Mismatch Unemployment*

Ayşegül Şahin

Federal Reserve Bank of New York

Joseph Song

Columbia University

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: DECEMBER 2013

Abstract

We develop a framework where mismatch between vacancies and job seekers across sectors translates into higher unemployment by lowering the aggregate job-finding rate. We use this framework to measure the contribution of mismatch to the recent rise in U.S. unemployment by exploiting two sources of cross-sectional data on vacancies: JOLTS and HWOL (a new database covering the universe of online U.S. job advertisements). Our calculations indicate that mismatch across industries and 3-digit occupations explains at most 1/3 of the total observed increase in the unemployment rate. Occupational mismatch has become especially more severe for college graduates, and in the West of the United States. Geographical mismatch unemployment plays no apparent role.

*We are especially grateful to June Shelp, at The Conference Board, for her help with the HWOL data. 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 U.S. unemployment rate rose from an average value of 4.6% in 2006 to its peak of 10% in October 2009, as the economy experienced the deepest downturn in the postwar period. Three years after its peak, the unemployment rate still hovered above 8%. This persistently high rate has sparked a vibrant debate among economists and policymakers. The main point of contention is the nature of these sluggish dynamics and, therefore, the appropriate policy response.

A deeper look at worker flows into and out of unemployment reveals that, while the inflow rate quickly returned to its pre-recession level, the job-finding rate is still substantially below what it was in 2006. Any credible explanation for the recent dynamics in unemployment must therefore operate through a long-lasting decline in the outflow rate. One such theory is that the recession has produced a severe sectoral mismatch between vacant jobs and unemployed workers: idle workers are seeking employment in sectors (occupations, industries, locations) different from those where the available jobs are. Such misalignment between the distribution of vacancies and unemployment would lower the aggregate job-finding rate.

The mismatch hypothesis is qualitatively consistent with three features of the Great Recession. First, in the period 2009-2012, the U.S. Beveridge curve (i.e., the empirical relationship between aggregate unemployment and aggregate vacancies) has displayed a marked outward movement indicating that, for a given level of vacancies, the current level of unemployment is higher than that implied by the last decade of historical data.[1] Put differently, aggregate matching efficiency has declined.[2] Second, around half of the job losses in this downturn were concentrated in construction and manufacturing.[3] To the extent that the unemployed in these battered sectors do not search for (or are not hired in) jobs in the sectors which largely weathered the storm (e.g., health care), mismatch would arise across occupations and

[1]See, among others, Elsby, Hobijn, and Şahin (2010), Hall (2010), and Daly, Hobijn, Şahin, and Valletta (2012). According to these studies, at the current level of vacancies, the pre-recession U.S. unemployment-vacancies relationship predicts an unemployment rate between 2 and 3 percentage points lower than its current value.

[2]According to Barlevy (2011) and Veracierto (2011), the size of this drop from its pre-recession level is between 15% and 30%, depending on the exact methodology used in the calculation.

[3]According to the Current Employment Statistics (CES), also known as the establishment survey, payroll employment declined by 7.4 million during the recession and construction and manufacturing jointly accounted for 54% of this decline.

industries. Third, house prices experienced a sharp fall, especially in certain regions (see e.g., Mian and Sufi, 2011). Homeowners who expect their local housing markets to recover may choose to forego job opportunities in other locations to avoid large capital losses from selling their house. Under this "house-lock"conjecture, mismatch between job opportunities and job seekers would arise mostly across locations.

In this paper, we develop a theoretical framework to conceptualize the notion of mismatch unemployment, and use it to measure how much of the recent rise in the U.S. unemployment rate is attributable to mismatch across sectors. We envision the economy as comprising a large number of distinct labor markets or sectors (e.g., segmented by industry, occupation, geography, or a combination of these attributes). Each labor market is frictional, i.e., its hiring process is governed by a matching function. To assess the existence of mismatch in the data, we ask whether, given the observed distribution of productive efficiency, matching efficiency, and vacancies across labor markets in the economy, unemployed workers are "misallocated," i.e., they search in the wrong sectors. 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 is the one that would be selected by *a planner who faces no impediment in moving idle labor across sectors, except for the within-market matching friction*. We show that optimality for this planner dictates that (productive and matching) efficiency-weighted vacancy-unemployment ratios be equated across sectors. By manipulating the planner's optimality condition, we construct a mismatch index that measures the fraction of hires lost every period because of misallocation of job seekers. Through this index, we can quantify how much lower the unemployment rate would be in the absence of mismatch. The difference between the observed unemployment rate and this counterfactual unemployment rate is *mismatch unemployment*.[4] As we explain in detail in the paper, choosing as benchmark the allocation of a planner who can shuffle labor across sectors at no cost has the implication that our estimates of sectoral mismatch are an upper bound.

Our measurement exercise requires disaggregated data on unemployment and vacancies. The standard micro data sources for unemployment and vacancies are, re-

[4]Our focus is on mismatch unemployment intended as unemployed searching in the "wrong" sector. A separate literature uses the term "mismatch" to denote the existence of employed individuals working on the "wrong" job—meaning a sub-optimal joint distribution of worker skills and firm's capital. See, for example, Eeckhout and Kircher (2011).

spectively, the Current Population Survey (CPS) and the Job Openings and Labor Turnover Survey (JOLTS). Unfortunately, JOLTS only allows disaggregation of vacancies into very broad geographical areas (4 Census regions) and 17 industries that roughly coincide with 2-digit NAICS classification.[5] In this paper, we introduce a new source of micro data, the Conference Board's Help Wanted OnLine (HWOL) database, designed to collect the universe of unique online job advertisements in the U.S. economy. Through this novel data set, we are able perform our empirical analysis at the 2- and 3-digit occupational level, at a more detailed geographical level (states and counties), and even by defining labor markets as a combination of occupation and location.[6]

Our empirical analysis yields indicates no significant role for geographical mismatch between unemployed workers and job vacancies across U.S. states or counties. Mismatch at the industry and 2- and 3-digit occupation level increased markedly during the recession but declined steadily throughout 2010, an indication of a countercyclical pattern in mismatch. A similar, but milder, hump shape in mismatch is observed around the 2001 recession. In line with this result, Barnichon and Figura (2013) document that aggregate matching efficiency has been strongly procyclical over the period 1976-2012.

We calculate that an additional four percent of monthly hires were lost during the Great Recession because of the misallocation of vacancies and job seekers across occupations and industries. As a result, our counterfactual analysis indicates that mismatch unemployment at the industry level can account for 0.75 percentage points out of the 5.4 percentage point total increase in the U.S. unemployment rate from 2006 to October 2009. At the 3-digit occupation level, the contribution of mismatch unemployment rises to 1.6 percentage points. When we compute 2-digit occupational mismatch separately for different education groups and different Census regions, we find its contribution to the observed increase in the unemployment rate is the largest among college graduates and for the West of the U.S., and it is the smallest among high-school dropouts and in the North-East.

The Great Recession coincided with an increase in the number of workers who stopped actively searching for jobs because of a "discouragement effect". We verified

[5] See Table C1 in the Appendix for a complete list of industries in the JOLTS.

[6] The HWOL micro data would allow an even more disaggregated analysis. The binding constraint is the small sample size of unemployed workers in the monthly CPS.

that, when we add these discouraged workers (who can be thought of job seekers with low search intensity) to the unemployed job-seeker counts by occupation, our conclusions are largely unaffected.

In an extension of the baseline analysis, we allow the misallocation of unemployed workers across sectors to also affect the vacancy creation decisions of firms: since the presence of job-seekers in declining sectors makes it easier to fill jobs in those sectors, it distorts firms' incentives in the direction of, inefficiently, creating vacancies in the wrong markets. This channel depresses aggregate vacancy creation relative to the planner's solution, giving a further boost to mismatch unemployment. This amplification can be very strong if the vacancy creation cost is close to linear, but for specifications of this cost function in line with the existing literature (i.e., closer to quadratic) the amplification is moderate. When this additional force is factored into our counterfactuals, the contribution of mismatch to the observed rise in the unemployment rate grows by a maximum of half of a percentage point.

With all the necessary caveats, discussed throughout the paper, our study indicates that, at the analyzed level of disaggregation, sectoral mismatch can explain at most 1/3 of the recent rise in the U.S. unemployment rate since from early 2006 to the end of 2009, the period of the sharp drop in the average job finding rate.

The model underlying our measurement exercise is a multi-sector version of the standard aggregate search/matching model (Pissarides, 2000). Within this class, the closest paper to ours is Jackman and Roper (1987): in a static matching model with many sectors, they show that distributing unemployment across sectors so that sectoral labor-market tightnesses are equalized maximizes aggregate hires, and they propose the use of mismatch indexes to summarize deviations from this allocation.[7] At that time, economists were struggling to understand why high unemployment was

[7]This idea goes back, at least, to Mincer (1966, page 126) who writes: "To detect the existence, degree, and changes in structural unemployment, (U, V) maps may be constructed for disaggregations of the economy in the cross-section, by various categories, such as industry, location, occupation, and any other classification of interest. For example, each location is represented by a point in the (U, V) map, and a scatter diagram showing such information for all labor markets may show a clear positive correlation. This would indicate that unemployment is largely nonstructural with respect to location, that is to say, that adjustments require movements within local areas rather than the more difficult movements between areas. In contrast, a negative relation in the scatter would indicate the presence of a structural problem. The scatters may, of course, show identifiable combinations of patterns. Observations of changes in these cross sectional patterns over time will show rotations and shifts, providing highly suggestive leads for diagnoses of the changing structure of labor supplies and demands."

so persistent in many European countries.[8] Padoa-Schioppa (1991) contains a number of empirical studies for various countries and concludes that mismatch was not an important explanation of the dynamics of European unemployment in the 1980s. Our paper contributes to reviving this old literature by extending it in several directions: (i) we develop a dynamic, stochastic environment with numerous sources of heterogeneity, (ii) we develop a framework to construct counterfactual measures of unemployment, absent mismatch, (iii) we incorporate the effect of misallocation on vacancy creation, and (iv) we perform our measurement at a much more disaggregated level, thanks to new micro data. Through this novel data source, we document new facts concerning changes in the correlation of vacancy and unemployment shares across sectors of the economy, and show that these facts are informative about the extent of sectoral mismatch, in this class of search/matching models.

Shimer (2007) proposed an alternative environment to measure mismatch between firms and workers across labor markets. The crucial difference between these two models is the notion of a vacancy or, equivalently, at which point of the meeting process vacancies are measured. The notion of vacancy we adopt is common to the entire search/matching approach to unemployment. Here, firms desiring to expand post vacancies: a vacancy is a manifestation of a firm's *effort to hire*. In Shimer's model, firms unsuccessful in meeting workers are left with idle jobs: a vacancy is therefore a manifestation of a firm's *failure to hire*. Both notions are theoretically correct. Since both models are parameterized using the same micro-data on vacancies, the key question is whether existing job-openings data from JOLTS and HWOL are more likely to represent firms' hiring effort or hiring failure. The short duration of job openings in JOLTS (2-4 weeks according to Davis, Faberman, and Haltiwanger, 2010) seems somewhat more consistent with the former view, but better data is needed to shed light on this critical point.

The remainder 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 unemployment rate counterfactuals. Here, we also discuss in some depth the interpretation of our measure of mismatch. Section 4 describes the data.

[8]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 evidence of shifts in the Beveridge curve for some countries.

Section 5 performs the empirical analysis. Section 6 analyzes the case in which mismatch also affects vacancy creation. In Section 7 we verify the robustness of our results to measurement error in unemployment and vacancy counts, and to specification error in the matching function. Section 8 concludes. Appendix A contains the proofs of our theoretical results, Appendix B contains more detail about the data and our measurement exercise, and Appendix C contains additional figures and tables.

2 Environment and planner problem

We begin by describing our economic environment and deriving the planner's optimal allocation rule of unemployed workers across sectors—the crucial building block of our theoretical analysis. Throughout these derivations, we maintain the assumption that the evolution of the vacancy distribution is exogenous. We relax this assumption in Section 6.

2.1 Benchmark environment

Time is discrete and indexed by t. 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_{it}), arise exogenously across sectors.[9] The economy is populated by a measure one of risk-neutral individuals who can be either employed in sector i (e_{it}) or unemployed and searching in sector i (u_{it}). Therefore, $\sum_{i=1}^{I}(e_{it}+u_{it})=1$. On-the-job search is ruled out and an unemployed worker, in any given period, can search for vacancies in one sector only. For the time being, we also rule out non-participation, but in the next section we relax this restriction.

Labor markets are frictional: new matches, or hires, (h_{it}) between unemployed workers (u_{it}) and vacancies (v_{it}) in market i are determined by the matching function $\Phi_t \phi_{it} m(u_{it}, v_{it})$, with m strictly increasing and strictly concave in both arguments and homogeneous of degree one in (u_{it}, v_{it}). The term $\Phi_t \phi_{it}$ measures matching efficiency (i.e., the level of fundamental frictions) in sector i, with Φ_t denoting the

[9] We explain in Section 6 that assuming that vacancies are exogenous is equivalent to a model where the job creation margin is endogenous, and the elasticity of the cost of creating vacancies is infinitely large.

aggregate component and ϕ_{it} the idiosyncratic sectoral-level component. The number of vacancies and matching efficiency are the only two sources of heterogeneity across sectors in our baseline model.

All existing matches produce Z_t units of output in every sector. Matches are destroyed exogenously at rate Δ_t, also common across sectors. Aggregate shocks Z_t, Δ_t, and Φ_t, and the vector of vacancies $\mathbf{v}_t = \{v_{it}\}$ are drawn from the conditional distribution functions $\Gamma_{Z,\Delta,\Phi}\left(Z_{t+1}, \Delta_{t+1}, \Phi_{t+1}; Z_t, \Delta_t, \Phi_t\right)$ and $\Gamma_{\mathbf{v}}\left(\mathbf{v}_{t+1}; \mathbf{v}_t, Z_t, \Delta_t, \Phi_t\right)$. The notation shows that we allow for autocorrelation in $\{Z_t, \Delta_t, \Phi_t, \mathbf{v}_t\}$, and for correlation between vacancies and all the aggregate shocks. The sector-specific matching efficiencies ϕ_{it} are independent across sectors and are drawn from $\Gamma_\phi\left(\phi_{t+1}; \phi_t\right)$, where $\phi_t = \{\phi_{it}\}$. The vector $\{Z_t, \Delta_t, \Phi_t, \mathbf{v}_t, \phi_t\}$ takes strictly positive values.

