1 - \sum_{i=1}^I \left(\frac{\phi_i}{\bar{\phi}_t} \right) \left(\frac{v_{it}}{v_t} \right)^\alpha \left(\frac{u_{it}}{u_t} \right)^{1-\alpha}. \quad (23)$$

3.3 Mismatch indexes with heterogeneous matching and productive efficiency

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

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

The \mathcal{M}_{xt}^u index. Following the same steps used for the derivation of $\mathcal{M}_{\phi t}^u$, it is easy to see that the \mathcal{M}_{xt}^u index is

$$\mathcal{M}_{xt}^u = \frac{u_t^M}{u_t} = \frac{1}{2} \sum_{i=1}^I \left| \frac{u_{it}}{u_t} - \left(\frac{x_i}{\bar{x}_t} \right)^{\frac{1}{\alpha}} \cdot \frac{v_{it}}{v_t} \right| \quad (25)$$

where

$$\bar{x}_t = \left[\sum_{i=1}^I x_i^{\frac{1}{\alpha}} \left(\frac{v_{it}}{v_t} \right) \right]^\alpha \quad (26)$$

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

The \mathcal{M}_{xt}^h index. The highest number of hires that can be obtained by optimally allocating the available unemployed workers is still given by equation (21). Substituting the optimality condition (24) in equation (21), the optimal number of new hires is $h_t^* = \Phi_t \bar{\phi}_{xt} v_t^\alpha u_t^{1-\alpha}$, where

$$\bar{\phi}_{xt} = \bar{x}_t \cdot \frac{\sum_{i=1}^I \left(\frac{1}{z_i} \right) x_i^{\frac{1}{\alpha}} \left(\frac{v_{it}}{v_t} \right)}{\sum_{i=1}^I x_i^{\frac{1}{\alpha}} \left(\frac{v_{it}}{v_t} \right)},$$

and note that, if z_i is constant across markets, $\bar{\phi}_{xt} = \bar{\phi}_t$. Since total new hires are given by (22), we obtain the counterpart of (23)

$$\mathcal{M}_{xt}^h = 1 - \sum_{i=1}^I \left(\frac{\phi_i}{\bar{\phi}_{xt}} \right) \left(\frac{v_{it}}{v_t} \right)^\alpha \left(\frac{u_{it}}{u_t} \right)^{1-\alpha}, \quad (27)$$

which measures the fraction of hires lost because of mismatch at date t .

In what follows, we also use the notation \mathcal{M}_{zt}^u and \mathcal{M}_{zt}^h to denote mismatch indexes for an economy where there is productivity heterogeneity but all markets have the same matching efficiency Φ_t .

3.4 Counterfactual analysis

With longitudinal data on $\{h_{it}, u_{it}, v_{it}\}$ for various sectors $i = 1, 2, \dots, I$ and dates $t = 1, 2, \dots, T$, assuming a Cobb-Douglas functional form for the matching function, we can consistently estimate the vacancy share α and the vector of sector-specific matching efficiencies $\{\phi_i\}$. Section 4 below illustrates this procedure in detail. Suppose the available data also allow to determine average productivity of labor $\{z_i\}$ in each sector. It is immediate to see that these are all the necessary ingredients to construct time series for both the \mathcal{M}_t^u and the \mathcal{M}_t^h indexes. This second group of indexes is especially useful for our counterfactuals.

Shifts in the Beveridge curve Consider the general case with both matching and productivity heterogeneity. By comparing optimal hires $h_t^* = \Phi_t \bar{\phi}_{xt} v_t^\alpha u_t^{1-\alpha}$ to actual hires (22), it can be seen that the aggregate matching function can be written as

$$h_t = \underbrace{(1 - \mathcal{M}_{xt}^h)}_{A_t} \cdot \bar{\phi}_{xt} \cdot \Phi_t \cdot v_t^\alpha u_t^{1-\alpha}, \quad (28)$$

which highlights that a shift in the aggregate matching function could have three separate sources: 1) a change in misallocation of idle labor across markets through \mathcal{M}_{xt}^h ; 2) reallocation of the demand for labor across sectors that, through a composition effect in the vacancy distribution, change the average value of matching efficiency $\bar{\phi}_{xt}$; 3) a change in matching efficiency common across all markets (Φ_t). We denote the product of all these three components as A_t .

We can use aggregate data on hires h_t , vacancies v_t and unemployment u_t , together with our estimate for α to calculate the total shift in the matching function

$$\log A_t = \log h_t - \alpha \log v_t - (1 - \alpha) \log u_t. \quad (29)$$

Since, by definition,

$$\log A_t = \log (1 - \mathcal{M}_{xt}^h) + \log \bar{\phi}_{xt} + \log \Phi_t, \quad (30)$$

and since we have independent measures of \mathcal{M}_{xt}^h and $\bar{\phi}_{xt}$, we can decompose the total shift in the aggregate Beveridge curve into these three components. In particular, we can assess how much of the observed shift is due to a change in mismatch.

Counterfactual unemployment To fix ideas about the impact of mismatch on equilibrium unemployment, recall that in steady state $u = s/(s + f)$ where s denotes the aggregate separation rate and $f \equiv h/u$ the aggregate job finding rate.¹⁴ A worse misallocation of unemployed workers across labor markets lowers hires and the job finding rate. A smaller job finding rate implies a higher unemployment rate.

There is an additional way in which the level of mismatch affects the unemployment rate, through the change in separation rate s . It is easy to see that

$$\frac{\partial u}{\partial s} = \frac{f}{(s + f)^2} > 0 \quad \text{and} \quad \frac{\partial^2 u}{\partial s \partial f} = \frac{s - f}{(s + f)^3} < 0,$$

where the second inequality holds for plausible parameterizations (where $s < f$). In other words, a rise in s will have a larger impact on unemployment in an economy with more mismatch (lower f). Intuitively, in such an economy it takes longer to reabsorb separating workers.

This discussion suggests the following strategy to construct a counterfactual unemployment rate absent mismatch, i.e., the purely *frictional* unemployment rate solving the problem of a planner who allocates workers to search always in the right sector. If, using (28), we let

$$f_t = \frac{h_t}{u_t} = (1 - \mathcal{M}_{xt}^h) \cdot \bar{\phi}_{xt} \cdot \Phi_t \cdot \left(\frac{v_t}{u_t}\right)^\alpha$$

be the actual aggregate job finding rate at date t , then the optimal job finding rate (without mismatch) is

$$f_t^* = \frac{h_t^*}{u_t^*} = \bar{\phi}_{xt} \cdot \Phi_t \cdot \left(\frac{v_t}{u_t^*}\right)^\alpha = \frac{f_t}{(1 - \mathcal{M}_{xt}^h)} \left(\frac{u_t}{u_t^*}\right)^\alpha.$$

Therefore, given an initial value for u_0^* (for example, the steady state value $s_0/(f_0^* + s_0)$), the counterfactual frictional unemployment rate can be obtained by iterating over the equation

$$u_{t+1}^* = s_t + (1 - s_t - f_t^*) u_t^*.$$

The difference between Δu and Δu^* over a given period of time measures the change in unemployment due to mismatch in the labor market.

