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: JULY 2011

PRELIMINARY

Abstract

This paper measures mismatch between job-seekers and vacancies 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 between sectors. The planner's optimal allocation is dictated by a "generalized Jackman-Roper condition" where (productive and matching) efficiency-weighted vacancy-unemployment ratios are equated across sectors. We develop this condition into mismatch indexes that allow us to quantify how much of the recent rise in U.S. unemployment is due 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. Higher mismatch across industries and occupations accounts for 0.8 to 1.4 percentage points of the recent rise in the unemployment rate, whereas geographical mismatch plays no role. We find that the role of mismatch in explaining the increase in unemployment varies considerably by education. Occupational mismatch explains a substantial fraction of the rise in unemployment (one third) for highly educated workers while it is quantitatively less important for less educated workers.

*We wish to thank Björn Brügemann, Mark Gertler, John Haltiwanger, Rob Shimer, and seminar participants at Bocconi, Chicago Fed, the 7th ECB/CEPR Labour Market Workshop, NBER Summer Institute, Maryland, Riksbank, Rochester, SF Fed, Sciences Po, SUNY at Albany, the Philadelphia Fed-Kiel Institute "Labor Market After the Great Recession" Conference, Universitat Autonoma de Barcelona, and Western Ontario for very helpful comments. Scott Nelson and Victoria Gregory provided excellent research assistance. 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 accept 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 task requirements 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 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 portion 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, generally referred to as "house lock," appears consistent with recent data that show 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 relationship 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 be-

[1] Various 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.

[2] This observation has been emphasized before by Davis, Faberman, and Haltiwanger (2010), Elsby, Hobijn, and Şahin (2010), Hall (2010), and others.

tween unemployment and vacancies across labor markets is one 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, occupation, skill or education, 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 fea-

[3]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.

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

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

ture 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 indexes. 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 rate 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 much lower equilibrium unemployment would be in the absence of mismatch, one needs to understand how the job-finding rate would change. 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 2.3.

We apply our analysis to the U.S. labor market and construct measures of mismatch across industries, occupations and geographic areas using vacancy data from the Job Openings and Labor Turnover Survey (JOLTS) and from the Conference Board's Help Wanted OnLine (HWOL) database, and unemployment data from the Current Population Survey (CPS).[6] We find that mismatch at the industry and occupation level increased during the recession and started to come down in 2010; an indication of a cyclical pattern for mismatch. Our calculations show that mismatch accounted for at most 0.8 to 1.4 percentage points of the total increase in the unemployment rate from the start of the recession to 2010 (around 5 percentage points). We also calculate geographic mismatch measures

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

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.[7] When we perform our study of occupational mismatch separately for different educational groups, we find that the portion of the rise in unemployment explained by the rise in mismatch rises steeply with education. This result is consistent with the view that the human capital of the highly educated is more specialized.

Our paper relates to an old, mostly empirical, literature that popularized the idea of mismatch (or what used to be called '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.[8] 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 and, hence, can be a source of variation in the job finding rate. Our approach is different and our scope broader, but we also show that fluctuations in mismatch act as shifts in the aggregate matching function.

At a more theoretical level, Shimer (2007a) and Mortensen (2009) were the first to develop the idea that an economy with many separate labor markets, and misallocation of job-seekers and vacancies across markets, could be empirically consistent with the aggregate Beveridge curve. In this set-up, workers are assigned randomly to markets. Alvarez and Shimer (2010), Birchenall (2010), Carrillo-Tudela and Visscher (2010), and Hertz and Van Rens (2011) have all proposed dynamic models with explicit mobility decisions across labor markets where unemployed workers, in equilibrium, may be mismatched. While less amenable to measurement than our framework, these models are better suited to study the deeper causes of mismatch.

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 compute our counterfactuals. Section 4 addresses some measurement issues. Section 5 describes the data. Section 6 performs the empirical analysis. Section 7 concludes.

[7] See, for example, Molloy, Smith, and Wozniak (2010) and Schulhofer-Wohl (2010).

[8] 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; 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.[9]

Time is discrete. The economy is comprised 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 $\mathbf{e} = \{e_1, ...e_I\}$ across markets and the total number of unemployed workers u are 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

[9] 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.

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's employment distribution $\{e_i'\}$. Finally, labor force decisions for next period are taken. Given ℓ' and $\{e_i'\}$, the stock of unemployed workers u' for next period is also 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:

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

The per period net output for the planner is equal 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's 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

$$\gamma Z z_i \Phi \phi_i m_u \left(\frac{v_i}{u_i}\right) + \beta \mathbb{E}\left[-V_u'(\cdot) + V_{e_i}'(\cdot)\right](1-\delta)\Phi\phi_i m_u \left(\frac{v_i}{u_i}\right) = \mu, \tag{7}$$

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

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

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

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 that 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:

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

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

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

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.[10] 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 matched and 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

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

[10] 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 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

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

Let us verify the conjecture:

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

which confirms the conjecture, since $\mathbb{E}\left[\Psi(Z',\delta',\Phi')\right]$ 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

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

and rearranging:

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

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:

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

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 vacancies and matching and productive efficiency 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 Extensions

Heterogeneous destruction rates. We now relax the assumption that the destruction rate δ is common across sectors. Denote the idiosyncratic component of the exogenous destruction rate in sector i as δ_i. To simplify the exposition, we set $\gamma = 1$ and assume that $\{Z, z_i, \delta, \delta_i\}$ all follow independent unit root processes. The envelope condition (9) becomes

$$V_{e_i} = Zz_i + \beta(1-\delta)(1-\delta_i)\mathbb{E}\left[V_{e_i}'\right].$$

Solving forward, and using the unit root assumption, we arrive at:

$$V_{e_i} = \frac{Zz_i}{1-\beta(1-\delta)(1-\delta_i)}$$

which, substituted into the (appropriately modified) equation (13) yields

$$Zz_i\Phi\phi_i m_u\left(\frac{v_i}{u_i}\right)+\frac{\beta\left(1-\delta\right)\left(1-\delta_i\right)}{1-\beta\left(1-\delta\right)\left(1-\delta_i\right)}Zz_i\Phi\phi_i m_u\left(\frac{v_i}{u_i}\right)=\mu.$$

Rearranging, we arrive at a modified 'generalized Jackman-Roper condition' where the planner equalizes

$$\frac{z_i\phi_i}{1-\beta\left(1-\delta\right)\left(1-\delta_i\right)}m_u\left(\frac{v_i}{u_i^*}\right) \tag{15}$$

across sectors. The new term captures the fact that the expected output of an unemployed in sector i is discounted differently by the planner in different sectors because of the heterogeneity in destruction rates.

Heterogeneous sensitivities to aggregate shocks. In our baseline model, one of the sources of reallocation of labor is sector-specific labor demand shifts (z_i). In a classic paper disputing Lilien's (1982) sectoral shift theory of unemployment, Abraham and Katz (1986) argue that, empirically, sectoral employment movements appear to be driven by aggregate shocks with different sectors having different sensitivities to the aggregate cycle. We show here that, under this alternative interpretation of what drives sectoral labor demand, our key result goes through under a minimal set of additional assumptions.

Let $Zz_i = Z^{\eta_i}$ where η_i is a sector specific parameter measuring the sensitivity of output in sector i to the aggregate shock Z. Let $\ln Z'$ follow a unit root process, with conditional distribution $N\left(\ln Z-\sigma^2/2,\sigma^2\right)$. Note that $\mathbb{E}\left[Z'z_i'\right]=Z^{\eta_i}\exp\left(\eta_i\left(\eta_i-1\right)\frac{\sigma^2}{2}\right)$. Using this result in the envelope condition (9) yields

$$V_{e_i}=\frac{Z^{\eta_i}\exp\left(\eta_i\left(\eta_i-1\right)\frac{\sigma^2}{2}\right)}{1-\beta\left(1-\delta\right)}$$

which, substituted into the (appropriately modified) equation (13), yields

$$Z^{\eta_i}\Phi\phi_i m_u\left(\frac{v_i}{u_i}\right)+\beta\left(1-\delta\right)\frac{\exp\left(\eta_i\left(\eta_i-1\right)\frac{\sigma^2}{2}\right)}{1-\beta\left(1-\delta\right)}Z^{\eta_i}\Phi\phi_i m_u\left(\frac{v_i}{u_i}\right)=\mu.$$

Rearranging, we obtain yet another modified 'generalized Jackman-Roper condition' where the planner equalizes

$$Z^{\eta_i}\phi_i m_u\left(\frac{v_i}{u_i^*}\right)\left[1+\beta\left(1-\delta\right)\left(\exp\left(\eta_i\left(\eta_i-1\right)\frac{\sigma^2}{2}\right)-1\right)\right] \tag{16}$$

across sectors. Given estimates of $\{\eta_i\}$ and of the variance of the aggregate shock, the expression above can be easily computed.

2.3 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 (plan-

ner'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 U.S. 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.[11] Within our approach, we can still learn about the deep sources of mismatch by examining how mismatch varies as we use different definitions of sectors (occupation, industry, location, education).

[11]For 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.