Within each period, events unfold as follows. At the beginning of the period, the aggregate shocks (Z_t, Δ_t, Φ_t), vacancies $\mathbf{v}_t$, and matching efficiencies ϕ_t are observed. At this stage, the distribution of active matches $\mathbf{e}_t = \{e_{1t}, \ldots, e_{It}\}$ across markets (and hence the total number of unemployed workers u_t) is also given. Next, unemployed workers are allocated to market i without any impediment to labor mobility. Once the unemployed workers are allocated, the matching process takes place and $h_{it} = \Phi_t \phi_{it} m\left(u_{it}, v_{it}\right)$ new hires are generated in each market. Production occurs in the e_{it} (pre-existing) plus h_{it} (new) matches. Finally, a fraction Δ_t of matches are destroyed exogenously in each market i, determining next period's employment distribution $\{e_{i,t+1}\}$ and stock of unemployed workers u_{t+1}.

Planner's solution In Appendix A.1 we prove that the planner's optimal rule for the allocation of unemployed workers across sectors can be written as

$$\phi_{1t} m_{u_1}\left(\frac{v_{1t}}{u^*_{1t}}\right) = \ldots = \phi_{it} m_{u_i}\left(\frac{v_{it}}{u^*_{it}}\right) = \ldots = \phi_I m_{u_I}\left(\frac{v_{It}}{u^*_{It}}\right), \tag{1}$$

where m_{u_i} is the derivative of the m function with respect to u_i, and where we have used the "*" to denote the planner's allocation. This condition states that the planner allocates more job seekers to those labor markets with more vacancies and higher matching efficiency until their marginal contribution to the hiring process is equalized across markets.[10]

[10] In equation (1), the derivative of the sector-specific matching function m is written as a function of sectoral market tightness only (with a slight abuse of notation) because of its CRS specification.

2.2 Heterogeneous productivities and job destructions

We now allow for sector-specific shocks to productivity and destruction rates that are uncorrelated across sectors and independent of the aggregate shocks Z_t and Δ_t. Note that when productivity is heterogeneous across sectors, maximizing aggregate output in the planner problem is no longer equivalent to maximizing employment.

In the derivations below, we first keep worker separations exogenous. Next, we allow the planner to choose whether to endogenously dissolve some existing matches and show that, under some conditions, it never chooses to do so. Throughout this extension, we also allow the planner to choose the size of the labor force.

2.2.1 Exogenous separations

Let labor productivity in sector i at date t be given by $Z_t z_{it}$, where each component z_{it} is strictly positive, i.i.d. across sectors and independent of Z_t. Similarly, denote the idiosyncratic component of the exogenous destruction rate in sector i as δ_{it}. Then, the survival probability of a match is $(1-\Delta_t)(1-\delta_{it})$. It is convenient to proceed under the assumption that $\{Z_t, 1-\Delta_t, z_{it}, 1-\delta_{it}\}$ are all positive martingales, which amounts to simple restrictions on the conditional distributions $\Gamma_{Z,\Delta,\Phi}, \Gamma_z$, and Γ_δ.[11] All the non-employed individuals produce output ζZ_t (which can be interpreted as home-production or the value of leisure). In addition, the unemployed incur a disutility cost of search.

Appendix A.2 proves that the planner's optimal allocation rule of unemployed workers equates

$$\frac{z_{it}-\zeta}{1-\beta(1-\Delta_t)(1-\delta_{it})}\phi_{it} m_{u_i}\left(\frac{v_{it}}{u^*_{it}}\right) \tag{2}$$

across markets. This rule establishes that the higher vacancies, matching efficiency, and expected discounted productive efficiency in market i, the more unemployed workers the planner wants searching in that market. In particular, expected output of an unemployed worker searching in sector i (net of the opportunity cost of employ-

[11] As we explain in Appendix A.2, the martingale assumption is convenient to solve forward, in closed form, the expected marginal value of an employed worker in sector i. A closed form solution can also be obtained if the components of the vector $\{Z_t, 1-\Delta_t, z_{it}, 1-\delta_{it}\}$ follow an AR(1) process. However, the derivations are more convoluted, and we do not make use of this more general assumption in the empirical analysis because our variables are well represented, statistically, by martingales. We have not attempted to solve the model under other stochastic processes.

ment ζ) is discounted differently by the planner in different sectors because of the heterogeneity in the expected duration of matches.

2.2.2 Endogenous separations

We now allow the planner to move workers employed in sector i into unemployment or out of the labor force, before choosing the size of the labor force for next period.

In Appendix A.3 we demonstrate that, if the planner always has enough individuals to pull into (out of) unemployment from (into) out of the labor force, it will never choose to separate workers who are already matched and producing. The planner's allocation rule remains exactly as in equation (2) and all separations are due to exogenous match destructions.

2.3 Heterogeneous sensitivities to the aggregate shock

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. Here we derive the planner allocation rule (2) under this alternative interpretation of the source of sectoral labor demand shifts.

Let productivity in sector i be $z_{it} = Z_t^{\eta_i}$ where η_i is a parameter measuring the elasticity of sectoral productivity to the aggregate shock Z with mean normalized to one. Let $\log Z_t$ follow a unit root process with innovation ϵ_t distributed as a $N(-\sigma_\epsilon/2, \sigma_\epsilon)$. In Appendix A.4, we show that the planner will allocate unemployed workers to equalize

$$\left[\frac{Z_t^{\eta_i-1}}{1-\beta(1-\Delta_t)(1-\delta_{it})\Omega_i} - \frac{\zeta}{1-\beta(1-\Delta_t)(1-\delta_{it})}\right]\phi_{it} m_{u_i}\left(\frac{v_{it}}{u_{it}^*}\right) \quad (3)$$

across sectors, where $\Omega_i \equiv \exp\left(\eta_i(\eta_i-1)\frac{\sigma_\epsilon}{2}\right)$. The new term Ω_i captures that the drift in future productivity in sector i varies proportionately with η_i because of the log-normality assumption. In essence, this sectoral drift changes the effective rate at which the planner discounts future output in that sector.

Understanding the nature of sectoral fluctuations exceeds the scope of this paper. A comparison of equations (2) and (3) reveals that the main lesson of this generalization is that our approach is valid under alternative views of what drives sectoral fluctuations: different views lead to different measurements of the sectoral component of productivity in the planner's allocation rule.

3 Mismatch index and mismatch unemployment

We now use the planner's allocation rule to derive an index measuring the severity of labor market mismatch between unemployed workers and vacancies. This mismatch index quantifies the fraction of hires lost because of misallocation, i.e., $(1 - h_t/h_t^*)$, where h_t denotes the observed aggregate hires and h_t^* the planner's hires. Next, we describe how this index allows to construct counterfactuals to measure the mismatch component of equilibrium unemployment.

From this point onward we must state an additional assumption, which is well supported by the data, as we show below: the sectoral matching function $m(u_{it}, v_{it})$ is Cobb-Douglas, i.e.,

$$h_{it} = \Phi_t \phi_{it} v_{it}^{\alpha} u_{it}^{1-\alpha}, \tag{4}$$

where h_{it} are hires in sector i at date t, and $\alpha \in (0, 1)$ is the vacancy share common across all sectors (in Section 7.6, we allow α to vary across sectors).

3.1 Mismatch index

From (4), summing across markets, the aggregate number of new hires can be expressed as:

$$h_t = \Phi_t v_t^{\alpha} u_t^{1-\alpha} \left[\sum_{i=1}^{I} \phi_{it} \left(\frac{v_{it}}{v_t} \right)^{\alpha} \left(\frac{u_{it}}{u_t} \right)^{1-\alpha} \right]. \tag{5}$$

The optimal number of hires that can be obtained by the planner allocating the u_t available unemployed workers across sectors is

$$h_t^* = \Phi_t v_t^{\alpha} u_t^{1-\alpha} \left[\sum_{i=1}^{I} \phi_{it} \left(\frac{v_{it}}{v_t} \right)^{\alpha} \left(\frac{u_{it}^*}{u_t} \right)^{1-\alpha} \right]. \tag{6}$$

Consider first the benchmark environment of Section 2.1. The optimality condition (1) dictating how to allocate unemployed workers between market i and market j

implies:

$$\frac{v_{it}}{u_{it}^*} = \left(\frac{\phi_{jt}}{\phi_{it}}\right)^{\frac{1}{\alpha}} \frac{v_{jt}}{u_{jt}^*}. \tag{7}$$

Substituting the optimality condition (7) in equation (6) , the optimal number of new hires becomes $h_t^* = \bar{\phi}_t \Phi_t v_t^\alpha u_t^{1-\alpha}$, where $\bar{\phi}_t = \left[\sum_{i=1}^{I} \phi_{it}^{\frac{1}{\alpha}} \left(\frac{v_{it}}{v_t}\right)\right]^\alpha$, a CES aggregator of the sector-level matching efficiencies weighted by their vacancy share. Therefore, we obtain the following expression for the mismatch index:

$$\mathcal{M}_{\phi t} = 1 - \frac{h_t}{h_t^*} = 1 - \sum_{i=1}^{I} \left(\frac{\phi_{it}}{\bar{\phi}_t}\right) \left(\frac{v_{it}}{v_t}\right)^\alpha \left(\frac{u_{it}}{u_t}\right)^{1-\alpha}. \tag{8}$$

$\mathcal{M}_{\phi t}$ measures the fraction of hires lost in period t because of misallocation. This index answers the question: if the planner had u_t available unemployed workers and used its optimal allocation rule, how many additional jobs would it be able to create? These additional hires are generated because, by better allocating the *same number* of unemployed, the planner can increase the aggregate job-finding rate and achieve more hires compared to the equilibrium (the "direct effect" of mismatch). It is useful to note that, in addition to this direct effect, u_t^* is in general lower than u_t which, for any given allocation rule, translates into a higher aggregate job-finding rate and more hires (the "feedback" effect of mismatch). $\mathcal{M}_{\phi t}$ measures only the direct effect of mismatch on hires, but the counterfactual of Section 3.2 fully incorporates the feedback effect as well.[12]

From (8) and (5) one can rewrite the aggregate matching function as

$$h_t = (1 - \mathcal{M}_{\phi t})\, \bar{\phi}_t \Phi_t v_t^\alpha u_t^{1-\alpha} \tag{9}$$

which makes it clear that higher mismatch lowers the (measured) aggregate efficiency of the matching technology and reduces the aggregate job-finding rate because some

[12]Dickens (2010) and Lazear and Spletzer (2012) use an alternative index proposed by Mincer (1966). In a previous version of this paper, we also reported results based on this index and argued that it is much less useful than the one we propose here because it only quantifies the number of job-seekers searching in the wrong sectors, but not how such misallocation lowers the job-finding rate and raises unemployment. In addition, the analysis in these papers does not allow for heterogeneity in productive and matching efficiency, a key determinant of the optimal allocation of job-seekers across labor markets.

unemployed workers search in the wrong sectors (those with relatively few vacancies). The term $\bar{\phi}_t$ can also contribute to a reduction in aggregate matching efficiency when the vacancy shares of the sectors with high ϕ fall.

In Appendix A.5, we show three useful properties of the index. First, $\mathcal{M}_{\phi t}$ is between zero (no mismatch) and one (maximal mismatch). Second, the index is invariant to "pure" aggregate shocks that shift the total number of vacancies and unemployed up or down, but leave the vacancy and unemployment shares across markets unchanged. Third, $\mathcal{M}_{\phi t}$ is increasing in the level of disaggregation. This last property suggests that every statement about the role of mismatch should be qualified with respect to the degree of sectoral disaggregation used.

Consider now the economy of Section 2.2, where labor markets also differ in their level of productive efficiency. It is useful to define "overall market efficiency" as $x_{it} \equiv \phi_{it}\left(z_{it}-\zeta\right)/\left[1-\beta\left(1-\Delta_t\right)\left(1-\delta_{it}\right)\right]$. Following the same steps, we arrive at the index

$$\mathcal{M}_{xt} = 1 - \sum_{i=1}^{I} \left(\frac{\phi_{it}}{\bar{\phi}_{xt}}\right)\left(\frac{v_{it}}{v_t}\right)^{\alpha}\left(\frac{u_{it}}{u_t}\right)^{1-\alpha}, \tag{10}$$

where

$$\bar{\phi}_{xt} = \sum_{i=1}^{I} \phi_{it}\left(\frac{x_{it}}{\bar{x}_t}\right)^{\frac{1-\alpha}{\alpha}}\left(\frac{v_{it}}{v_t}\right), \text{ with } \bar{x}_t = \left[\sum_{i=1}^{I} x_{it}^{\frac{1}{\alpha}}\left(\frac{v_{it}}{v_t}\right)\right]^{\alpha}. \tag{11}$$

$\bar{\phi}_{xt}$ is an aggregator of the market-level overall efficiencies weighted by their vacancy share.[13]

In the absence of heterogeneity with respect to matching efficiency, productivity, or job destruction, the index becomes $\mathcal{M}_t = 1 - \sum_{i=1}^{I}\left(\frac{v_{it}}{v_t}\right)^{\alpha}\left(\frac{u_{it}}{u_t}\right)^{1-\alpha}$. In what follows, we will also use the notation $\mathcal{M}_{zt}$ and $\mathcal{M}_{\delta t}$ to denote mismatch indexes for an economy where the only source of heterogeneity is productivity and job destruction rates, respectively.

Finally, the notation $\mathcal{M}_t^{AK}$ is used to denote the indexes calculated following the Abraham-Katz view of sectoral fluctuations. The only difference with $\mathcal{M}_{xt}$ is the

[13] Since the planner now maximizes output (and not employment), theoretically this index could be negative. An index which measures the fraction of *output* (instead of hires) lost to misallocation can be easily computed by weighting the gap between actual and planner's hires in each sector by sectoral productivity, and it is always positive.

definition of overall market efficiency x_{it}.