Notice that this strategy assumes that the sequences for $\{s_t\}$ and $\{v_t\}$ are taken from the data (i.e., are the same in the equilibrium and in the counterfactual). This is consistent with the theoretical model where vacancies are exogenous to the planner and separations equal exogenous match destructions (voluntary quits are zero for the planner).

¹⁴We calculate the aggregate separation rate and the job-finding rate f using the methodology described in Shimer (2005). Consequently f includes transitions into nonparticipation as well as employment. We apply our correction to this total outflow rate and do not make a distinction between flows depending on their destination. As Shimer (2007b) shows in his Figure 4, the ratio of unemployment-to-employment flow rate to the unemployment-to-nonparticipation flow rate is very stable over the business cycle. Thus, our counterfactual gives us an upper bound on the effect of mismatch on the job-finding rate but does not cause a cyclical bias on the effect of mismatch on the unemployment rate.

3.4.1 Measurement error in vacancies

Suppose that true vacancies (V_{it}) in market i are a factor $\mu_i^{\frac{1}{\alpha}}$ of the observed vacancies (v_{it}), i.e., $V_{it} = v_{it}\mu_i^{\frac{1}{\alpha}}$. Since this problem appears to be less severe for unemployment and hires data, we assume that there is no measurement error in these variables (or measurement error is constant across sectors). For simplicity, consider the economy without heterogeneity in productive or matching efficiency. The true mismatch index is

$$\mathcal{M}_{\mu t}^u = \frac{1}{2} \sum_{i=1}^I \left| \frac{U_{it}}{U_t} - \frac{V_{it}}{V_t} \right| = \frac{1}{2} \sum_{i=1}^I \left| \frac{u_{it}}{u_t} - \frac{v_{it}\mu_i^{\frac{1}{\alpha}}}{\sum_{i=1}^I v_{it}\mu_i^{\frac{1}{\alpha}}} \right|,$$

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

$$\begin{aligned} \mathcal{M}_{\mu t}^u &= \frac{1}{2} \sum_{i=1}^I \left| \frac{u_{it}}{u_t} - \frac{\mu_i^{\frac{1}{\alpha}}}{\sum_{i=1}^I \mu_i^{\frac{1}{\alpha}} \left(\frac{v_{it}}{v_t}\right)} \cdot \frac{v_{it}}{v_t} \right| \\ &= \frac{1}{2} \sum_{i=1}^I \left| \frac{u_{it}}{u_t} - \left(\frac{\mu_i}{\bar{\mu}}\right)^{\frac{1}{\alpha}} \cdot \frac{v_{it}}{v_t} \right| \end{aligned} \quad (31)$$

where

$$\bar{\mu} = \left[\sum_{i=1}^I \mu_i^{\frac{1}{\alpha}} \left(\frac{v_{it}}{v_t}\right) \right]^{\alpha}.$$

Similarly, the true \mathcal{M}_t^h index is

$$\begin{aligned} \mathcal{M}_{\mu t}^h &= 1 - \sum_{i=1}^I \left(\frac{V_{it}}{V_t}\right)^{\alpha} \left(\frac{U_{it}}{U_t}\right)^{1-\alpha} = 1 - \sum_{i=1}^I \left(\frac{v_{it}\mu_i^{\frac{1}{\alpha}}}{\sum_{i=1}^I v_{it}\mu_i^{\frac{1}{\alpha}}}\right)^{\alpha} \left(\frac{u_{it}}{u_t}\right)^{1-\alpha} \\ &= 1 - \sum_{i=1}^I \left(\frac{\mu_i}{\bar{\mu}}\right)^{\alpha} \left(\frac{v_{it}}{v_t}\right)^{\alpha} \left(\frac{u_{it}}{u_t}\right)^{1-\alpha}. \end{aligned}$$

Is it possible to identify measurement error in vacancies μ_i in each sector? With a 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

$$h_{it} = \Phi_t \cdot \mu_i \cdot v_{it}^{\alpha} u_{it}^{1-\alpha}$$

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 precisely the measurement error in vacancies μ_i . Given an estimate of α , one can therefore obtain an estimate of μ_i , precisely as we propose to estimate ϕ_i . To sum up, sectors where vacancies are especially underreported (i.e., $\mu_i \gg 1$) will look like sectors with higher matching efficiency.

4 Data description

We begin this section by describing the data sources. Next we analyze the issue of specification of the matching function at the sectoral level.

4.1 Data Description

Throughout our analysis, we focus on two definitions of labor markets: the first is a broad industry classification and the second is given by the Census regions.¹⁵ The first definition allows us to study skill mismatch while the second is useful to examine geographic mismatch. As we have discussed earlier, our analysis requires detailed information about vacancies, hires, unemployment, and productivity across different labor markets. Vacancy and hire data come from the Job Openings and Labor Turnover Survey (JOLTS) which provides survey-based measures of job openings and hires at a monthly frequency for seventeen industry classifications and four Census regions.¹⁶ Similarly, we calculate unemployment counts from the CPS for the same industry classifications and regions.¹⁷

Computation of mismatch indexes with heterogeneous productive and matching efficiency requires estimates of labor-market specific productivities and matching efficiencies. For the first, we use average hourly earnings from the Current Employment Statistics (CES) with the exception of the government sector. To make definitions of sectors consistent across the CES and JOLTS, we aggregate up the earnings data for some sectors by weighting earnings by employment.¹⁸ For the government sector, we calculate average hourly earnings from the May Outgoing Rotation survey of the CPS. We calculate average hourly earnings using total weekly earnings and hours worked in a week for full time workers.

The calculation of market-specific match efficiency parameters, ϕ_i , is more involved. We use hires and vacancies from the JOLTS and unemployment from the CPS and estimate ϕ_i at the sectoral level. We describe the details below.

4.1.1 Alternative sources of vacancy data

In future work we will also conduct our mismatch analysis for occupations and at a much finer level of geographic detail than the Census regions using vacancy data from the Help Wanted OnLine (HWOL)

¹⁵These are data constrained imposed by JOLTS data. As we explain below, using HWOL data on ads will permit much more disaggregation.

¹⁶For more details on the JOLTS, see <http://www.bls.gov/jlt/>.

¹⁷Note that industry affiliations are not available for all unemployed workers in the CPS. From 2000-2010, on average about 13.3% of unemployed do not have industry information. Some of these workers have never worked before and some are self-employed.