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 the recent rise in U.S. 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

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

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

$$\mathcal{M}_t^u = \frac{u_t^M}{u_t} = \frac{1}{2}\sum_{i=1}^{I}|\frac{u_{it}}{u_t} - \frac{v_{it}}{v_t}|. \tag{17}$$

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 searched 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:

$$h_t = \Phi_t v_t^{\alpha} u_t^{1-\alpha} \cdot \left[\sum_{i=1}^{I} \left(\frac{v_{it}}{v_t} \right)^{\alpha} \left(\frac{u_{it}}{u_t} \right)^{1-\alpha} \right]. \tag{18}$$

The first term in (18) 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:

$$\mathcal{M}_t^h = 1 - \frac{h_t}{h_t^*} = 1 - \sum_{i=1}^{I} \left(\frac{v_{it}}{v_t} \right)^{\alpha} \left(\frac{u_{it}}{u_t} \right)^{1-\alpha}. \tag{19}$$

The index $\mathcal{M}_t^h$ measures precisely what fraction of hires is lost because of misallocation.[12] It is easy to see that $\mathcal{M}_t^h \leq 1$. To show that $\mathcal{M}_t^h \geq 0$, note that

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

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 increasing 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

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

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

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

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

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.

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

$$u_{it}^* = \phi_i^{\frac{1}{\alpha}}\cdot\left(\frac{v_{it}}{v_{\phi t}}\right)\cdot u_t. \tag{20}$$

Recall that the share of unemployed workers searching in the wrong sector is $u_t^M = \frac{1}{2}\sum\limits_{i=1}^{I}|u_{it}-u_{it}^*|$. Substituting the expression for u_{it}^* from (20) 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

$$\mathcal{M}^u_{\phi t} = \frac{u^M_t}{u_t} = \frac{1}{2}\sum_{i=1}^{I} |\frac{u_{it}}{u_t} - \left(\frac{\phi_i}{\bar{\phi}_t}\right)^{\frac{1}{\alpha}} \cdot \frac{v_{it}}{v_t}| \tag{21}$$

where

$$\bar{\phi}_t = \left[\sum_{i=1}^{I} \phi_i^{\frac{1}{\alpha}} \left(\frac{v_{it}}{v_t}\right)\right]^{\alpha} \tag{22}$$

is a CES aggregator of the market-level matching efficiencies weighted by their vacancy share. The index in (21) is similar to the index (17) 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 $\mathcal{M}^h_{\phi t}$ index. The optimal aggregate number of hires is

$$h^*_t = \Phi_t v_t^{\alpha} u_t^{1-\alpha} \left[\sum_{i=1}^{I} \phi_i \left(\frac{v_{it}}{v_t}\right)^{\alpha} \left(\frac{u^*_{it}}{u_t}\right)^{1-\alpha}\right]. \tag{23}$$

Substituting the optimality condition (20) in equation (23), 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 (22). Similarly, we can define the total number of observed new hires as

$$h_t = \Phi_t v_t^{\alpha} u_t^{1-\alpha} \left[\sum_{i=1}^{I} \phi_i \left(\frac{v_{it}}{v_t}\right)^{\alpha} \left(\frac{u_{it}}{u_t}\right)^{1-\alpha}\right], \tag{24}$$

and hence the counterpart of (19) in the heterogeneous markets case becomes

$$\mathcal{M}^h_{\phi t} = 1 - \frac{h_t}{h^*_t} = 1 - \sum_{i=1}^{I} \left(\frac{\phi_i}{\bar{\phi}_t}\right) \left(\frac{v_{it}}{v_t}\right)^{\alpha} \left(\frac{u_{it}}{u_t}\right)^{1-\alpha}. \tag{25}$$

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:

$$\frac{v_{1t}}{u^*_{1t}} = \left(\frac{x_i}{x_1}\right)^{\frac{1}{\alpha}} \cdot \frac{v_{it}}{u^*_{it}}. \tag{26}$$

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

$$\mathcal{M}^u_{xt} = \frac{u^M_t}{u_t} = \frac{1}{2}\sum_{i=1}^{I} |\frac{u_{it}}{u_t} - \left(\frac{x_i}{\bar{x}_t}\right)^{\frac{1}{\alpha}} \cdot \frac{v_{it}}{v_t}| \tag{27}$$

where

$$\bar{x}_t = \left[\sum_{i=1}^{I} x_i^{\frac{1}{\alpha}} \left(\frac{v_{it}}{v_t} \right) \right]^{\alpha} \tag{28}$$

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 (23). Substituting the optimality condition (26) in equation (23), 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 (24), we obtain the counterpart of (25)

$$\mathcal{M}_{xt}^h = 1 - \sum_{i=1}^{I} \left(\frac{\phi_i}{\bar{\phi}_{xt}} \right) \left(\frac{v_{it}}{v_t} \right)^{\alpha} \left(\frac{u_{it}}{u_t} \right)^{1-\alpha}, \tag{29}$$

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 $i = 1, 2, ..., I$ and dates $t = 1, 2, ..., T$, and 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 5 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.

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.[13] A worse misallocation of unemployed workers

[13] We 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 assumption does not cause a cyclical bias on the effect of mismatch on the unemployment rate.

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. By comparing optimal hires $h_t^* = \Phi_t \bar{\phi}_{xt} v_t^{\alpha} u_t^{1-\alpha}$ to actual hires, we can write $h_t = \left(1 - \mathcal{M}_{xt}^h\right) \cdot \bar{\phi}_{xt} \cdot \Phi_t \cdot v_t^{\alpha} u_t^{1-\alpha}$. If, using this equation, we let

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

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

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

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

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

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 vacancy creation and separations are exogenous (recall that voluntary quits are zero for the planner).

4 Measurement Issues Related to Unemployment and Vacancies

In this section we discuss two measurement issues related to unemployment and vacancy statistics we use in our empirical analysis: 1) inferring the labor markets that unemployed workers are searching in, and 2) correcting for unreported vacancies.

4.1 Adjustment of unemployment count

In the baseline analysis of Section 2, we classify an unemployed worker as unemployed in sector i if her last job was in sector i. Unfortunately, CPS collects no information on where the worker is directing her search. However, using the panel dimension of CPS, it is possible to observe, for unemployed who find jobs from one month to the next, in which sector they were reemployed. We show below that, under some assumptions, this is enough to infer where they were searching.

Consider an economy with I sectors. Let u_i be the unemployed worker whose last job is in sector i, and u_i^* be the true number of unemployed actually searching in sector i. Finally, let u_i^j be the number of unemployed whose last job is in sector i but searching in sector j. By definition, we have

$$u_i = \sum_{j=1}^{I} u_i^j.$$

Suppose we observe h_i^j, the number of unemployed workers hired in sector j whose last job was in sector i. Let the total number of hires in sector j be h^j. Assume that the job finding rate in sector j is the same for all unemployed, independently of the sector of provenance, except if their previous job was in that same sector. Then:

$$\frac{h_i^j}{u_i^j} = \xi_i \frac{h^j}{u_j^*} \text{ for all } i = 1, ..., I$$

where $\xi_i = \xi > 1$ for $i = j$, and $\xi_i = 1$ otherwise. Rearrange the above equation as

$$u_i^j = \frac{1}{\xi_i} \left(\frac{h_i^j}{h_j} \right) u_j^* \tag{30}$$

and sum across all j to obtain the I linear equations

$$u_i = \frac{1}{\xi_i} \sum_{j=1}^{I} \left(\frac{h_i^j}{h_j} \right) \cdot u_j^*, \text{ for all } i = 1, ..., I,$$

in the $(I + 1)$ unknowns $\left\{u_j^*, \xi\right\}$. The last equation needed to make the system determinate is the "aggregate consistency" condition

$$\sum_{j=1}^{I} u_j^* = \sum_{j=1}^{I} u_j, \tag{31}$$

stating that the true distribution of unemployed across sectors must sum to the observed total number of unemployed.

Even though this exactly identified system of linear equations has a unique solution, we have no guarantee that this solution is a nonnegative vector. We return on this point in the empirical section.

4.2 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

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

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

$$\begin{aligned} \mathcal{M}_{\mu t}^u &= \frac{1}{2}\sum_{i=1}^{I} |\frac{u_{it}}{u_t} - \frac{\mu_i^{\frac{1}{\alpha}}}{\sum_{i=1}^{I} \mu_i^{\frac{1}{\alpha}} \left(\frac{v_{it}}{v_t}\right)} \cdot \frac{v_{it}}{v_t}| \\ &= \frac{1}{2}\sum_{i=1}^{I} |\frac{u_{it}}{u_t} - \left(\frac{\mu_i}{\bar{\mu}}\right)^{\frac{1}{\alpha}} \cdot \frac{v_{it}}{v_t}| \end{aligned} \tag{32}$$

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

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

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., $\mu_i >> 1$) will look like sectors with higher matching efficiency.

5 Data and Sectoral Matching Functions

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

5.1 Data Description

Throughout our analysis, we focus on three definitions of labor markets: the first is a broad industry classification, the second is a broad (2-digit) occupation classification, and the third is a geographic classification, based on U.S. states. The first two definitions allow us to study skill mismatch while the last one is used to examine geographic mismatch. In addition, we also study mismatch within and across four education categories, based on educational attainment.

As we have discussed earlier, our analysis requires detailed information about vacancies, hires, unemployment, and productivity across different labor markets. Vacancy and hire data at the industry level 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.[14] The JOLTS also provides limited geographic information, enabling us to study mismatch across four broad Census regions. At the occupation and state level we use vacancy data from the Help Wanted OnLine (HWOL) dataset provided by The Conference Board (TCB). The HWOL data also allow us to classify vacancies by education level. We describe these data in more detail below. With regard to the unemployed, we calculate unemployment counts from the CPS for the same industry, occupation, geography and education classifications that we use for vacancies.[15]

Computation of mismatch indexes with heterogenous productive and matching efficiency requires estimates of labor-market specific productivities, matching efficiencies, and shares of the matching function. We compute these parameters at the industry level. As a proxy for productivity, 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.[16] 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.