3.2 Mismatch unemployment

The mismatch index allows us to construct the counterfactual unemployment rate, u_t^*, in the absence of mismatch. Using (10), the actual aggregate job-finding rate in the economy at date t can be written as

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

Let u_t^* be counterfactual unemployment under the planner's allocation rule. The optimal number of hires at date t when u_t^* unemployed workers are available to be allocated across sectors is $\bar{\phi}_{xt}\Phi_t v_t^{\alpha}(u_t^*)^{1-\alpha}$. Therefore, the optimal job-finding rate (in absence of mismatch) is

$$f_t^* = \bar{\phi}_{xt}\Phi_t\left(\frac{v_t}{u_t^*}\right)^{\alpha} = f_t \cdot \underbrace{\frac{1}{(1-\mathcal{M}_{xt})}}_{\text{Direct Effect}} \cdot \underbrace{\left(\frac{u_t}{u_t^*}\right)^{\alpha}}_{\text{Feedback}} \tag{12}$$

There are two sources of discrepancy between counterfactual and actual job-finding rate. The first term in (12) captures the fact that a planner with u_t available job-seekers to move across sectors would achieve a better allocation and a higher job-finding rate. This effect, which we call the "direct" misallocation effect, is summarized by the mismatch index, as explained. The second term captures a "feedback" effect of misallocation: no mismatch means lower unemployment $(u_t^* < u_t)$ which, in turn, increases the probability of meeting a vacancy for job-seekers. This feedback effect can cause mismatch unemployment to remain above average for some time even if $\mathcal{M}_{xt}$ quickly reverts to its average after an increase, because it takes time for the additional unemployed to be reabsorbed. This is a pattern we observe in our empirical analysis.

Given an initial value for u_0^*, the dynamics of the counterfactual unemployment rate can be obtained by iterating forward on equation

$$u_{t+1}^* = s_t + (1 - s_t - f_t^*)\,u_t^*, \tag{13}$$

where s_t is the separation rate. Our strategy takes the sequences for separation rates $\{s_t\}$ and vacancies $\{v_t\}$ directly from the data when constructing the counterfactual sequence of $\{u_t^*\}$ from (13), an approach consistent with the theoretical model where vacancy creation and separations are exogenous to the planner. The gap between actual unemployment u_t and counterfactual unemployment u_t^* is mismatch unemployment.

In the next section we briefly discuss our methodology and the proper interpretation of our measure of mismatch unemployment. In the rest of the paper we apply this methodology to quantify the contribution of mismatch to the recent rise in the aggregate U.S. unemployment rate.

3.3 Interpretation of our measure of mismatch

Formalizing mismatch unemployment as "distance from a benchmark allocation," as we do, follows, in essence, the same insights of the vast literature on misallocation and productivity (Lagos, 2006; Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009; Moll, 2011; Jones, 2013). Our implementation has two distinctive features. First, we do not need to solve for equilibrium allocations (and, hence, make specific assumptions about firms' and workers' behavior, their information set, price determination, etc.). We simply take the empirical joint distribution of unemployment and vacancies across sectors as the equilibrium outcome.[14] Second, we construct the counterfactual distribution (in absence of mismatch) from a simple planner's problem which can be solved analytically. The strength of these two combined features is that finer disaggregation in the available micro data poses no threat to the feasibility of the exercise. The approach we propose is robust and easily implementable, even with a high number of labor markets, and multiple sources of heterogeneity, idiosyncratic shocks, and aggregate fluctuations.

Our methodology yields a measure of *mismatch across sectors* (defined by the jointly observable characteristics of job vacancies and unemployed job seekers), not within sectors. Put differently, concluding that mismatch plays a small role at the level of 2-digit occupations does not necessarily rule out its importance at the 3-

[14]The extension to endogenous vacancy requires a minimal set of, mostly standard, assumptions that are discussed in Section 6.

or 5-digit level.[15] It follows that, when quantifying the contribution of mismatch unemployment through our approach, it is important to clearly specify the level of disaggregation of the analysis. Moreover, our measure of mismatch captures the sectoral misallocation between job vacancies and *unemployed job seekers*. We therefore abstract from another class of job-seekers, employed workers who search on the job. We conjecture that, in the generalized planner problem where the planner can let employed workers search for jobs in more productive sectors, optimality would push the planner towards equalizing the (efficiency-weighted) ratio of vacancies to total job seekers (employed and not).[16] In Section 7.4, we verify that mismatch between vacancies and unemployment behaves very similarly to an index that also includes, among the job seekers, employed workers who report to search on the job.

The empirical method we have developed allows us to learn about the relative importance of different dimensions of mismatch by partitioning the labor market based on several characteristics (e.g., industry, occupation, education, geography). Studying how mismatch, and its dynamics, vary across these dimensions is surely informative about the forces at work in the economy. However, our methodology is not well suited to separately quantify the deeper sources of misallocation. This task requires specifying and solving a fully structural equilibrium model which, at the level of generality of our analysis, would be computationally unfeasible. Factors explaining the discrepancy between the empirical and planner's distribution of unemployment across sectors –that these structural models should incorporate– include moving (e.g., retraining or migration) costs, relative wage rigidity, risk-aversion and imperfect insurance, or certain government policies that may hamper the reallocation of idle labor from shrinking to expanding sectors. Since moving costs are a characteristic of the physical environment which would also feature in a planner's problem,

[15] This caveat applies even at a very high level of disaggregation. Observing a high number of vacancies for Web Developers (a 5-digit occupation) in Santa Clara county, and a high number of job-seekers in that same labor market would be interpreted as a sign of low mismatch across narrowly defined sectors. However, a situation where those same job-seekers do not have the technical knowledge required by the employers to staff their vacancies (e.g., the technology has changed and the skills of the unemployed have become obsolete), is a form of "skill mismatch."

[16] The other forces affecting the planner's solution depend on the details of how on-the-job search is modelled. For example, the degree of substitutability of unemployed and employed job-seekers in the matching function determines the congestion effect that the marginal job-seeker of one type imposes on the other type. Whether on-the-job search is costless, or has a cost in terms of foregone output or disutility, will determine the fraction of employed workers searching and their target sectors.

whereas our benchmark planner's allocation is derived under costless between-sector mobility, our calculations on the role of mismatch have the nature of an upper bound. The analysis of Herz and van Rens (2011) suggests that, among the sources of mismatch, relative wage rigidity (across locations and industries) is vastly more important than moving costs. In light of their finding, our planner problem may provide a tight upper bound.

4 Data

We focus on three definitions of labor markets: the first is a broad industry classification. The second is an occupation classification, based on both the 2-digit and 3-digit Standard Occupational Classification (SOC) system.[17]. The third is a geographic classification, based on U.S. counties and metropolitan areas (MSA's).[18]

To be empirically viable, our methodology calls for: (i) sectoral data on vacancies, unemployment, and the vacancy share of the matching function for the $\mathcal{M}$ index; (ii) data in (i) plus market-specific matching efficiency parameters for the $\mathcal{M}_{\phi t}$ index; and data in (ii) plus information on productive efficiency (productivity and separation rates) by sector for the $\mathcal{M}_{xt}$ index and its corresponding counterfactual. Deriving market-specific matching efficiencies, as well as the vacancy share, involves estimating matching functions and, therefore, requires data on hires.

4.1 Vacancies from the JOLTS and the HWOL

At the industry level, we use vacancy data from the Job Openings and Labor Turnover Survey (JOLTS), which provides survey-based measures of job openings and hires at a monthly frequency, starting from December 2000, for seventeen industries roughly corresponding to the 2-digit NAICS classification.[19] At the occupation and county

[17]See Tables C1-C3 in Appendix B for a list of industries and occupations used in the empirical analysis. In total, there are 22 2-digit SOC's and 93 3-digit SOC's. We use all the 2-digit categories with the exception of Farming, Fishing, and Forestry. We exclude 3-digit SOC's exhibiting fewer than 10 observations in the CPS unemployment counts at least once in the sample period. These small cells account for 60% of the 3-digit SOC's, but represent only 15.6% of unemployed workers in the CPS.

[18]We focus on counties whose population is at least 50,000 and group together counties in the same metropolitan area. This procedure gives a total of 280 local labor markets.

[19]Since the JOLTS is a well known and widely used survey, we do not provide further details. For more information, see http://www.bls.gov/jlt/. See also Faberman (2009).

level, we use vacancy data from the Help Wanted OnLine (HWOL) dataset provided by The Conference Board (TCB). This is a novel data set containing the universe of online advertised vacancies posted on internet job boards or in newspaper online editions. It covers roughly 16,000 online job boards and provides detailed information about the characteristics of advertised vacancies for three to four million unique active ads each month.[20] The HWOL database started in May 2005 as a replacement for the Help-Wanted Advertising Index of print advertising maintained by TCB.[21]

Each observation in the HWOL database refers to a unique ad and contains information about the listed 6-digit occupation, the geographic location of the advertised vacancy down to the county level, whether the position is for full-time, part-time, or contract work (essentially self-employed contractors or consultants: e.g., computer specialists, accountants, auditors), the education level required for the position, and the hourly and annual mean wage.[22] For 57% of ads we also observe the industry NAICS classification. The majority of online advertised vacancies are posted on a small number of job boards: about 60% of all ads appear on five job boards.[23]

It is worth mentioning some measurement conventions in the HWOL data: first, the same ad can appear on multiple job boards. To avoid double-counting, TCB uses a sophisticated unduplication algorithm that identifies unique advertised vacancies on the basis of the combination of company name, job title/description, city or state. Second, there are some cases in which multiple locations (counties within a state) are listed in a given ad for a given position. TCB follows the rule that if the counties are in the same state or MSA the position is taken to represent a single vacancy, but if they appear in different MSA's and in different states they reflect distinct vacancies. In addition, the dataset records one vacancy per ad. There is a small number of cases in which multiple positions are listed, but the convention used is one vacancy per ad.

[20]The data are collected for The Conference Board by Wanted Technologies. For detailed information on survey coverage, concepts, definitions, and methodology see the Technical Notes at http://www.conference-board.org/data/helpwantedonline.cfm

[21]Our empirical analysis covers the December 2000-June 2011 period for the JOLTS, and May 2005-June 2011 for the HWOL.

[22]The education and wage information is imputed by TCB. Education is imputed from BLS data on the education content of detailed 6-digit level occupations. Wages are imputed using BLS data from the Occupational Employment Statistics (OES), based on the occupation classification. For a subset of the ads we also observe the sales volume and the number of employees of the company, as well as the actual advertised salary range, but in this paper we do not attempt to use this additional information.

[23]The five largest job boards are: CareerBuilder, Craigslist, JOBcentral, Monster, and Yahoo!HotJobs.

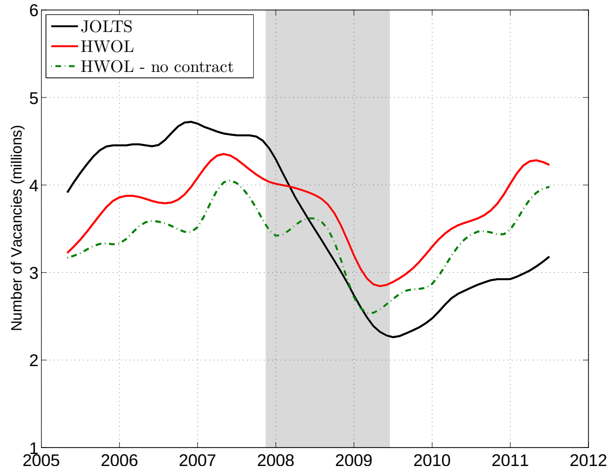


Figure 1: Comparison between the JOLTS and the HWOL (The Conference Board Help Wanted OnLine Data Series) aggregate time series.

More importantly, the growing use of online job boards over time may induce a spurious upward trend. Figure 1 plots JOLTS vacancies and HWOL ads at the national level. The total count of active vacancies in HWOL is below that in JOLTS until the beginning of 2008 and is above from 2008 onwards, a pattern which may reflect the increasing penetration of online job listings over time. Nevertheless, the average difference between the two aggregate series is only about 16% of the JOLTS total, and the correlation between the two aggregate series is about 0.65. To the extent that this trend towards online recruitment does not differ too much across sectors, our calculations are not affected. In Section 7.5, we propose a reweighing scheme for HWOL that aligns it more closely to JOLTS and show that our findings remain robust. We report additional detailed comparisons between the JOLTS and HWOL vacancy series in Appendix B.1.

4.2 Unemployment from the CPS

We calculate unemployment counts from the Current Population Survey (CPS) for the same industry and occupation classification that we use for vacancies.[24] For geography, we use the Local Area Unemployment Statistics (LAUS) which provides

[24]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. Only about 1.5% of unemployed are missing occupation information. Some of these workers have never worked before and some are self-employed.

monthly estimates of total unemployment at the county and MSA level.[25] The CPS reports the industry and occupation of unemployed workers' previous jobs. We begin by assuming that all unemployed workers search only in the sector that they had last worked in. We relax this assumption in Section 7. The small sample size of the CPS limits the level of disaggregation of our analysis, and prevents us from using HWOL ads data to their full effect.[26]

4.3 Matching functions

To compute market-specific matching efficiency parameters, ϕ_i, and vacancy share α, we estimate aggregate and sector-specific (constant-returns to scale) matching functions using various specifications, estimation methods, and data sources. In particular, we follow Borowczyk-Martins et al. (2012) in dealing with the well known endogeneity issues in matching function estimation. Appendix B.2 contains a detailed description of our methodology and results.

Our findings (see Table C4 in Appendix C) indicate that a value of the vacancy share $\alpha = 0.5$ is appropriate. This value is roughly in the middle of the range of estimates used in other recent papers in the matching literature.[27] Moreover, our mismatch indices are typically highest for $\alpha = 0.5$; therefore, this value is consistent with the spirit of reporting an upper bound for mismatch unemployment. Tables C6-C8 in Appendix C contain estimates of sector-specific matching efficiencies.

4.4 Productive efficiency

We use various proxies for productivity, depending on data availability. At the industry level, we compute labor productivity by dividing value added for each industry from the Bureau of Economic Analysis (annual data) by average employment in that industry from the Establishment Survey.[28] At the occupation level, for lack of a better proxy, we use annual data on average hourly wages from the Occupational

[25] See http://www.bls.gov/lau/ for more information on LAUS.