¹⁸In particular, we aggregate up the earnings data for “transportation and warehousing” and “utilities” into one sector by weighting earnings by employment. “Financial Activities” is broken down into 6 disaggregate sectors which we also aggregate up the same way. Earnings data are not reported for the education sector separately. We use earnings for the “education and health” and “health” sectors to back out the earnings data for education.

		CES	Cobb Douglas
σ	Point estimate	-0.074	
	95% Conf. Interval	(-0.267, 0.081)	-
α	Point estimate	0.512	0.532
	95% Conf. Interval	(0.466, 0.551)	(0.514, 0.553)
Φ	Point estimate	0.939	0.943
	95% Conf. Interval	(0.922, 0.958)	(0.925, 0.962)

Table 1: CES vs. Cobb Douglas

dataset provided by The Conference Board (TCB). This is a novel data series that covers the universe of online advertised vacancies posted on internet job boards or on newspaper online editions.¹⁹ The HWOL data base started in May 2005 as a replacement for the Help-Wanted Advertising Index of print advertising maintained by TCB. It covers roughly 1,200 online job boards and provides detailed information about the characteristics of advertised vacancies for several million active ads each month. When the same ad for a given position is posted on multiple job boards, an unduplication algorithm is used that identifies unique advertised vacancies on the basis of (company name, job title/description, city or State).

Each observation in the HWOL data base refers to a unique ad and contains information about the listed occupation at the 6-digit level, the geographic location of the advertised vacancy down to the county level, whether the position is full-time or part-time, the education level of the position, and the hourly and annual mean wage (from BLS data on Occupational Employment Statistics (OES), based on the occupation classification). For a subset of ads we also observe the industry NAICS classification, the sales volume and number of employees of the company, and the advertised salary.

The aggregate trends from the HWOL data base are roughly consistent with those from the JOLTS data. We will use HWOL data to construct mismatch indexes by 2-digit occupation, by State, and by 2-digit occupations times Census divisions. Given the level of detail of the vacancy information from HWOL, the limitations in constructing mismatch indexes arise from the unemployment side because of the relatively small size of the CPS. We also plan to request unemployment insurance data (UI records) from individual states to conduct a more detailed analysis of mismatch for selected states.

4.2 Matching function specification

We start by showing that a matching function with unit elasticity is a reasonable representation of the hiring process at the sectoral level. Using the JOLTS data for the 2-digit definition of sectors (17 industries) and the period December 2000-December 2010, we estimate the parameters of the

¹⁹The data are collected for The Conference Board by Wanted Technologies.

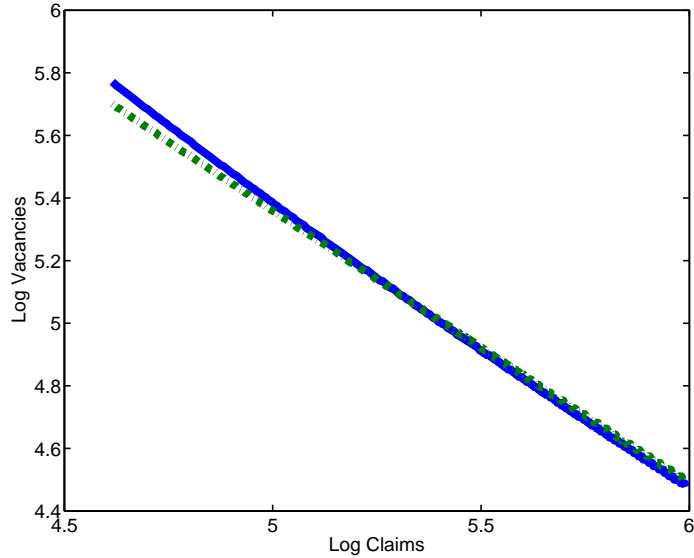


Figure 1: Iso Matching Curves for CES (Solid) and Cobb-Douglas (Dashed)

following CES matching function via minimum distance:²⁰

$$\ln \left(\frac{h_{it}}{u_{it}} \right) = \ln \Phi + \frac{1}{\sigma} \ln \left[\alpha \left(\frac{v_{it}}{u_{it}} \right)^\sigma + (1 - \alpha) \right]. \quad (32)$$

Recall that $\sigma \in (-\infty, 1)$ with $\sigma = 0$ in the Cobb-Douglas case.²¹ As the left column of Table 1 indicates, we find that $\hat{\sigma} = -0.074$ implying an elasticity around 0.93, hence only slightly smaller than the Cobb-Douglas benchmark. Moreover, $\hat{\sigma}$ is not significantly different than zero at the 5% significance level. The right panel of Table 1 reports estimation results for the Cobb-Douglas case (i.e., imposing the constraint $\hat{\sigma} = 0$). The results indicate that there is no statistically significant difference in the estimates $(\hat{\alpha}, \hat{\Phi})$ between the CES and the Cobb-Douglas case; therefore the latter specification is a good approximation for the matching function at this level of aggregation. Figure 1 plots the iso-matching curves for the CES and the Cobb-Douglas specifications 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 and use a Cobb-Douglas matching function throughout the paper.

The next step is to estimate the parameters of the matching function that are required for computing mismatch indexes. We start by estimating an aggregate matching function of the form

$$\ln \left(\frac{h_t}{u_t} \right) = \ln \Phi_t + \alpha \ln \left(\frac{v_t}{u_t} \right)$$

²⁰Note 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.

²¹We use simulated annealing to minimize the minimum distance criterion to ensure that we obtain a global minimum. 95% confidence intervals are computed via bootstrap methods.

	Full Sample		Truncated Sample	
	OLS	Fixed Effects	OLS	Fixed Effects
Aggregate	0.797 (0.014)	- -	0.611 (0.018)	- -
Aggregate (Quadratic Time Trend)	0.673 (0.011)	- -	0.691 (0.026)	- -
Industry	0.529 (0.009)	0.671 (0.012)	0.402 (0.007)	0.504 (0.010)
Industry (Quadratic Time Trend)	0.445 (0.009)	0.556 (0.008)	0.385 (0.013)	0.500 (0.012)

Note: Standard errors in parenthesis. Full sample: Dec 2000-Dec 2010. Truncated sample: Dec 2000 to Dec 2007.

Table 2: Estimates of the vacancy share α

where h_t is the number of matches, u_t is unemployment and v_t in the number of vacancies in month t . We use hires from the JOLTS as our measure of matches.²² Vacancies come from the JOLTS and aggregate unemployment numbers come from the CPS. The first row of Table 2 reports estimates of α for two sample periods. The estimate for α is 0.797 if we use our full sample which spans December 2000 to December 2010. When we constrain the sample to pre-recession data (December 2000 to December 2007), the estimate for α is lower at 0.611. As we have discussed earlier, there is potentially some time variation in ϕ . This is likely to cause a difference between the two estimates of α obtained with two different sample periods. To capture the time variation in Φ , we run a similar regression with a quadratic time trend: the results are reported in the second row of Table 2. With the quadratic time trend, estimates of α are much closer for the full sample and the pre-recession sample at around 0.67 – 0.69.