[14] For more details on the JOLTS, see http://www.bls.gov/jlt/.

[15] Note 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.

[16] In 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.

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

To calculate the adjusted unemployment counts described in Section 4, we use the semi-panel dimension of the CPS and follow the algorithm described in Hobijn (2011). For unemployed workers, the survey reports the industry of the workers previous job while for employed workers, the survey reports the industry of the current job. Since respondents in the CPS are interviewed for several consecutive months, given any two adjacent months, we can track unemployed workers who find new employment from one month to the next. Thus we can obtain two key facts about unemployed workers who find jobs: 1. the industry of the previous job prior to the workers unemployment spell; 2. the industry of the new job. We create annual transition rate matrices by aggregating monthly data and calculating a five year centered moving average for 2001-2010. We exclude any individuals (unemployed and employed workers) who do not have an industry classification. We then infer the number of job seekers in each industry using the method outlined in Section 4. In our calculation of unemployment counts, to guarantee a non-negative solution to the linear system, we have set all entries to zero in the transition matrices which accounted for less than 5% of hires in any given sector.

5.1.1 The online vacancy data

We conduct our mismatch analysis for 2-digit occupations, for the 50 U.S. states and by education levels using vacancy data from the Help Wanted OnLine (HWOL) 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.[17] 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 the combination 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).[18] 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.

[17] The data are collected for The Conference Board by Wanted Technologies.

[18] The education level is imputed by TCB based on BLS information on the education content of detailed 6-digit level occupations. We classify vacancies by education level using an algorithm that we describe in detail in Section 6.4 below.

Figure 1: Comparison Between JOLTS and HWOL. Top-left panel: Midwest, Top-right panel: Northeast, Bottom-left panel: West, Bottom-right panel: South.

The aggregate trends from the HWOL data base are roughly consistent with those from the JOLTS data: in Figure 1 we plot JOLTS vacancies and HWOL ads by Census region. At the national level, the total count of active vacancies in HWOL is slightly below that in JOLTS until the end of 2007, and is slightly above from 2008 onwards. This difference is most pronounced in the South, and may reflect the growing penetration of online job listings over time. The average difference between the two aggregate series is about 11% of the total. The correlation between the two aggregate series is very high, 0.91, indicating that the patterns over time are very similar.

The vast majority of online advertised vacancies is posted on a small number of job boards: about 70% of all ads appears on nine job boards;[19] about 60% is posted on only five job boards. It is worth mentioning some measurement issues in the HWOL data: first, as mentioned earlier, there seems to be a slight time trend in the time series for HWOL vacancies relative to JOLTS, perhaps reflecting the growing use of online job boards over time. This should not overly affect our indices given the very

[19]These are: "Absolutely Health Care", "Craigslist", "JOBcentral", "CareerBuilder", "Monster", "Yahoo!HotJobs", "Recruiter Networks", "Dice","DataFrenzy".

		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

high correlation between the two series. In future work we plan to perform some robustness checks restricting the sample to a subset of job boards that have been more stable over time, to mimic the JOLTS series more closely.

Secondly, the dataset records one vacancy per ad. There is a small number of cases in which multiple positions are listed, but the convention of one vacancy per ad is used for simplicity. Finally, there are some cases in which multiple locations (counties within a state) are listed in a given ad for a given position. However, this is not an issue for our analysis since we focus on states as the smallest unit of geographic analysis at present.

Currently, we use HWOL data to construct mismatch indexes by 2-digit occupation, by state, as well as within and across education groups. Given the richness of detail of the vacancy information contained in HWOL, the limitations in constructing finer mismatch indexes arise from the unemployment side because of the relatively small size of the CPS. In future work, we plan to use job seeker data (typically, from public career centers) in individual states to conduct a more detailed analysis of mismatch for selected states.

5.2 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 industries and the period December 2000-December 2010, we estimate the parameters of the following CES matching function via minimum distance:[20]

$$\ln\left(\frac{h_{it}}{u_{it}}\right) = \ln\Phi + \frac{1}{\sigma}\ln\left[\alpha\left(\frac{v_{it}}{u_{it}}\right)^{\sigma} + (1-\alpha)\right]. \tag{33}$$

[20]Note that JOLTS reports vacancies and hires on the last day of the month and the CPS reports the number of unemployed during the survey week, which is the week containing the 12th day of the month. To be consistent with the timing of the measurement of flows and stocks, we use unemployment and vacancy stocks in month $t-1$ and hires in month t in all regressions.

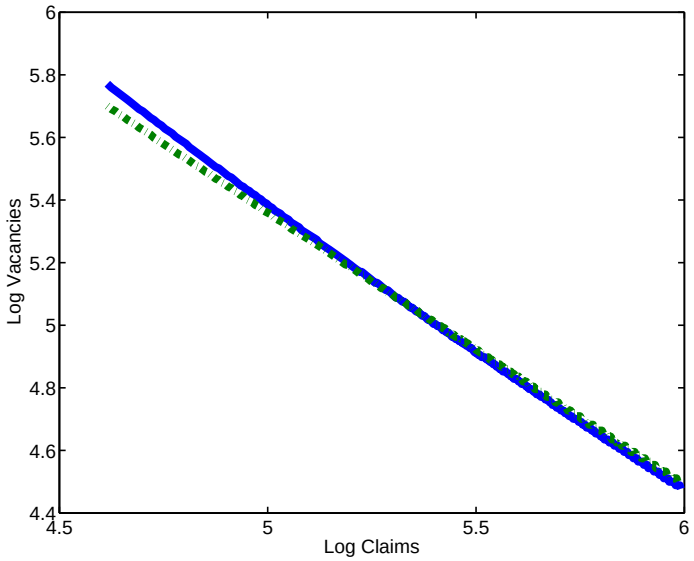

Figure 2: Iso-matching curves for CES (Solid) and Cobb-Douglas (Dashed)

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 2 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

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

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

[21] We 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.

[22] 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.

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

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.

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

$$\ln\left(\frac{h_{it}}{u_{it}}\right) = \ln \Phi_t + \ln \phi_i + \alpha \ln\left(\frac{v_{it}}{u_{it}}\right) \tag{34}$$

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.

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

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.[23]

Estimation of (34) 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 (e.g., informal referrals), as well as underreported vacancies as discussed in Davis, Faberman, and Haltiwanger (2010).[24]

5.3 A First Look At Mismatch

It is useful to examine the vacancy and unemployment shares of different sectors, occupations and geographic areas for a preliminary investigation of mismatch since these statistics are inputs into our mismatch indexes. If vacancy and unemployment shares of different labor markets do not vary over

[23]Estimates of α using HWOL vacancy data are roughly consistent with the ones obtained using JOLTS.

[24]Recall that in, Section 4.2, we showed that ϕ_i is proportional to underreported vacancies, when the latter are reported with error.

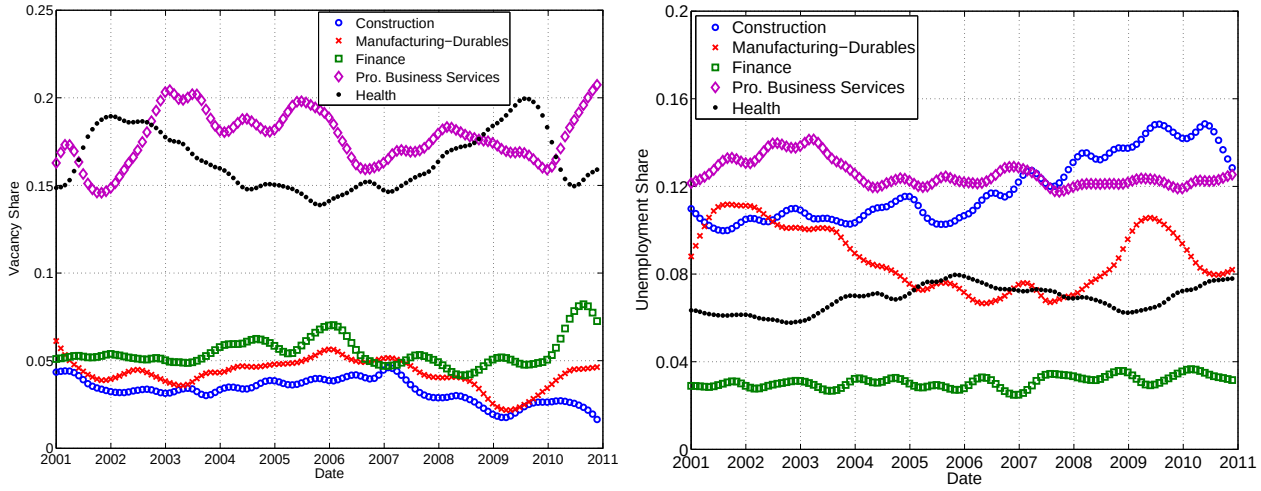

Figure 3: Vacancy and unemployment share by selected industry.