[26] The average number of unemployed in the CPS for the May 2005 to June 2011 period is 4,557 with a range of 2,808 to 12,436.

[27] A few examples are $\alpha = 0.5$ in Davis, Faberman, and Haltiwanger (2010), $\alpha = 0.28$ in Shimer (2005), $\alpha = 0.54$ in Mortensen and Nagypal (2007), α between 0.66 and 0.72 in Barnichon and Figura (2013).

[28] http://www.bea.gov/industry/

Employment Statistics (OES).[29] Similarly, at the county level, we use median weekly wage earnings from the Quarterly Census of Employment and Wages (QCEW).[30] We recognize that wage levels might be affected by factors other than productivity like unionization rates, compensating differentials, monopoly rents, etc. To partially address this issue, we normalize the average wage for each occupation to unity at the beginning of our sample and focus on relative wage movements over time. We also apply the same normalization to industry-level productivity measures for consistency.

We calculate job destruction rates at the industry level from the Business Employment Dynamics (BED) as the ratio of gross job losses to employment.[31] Since the BED is quarterly, we assume that the destruction rate is the same for the three months corresponding to a specific quarter and impute the corresponding monthly destruction rates. Because job destruction rates by occupation are not available, we compute the employment to unemployment transition rates by occupation in the last job from the CPS semi-panel. Figures C3 and C4 in Appendix C show the evolution of productivity and job destruction rates for selected industries and occupations.

Finally, with respect to output from home-production for the non-employed, ζ, our quantitative analysis indicates that the impact of mismatch is the largest when $\zeta = 0$. In keeping with our "upper bound" nature of the measurement exercise, in our baseline calculations we use this value, but verify the robustness of our conclusions for a range of values for ζ between 0 and 0.25.

5 Results

We begin by documenting the dynamics of the cross-sectoral correlation between vacancy and unemployment shares, which anticipates some of our findings on the mismatch indexes. Next, we study industry-level, occupational-level, and geographical mismatch unemployment, in that order.

[29] See http://www.bls.gov/oes/

[30] See http://www.bls.gov/cew/

[31] See http://www.bls.gov/bdm/. We recognize this is an imperfect proxy for separations, but (i) monthly employment-unemployment transitions computed from CPS semi-panel at the industry level are much noisier, and (ii) during 2001-2010, only 16 pct of quits ends into unemployment, as opposed to 91 pct of layoffs (see Elsby et al., 2010).

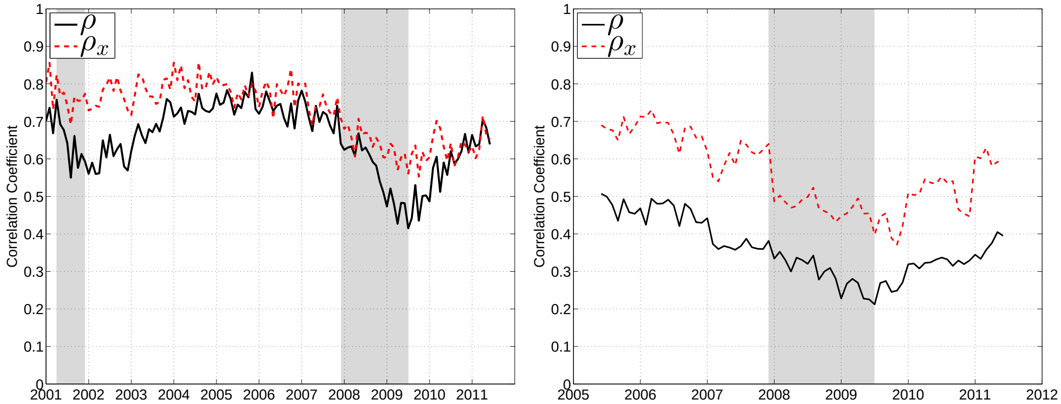


Figure 2: Correlation coefficient between u and v shares across industries (left panel) and two digit occupations (right panel).

5.1 Correlation between vacancy and unemployment shares

From our definition of mismatch, it is clear that there is a close association between mismatch indexes and the correlation between unemployment and vacancy shares across sectors. The planner's allocation rule implies a perfect correlation between unemployment shares and (appropriately weighted) vacancy shares. A correlation coefficient below one is a signal of mismatch, and a declining correlation is a signal of worsening mismatch. Figure 2 plots the time series of this correlation coefficient across industries (left panel) and occupation (right panel) over the sample period. For each case, we report two different correlation coefficients motivated by the definitions of the mismatch indexes we derived in Section 3: ρ: between (u_{it}/u_t) and (v_{it}/v_t), and ρ_x: between (u_{it}/u_t) and $(x_i/\bar{x}_t)^{\frac{1}{\alpha}}(v_{it}/v_t)$. The two series behave similarly. They drop sharply from early 2006 to mid 2009 and recover thereafter, indicating a rise in mismatch during the recession that is, however, relatively short-lived.

5.2 Industry-level mismatch

The left panel of Figure 3 plots $\mathcal{M}_t$ and $\mathcal{M}_{xt}$ across 2-digit industries.[32] This figure shows that, before the last recession (in mid 2006), the fraction of hires lost because of misallocation of unemployed workers across industries ranged from 2-3 percent

[32]All mismatch indexes throughout the paper are HP filtered to eliminate high frequency movements and better visualize the variation in the indexes. To facilitate the comparison across different definitions of labor markets, we plot all the mismatch indexes and mismatch unemployment rates using the same vertical distance on the y axis, 0.15 and 2.5 percentage points, respectively.

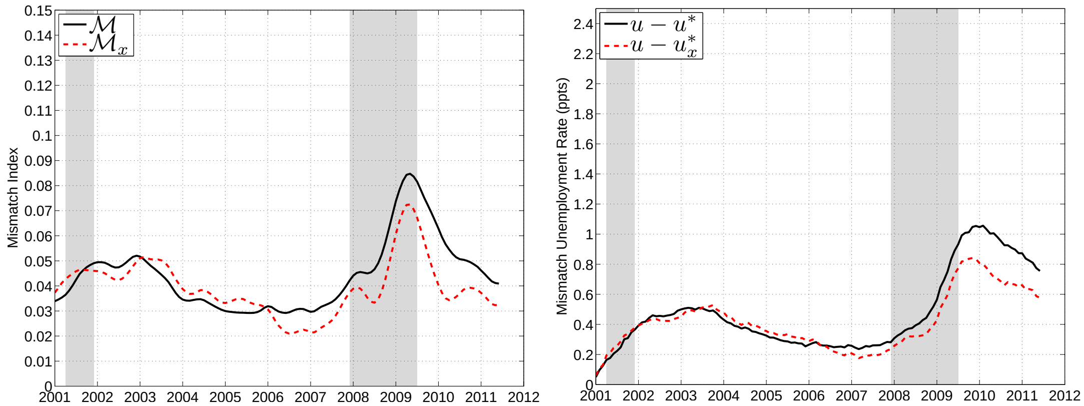


Figure 3: Mismatch index $\mathcal{M}_t$ and $\mathcal{M}_{xt}$ by industry (left panel) and the corresponding mismatch unemployment rates (right panel).

per month, depending on the index used. At the end of the recession, in mid 2009, it had increased to roughly 7-8 percent per month, and it has since dropped again to almost its pre-recession level. To sum up, both indexes indicate a sharp rise in mismatch between unemployed workers and vacant jobs across industries during the recession, and a subsequent fairly rapid decline.[33]

How much of the observed rise in the unemployment rate can be explained by mismatch? Table 1 shows the change in mismatch unemployment between the average of 2006 and October 2009.[34] The main finding is that worsening mismatch across these seventeen industries explains (depending on the index used) between 0.59 and 0.75 percentage points of the rise in U.S. unemployment from 2006 to its 2009 peak, i.e., at most 14 percent of the increase. The right panel of Figure 3 shows

[33]To shed more light on the dynamics of the mismatch index, it is useful to examine the evolution of vacancy and unemployment shares of different industries, i.e., the individual components of the index. In Figure C5, we plot the vacancy and unemployment shares for a selected set of industries using the JOLTS definition in Appendix C. 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 durable goods manufacturing went up while the unemployment share of the health sector decreased. Starting from 2010, sectoral unemployment and vacancy shares began to regress towards 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.

[34]The average unemployment rate was 4.6% in 2006 and 10.0% at its peak in October 2009, indicating a 5.4 percentage point increase. Throughout the paper we compare the average of 2006 with the unemployment peak (October 2009) when we discuss the role of mismatch in the increase in the unemployment rate.

	Index	$u_{06}-u^*_{06}$	$u_{10.09}-u^*_{10.09}$	$\Delta(u-u^*)$	$\Delta(u-u^*)/\Delta u$
Industry	$\mathcal{M}$	0.26	1.01	0.75	13.9%
	$\mathcal{M}_x$	0.24	0.84	0.59	11.0%
	$\mathcal{M}_x^{AK}$	0.28	0.89	0.61	11.2%
	$\mathcal{M}_x^{v^*}(\varepsilon=0.5)$	0.67	1.90	1.22	22.5%
	$\mathcal{M}_x^{v^*}(\varepsilon=1.0)$	0.35	1.24	0.90	16.6%
	$\mathcal{M}_x^{v^*}(\varepsilon=2.0)$	0.27	0.95	0.69	12.7%
2-digit Occ.	$\mathcal{M}$	0.85	2.00	1.15	21.3%
	$\mathcal{M}_x$	0.42	1.02	0.60	11.1%
	$\mathcal{M}_x^{v^*}(\varepsilon=0.5)$	1.08	2.60	1.52	28.1%
	$\mathcal{M}_x^{v^*}(\varepsilon=1.0)$	0.75	1.81	1.07	19.7%
	$\mathcal{M}_x^{v^*}(\varepsilon=2.0)$	0.58	1.41	0.83	15.3%
3-digit Occ.	$\mathcal{M}$	1.33	2.91	1.58	29.3%
	$\mathcal{M}_x$	0.79	1.73	0.94	17.4%
Routine/Cognitive	$\mathcal{M}^{RC}$	0.41	1.07	0.67	12.3%
County	$\mathcal{M}$	0.32	0.46	0.14	2.6%
	$\mathcal{M}_z$	0.32	0.45	0.14	2.5%
2-digit $\times$ division	$\mathcal{M}$	0.81	1.71	0.90	16.9%
2-digit	$\mathcal{M}$	0.68	1.53	0.85	16.0%

Table 1: Changes in mismatch unemployment at the industry, occupation, and county levels. All the differences are calculated as the difference between October 2009 and the average of 2006. Note that $\Delta u = 5.4$ percentage points. All calculations are monthly, except for the last two lines which are quarterly.

mismatch unemployment (i.e., the difference between the actual and the counterfactual unemployment rates) at the industry level for the 2001-2011 period, computed as described in Section 3.2. Mismatch unemployment has declined since early 2010, but it remains above its pre-recession levels. Figure C6 in Appendix C shows mismatch indexes with one source of heterogeneity at a time, $\mathcal{M}_\phi, \mathcal{M}_z, \mathcal{M}_\delta$, and the corresponding mismatch unemployment rates. The results are very similar.

In Section 2.3, we have shown how the planner's allocation rule changes under the alternative Abraham-Katz interpretation of sectoral employment movements. As Table 1 shows, the corresponding index $\mathcal{M}_{xt}^{AK}$ implies a contribution of mismatch unemployment similar to the benchmark.[35]

Table C9 and Figures C8-C11 in Appendix C contain a sensitivity analysis on industry-level mismatch with respect to (i) values of α ranging from 0.3 to 0.7; (ii)

[35]Figure C7 in Appendix C shows the mismatch index and the corresponding mismatch unemployment computed using the benchmark specification and this alternative interpretation.

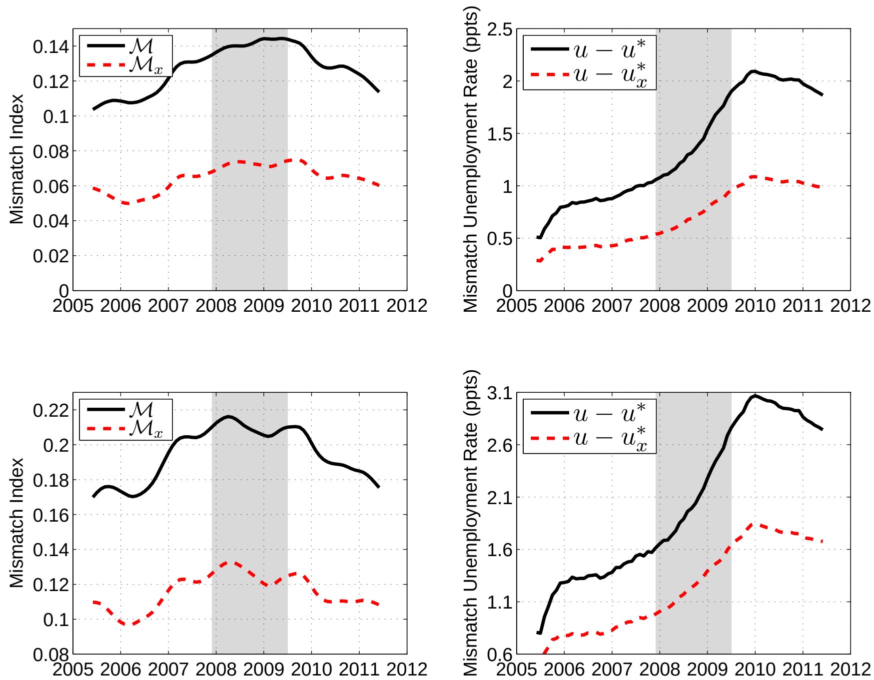


Figure 4: Mismatch indexes $\mathcal{M}_t$ and $\mathcal{M}_{xt}$ by 2-digit occupation (upper left panel) and 3-digit occupation (lower left panel). Corresponding mismatch unemployment rates for 2-digit (upper right panel) and 3-digit occupations (lower right panel).

alternative estimates of matching efficiency ϕ_i's which are separately estimated for the periods before and after the recession;[36] (iii) values of the home-production flow ζ ranging between 0 and 0.25 of aggregate productivity; (iv) using hires data from the CPS instead of the JOLTS; and (v) using HWOL vacancy data by industry instead of the JOLTS. The results are very robust: the contribution of (2-digit) industry-level mismatch to the rise in the unemployment rate around the Great Recession varies between 0.5 and one percentage points.