In addition to the aggregate regressions, we also exploit industry-level data on hiring, vacancies and unemployment and estimate the following regression

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

for both our full and pre-recession samples. We constrain Φ_t to be the same across sectors and allow for a quadratic time trend to control for time variation. The results are reported in the last two rows of Table 2, in the columns labeled “OLS”. The estimates of α are lower than the ones estimated by the aggregate regression varying between 0.38 and 0.53. As in the case of aggregate regressions, allowing for time variation lowers the estimate of α .

²²An 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.

Sector	ϕ
Arts	1.50
Construction	1.46
Mining	1.37
Accommodations	1.32
Retail	1.25
Professional Business Services	1.19
Real Estate	1.15
Wholesale	1.05
Other	0.98
Transportation and Utilities	0.98
Manufacturing - Nondurables	0.83
Education	0.82
Health	0.80
Government	0.74
Manufacturing - Durables	0.71
Finance	0.69
Information	0.63

Table 3: Industry-Specific Matching Efficiencies

Finally, we allow for match efficiencies to vary across sectors and estimate:

$$\ln \left(\frac{h_{it}}{u_{it}} \right) = \ln \Phi_t + \ln \phi_i + \alpha \ln \left(\frac{v_{it}}{u_{it}} \right) \quad (33)$$

The estimation results are reported in the last two rows of Table 2, in the columns labeled “Fixed Effects”. In these cases, estimates of α vary between 0.50 – 0.67 with higher estimates when we use the full sample.

To summarize, our analysis shows that it is important to control for time and sectoral variation in (ϕ_i) . In light of our analysis, we choose $\alpha = 0.60$ throughout the paper and provide some sensitivity analysis to the choice of the vacancy share value.

Estimation of (33) also provides us with sector-specific estimates of match efficiency (ϕ_i) . These estimates are reported in Table 3. Industry-specific match efficiency estimates (ϕ_i) vary considerably and are between 0.63 to 1.5. Among the industries, education, health, finance, and information stand out as low-efficiency sectors while construction stands out as a high efficiency sector. 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) . High efficiency might also be an outcome of different hiring practices in different industries, as well as underreported vacancies as discussed in Davis, Faberman, and Haltiwanger (2010).²³

²³Recall that in, Section 3.4.1, we showed that ϕ_i is proportional to underreported vacancies, when the latter are reported with error.

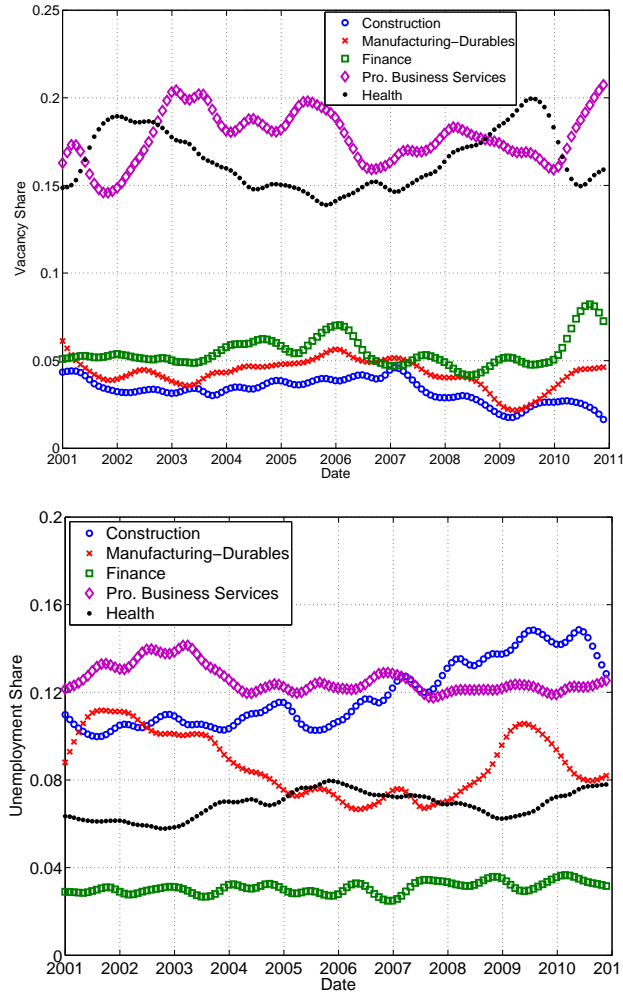


Figure 2: Vacancy and Unemployment Share by Sector

4.3 A First Look At Mismatch

It is useful to examine the vacancy and unemployment shares of different sectors and regions for a preliminary investigation of mismatch since these statistics are inputs into our mismatch indexes. If vacancy and unemployment shares of different sectors and regions do not vary over time, there is little room for mismatch to play an important role in the increase in the unemployment rate. To examine this issue, we plot the vacancy and unemployment shares for a selected set of sectors. As Figure 2 shows, the shares have been relatively flat in the 2004-2007 period. However, starting in 2007, vacancy shares started to change noticeably. Construction and durable goods manufacturing were among the sectors which experienced a decline in their vacancy shares while the health sector saw its vacancy share increase. Concurrently, unemployment shares of construction and durables good manufacturing went up while the unemployment share of the health sector decreased. Interestingly starting from 2010, unemployment and vacancy shares of sectors began to normalize and almost went

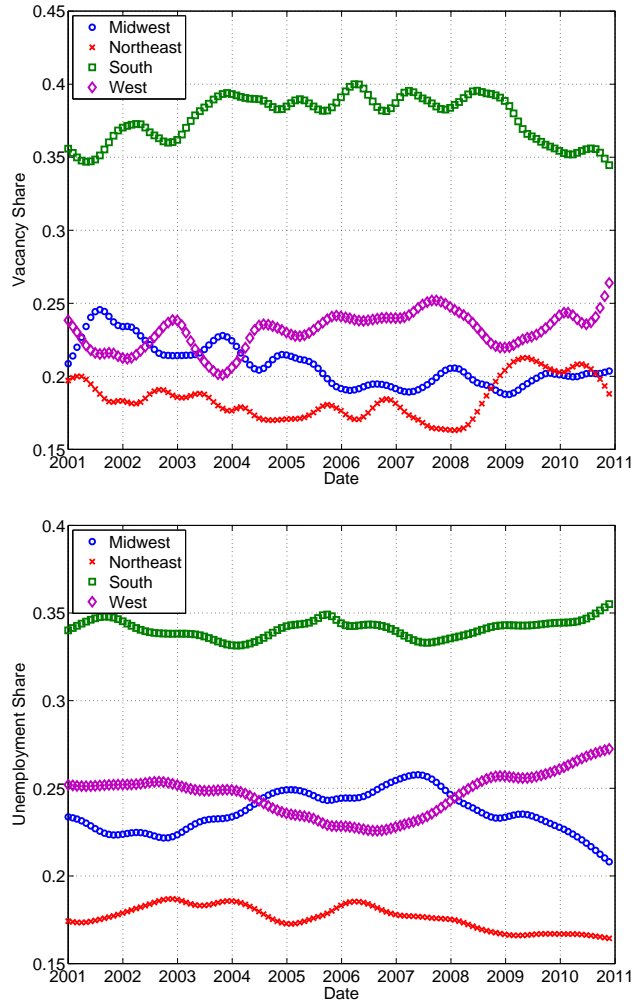


Figure 3: Vacancy and Unemployment Share by Census Region

back to their pre-recession levels with the exception of the construction sector. The vacancy share of the construction sector remains well below its pre-recession level.