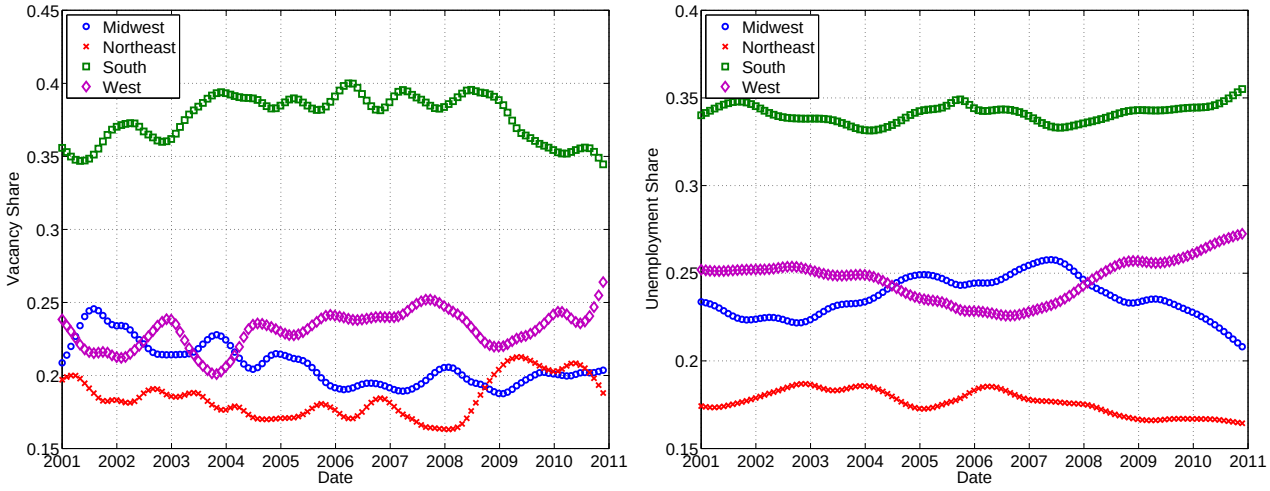

Figure 4: Vacancy and unemployment shares by Census Regions.

time, there is little room for mismatch to play an important role in the increase in the unemployment rate. To examine this issue, we first plot the vacancy and unemployment shares for a selected set of industries using the JOLTS definition. As Figure 3 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 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.

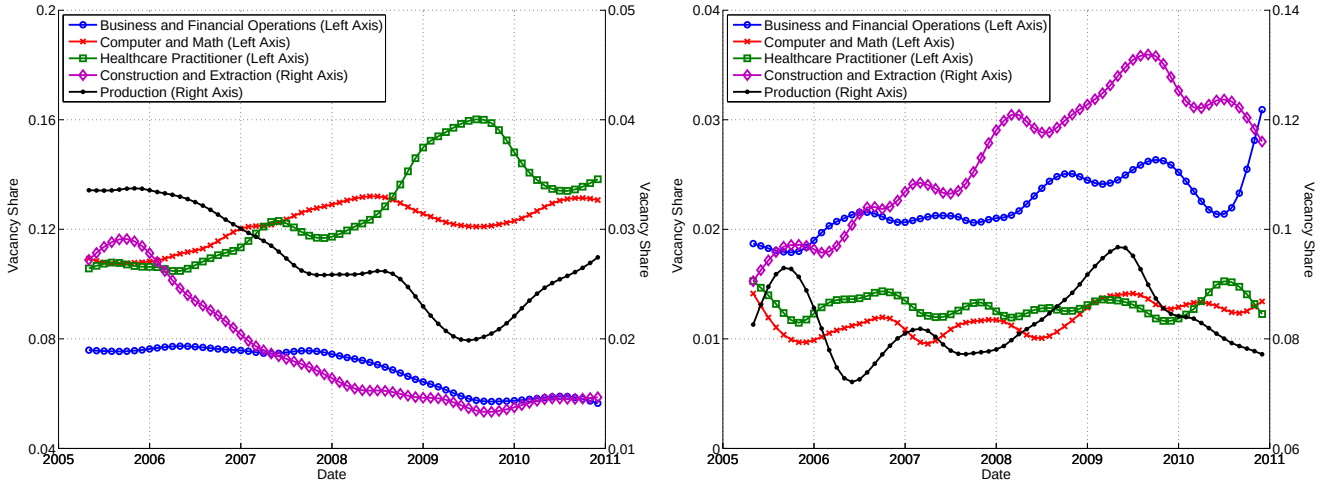

Figure 5: Vacancy and unemployment shares by selected occupations.

Figure 4 also shows the behavior of vacancy and unemployment shares by Census region. The West experienced an increase in its unemployment share and a mild decline in its vacancy share coinciding with the recession. The Midwest fared relatively better with a slight decline in its unemployment share and an increase in its vacancy share. The Northeast and the South also show some downward movements in their vacancy share and unemployment share, respectively. The figures suggest that there is less variation in shares by region than by industry, potentially suggesting a less important role for geographic mismatch relative to skill mismatch.

Now turning to the HWOL data, we plot the vacancy and unemployment shares for a selected set of occupations and U.S. states. Figure 5 shows the unemployment and vacancy shares of selected 2-digit occupations. As the figure indicates, the shares have changed noticeably during the most recent downturn. Business and financial operations, production and construction/extraction were among the occupations which experienced a decline in their vacancy shares and an increase in their unemployment shares. Concurrently, vacancy shares of healthcare practitioner and computer and math occupations went up. Starting from 2010, similar to the JOLTS data, unemployment and vacancy shares began to normalize. For some occupations these shares almost went back to their pre-recession levels (for example production) while for others (for example construction and extraction) the shares are still considerably different from their pre-recession levels. These patterns suggest that skill mismatch measured at the occupation level may have increased during the recession, but started to revert back as the recovery in the labor market began.

Figure 6 shows the behavior of vacancy and unemployment shares for a selection of U.S. states. California and Florida were hit hard by the recession, as reflected by the decline in their vacancy shares and the notable increase in their unemployment shares. As one might expect, California experienced a drastic deterioration of labor market conditions: California's vacancy share went down from over 15% to 11% and its unemployment share went up by 4 percentage points, from around

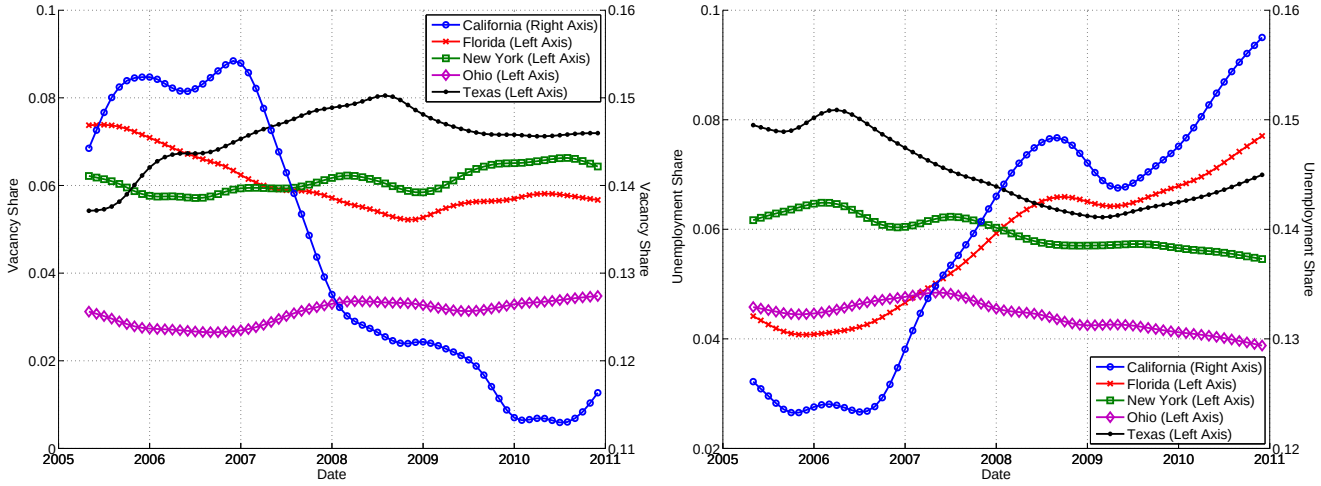

Figure 6: Vacancy and unemployment shares by selected states.

12% to almost 16%. New York, Ohio and especially Texas fared relatively better. Unemployment and vacancy shares still seem quite different from their pre-recession levels: this may be potentially due to a differential geographic impact of the recession as well as to other long-run differences in regional trends.

6 Empirical results

This section collects the results of our empirical analysis of mismatch by industry, occupation, Census region, U.S. state, and education. We also perform the counterfactual exercises described in Section 3.4.