5.3 Occupation-level mismatch

Figure 4 plots the $\mathcal{M}_t$ and $\mathcal{M}_{xt}$ indexes (left panels) and the resulting mismatch unemployment (right panels) for 2 and 3-digit SOC's. $\mathcal{M}_t$ index for 2-digit occupations rises by almost 4 percentage points. Similar to the pattern observed for industries, the rise in mismatch leads the recession by over a year. As seen in the figure and in

[36]We denote this index as $\mathcal{M}_x^{break}$.

Table 1, based on the $\mathcal{M}_t$ index, around 1.1 percentage points (or 21%) of the recent surge in U.S. unemployment can be attributed to occupational mismatch measured at the 2-digit occupation level. At the 3-digit level, the portion of the increase in unemployment attributable to mismatch is around 1.6 percentage points (or roughly 29% of the rise in the unemployment rate).[37]

The $\mathcal{M}_{xt}$ index is lower than the $\mathcal{M}_t$ index and features a smaller rise, implying around 2% of additional hires lost because of mismatch. This index suggests that between 0.6 and 0.9 percentage points of the rise in the unemployment rate (or between 11% and 17% of the increase) was due to mismatch at the 2-digit and 3-digit SOC levels, respectively. Therefore, similar to what we found for industries, the index that accounts for heterogeneity in matching and productive efficiency across occupations, implies a smaller role for mismatch unemployment.[38]

Table C10 and Figures C14-C15 in Appendix C contains a sensitivity analysis on occupational-level mismatch at the 2-digit level with respect to (i) the value of α; (ii) alternative estimates of matching efficiency ϕ_i's which are separately estimated for the periods before and after the recession; and (iii) values of the home-production flow ζ ranging between 0 and 0.25 of aggregate productivity. Our findings remain robust to these alternative specifications.

5.3.1 The role of job polarization for occupational mismatch

Job polarization refers to the increasing concentration of employment in the highest- and lowest-wage occupations, with job opportunities in middle-skill occupations disappearing, as documented by Acemoglu and Autor (2011). To capture the effect of job polarization on mismatch, we classify 2-digit occupations into four categories: routine cognitive, routine manual, non-routine cognitive, and non-routine manual.

[37]Figure C12 in Appendix C shows the unemployment and vacancy shares of selected 2-digit SOC's, i.e., the individual components of the index. 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 that experienced a decline in their vacancy shares and an increase in their unemployment shares. Concurrently, vacancy shares of health-care practitioner and sales and related occupations went up and the corresponding unemployment shares declined. Starting from 2010, similar to the JOLTS data, unemployment and vacancy shares began to normalize.

[38]Figure C13 in Appendix C shows mismatch indexes with one source of heterogeneity at a time, $\mathcal{M}_\phi, \mathcal{M}_z, \mathcal{M}_\delta$. The corresponding mismatch unemployment rates at the 2-digit occupation level are reported in Table C10.

We call this classification "Routine/Cognitive" and denote the corresponding mismatch index with $\mathcal{M}^{RC}$.[39] Figure C16 in Appendix C contrasts the unadjusted mismatch index across these four occupation groups against the index calculated at the 2-digit level, and reports the implied path for mismatch unemployment. Our findings are summarized in Table 1. The lower level of the index suggests additional mismatch within these four broad categories. Despite the gap in the level of the two indices, the dynamics of the $\mathcal{M}^{RC}$ index are similar to those of the mismatch index computed using all 2-digit occupations. In essence, the vacancy (unemployment) share dropped (rose) faster for routine manual occupations relative to the other groups, accounting for at least half of the increase in mismatch unemployment across the twenty-one 2-digit occupations.

Jaimovich and Siu (2012) link the job polarization hypothesis to jobless recoveries by analyzing employment changes during recessions and recoveries across these occupational groups. They show that employment declined more in routine occupations during the most recent downturn, in line with the increase in mismatch during the recession. They also show that employment remained stagnant in all occupational categories during the recovery, which is consistent with the decline in mismatch after the recession.

5.3.2 Occupational mismatch within education groups and within regions

Is occupational mismatch a more relevant source of unemployment dynamics for less skilled or for more skilled workers? A priori, the answer is ambiguous: more education means more adaptability, but also more specialized knowledge. To address this question, we define four education categories (less than high school diploma, high school diploma or equivalent, some college or Associate's degree, Bachelor's degree or higher) and analyze mismatch by 2-digit occupation *within* each of these four education groups.

The CPS provides information on the education level of the unemployed. Recall that each job listing recorded in HWOL reports its 6-digit occupation. The BLS provides information on the distribution of workers employed in each 6-digit occupation

[39]We classify occupations at the 2-digit level instead of directly using Acemoglu and Autor's classification. While their way of classifying occupations is more detailed, our classification broadly captures this distinction and is more comparable with the rest of our analysis. See Table C2 in Appendix C for our classification of occupations into these four groups.

	$u_{06} - u^*_{06}$	$u_{10.09} - u^*_{10.09}$	$\Delta(u - u^*)$	Δu	$\Delta(u - u^*)/\Delta u$
Less than HS	0.71	1.69	0.98 ppts	8.5 ppts	11.5%
HS Degree	0.60	1.50	0.89 ppts	6.9 ppts	12.9%
Some College	0.71	1.68	0.97 ppts	5.3 ppts	18.2%
College Degree	0.38	1.03	0.65 ppts	2.7 ppts	23.9%

Table 2: Changes in mismatch unemployment across 2-digit occupations for different education groups using $\mathcal{M}_t$. All the changes are calculated as the difference between October 2009 and the average of 2006. Note that $\Delta u = u_{10.09} - u_{06}$ and that Δu varies by education.

broken down by their educational attainment.[40] We allocate the total 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.[41] Finally, we aggregate up to the 2-digit level to obtain vacancy counts for each occupation by education cell. The implicit assumption we make in using the BLS information is that the educational requirement of newly created vacancies, for each occupation, is equal to the educational content in the existing jobs for that same occupation.

The counterfactual exercises summarized in Table 2 reveal a 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 8.5 percentage point increase in the unemployment rate of that group. For high school graduates, mismatch explains 0.89 (13%) out of the 6.9 percentage point increase in unemployment. For those with some college, mismatch explains about 1.0 (18%) out of a 5.3 percentage point rise in unemployment, and for college graduates 0.65 (24%) 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

[40]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.

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

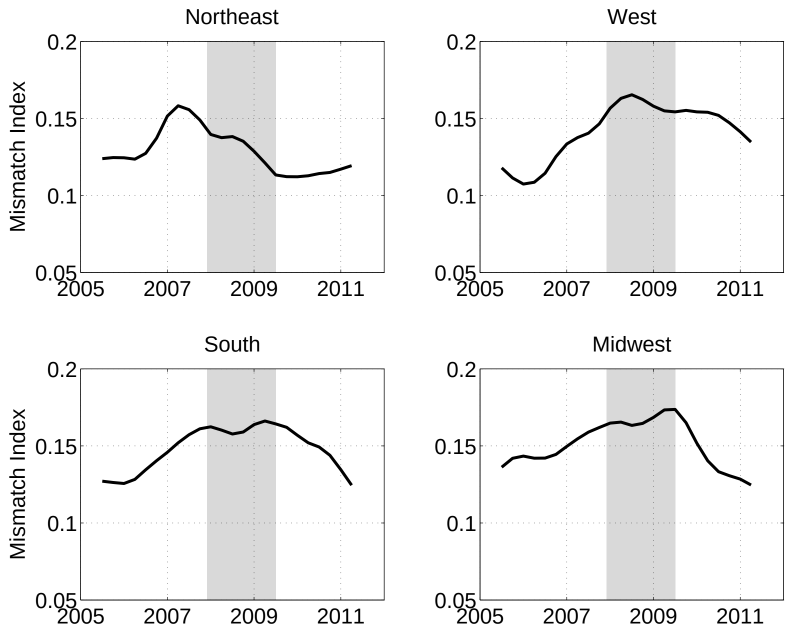


Figure 5: 2-digit occupational mismatch indexes $\mathcal{M}_t$ in the four U.S. Census regions.

to roughly one quarter of the increase for each group.[42]

Looking at occupational mismatch separately for each of the four U.S. Census regions (Figure 5) reveals that the only region where our index is still significantly above its pre-recession level is the West, i.e. the region where the fall in house prices and the rise in unemployment were the sharpest.

5.4 Geographical mismatch

We perform our geographical analysis on mismatch across U.S. counties using the HWOL data on online job ads coupled with LAUS data on the unemployed.

Figure 6 shows the indexes $\mathcal{M}_t$ and $\mathcal{M}_{zt}$ and the corresponding mismatch unemployment rates. We find that geographic mismatch is very low (about 1/10 of the 2-digit occupation index, even though the number of sectors is 10 times higher) and is essentially flat over the sample period. These two results are interesting because they indicate that (i) the rise of the index with the number of sectors, and (ii) its counter-cyclicality are not mechanical features of our methodology, but they depend

[42]Figures C17 in Appendix C plots mismatch indexes within each broad education category. The index for college graduates is the only one which is still significantly above its 2006 level.

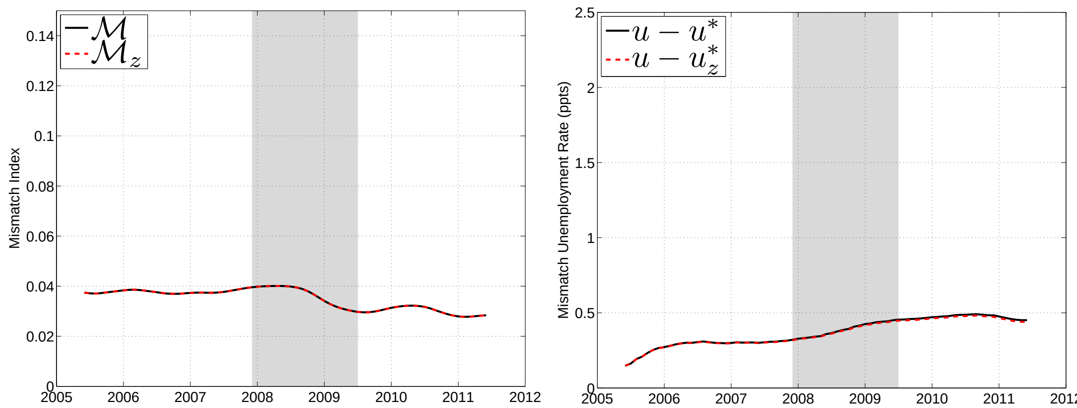


Figure 6: Geographical mismatch indexes $\mathcal{M}_t$ and $\mathcal{M}_{zt}$ by county (left panel) and corresponding mismatch unemployment rates (right panel).

on how the equilibrium distribution of unemployment and vacancies varies (i) across labor markets and (ii) evolves over the cycle.

Unsurprisingly, the rise in mismatch unemployment implied by this index is around one tenth of a percentage point, implying that geographical mismatch—across U.S. counties and MSAs—played a negligible role in the recent dynamics of U.S. unemployment. This finding is consistent with other recent work that investigated the link between housing market and labor market using different methods (see, e.g., Schulhofer-Wohl, 2010; Farber, 2012; Karahan and Rhee (2012); Kothari, Saporta-Ecksten, and Yu, 2013).[43]

We also examine mismatch across labor markets jointly defined by occupation and location. Because of the small sample size of the CPS, we define sectors as the combination of 2-digit occupations and the nine Census divisions, and perform our analysis at the quarterly frequency. Both mismatch index and mismatch unemployment are very similar to those computed at the 2-digit occupation level.[44]

5.5 Is the Great Recession different from the 2001 recession?

At the industry-level, the sample is long enough to allow a comparison of mismatch unemployment in the Great Recession to that of the 2001 recession. Figures 2 and

[43] We also compute geographic mismatch for the 50 U.S. states using the HWOL data on online job ads coupled with CPS data on the unemployed. The JOLTS provides limited geographic information, enabling us to study mismatch only across the four broad Census regions. Our conclusions from these state- and region-based analyses are fully aligned with the county-based study.

[44] See Figure C18 in Appendix C and Table 1.

3 show that the fall in the cross-sectoral unemployment-vacancy correlation and the rise in our mismatch index is common to the last two downturns. In Table C11 in Appendix C we report our calculations on the role of mismatch unemployment in 2001. We find that worsening mismatch accounted for a larger portion of the (smaller) rise in unemployment in the 2001 recession (23% instead of 11-14%). This finding echoes the fact that the dynamics of employment for different occupational groups were much more asymmetric in 2001 than in 2008 (Jaimovich and Siu, 2012).

6 Endogenous vacancy distribution

In this section, we relax the assumption of exogeneity of the distribution of vacancies maintained so far. Why would endogenizing vacancies affect our calculations? If, in equilibrium, too many job-seekers search in the sectors with low matching and productive efficiency, private firms' job creation decisions are distorted: an excessive number of vacancies will be posted in those sectors (because of the higher probability of recruitment) compared to the choice of a planner who allocates vacancies and job seekers based on relative efficiency across sectors. The result is a lower number of aggregate vacancies and a lower aggregate job-finding rate in equilibrium—another "feedback" effect of mismatch stemming, this time, from the vacancy side.

We begin by stating some additional assumptions on the equilibrium data generating process required to identify the shocks to the vacancy creation cost. These cost-shocks are needed to compute the planner's counterfactual vacancy distribution. We then proceed to formally explain this additional feedback effect of mismatch. Finally, we present our findings. Appendix A.6 contains more details on all the derivations.