Figure 3 also shows the behavior of vacancy and unemployment shares by region. The West experienced an increase in its unemployment share and a mild decline in its vacancy share while the South saw a decline in its vacancy share. The Midwest and the Northeast fared relatively better with declines in their unemployment shares and slight increases in their vacancy shares. The figures suggest that there is less variation in shares by region potentially suggesting a less important role for geographic mismatch relative to skill mismatch.

5 Results based on JOLTS vacancy data

We present a first set of results on mismatch unemployment across the four Census regions and the 17 industries classified in JOLTS.

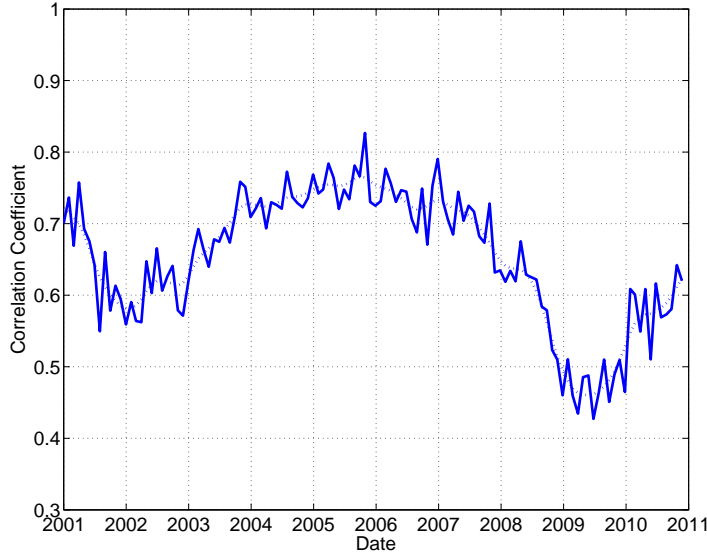


Figure 4: Correlation Coefficient Between u and v Shares

5.1 Industry-level mismatch

From our definition of mismatch in the labor market, it is clear that there is a close association between mismatch indexes and the correlation between unemployment and vacancy shares across sectors. Figure 4 plots the time series of this correlation coefficient across industries over the sample period. The coefficient drops from 0.75 in mid 2006 to 0.45 in mid 2009, and recovers thereafter, indicating a rise in mismatch during the recession. We should expect the mismatch indexes to show a similar pattern.

The top panel of Figure 5 plots the \mathcal{M}_t^u indexes in their various versions described in Section 3: the plain index, \mathcal{M}_t^u , the one adjusted for heterogeneity in matching efficiency, $\mathcal{M}_{\phi t}^u$, the one adjusted for heterogeneity in productivity $\mathcal{M}_{z t}^u$, and, finally, the one modified to account for both sources, $\mathcal{M}_{x t}^u$. All the adjusted indexes appear as shifted versions of the plain index and paint a consistent picture: the fraction of unemployed workers misallocated, i.e., searching in the wrong sector, increased by 10 percentage points from early 2007 to mid 2009, and then dropped somewhat but remained at a higher level than its pre-recession level. Figure 6 plots some of the key components of this index, i.e., the difference between unemployment and vacancy shares in various sectors. It clearly displays how, over the 2006-2009 period, this gap grew in construction and manufacturing (and in finance somewhat) and fell in the health sector. These diverging trends explain the rise in industry-level mismatch.

Turning to the index \mathcal{M}_t^h measuring the fraction of hires lost because of the misallocation of unemployed workers across industries, bottom panel of Figure 5 shows that, before the last recession, this fraction ranged from 1 to 3 percent per month, depending on the index used. At the end of the recession, in mid 2009, it had increased to 4-8 percent per month, and then it dropped again. To sum up, both \mathcal{M}_t^u and \mathcal{M}_t^h indicate a rise in mismatch between unemployed workers and vacant jobs

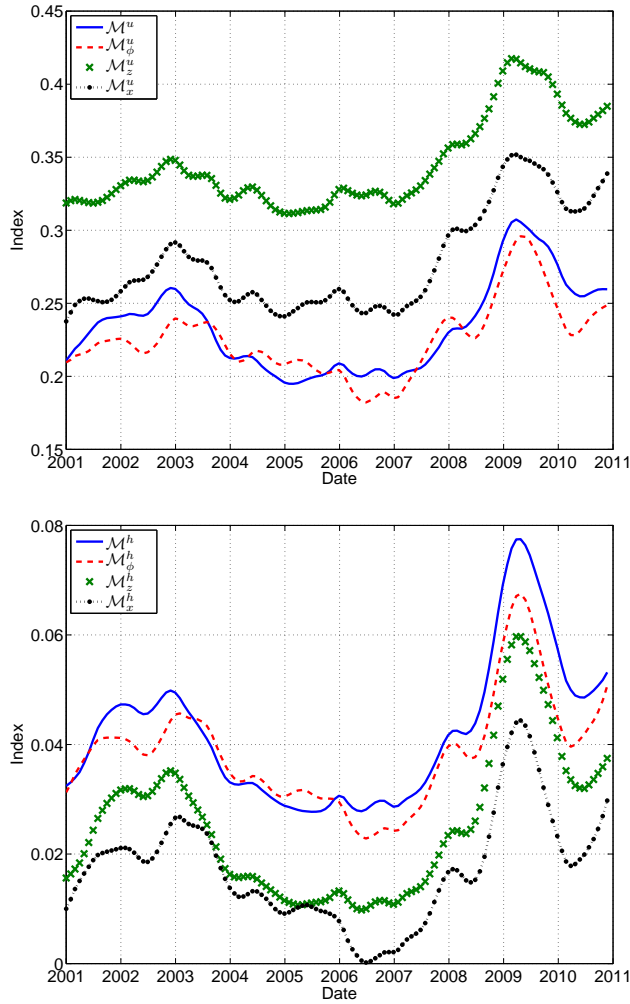


Figure 5: Mismatch Index \mathcal{M}_t^u (top panel) and \mathcal{M}_t^h (bottom panel) by Industry

across industries during the recession, and a subsequent fairly rapid decline.