6.1 Industry-level mismatch

We present a first set of results on mismatch unemployment across the 17 industries classified in JOLTS. 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 7 plots the time series of this correlation coefficient across industries over the sample period. In particular, we report three different correlation coefficients motivated by the definitions of the mismatch indexes we derived in Section 3: 1. ρ: between (u_{it}/u_t) and (v_{it}/v_t); 2. ρ_ϕ: between (u_{it}/u_t) and $(\phi_i/\bar{\phi}_t)^{\frac{1}{\alpha}}(v_{it}/v_t)$, and 3. ρ_z: between (u_{it}/u_t) and $(z_i/\bar{z}_t)^{\frac{1}{\alpha}}(v_{it}/v_t)$. The basic correlation coefficient (ρ) 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 left panel of Figure 8 plots the $\mathcal{M}_t^u$ indexes in their various versions described in Section 3:

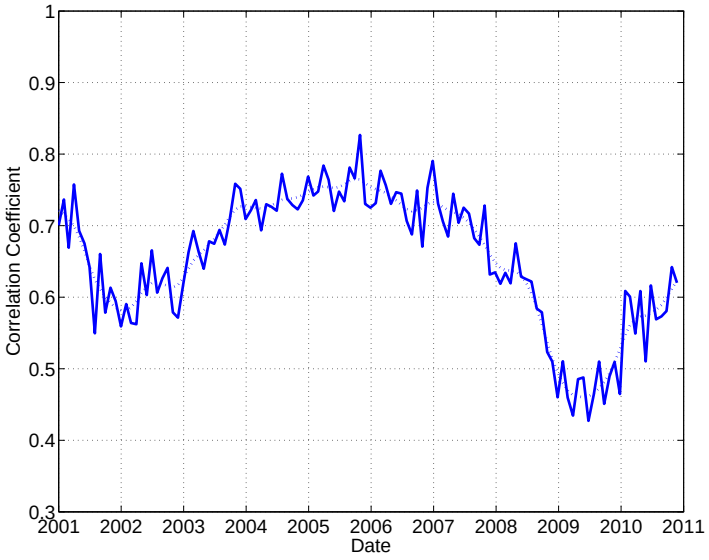

Figure 7: Correlation coefficient between u and v shares.

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 $\mathcal{M}_{zt}^u$, and, finally, the one modified to account for both sources, $\mathcal{M}_{xt}^u$. 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 about ten percentage points from early 2007 to mid 2009, and then dropped somewhat but remained at a higher level than its pre-recession level.

Turning to the index $\mathcal{M}_t^h$ measuring the fraction of hires lost because of the misallocation of unemployed workers across industries, the right panel of Figure 8 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 across industries during the recession, and a subsequent fairly rapid decline.

The four panels of Figure 9 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 0.8 percentage points of the five percentage point rise in U.S. unemployment, depending on the index used, i.e., at most 16 percent of the increase.[25]

As we have discussed in Section 3.4, an increase in mismatch causes the job-finding rate to decline. This decline in the job-finding rate has a direct impact on the unemployment rate. In addition to this direct effect, a rise in the separation rate has a larger impact on unemployment in an economy with

[25]Note that the average unemployment rate was 4.6% in 2006 and 9.6% in 2010, indicating a five percentage point increase. Throughout the paper we compare the average of 2006 with the average of 2010 when we discuss the role of mismatch in the increase in the unemployment rate.

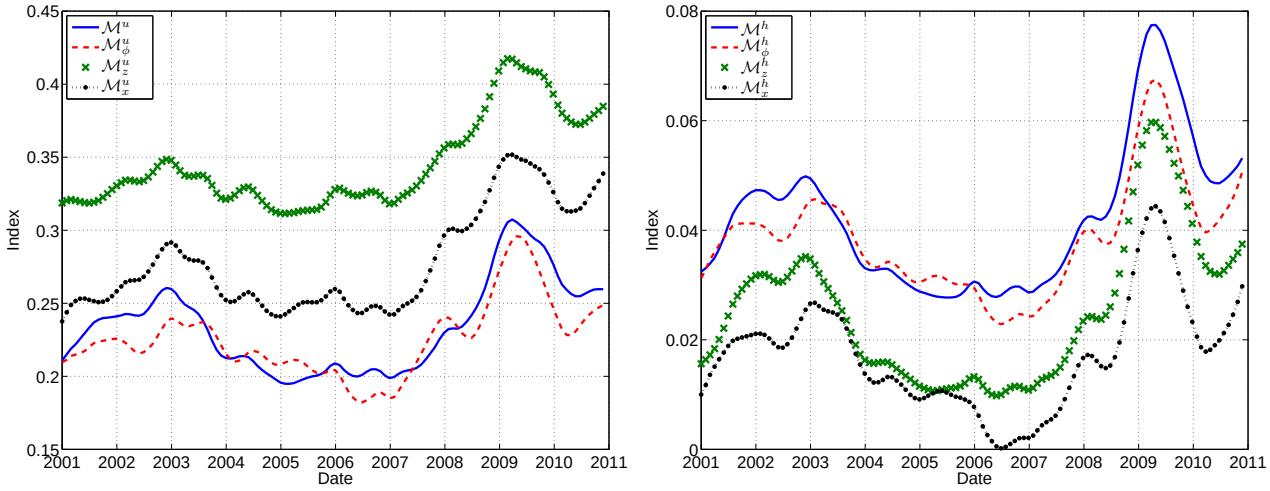

Figure 8: Mismatch index $\mathcal{M}_t^u$ (top panel) and $\mathcal{M}_t^h$ (bottom panel) by industry.

higher mismatch. Since during the most recent recession the separation rate increased considerably (from 2.1% in Dec 2006 to 2.8% in Dec 2008), some of the effect that we see on the unemployment rate is due to the increase in the separation rate. To isolate the effect of increased separations, we calculate some additional counterfactuals. In these counterfactuals, we freeze mismatch at its pre-recession level and let the separation rate vary as it did in the data. We find that around half of the increase in mismatch unemployment is due to the increase in mismatch and the other half is due to the interaction of mismatch with raising separations for $\mathcal{M}_t^h$ and $\mathcal{M}_{\phi t}^h$ counterfactuals. For the other two indices, the direct effect of mismatch is more modest.

6.1.1 Industry-level Mismatch with Adjusted Unemployment Counts

The empirical results we presented above assume that each unemployed worker is searching in the same industry of her previous job. We relax this assumption and infer the number of job seekers in each industry using the method outlined in Section 4 and the data described in Section 5. The left panel of Figure 10 shows the mismatch index $\mathcal{M}_t^h$ calculated using the adjusted unemployment counts as well the baseline $\mathcal{M}_t^h$ index. The adjustment causes the level of the index to increase by about 0.01 to 0.04. We also compute the counterfactual unemployment rate corresponding to the adjusted index as shown in the right panel of Figure 10. Not surprisingly, the counterfactual unemployment rate implied by the adjusted counts is lower than our baseline case, however in terms of accounting for the increase in the unemployment rate both indexes have remarkably similar quantitative implications. According to both indexes, 0.8 percentage points of the five percentage point rise in U.S. unemployment is due to industry-level mismatch.

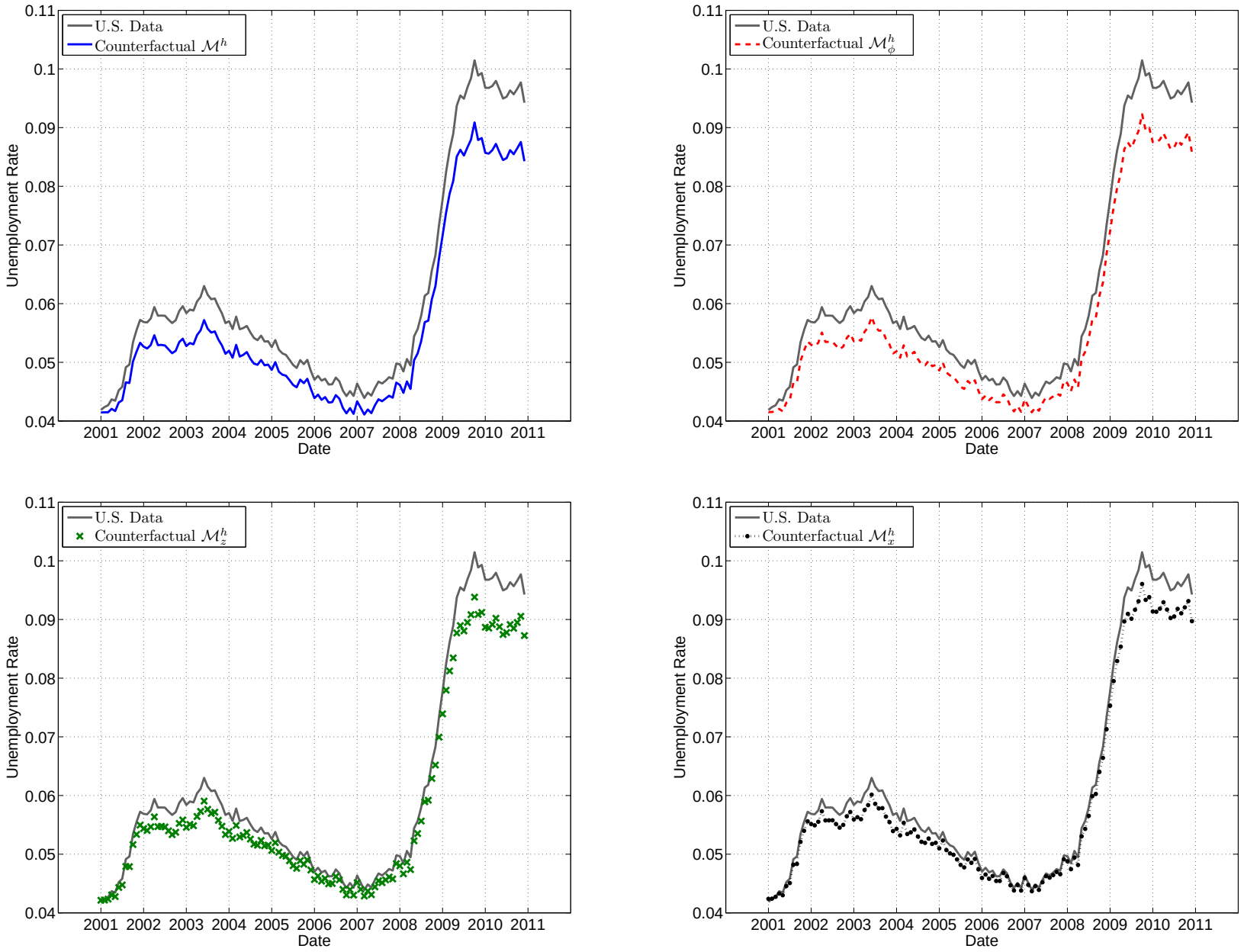

Figure 9: Counterfactual unemployment rates: Industry.