6.1 Measurement of the vacancy creation cost

Let the cost, in terms of final good, of creating v_{it} vacancies in sector i at date t be

$$K_{it}\left(v_{it}\right) = \kappa_{it}^{\varepsilon} \cdot \frac{v_{it}^{1+\varepsilon}}{1+\varepsilon}, \quad \text{with } \varepsilon \in \left(0, \infty\right). \tag{14}$$

With this isoelastic specification, ε measures the elasticity of the vacancy creation cost, i.e., how the (log of the) the marginal cost increases with the (log of the) number of vacancies.[45] The random variable κ_{it} shifts the cost of vacancy creation across sectors and over time. We let κ_{it} be independent of the other idiosyncratic shocks, and denote its conditional distribution as Γ_κ. The choice of how many vacancies to post takes place after observing sectoral and aggregate states, but before the allocation of unemployment across sectors.

Up to this point, we could conduct our analysis without modeling the behavior and choices of firms and workers in equilibrium. However, the measurement of $\{\kappa_{it}\}$ requires imposing a minimal amount of structure on the equilibrium data generating process. Three assumptions suffice: (1) free entry of vacancies in each sector; (2) a bargaining protocol between firms and workers such that the firm obtains a share λ, and the worker a share $(1-\lambda)$, of the expected discounted output flow—in particular, outside options do not matter for the bargaining outcome (as in Shaked and Sutton, 1984; Acemoglu, 1996); and (2) no within-market congestion externality, in the spirit of Hosios (1990).[46]

Free entry is the standard condition determining vacancies in this class of matching models. The choice of bargaining protocol is convenient because it enables us to remain agnostic about the equilibrium value of unemployment for a worker—therefore reducing to a bare minimum the structure needed on the equilibrium model. The Hosios condition isolates mismatch unemployment as the unique source of discrepancy between the efficient and equilibrium distributions of vacancies.

Under assumptions (1) and (2), the equilibrium condition in the economy of Section 2.2 with heterogeneity in $\{\phi_{it}, z_{it}, \delta_{it}, \kappa_{it}\}$ is:

$$\kappa_{it}^{\varepsilon}\left(v_{it}\right)^{\varepsilon} = \Phi_t \phi_{it}\left(\frac{u_{it}}{v_{it}}\right)^{1-\alpha} \lambda \frac{Z_t z_{it}}{1-\beta\left(1-\Delta_t\right)\left(1-\delta_{it}\right)} \tag{15}$$

stating that the marginal cost of a vacancy in sector i (the left hand side), also heterogeneous across sectors, is equated to its expected marginal gain for the firm (the

[45] Because of constant returns in the sector-specific matching function, it is the convexity of the cost function that prevents concentrating all vacancies and unemployed workers in the sector with the highest efficiency. We follow the convention, common in this literature, that this cost has to be paid every period the vacancy is maintained open.

[46] The extensive form game corresponding to this bargaining outcome is spelled out in Acemoglu (1996, Appendix 1). The key assumption is that if, once the pair is formed, a party wants to quit the bargaining, it can rematch within the period within the same sector (i.e., with an identical partner) by paying a small fixed cost.

right hand side). Note that the individual firm takes the sectoral meeting probability as given. Note also that, as $\varepsilon \to \infty$, $v_{it} = 1/\kappa_{it}$, i.e., vacancies are exogenously determined. This special case corresponds to the economy of Section 2.

All variables in condition (15) are observable, except for κ_{it} and ε. For a given value of the elasticity ε, we derive the sequence for κ_{it} that makes that condition hold exactly at every date t in each sector i. This strategy amounts to attributing, residually, fluctuations in vacancies to variation in the cost of job creation, once exogenous variation in productivity and separation rates (both observable) have been accounted for.[47] Then, we can use this cost sequence in the planner's vacancy creation condition to compute the planner's distribution of vacancies.

6.2 Comparison between equilibrium and planner FOCs

In Appendix A.6, we show that the planner problem of Section 2.2, augmented with a vacancy creation decision where the planner faces the cost function (14), yields the first-order condition

$$\kappa_{it}^{\varepsilon}\left(v_{it}^{*}\right)^{\varepsilon} = \Phi_t \phi_{it}\left(\frac{u_{it}^{*}}{v_{it}^{*}}\right)^{1-\alpha} \alpha \frac{Z_t z_{it}}{1-\beta\left(1-\Delta_t\right)\left(1-\delta_{it}\right)} \tag{16}$$

equating the marginal cost of a vacancy to its marginal gain, in turn equal to the expected discounted value of output conditional on matching, times the marginal effect of an additional vacancy on the probability of meeting an unemployed worker allocated to sector i.[48]

A comparison of equations (15) and (16) is instructive. Imposing the Hosios condition $\lambda = \alpha$ in (15), within-market congestion externalities are ruled out and the

[47] It is well known that productivity shocks alone are unable to explain fluctuations in vacancies in a matching model with standard parameterization (Shimer, 2005). Investigating the fundamental sources of vacancy fluctuations is beyond the scope of this paper. We limit ourselves to point out that recent papers (e.g., Petrosky-Nadeau, 2013) have emphasized the role of credit shocks and asymmetric information in lending for the observed collapse of job creation during the last recession. In these models, this mechanism works through the free entry condition, precisely as a source of fluctuations in κ_{it}. A planner subject to the same asymmetric information would face the same fluctuations in κ_{it}.

[48] For ease of exposition, in equation (16) we have already set the flow output from non-employment ζ to zero, since this is the value we use in the quantitative analysis (to facilitate the comparison with the baseline model). Recall that in the model with exogenous vacancies we used $\zeta = 0$ because we found that it is the value that maximizes the role of mismatch. All the derivations in Appendix A.6 are obtained for the general case $\zeta \geq 0$.

only reason why equilibrium vacancies in sector i differ from their efficient counterpart is that the number of unemployed workers is the "wrong" one, i.e., the only reason is mismatch unemployment. If in equilibrium an excessive number of unemployed workers search for jobs in declining sectors, firms would create more vacancies than the planner in those sectors, amplifying the initial source of misallocation. Combining equations (15) and (16), we therefore arrive at the relationship

$$\frac{v_{it}}{v_{it}^*} = \left(\frac{u_{it}}{u_{it}^*}\right)^{\frac{1-\alpha}{1-\alpha+\varepsilon}}$$

which demonstrates that the extent to which mismatch unemployment, i.e. deviations of u_{it} from u_{it}^*, translate into misallocation of vacancies in equilibrium (i.e., deviation of v_{it} from v_{it}^*) depends on the value of the elasticity ε. If the marginal cost function is steep (ε high), large differences in the ratio (u_{it}/u_{it}^*) and, therefore, in meeting probabilities and expected output gains, translate into small differences in the ratio (v_{it}/v_{it}^*) . In this case, the planner's vacancies are close to equilibrium vacancies, as assumed in our benchmark analysis. If, instead, ε is close to zero, the misallocation of unemployed workers across sectors translates "one for one" into the distribution of vacancies.

In Appendix A.6, we lay out a simple algorithm to compute the planner's optimal allocation of vacancies across sectors $\{v_{it}^*\}$, and we explain how to modify the calculation of counterfactual unemployment to take into account this additional margin of choice for the planner. It is instructive to examine the relationship between the planner and the equilibrium aggregate job-finding rate in this economy:

$$f_t^* = f_t \cdot \underbrace{\frac{1}{(1-\mathcal{M}_{xt})}}_{\text{Direct Effect}} \cdot \underbrace{\left(\frac{u_t}{u_t^*}\right)^{\alpha}}_{\text{Feedback through } u} \cdot \underbrace{\left[\left(\frac{\bar{\phi}_{xt}^*}{\bar{\phi}_{xt}}\right) \cdot \left(\frac{v_t^*}{v_t}\right)^{\alpha}\right]}_{\text{Feedback through } v}, \tag{17}$$

where $\bar{\phi}_{xt}$ is given by equation (11) and $\bar{\phi}_{xt}^*$ is the same aggregator, but with the planner's vacancy shares v_{it}^*/v_t^* instead of the observed shares. Compared to (12), the equation above features an additional feedback effect of mismatch that operates through vacancies and has two components. Mismatch reduces the aggregate job-finding rate by (i) distorting the distribution of vacancy shares across sectors (the first term in the square brackets) , and (ii) lowering total vacancies (the second term).

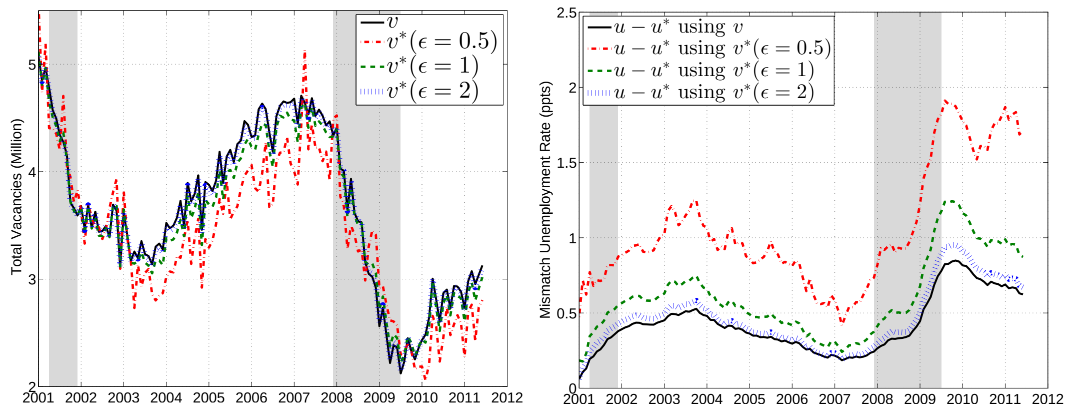


Figure 7: Aggregate vacancies and (left panel) and corresponding mismatch unemployment rates (right panel) at the industry level using endogenous vacancies specification with JOLTS.

6.3 Results

The first challenge we face is to choose a value for the marginal cost elasticity ε. Here, we rely on the existing literature. Merz and Yashiv (2007) specify a cost function where the argument is hires, and estimate an elasticity of 2.40 on aggregate US time series. Given a Cobb-Douglas specification for the matching function and a value for $\alpha = 0.5$, their estimate translates into an elasticity with respect to vacancies of 1.20. Coşar, Guner, and Tybout (2011) use establishment-level data for Colombia and estimate ε to be 1.085. Lise and Robin (2013) report an estimate of ε of 1.12 based on aggregate US time series. In all these papers, the identification of ε comes from the response of vacancies and employment changes to productivity shocks, and ε is precisely estimated. We conclude that existing estimates of ε, at various level of disaggregation, are quite tightly centered around one.

Given ε, we can estimate the sector-specific vacancy cost creation vector $\{\kappa_{it}\}$. Our estimates of vacancy costs κ_{it} increase for almost all industries and occupations during the recession, therefore contributing to the observed drop in vacancies. Figure C19 in Appendix C plots the estimated sequences of κ_{it} in some selected industries and occupations for the case $\varepsilon = 1$. Next, we compute the distribution of planner's vacancies and the implied planner's aggregate job-finding rate with endogenous vacancies (17), which we then feed into the law of motion for the unemployment rate to perform our counterfactual exercise.

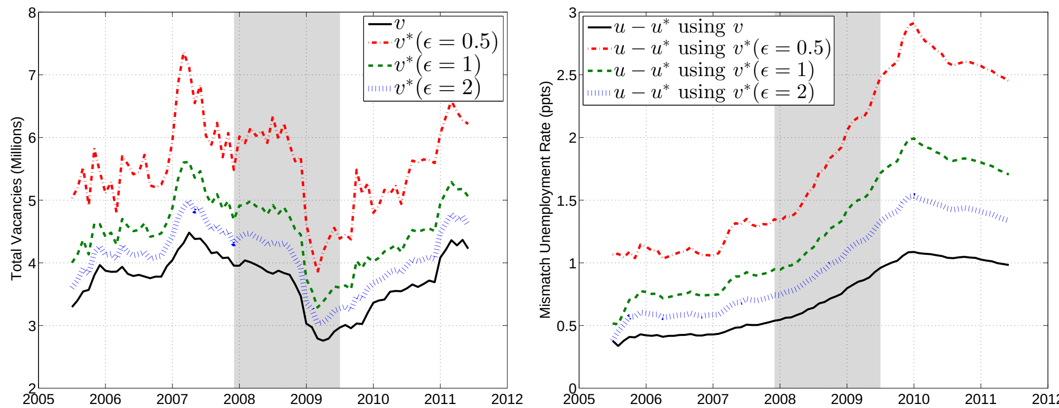


Figure 8: Aggregate vacancies and (left panel) and corresponding mismatch unemployment rates (right panel) at the occupation level using endogenous vacancies specification with the HWOL (The Conference Board Help Wanted OnLine Data Series).

Table 1 summarizes the results.[49] We first present our analysis by industry. Figure 7 (left panel) plots aggregate vacancies v_t^* in the planner's economy for different values of ε. The main result is that quantitatively significant deviations between v_t^* and v_t (the data) occur only for low values of the cost elasticity ε. For $\varepsilon \geq 1$, planner and equilibrium vacancies line up closely. This finding is reflected into the calculation of mismatch unemployment (right panel). For $\varepsilon = 1$, with endogenous vacancy creation, mismatch unemployment rises by 0.9 percentage points between 2006 and October 2009, i.e., only an additional 0.3 percentage points relative to the exogenous vacancy calculation. For $\varepsilon = 0.5$, mismatch unemployment is generally higher, but its increase between 2006 and October 2009 is still about 1.2 percentage points—not far from the case of unit elasticity.

Turning to occupations, for $\varepsilon = 1$, planner and equilibrium vacancies line up fairly closely and, as Figure 8 indicates, the contribution of mismatch unemployment to the rise in the U.S. unemployment rate between 2006 and October 2009 is 1.1 percentage points. For $\varepsilon = 0.5$, it increases up to 1.5%, or 28% of the total rise in unemployment.

To summarize, as expected, the contribution of mismatch unemployment is larger when the distribution of vacancies is endogenized. Nevertheless, our results of Section 5 derived under exogenous vacancies (or infinite marginal cost elasticity) are close to those obtained from the model with endogenous vacancy creation and uni-

[49]The indexes computed with endogenous vacancies have superscript v^*.

tary marginal cost elasticity, a specification supported by existing estimates. Our calculations also show that mismatch could have played a major role in the recent rise of unemployment, by dampening aggregate vacancy creation, only if one is willing to maintain that the cost elasticity is below 1/2. While our current knowledge suggests that such a range is not too plausible, the number of available empirical estimates of this parameter is still small, so more research is needed to firmly establish this inference.