We now turn to the counterfactual exercises described in Section 3.4. Figure 7 illustrates the results of the counterfactual on the shift in the aggregate Beveridge curve. The dashed line represents the empirical shift estimated from the data, i.e., the term $\log A_t$ of equation (29). The other two lines are the shifts induced by mismatch computed based on the indexes \mathcal{M}_t^h and \mathcal{M}_{xt}^h . The message of this exercise is that only a minimal part of the observed decline in the aggregate job finding rate is due to a rise in mismatch unemployment across industries.

The four panels of Figure 8 contain the observed unemployment rate and the counterfactual unemployment rates constructed following the strategy of Section 3.4. The main finding is that worsening mismatch across industries explains between 0.4 and 0.7 percentage points of the 5 percentage point rise in US unemployment since the beginning of 2007, depending on the index used, i.e., at most 14 percent of the increase.

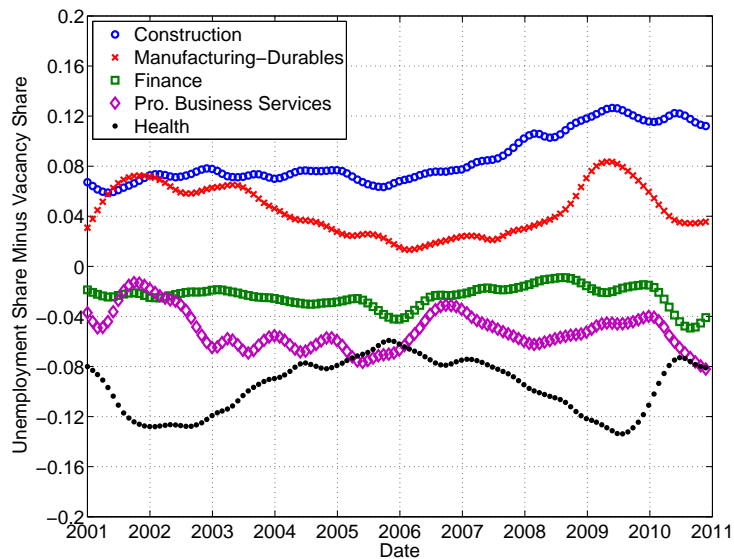


Figure 6: Key Components of \mathcal{M}_t^u

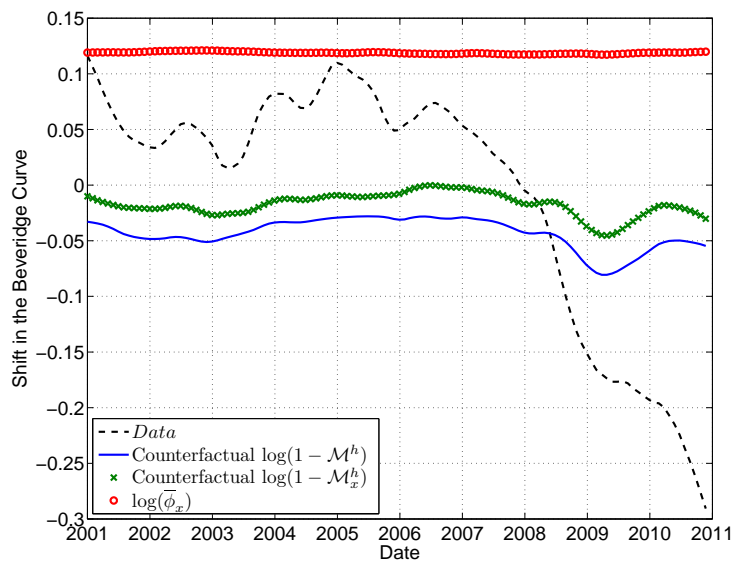