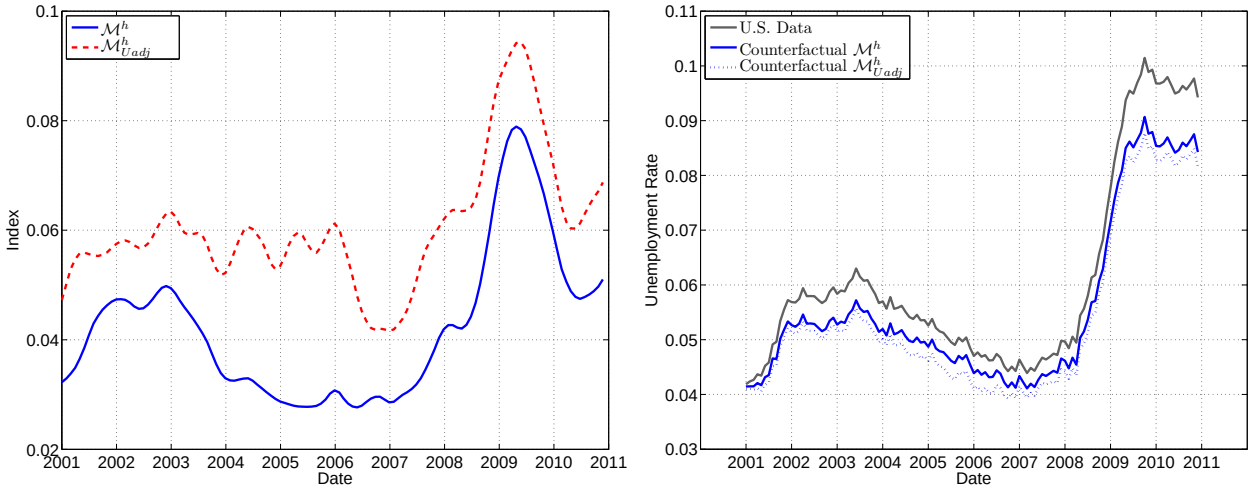

Figure 10: Mismatch index $\mathcal{M}_t^h$ by industry with unadjusted and adjusted unemployment counts (left panel) and corresponding counterfactuals (right panel).

6.1.2 Industry-level Mismatch with Heterogeneous Destruction Rates

In progress.

6.2 Occupational-level mismatch

We now present our results on mismatch unemployment across two-digit occupations based on HWOL job advertisement and CPS unemployment data. Recall that the HWOL ads data begin in May 2005.

The top-left panel of Figure 11 plots the correlation between vacancy and unemployment shares across 2-digit SOC's. As for the industry-level analysis, we document a significant decline in the correlation, by about 0.2 from 2006 to 2009. This fall in the correlation is the counterpart of an increase in mismatch indexes (top-right and bottom-left panels). The $\mathcal{M}_t^h$ index rises by 0.04 over the same period, i.e., the fraction of monthly hires lost because of occupational mismatch grew by 4% over that period. This rise is higher than the increase in mismatch documented at the industry level. Moreover, the level of the index is substantially higher.

Comparing the actual unemployment rate to the counterfactual unemployment in absence of mismatch (bottom-right panel of Figure 11), we conclude that around 1.4 percentage points of the recent surge in US unemployment (or around one quarter) can be attributed to occupational mismatch.

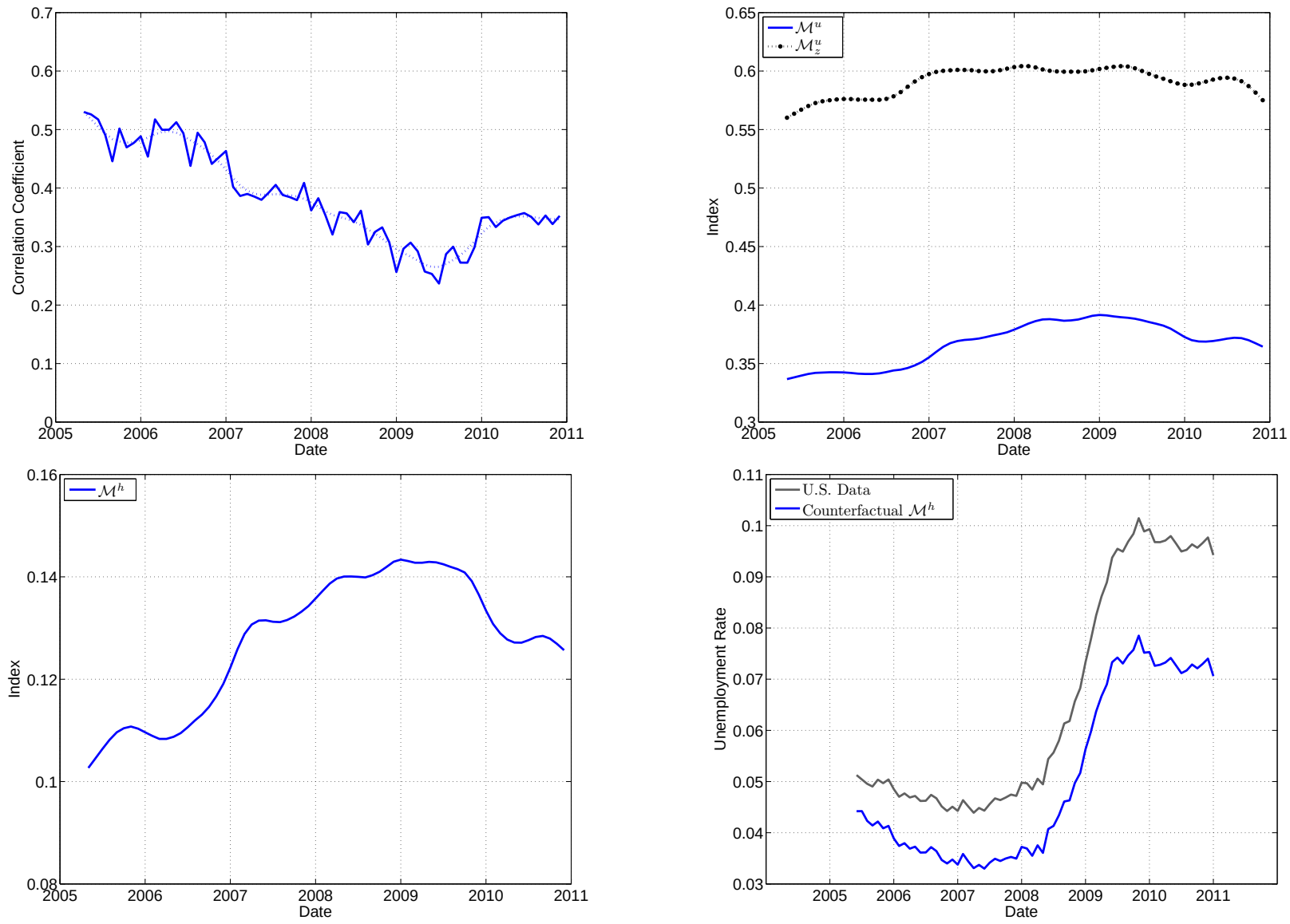

Figure 11: Top-left panel: Correlation coefficient between vacancy and unemployment shares across occupations. Top-right panel: Mismatch index $\mathcal{M}^u_t$. Bottom-left panel: Mismatch index $\mathcal{M}^h_t$. Bottom-right panel: Counterfactual unemployment rate: occupation

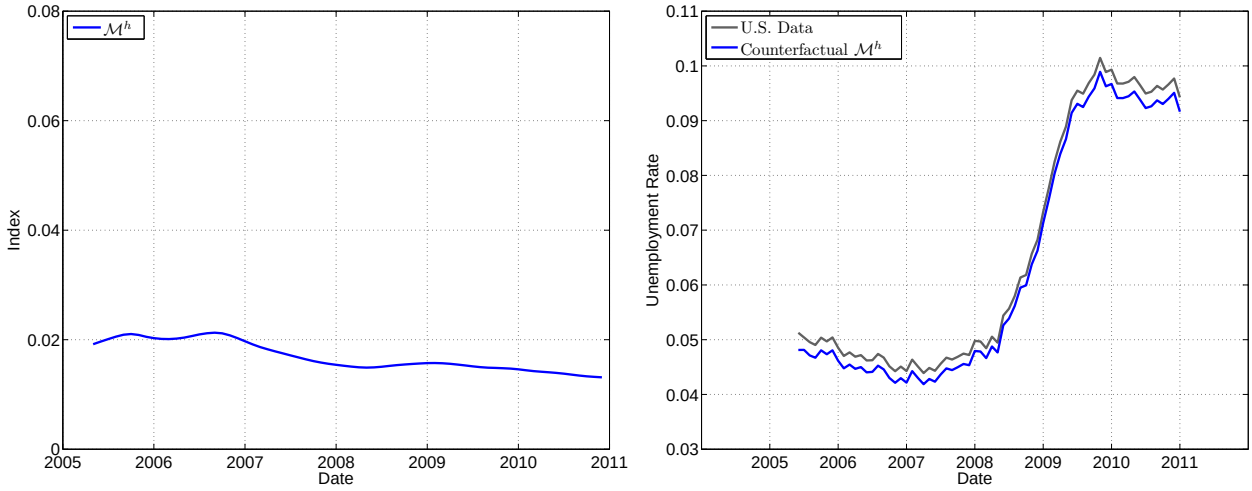

Figure 12: Mismatch index $\mathcal{M}_t^h$ by state (left panel) and corresponding counterfactual unemployment (right panel).