7 Robustness on inputs and specification of the matching function

The matching function is a key ingredient of our analysis. In this section we investigate a number of potential concerns that relate to the measurement of its inputs (job seekers and job vacancies) and to its specification.

Our unemployment counts for industry and occupation are calculated from the CPS samples. We explore whether this random sampling can generate a bias in our mismatch index. With respect to job seekers, we have assumed that each unemployed worker is searching in the same industry or occupation as the one where she was last employed. Here, we correct our index for the direction of search based on observed unemployment-employment transitions. Since the focus of our study is on mismatch *unemployment*, so far we have only included unemployed workers among job-seekers in all our calculations. It is useful to ask whether our findings are robust to broader definitions of job-seekers which includes (i) discouraged workers, and (ii) employed workers searching on the job. The HWOL data on aggregate vacancies show a stronger upward trend than their JOLTS counterpart. If this trend is uneven across sectors, it may bias our mismatch measures. Here we assess the magnitude of this bias. Finally, we have assumed that the input shares of vacancies and unemployment $(\alpha, 1-\alpha)$ are constant across sectors. This assumption is crucial for maintaining tractability, but the model can be solved numerically with heterogeneous shares to confirm this restriction does not drive our findings. The results of this sensitivity analysis are summarized in Table 3 .

Since the endogenous vacancy creation margin did not substantially affect our results, in this section we use the baseline model with an exogenous distribution of vacancies. With the exception of the adjustment for the direction of search (done both at the industry and occupation level) and and heterogeneity in input shares –which requires a long time series and is therefore done at the industry level– we perform our sensitivity analysis for 2-digit occupations. Finally, we use $\mathcal{M}_t$ (the index unadjusted for heterogeneity) since, as clear from Table 1, it is the one that leads to the largest role for mismatch.

7.1 Sampling error as a potential source of bias

In Section 5 we documented a positive correlation between unemployment and vacancy shares across industries and occupations. Under this scenario, classic measurement error in sectoral unemployment counts may lead to an upward bias in our mismatch index because it artificially lowers the cross-sectoral correlation between vacancy and unemployment shares towards zero (an example of "division bias").

To assess the size of the bias, we draw 5,000 independent samples, with replacement, from our CPS data at the 2-digit occupation level. Each bootstrapped sample is of the same size as the original CPS sample.[50] For each sample, we compute the mismatch index. The mean index computed from the resulting sampling distribution is virtually identical to our point estimate, suggesting that this potential source of bias is quantitatively negligible. With the sampling distribution in hand, we are also able to compute confidence intervals for the mismatch index and for mismatch unemployment. The 95% confidence band is around 0.2 percentage points for both variables, thereby confirming that our estimates are quite precise. See Figure C20 in Appendix C.

7.2 Adjustment for direction of search

We now relax the assumption that unemployed workers search in their last sector of employment, and propose an alternative calculation of the number of job-seekers in each industry or occupation by exploiting the semi-panel dimension of the CPS. Respondents in the CPS are interviewed for several consecutive months and we can track unemployed workers who find new employment from one month to the next and

[50]We did it in two ways: (i) using the unweighted microdata from the CPS, and (ii) using the population weights in the CPS. Results are almost unchanged.

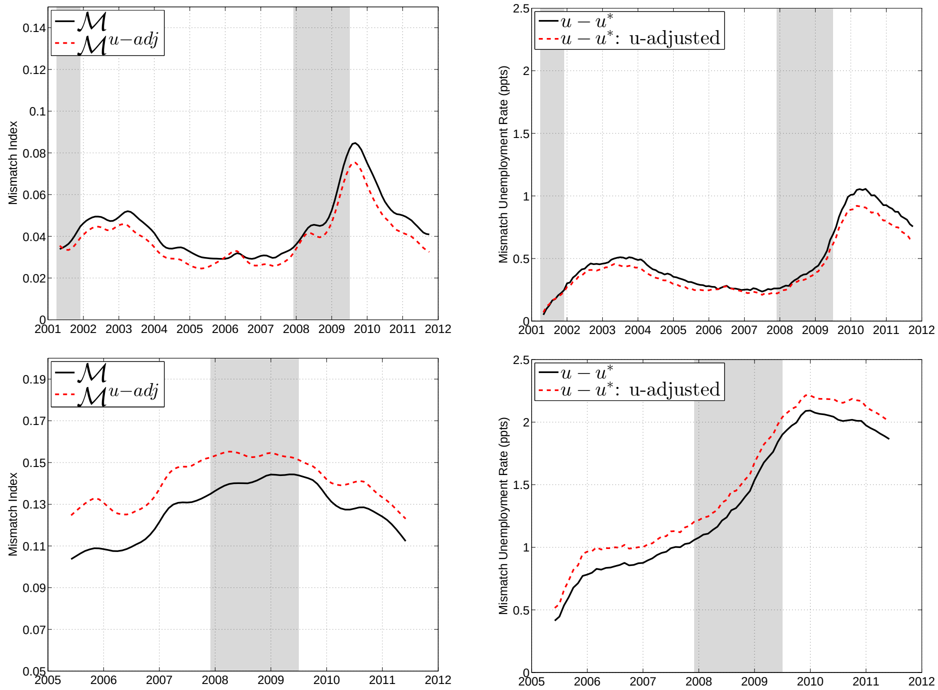


Figure 9: Mismatch index with unadjusted ($\mathcal{M}$) and adjusted ($\mathcal{M}^{u-adj}$) unemployment counts by industry (top-left panel) and corresponding mismatch unemployment (top-right panel). Mismatch index with adjusted and unadjusted unemployment counts by occupation (bottom-left panel) and corresponding mismatch unemployment (bottom-right panel).

record: (i) industry/occupation of the job prior to the worker's unemployment spell; (ii) industry/occupation of the new job. We then create annual transition matrices (from sector i to sector j) by aggregating monthly flows, as in Hobijn (2012). We then infer the number of job seekers in each sector using a simple statistical algorithm, whose key assumption is that every unemployed searching for a job in sector j has the same probability of being hired, independently of the sector of origin, except when coming from sector j itself in which case she is allowed to have a higher job-finding rate. The method is outlined in detail in Appendix B.3.[51]

We first report our results by industry. The top-left panel of Figure 9 shows the mismatch index calculated using the adjusted unemployment counts, which we call $\mathcal{M}_t^{u-adj}$, as well the unadjusted $\mathcal{M}_t$ index. The adjustment causes the level of the

[51]Figures C21 and C22 plot the adjusted and unadjusted unemployment counts for some selected industries and occupations. As expected, for example, this correction reduces the number of unemployed workers searching in construction and increases that of those seeking jobs in healthcare.

	$u_{06} - u^*_{06}$	$u_{10.09} - u^*_{10.09}$	$\Delta(u - u^*)$	$\Delta(u - u^*)/\Delta u$
$\mathcal{M}$	0.85	2.00	1.14	21.3%
$\mathcal{M}^{u-adj}$	0.84	2.00	1.16	21.4%
$\mathcal{M}^{v-adj}$	0.92	2.12	1.19	22.1%
$\mathcal{M}^D$ (all D in U)	0.92	2.03	1.11	20.6%
$\mathcal{M}^D$ (D from constr. and prod. in U)	1.06	2.33	1.27	23.4%
$\mathcal{M}^E$ (E: weighted by search time)	0.78	1.90	1.13	20.9%
$\mathcal{M}^E$ (E: fraction searching)	0.79	1.97	1.18	21.8%

Table 3: Changes in mismatch unemployment across 2-digit occupations using the baseline index $\mathcal{M}_t$ with different adjustments. The first adjustment for discouraged (D) workers counts all discouraged workers as unemployed (U) while the second one only counts discouraged workers from construction and production as unemployed. The first adjustment for employed (E) job seekers is done by using the time used for job search by the employed relative to the unemployed while the second adjustment assumes that all employed who report positive search time are counted as unemployed. All the changes are calculated as the difference between October 2009 and the average of 2006.

index to decrease somewhat during the sample period. When using the adjusted counts, 0.65 percentage points of the roughly 5.4 percentage point rise in the U.S. unemployment rate is due to industry-level mismatch, compared to 0.75 percentage points without the adjustment (top-right panel).

The bottom row of Figure 9 reports our analysis by occupation. Again, both the adjusted $\mathcal{M}_t^{u-adj}$ index and mismatch unemployment track their counterparts without adjustment. In contrast to the industry-level analysis, the adjusted index for occupations is slightly higher than in the baseline case. However, quantitatively, the contribution of mismatch to the rise in the U.S. unemployment rate is virtually the same when using adjusted unemployment counts by occupation.

The key reason why our findings are robust to this adjustment is that the estimated transition matrices by industry and occupation reveal that the bulk of the unemployed workers keeps searching in the sector of their previous employment. Table 3 summarizes these results.

7.3 Adjustment for discouraged workers

According to the CPS, an individual is unemployed if she does not have a job, has actively looked for employment in the past four weeks and is currently available to work. However, it is possible that some workers become discouraged from unsuccessful job search and reduce their search intensity enough to be classified as out of the labor force in the official statistics. This grey area between unemployment and non-participation is occupied by "discouraged workers."[52]

If workers from certain occupations are more likely than others to become discouraged (and exit from unemployment) or remain discouraged (and delay re-entry into the unemployment ranks), our mismatch measures—based on the official unemployment counts—may be biased. For example, if most of the discouraged workers who dropped out of the labor force during the last recession originate from the construction sector, then the number of unemployed would be an under-estimate of the true number of potential job seekers in the construction sector. In this example, actual mismatch would be larger than what we measure when including only the unemployed among the job-seekers. However, if the number of discouraged workers across sectors is roughly proportional to that of the unemployed, then the effect of this adjustment would be minor.

To correct for this potential source of bias, we count workers in the CPS classified as "discouraged not in the labor force" (D), record their previous occupation, and add them to the corresponding unemployment stock, month by month, for the entire sample period.[53] Table C12 in Appendix C reveals that, on average, the distributions of discouraged and unemployed workers are strikingly similar across occupations—the correlation is around 0.95. As a consequence, including discouraged workers affects the job-seeker shares of different occupations only marginally. As Table 3 and Figure 10 show, the difference between the modified mismatch index, which we call $\mathcal{M}^D$, and the original index is quantitatively insignificant.

Next, to maximize the potential impact of such correction, we only count discour-

[52]The CPS classifies as discouraged workers those individuals "not in the labor force who want and are available for a job and who have looked for work sometime in the past 12 months (or since the end of their last job if they held one within the past 12 months), but who are not currently looking because they believe there are no jobs available or there are none for which they would qualify."

[53]The information about previous occupation of discouraged workers is incomplete. We therefore compute the distribution of previous occupations and we impute it (as if that was a random subsample) to the entire sample from the sub-sample of discouraged workers for which we have this information (around 10% of the total). We have also tried an alternative strategy where we identified those workers who flowed from unemployment to discouragement between month t and $t + 1$ and we added them back to the unemployment pool in the occupation of origin at month t. Results are similar with both methods, but the effect of this adjustment is larger for the first strategy and, hence, in what follows we only report results for that case.

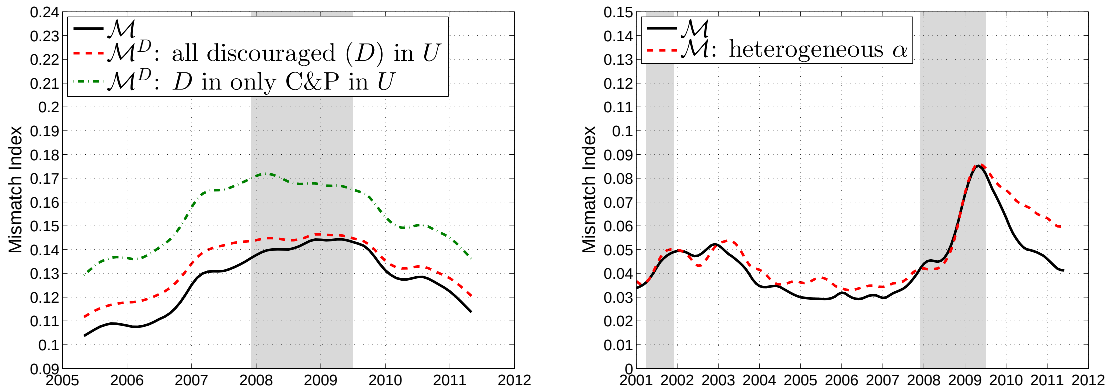


Figure 10: Mismatch index $\mathcal{M}_t^D$ by occupation including all discouraged (D) workers and only discouraged workers in Construction and Production (C&P) in unemployment (U) (left panel). Mismatch index $\mathcal{M}_t$ by industry allowing for heterogenous vacancy share parameter α across industries (right panel).

aged workers in construction and production related occupations (mostly manufacturing) as unemployed, the occupation groups with the largest increase (decrease) in their unemployment (vacancy) share. Once again, the adjustment has small effects: the contribution of mismatch to the rise in the unemployment rate is 23.4% as opposed to 21.3% in the baseline case. All these results are reported in Table 3, and Figure C23 in Appendix C shows the plot of mismatch unemployment.

7.4 Adjustment for employed job-seekers

Since the CPS does not have any information on job search behavior of employed workers, we use the American Time Use Survey (ATUS) to impute the number of employed job seekers in each sector. The ATUS reports the amount of time respondents devoted to various activities on the day preceding the day of the interview, including time spent on job search activities. In addition, it reports the individual's occupation (2-digit SOC) and her employment status. These data allow us to make an adjustment for on-the-job search. The correction is in the same spirit as the one for discouraged workers, i.e., broadening the notion of job seekers, and this modified index is called $\mathcal{M}^E$.

We implement two versions of this adjustment. First, we compute the ratio between the average time spent searching by employed workers in occupation i and that spent by the unemployed, and augment the job-seeker count in occupation i in month t with a number equal to that ratio times the CPS employment stock in this same occupation-month.[54] The shortcoming of this method is that, if employed workers allocate less time to job search because they are more effective, we would underestimate the contribution of employed job-seekers. In our second version, we compute the number of all the workers employed in occupation i who report any positive amount of time spent searching for another job and add it to the unemployment stock in occupation i.[55]

These modifications do not result in major changes in the distribution of job seekers across occupations and thus have very small effects on our mismatch measures.[56] The plot of the modified mismatch index is shown in Figure C24 in Appendix C.