Figure 7: Beveridge Curve Counterfactual

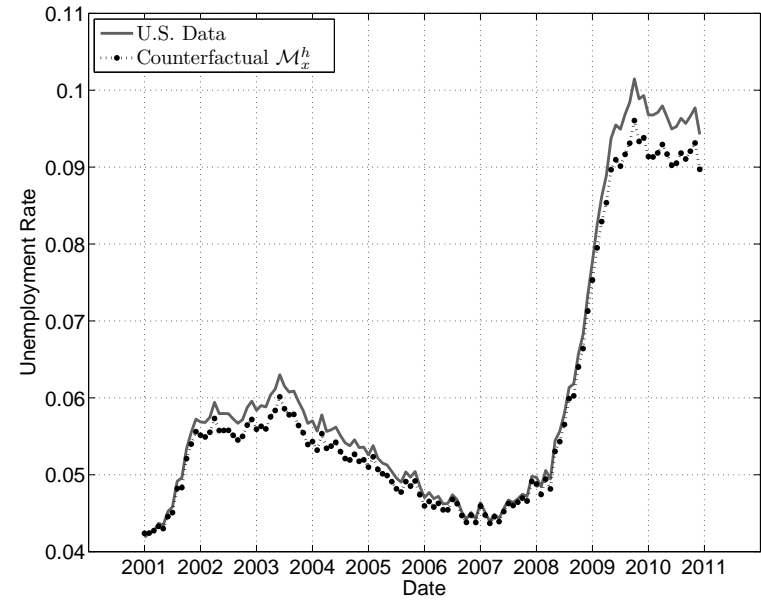
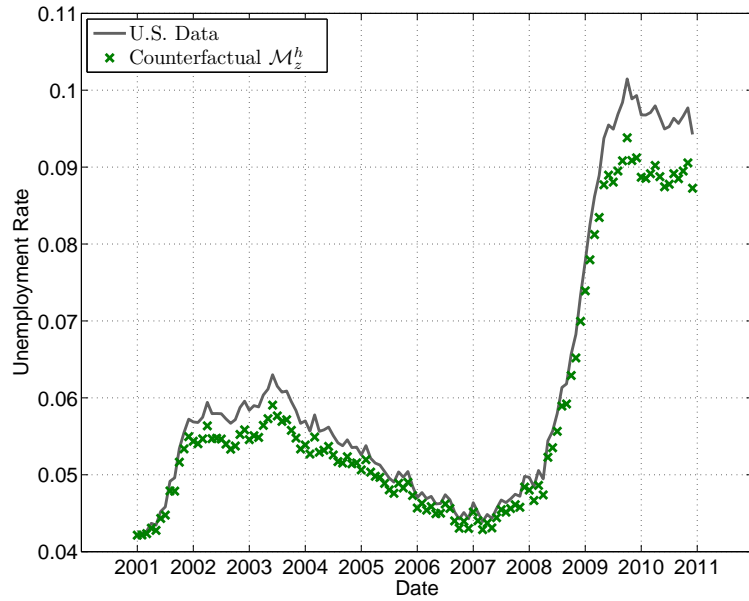
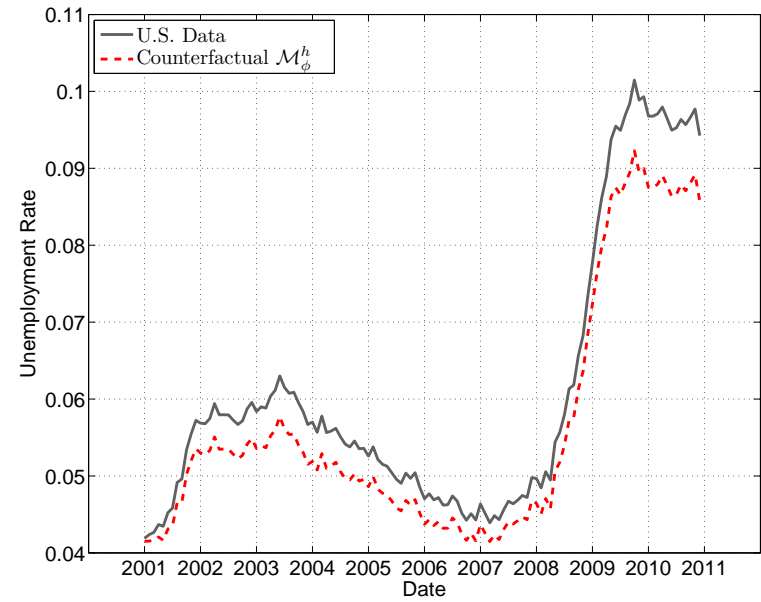
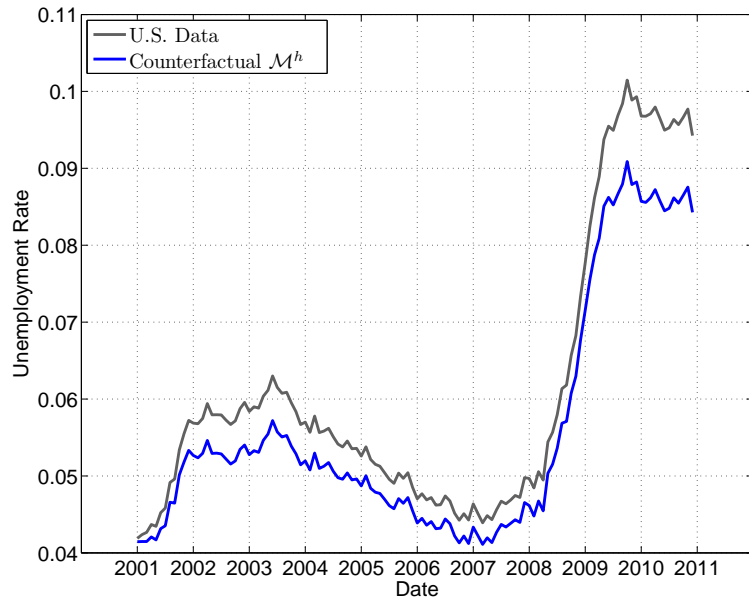


Figure 8: Counterfactual Unemployment Rates: Industry

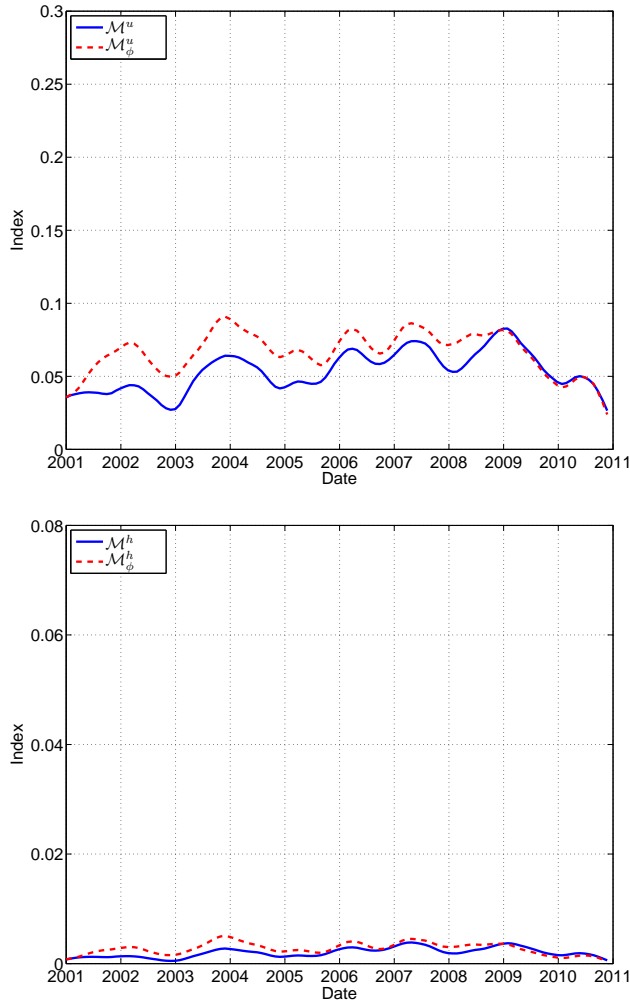


Figure 9: Mismatch Index \mathcal{M}_t^u (left panel) and \mathcal{M}_t^h (right panel) by Census Region

5.2 Geographical mismatch

Figure 9 plots the \mathcal{M}_t^u and \mathcal{M}_t^h mismatch indexes across the four Census regions, using the same scale as for the industry-level indexes. Regional mismatch is very low and does not show any significant trend. Unsurprisingly, the counterfactual unemployment computed based on these indexes (see Figure 10) is essentially on top of the actual series, meaning that geographical mismatch –across Census regions– plays no role in the recent dynamics of US unemployment.

6 Conclusion

We have developed a theoretical framework that gives rise to a well defined notion of mismatch in the labor market. We consider a dynamic stochastic economy with many distinct frictional labor markets,