6.3 Geographical mismatch

We now turn to geographical mismatch. We have two sources of the data to study misallocation of unemployed workers across geographical areas: vacancies by Census region and the more disaggregated measure of online job ads by state.

We first calculate mismatch indexes across the four Census regions and find that regional mismatch is very low and does not show any significant trend. The level of $\mathcal{M}_t^h$ for regional mismatch has been between 0.001 and 0.004. This is less than 10% of the level of the industry mismatch index which was between 0.04 to 0.08 for the same period. Unsurprisingly, the counterfactual unemployment computed based on regional indexes is essentially the same as the actual series, implying that geographical mismatch –across Census regions– plays no role in the recent dynamics of US unemployment.

Figure 12 shows the mismatch index $\mathcal{M}_t^h$ using HWOL vacancies across the 50 U.S. states and the corresponding counterfactual experiment. Our conclusions from the analysis of JOLTS data on Census regions are confirmed at this higher level of disaggregation: we find little evidence of an increase in geographical mismatch.

6.4 Mismatch across and within education groups

Finally, we present our analysis of mismatch by education level, focusing on two different exercises. First, we compute our mismatch indexes *across* the following four education categories: less than high school diploma; high school diploma or equivalent; some college and Associate's degree; Bachelor's degree or higher. In other words, we treat these four education groups as distinct labor markets for the purpose of constructing our mismatch indexes. Second, we analyze mismatch by 2-digit occupation

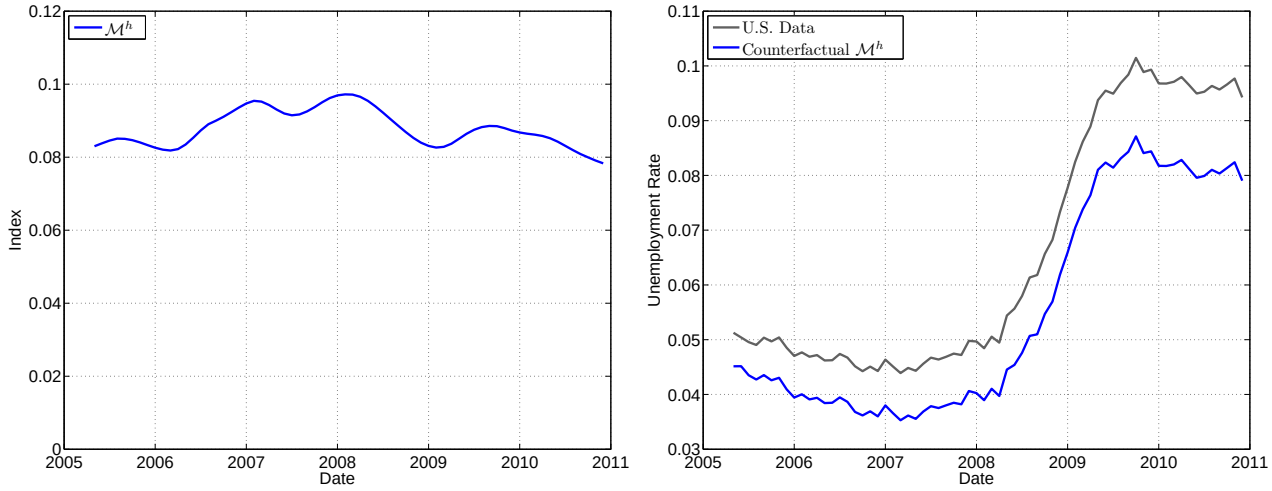

Figure 13: Mismatch index $\mathcal{M}_t^h$ across education groups (left panel) and corresponding counterfactual unemployment (right panel)

within these four education groups. This enables us to determine whether occupational mismatch has increased more or less for specific education categories.

The vacancy data for this analysis come from the HWOL series. As noted before, each ad recorded in HWOL constitutes an individual observation with a 6-digit occupation classification. We use this information, together with information from the BLS on the education content of 6-digit occupations, to construct vacancy counts for each 2-digit occupation by education level cell. In particular, the BLS provides information on the distribution of workers employed in each 6-digit occupation, broken down by their highest level of education attained.[26] We then allocate the count of vacancies from HWOL in a given month for a given 6-digit occupation to each of the four education groups we consider, proportionally to the educational attainment distributions from the BLS.[27] Finally, we aggregate up to the 2-digit occupation level to obtain vacancy counts for each occupation by education cell. The key assumption underlying this methodology is the educational content of new vacancies has not shifted significantly compared to the one of existing vacancies.

The results on mismatch across education groups are reported in Figure 13. The left panel plots the $\mathcal{M}_t^h$ mismatch index: it started rising in 2007, reached a peak in 2008, and has since declined to almost pre-recession levels. The counterfactual unemployment exercise is depicted in the right panel. It shows that the rise in mismatch across education groups explain about 0.7 percentage points of the rise in unemployment observed between 2006 and 2010.

[26]This information comes from the American Community Survey microdata from 2006-08. See the BLS website at http://www.bls.gov/emp/ep_table_111.htm; see also http://www.bls.gov/emp/ep_education_tech.htm for additional details.

[27]For robustness, we have also experimented with other allocation rules, for instance not imputing vacancies to an education level that accounts for less than 15% of the workers in a 6-digit SOC. The results of the mismatch analysis are very similar.

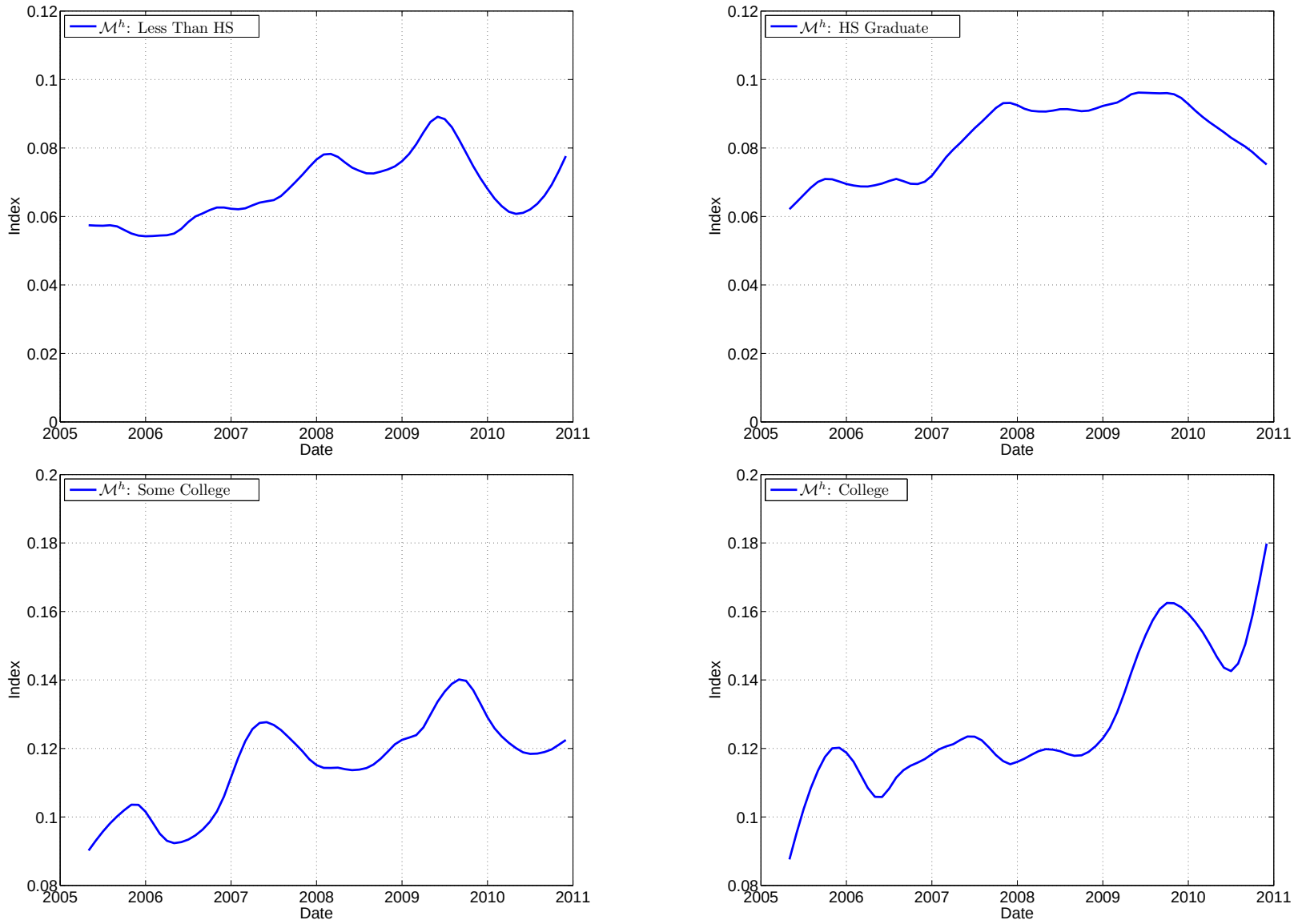

Figure 14: Mismatch index $\mathcal{M}_t^u$ by occupation within different education groups Top-left panel: Less than high school diploma. Top-right panel: High school diploma or equivalent. Bottom-left panel: Some college and Associate's degree. Bottom-right panel: Bachelor's degree or higher.