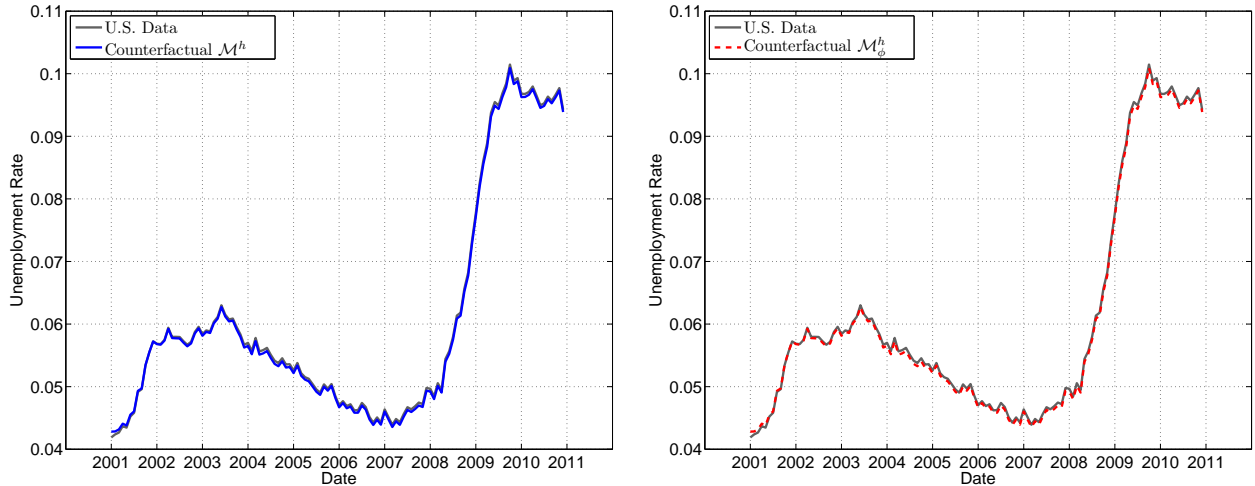


Figure 10: Counterfactual Unemployment Rates: Census Region

and compare the actual distribution of unemployment with the optimal allocation resulting from the solution to a planner’s problem. With the distribution of vacancies being determined exogenously every period, the planner maximizes utility over allocations of unemployment taking as given any search and matching frictions within each market, but assuming costless mobility of the unemployed across markets.

The solution to this planner’s problem constitutes in our view a clean benchmark to think about the extent of misallocation in the labor market. Our framework yields a set of generalized Jackman-Roper (JR) conditions in a dynamic setting with heterogeneous productivities and match efficiencies across markets. The generalized JR conditions can be easily used to construct mismatch indices that measure the fraction of unemployed searching in the “wrong” markets.

We also derive a second family of indices that capture the fraction of hires lost because of mismatch. These indices can be used to compute a counterfactual series for frictional unemployment in the absence of mismatch. They can also be used to decompose any observed shift of the Beveridge curve into a portion coming from changes in aggregate match efficiency and a portion arising from mismatch.

In the empirical part of the paper we use vacancy data by industry from JOLTS (and soon HWOL data by occupation and detailed geographic location) to compute our indices for the period 2000-2010. We find that the rise in mismatch at the sectoral level can explain less than one percentage point of the observed increase in the unemployment rate from the start of the recession to 2010. We also find that sectoral mismatch can explain only a small fraction of the observed shift in the Beveridge curve. Finally, we calculate geographic mismatch measures and find no role for geographic mismatch in explaining the increase in the unemployment rate.

In future work we plan to extend our framework to take explicitly into account moving and re-

training costs across markets (location, industries, occupations or any combinations thereof). This will allow us to construct measures of mismatch based on constrained efficient solutions to the planner's problem with various degrees of frictions across markets. A further extension of this framework is to model mismatch as an equilibrium outcome, to be able to study the impact of various kinds of shocks and distortions on mismatch.

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] Barnichon, R. and A. Figura (2011). "What Drives Matching Efficiency? A Tale of Composition and Dispersion," FRB Staff WP 2011-10.
- [4] Birchenall, J. A. (2010). "A Competitive Theory of Mismatch," mimeo, UCSB.
- [5] Carrillo-Tudela and L. Visscher (2010). "Search and Endogenous Reallocation over the Business Cycle," mimeo.
- [6] Davis, S. J., Faberman, R. J., Haltiwanger, J. C. (2010). "The Establishment-Level Behavior of Vacancies and Hiring," NBER Working Paper 16265.
- [7] Elsby, M., B. Hobijn, and A. Şahin (2010). "The Labor Market in the Great Recession," *Brookings Papers on Economic Activity*, Spring 2010, 1-48.
- [8] Ferreira, Fernando, J. Gyourko, and J. Tracy (2010). "Housing Busts and Household Mobility," *Journal of Urban Economics*, 68(1), 34-45.
- [9] Fujita, Shigeru (2011). "Effects of Extended Unemployment Insurance Benefits: Evidence from the Monthly CPS," Working Paper 10-35/R.
- [10] Hall, R.E. (2010). "The Labor Market in the Current Slump," mimeo.
- [11] Hsieh, C. and P. Klenow (2009). "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124, 1403-1448.
- [12] Hornstein, A., P. Krusell, and G. Violante (2007). "Technology-Policy Interaction in Frictional Labor Markets," *Review of Economic Studies*, 74, 1089-1124.
- [13] Jackman, R. and S. Roper (1987). "Structural Unemployment," *Oxford Bulletin of Economics and Statistics*, 49, 9-36.
- [14] Kocherlakota, N. (2010). "Inside the FOMC," Speech at Marquette, Michigan on August 17, 2010.
- [15] Lagos, R. (2006). "A Model of TFP," *Review of Economic Studies* 73, 983-1007.

- [16] Lilien, D.M. (1982). "Sectoral Shifts and Cyclical Unemployment," *Journal of Political Economy*, 90, 777-793.
- [17] Ljungqvist, L. and T. J. Sargent (1998). "The European Unemployment Dilemma," *Journal of Political Economy* 106, 514-550.
- [18] Molloy, R. C. Smith, and A. Wozniak (2010). "Internal Migration in the US: Updated Facts and Recent Trends," mimeo.
- [19] Mortensen, D.T. (2009). "Island Matching," *Journal of Economic Theory*, 144(6), 2336-2353.
- [20] Mortensen, D.T. and Pissarides, C.A. (1999). "Unemployment Responses to Skill-biased Technology Shocks: the Role of Labour Market Policy," *Economic Journal*, 109(455), 242-265.
- [21] Padoa-Schioppa F. (1991). Mismatch and Labour Mobility. Centre for Economic Policy Research.
- [22] Phelps, E. (2008). "U.S. Monetary Policy and the Prospective Structural Slump," Speech to the 7th Annual BIS Conference on Monetary Policy, Luzern, Switzerland.
- [23] Restuccia, D. and R. Rogerson (2008). "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 11, 707-720.
- [24] Şahin, A., J. Song, G. Topa and G. Violante (2011). "Does Mismatch Explain the Recent Rise in U.K. Unemployment?," mimeo.
- [25] Schulhofer-Wohl, S. (2010). "Negative Equity Does Not Reduce Homeowners' Mobility," mimeo.
- [26] Shimer R. (2005). "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, 95, 25-49.
- [27] Shimer R. (2007a). "Mismatch," *American Economic Review*, 97, 1074-1101.
- [28] Shimer R. (2007b). "Reassessing the Ins and Outs of Unemployment," mimeo.
- [29] Valletta, R. and K. Kuang (2010). "Extended Unemployment and UI Benefits," *FRBSF Economic Letter*, 2010-12.