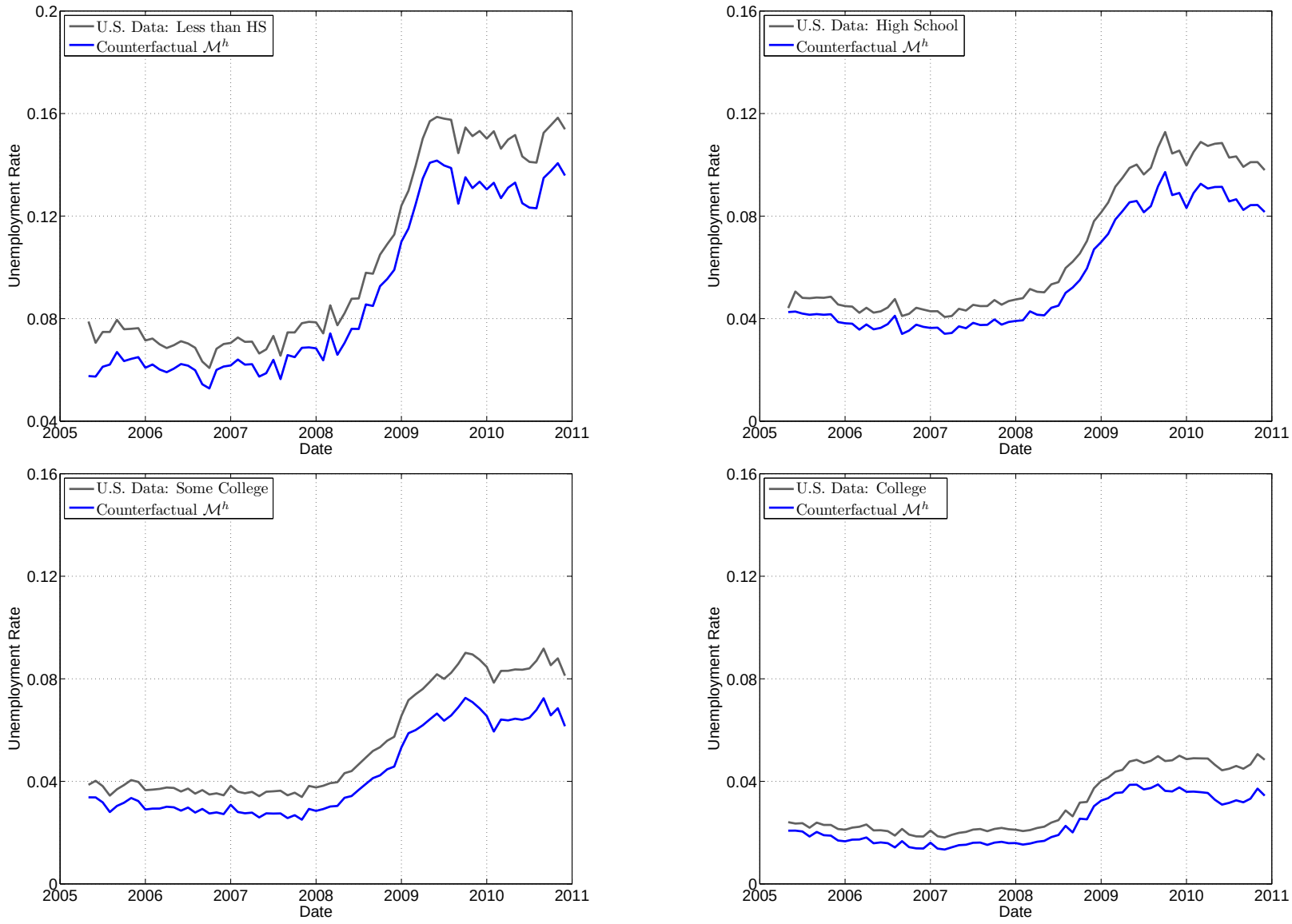

Figure 15: Counterfactual unemployment rate for different education groups. Top-left panel: Less than high school diploma. Top-right panel: High school diploma or equivalent. Bottom-left panel: Some college and Associate's degree. Bottom-right panel: Bachelor's degree or higher.

Figures 14 and 15 illustrate our findings on occupational mismatch within each broad education category. The $\mathcal{M}_t^h$ mismatch indexes are shown in Figure 14 and the counterfactual unemployment exercises in Figure 15. Notice that actual unemployment varies considerably across the four panels in Figure 15, since we are plotting unemployment for workers within each educational attainment group. Unemployment experiences differ greatly by education: for workers with less than high school, the unemployment rate rose from about 7% in 2006 to about 15% in 2010, an increase of about eight percentage points. The increase in unemployment rate over the same time period for high school graduates and those with some college was, respectively, 6 and 4.8 percentage points. For college graduates, the unemployment rate went from 2% to 4.7%, an increase of only 2.7 percentage points over the same period.

The occupational mismatch index rose within all four education groups, but more so in the some college and college categories. The counterfactual exercises reveal a very clear pattern: the contribution of occupational mismatch to the rise in unemployment between 2006 and 2010 grows as we move from the lowest to the highest education category. In particular, for the less than high school group, mismatch explains a little less than one percentage point (12%) of the eight percentage point increase in unemployment for that group. For high school graduates, mismatch explains 1.2 (20%) out of the six percentage point increase in unemployment. For those with some college, mismatch explains about 1.4 (29%) out of a 4.8 percentage point rise in unemployment, and for college graduates 0.9 (33%) out of the 2.7 percentage point observed increase. Thus, the fraction of the rise in unemployment that can be attributed to the rise in occupational mismatch increases monotonically with education from about one eighth to roughly one third.

7 Conclusion

We have developed a theoretical framework that gives rise to a well defined notion of mismatch between unemployment and vacancies across separate labor markets (sectors) of an economy. We model a dynamic stochastic economy with many distinct frictional labor markets, 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 output 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 of idle labor. This solution yields a set of Jackman-Roper (JR) conditions generalized to 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 1) the fraction of unemployed searching in the "wrong" markets, and 2) the fraction of hires lost be-

cause of mismatch. These latter indices can be used to compute a counterfactual series for frictional unemployment in the absence of mismatch.

In the empirical part of the paper we use vacancy data by industry, occupation, geographic area, and education to compute our indexes for the period 2000-2010. We find that the rise in mismatch at the industry, occupational, and education level can explain between 0.8 and 1.4 percentage points of the observed increase in the unemployment rate from the start of the recession to 2010. Our results indicate that the role of mismatch in explaining the increases in unemployment varies considerably by education. Occupational mismatch explains a substantial fraction of the rise in unemployment (one third) for high-educated workers while it is quantitatively less important for less-educated workers. Finally, we calculate geographic mismatch measures across U.S. states and find no role for geographic mismatch in explaining the increase in the unemployment rate.

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," *Econometrica*, 79:1, 75-122.

[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 C. and L. Visscher (2010). "Search and Endogenous Reallocation over the Business Cycle," mimeo.

[6] Davis, S. J., Faberman, R. J., and J.C. Haltiwanger. (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, F., J. Gyourko, and J. Tracy (2010). "Housing Busts and Household Mobility," *Journal of Urban Economics*, 68(1), 34-45.

[9] Fujita, S. (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] Hertz, B., and T. Van Rens (2011). "Structural Unemployment," mimeo.

[12] Hobijn, B. (2011). " Industry-Occupation Mix," mimeo.

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

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

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

[16] Kocherlakota, N. (2010). "Inside the FOMC," Speech at Marquette, Michigan on August 17, 2010.

[17] Lagos, R. (2006). "A Model of TFP," *Review of Economic Studies* 73, 983-1007.

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

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

[20] Molloy, R. C. Smith, and A. Wozniak (2010). "Internal Migration in the US: Updated Facts and Recent Trends," mimeo.

[21] Mortensen, D.T. (2009). "Island Matching," *Journal of Economic Theory*, 144(6), 2336-2353.

[22] Mortensen, D.T. and C.A. Pissarides (1999). "Unemployment Responses to Skill-biased Technology Shocks: the Role of Labour Market Policy," *Economic Journal*, 109(455), 242-265.

[23] Padoa-Schioppa F. (1991). Mismatch and Labour Mobility. Centre for Economic Policy Research. Cambridge University Press.

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

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

[26] Şahin, A., J. Song, G. Topa and G. Violante (2011). "Does Mismatch Explain the Recent Rise in U.K. Unemployment?," mimeo.

[27] Schulhofer-Wohl, S. (2010). "Negative Equity Does Not Reduce Homeowners' Mobility," mimeo.

[28] Shimer R. (2005). "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, 95, 25-49.

[29] Shimer R. (2007a). "Mismatch," *American Economic Review*, 97, 1074-1101.

[30] Shimer R. (2007b). "Reassessing the Ins and Outs of Unemployment," mimeo.

[31] Valletta, R. and K. Kuang (2010). "Extended Unemployment and UI Benefits," *FRBSF Economic Letter*, 2010-12.