7.5 Reweighing of HWOL vacancies

The two main concerns with the HWOL data are that (i) some sectors may systematically over- or under-use online recruitment tools compared to the aggregate and (ii) the upward trend in the penetration of online advertisement may be faster or slower in some sectors than others. To address these concerns, we reweight HWOL vacancy counts by occupation in order to match the total vacancy counts by industry and region in JOLTS, month by month. Appendix B.4 describes our approach in detail.

Table C13 in Appendix C reports the estimated weights by industry and region. A low (high) weight means that sector or region makes use of online recruitment boards more (less) than the aggregate economy. Our findings are quite intuitive: Finance, Real Estate, and Professional Services are among the most over-represented

[54] The ATUS has a considerably smaller sample size relative to the CPS, so we can only make this adjustment for each occupation by pooling all the years (2003-2011) together.

[55] For this extension, we do not perform a correction for the direction of search, as we did for unemployed job seekers. A recent paper by Hyatt and McEntarfer (2013) uses Longitudinal Employer-Household Dynamics (LEHD) data to calculate the industries of origin and destination of employment-to-employment flows. We calculated the correlation between the entries of their transition matrix and the entries of the one we estimated for unemployed workers in Section 7.2. This correlation is very high (0.96), suggesting that our correction for on-the-job search would be robust to a further correction for the direction of search, as the one proposed for the unemployed job-seekers.

[56] The correlation between the modified and original unemployment shares of occupations over time is between 0.987 and 0.997 with the first method and above 0.999 with the second method. The average absolute difference between the modified and the original index is 0.01 when we use the first method and 0.007 when we use the second method.

industries in online recruitment, and Accommodation, Government, and Construction among the most under-represented. Weights change somewhat over time, but the correlation between the 2005-06 and the 2010-11 weights is 0.90, indicating that the upward trend is quite common across sectors.

When we recompute the mismatch index using these reweighted vacancy counts by 2-digit occupation ($\mathcal{M}_t^{v-adj}$) we do find a slightly higher increase in occupational mismatch (see Figure C25 in the Appendix), but as can be seen in Table 3, the counterfactual exercise yields results similar to our baseline calculation with the raw HWOL data. Overall, these findings are encouraging and, over time, more will be learned about the virtues and limitations of this new data set. For the moment, one should bear in mind that results based on HWOL may be not as definitive as those based on JOLTS.[57]

7.6 Heterogeneous α across sectors

So far, we have assumed that the elasticity of hires to vacancies (α) in the matching function is the same for all sectors. Here we relax this assumption and follow the derivation in Appendix A.7 to numerically solve for the sectoral mismatch index and for mismatch unemployment, when α varies across sectors. We perform this analysis by industry because we need a long time series to precisely estimate α_i sector by sector, and JOLTS has over 50 data points more than HWOL. Table C14 in Appendix C reports the estimates of α_i and the implied new estimates of ϕ_i by industry. There is some variation in α_i across industries and, while most of these differences are statistically insignificant, there are sectors with large elasticities (e.g., Health and Government, where α_i is between 0.7-0.8) and others with elasticities half as large (e.g., Construction and Real estate, where α_i is between 0.35-0.4).

How much does this heterogeneity affect our estimates of mismatch unemployment at the industry level, relative to the homogeneous α case? Figure 10 shows that

[57] In a previous version of the paper (Şahin et al, 2012) we also address the issue that vacancies may be measured with error (in both JOLTS and HWOL), since not all hires occur through formal advertisement (see, e.g., Galenianos, 2012, for an analysis of hiring through referrals). We show that markets where vacancies are severely under-reported look like markets with higher matching efficiency, and argue that our calculations are still appropriate. Intuitively, it makes no difference to the planner whether ϕ_{it} is high in a sector because pure matching efficiency is high or because actual vacancies are larger than those formally advertised: in both cases, the planner would like to allocate many job-seekers to that sector.

the two mismatch indexes track closely each other until the end of the recession, but the index calculated allowing for heterogeneity in α declines more gradually afterwards. As a result, mismatch unemployment (displayed in Figure C26 in Appendix C) remains higher (but only by 0.2 percentage points) than its homogeneous α counterpart throughout 2010.

8 Conclusion

In this paper, we developed a framework to coherently define and measure mismatch unemployment. We use this framework to ask how much sectoral mismatch contributed to the dynamics of U.S. unemployment around the Great Recession. Our findings indicate that mismatch across counties, 2-digit industries, and 3-digit occupations explains around 1/3 of the recent rise in the U.S. unemployment rate. Our formalization of mismatch, and several choices made in our measurement exercise, mean that this estimate should be considered as an upper bound for each level of disaggregation we analyzed.

While, admittedly, our approach does not put us in the best position to separately identify the many potential causes of mismatch, we argued that analyzing different layers of disaggregation (e.g., occupation, industry, education, geography), as we do, is informative nevertheless. The absence of an increase in geographical mismatch casts doubts on the "house lock" hypothesis, a conclusion in line with existing research (e.g., Schulhofer-Wohl, 2010; Farber, 2012; Karahan and Rhee (2012); Kothari, Saporta-Ecksten, and Yu, 2013). The non-negligible role played by occupational mismatch, especially for high-skilled workers, leaves room for explanations based on labor demand shifts combined with human capital specialization, relative wage rigidity, and government policies. Kambourov and Manovskii (2009), Alvarez and Shimer (2010), Carrillo-Tudela and Visscher (2013), and Wiczer (2013), among others, have proposed equilibrium models where unemployed workers cumulate specific human capital and, in equilibrium, make explicit mobility decisions across distinct labor markets. Going forward, these frameworks should be, potentially, well suited to investigate the structural causes of mismatch unemployment, i.e., why job seekers search for jobs in the "wrong" sectors.

References

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

[2] Acemoglu, D. (1996). "A Microfoundation For Social Increasing Returns in Human Capital Accumulation," *Quarterly Journal of Economics*, vol. 111, 779-804.

[3] Acemoglu, D. and D. Autor (2011). "Skills, Tasks and Technologies: Implications for Employment and Earnings," in O. Ashenfelter and D. Card (Eds.), Handbook of Labor Economics, vol. 4B, Chapter 12, 1043-1171. Elsevier.

[4] Alvarez F. and R. Shimer (2010). "Search and Rest Unemployment," *Econometrica*, vol. 79:1, 75-122.

[5] Barlevy, G. (2011). "Evaluating the Role of Labor Market Mismatch in Rising Unemployment," Federal Reserve Bank of Chicago, Economic Perspectives, vol. 35, Third Quarter.

[6] Barnichon R. and A. Figura (2013). "Labor Market Heterogeneity and the Aggregate Matching Function," mimeo.

[7] Borowczyk-Martins, D., G. Jolivet, and F. Postel-Vinay (2013). "Accounting For Endogeneity in Matching Function Estimation," *Review of Economic Dynamics*, vol. 16(3), 440-451.

[8] Carrillo-Tudela C. and L. Visscher (2010). "Search and Endogenous Reallocation over the Business Cycle," mimeo.

[9] Coşar, K. , N. Guner, and J. Tybout (2011). "Firm Dynamics, Job Turnover, and Wage Distributions in an Open Economy," NBER Working Paper 16326.

[10] Daly, M., B. Hobijn, A. Şahin, and R. Valletta (2012). "A Rising Natural Rate of Unemployment: Transitory or Permanent?," *Journal of Economic Perspectives*, vol. 26(3), 3-26.

[11] Davis, S. J., Faberman, R. J., and J.C. Haltiwanger (2010). "The Establishment-Level Behavior of Vacancies and Hiring," *Quarterly Journal of Economics*. Forthcoming.

[12] Dickens, W. T. (2010). "Has the Recession Increased the NAIRU?", mimeo, Northeastern University and the Brookings Institution.

[13] Eeckhout, J. and P. Kircher (2011). "Identifying Sorting: In Theory," *Review of Economic Studies*, vol. 78(3), 872-906.

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

[15] Faberman, R. J. (2009). "Studying the Labor Market with the Job Openings and Labor Turnover Survey," in Producer Dynamics: New Evidence from Micro Data Volume, edited by T. Dunne, J. Jensen, and M. Roberts, U. of Chicago Press.

[16] Farber H. S. (2012). “Unemployment in the Great Recession: Did the Housing Market Crisis Prevent the Unemployed from Moving to Take Jobs?,” *American Economic Review: Papers & Proceedings*, vol. 102(3), 520525

[17] Galenianos, M. (2012). “Hiring Through Referrals,” mimeo.

[18] Hall, R.E. (2010). “The Labor Market in the Current Slump,” mimeo.

[19] Herz, B. and T. Van Rens (2011). “Structural Unemployment,” mimeo.

[20] Hobijn, B. (2012). “ Industry-Occupation Mix,” mimeo.

[21] Hosios, A. (1990). “On the Efficiency of Matching and Related Models of Search and Unemployment,” *Review of Economic Studies* vol. 57, 279-298.

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

[23] Hyatt, H. R., and E. McEntarfer (2013). “Job-to-Job Flows and the Business Cycle,” mimeo.

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

[25] Jaimovich, N. and H. Siu (2012). “The Trend is the Cycle: Job Polarization and Jobless Recoveries,” NBER Working Paper 18334.

[26] Jones C.I. (2013). “Misallocation, Economic Growth, and Input-Output Economics,” in Advances in Economics and Econometrics, Tenth World Congress, vol. II, Cambridge University Press.

[27] Kambourov, G. and I Manovskii (2009). “Occupational Mobility and Wage Inequality,” Review of Economic Studies, vol. 76(2), 731-759.

[28] Karahan, F. and S. Rhee (2012). “Geographical Reallocation and Unemployment during the Great Recession: The Role of the Housing Bust,” mimeo.

[29] Kothari, S., Saporta-Eksten, I. and E. Yu (2013). “The (Un)importance of Geographical Mobility in the Great Recession”, *Review of Economic Dynamics*, vol. 16(3), 553-563.

[30] Lagos, R. (2006). “A Model of TFP,” *Review of Economic Studies*, vol. 73, 983-1007.

[31] Lazear, E. and J. Spletzer (2012). “The United States Labor Market: Status Quo or A New Normal?,” mimeo.

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

[33] Lise, J. and J-M. Robin (2013) “The Macro-dynamics of Sorting between Workers and Firms, mimeo.

[34] Merz, M. and E. Yashiv (2007). "Labor and the Market Value of the Firm," *American Economic Review*, vol. 97, 1419-1431.

[35] Mian, A. and A. Sufi (2011). "What Explains High Unemployment? The Aggregate Demand Channel," mimeo.

[36] Mincer, J. (1966). "The Concept of Job Vacancies in a Dynamic Theory of the Labor Market," in The Measurement and Interpretation of Job Vacancies, Columbia University Press.

[37] Moll, B. (2011). "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?," mimeo.

[38] Mortensen, D.T. and Nagypal (2007). "More on Unemployment and Vacancy Fluctuations," *Review of Economic Dynamics*, vol. 10, 327-347.

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

[40] Petrosky-Nadeau, N. (2013). "Credit, Vacancies and Unemployment Fluctuations," *Review of Economic Dynamics.* Forthcoming.

[41] Pissarides C. (2000), Equilibrium Unemployment Theory (2nd edition), MIT Press, Cambridge, MA.

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

[43] Şahin, A., J. Song, G. Topa, and G.L. Violante (2012). "Mismatch Unemployment," NBER Working Paper 18265.

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

[45] Shaked, A. and J. Sutton (1984). "Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model," *Econometrica*, vol. 52, 13511364.

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

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

[48] Veracierto (2011). " Worker Flows and Matching Efficiency," Federal Reserve Bank of Chicago, Economic Perspectives, vol. 35, Fourth Quarter.

[49] Wiczer, D. (2013). "Long Term Unemployment: Attached and Mismatched?" mimeo.

Appendix Not For Publication

A Theoretical Appendix

This Appendix formally derives all the theoretical results of Sections 2 and 6. In what follows, we adopt a recursive formulation for all the planner's problems, and state them as dynamic-programming problems where the arguments of the planner's value function V are the relevant state variables. The prime symbol $(')$ is used to denote next-period values.

A.1 Benchmark environment

We solve the planner's problem of Section 2.1. The efficient allocation at any given date is the solution of the following planner's problem that we write in recursive form:

$$\begin{aligned}
V\left(\mathbf{e};\mathbf{v},\phi,Z,\Delta,\Phi\right) &= \max_{\{u_i\geq 0\}}\sum_{i=1}^{I} Z\left(e_i+h_i\right)+\beta\mathbb{E}\left[V\left(\mathbf{e}';\mathbf{v}',\phi',Z',\Delta',\Phi'\right)\right] \\
s.t. &: \\
\sum_{i=1}^{I}\left(e_i+u_i\right) &= 1 \qquad \text{(A1)}\\
h_i &= \Phi\phi_i m\left(u_i,v_i\right) \qquad \text{(A2)}\\
e_i' &= \left(1-\Delta\right)\left(e_i+h_i\right) \qquad \text{(A3)}\\
&\Gamma_{Z,\Delta,\Phi}\left(Z',\Delta',\Phi';Z,\Delta,\Phi\right),\ \Gamma_{\mathbf{v}}\left(\mathbf{v}';\mathbf{v},Z',\Delta',\Phi'\right),\Gamma_{\phi}\left(\phi';\phi\right) \qquad \text{(A4)}
\end{aligned}$$

The per period output for the planner is equal to $Z\left(e_i+h_i\right)$ in each market i. The first constraint $(A1)$ states that the planner has $1-\sum_{i=1}^{I}e_i$ unemployed workers available to allocate across sectors. Equation $(A2)$ states that, once the allocation $\{u_i\}$ is chosen, the frictional matching process in each market yields $\Phi\phi_i m\left(u_i,v_i\right)$ new hires which add to the existing e_i active matches. Equation $(A3)$ describes separations and the determination of next period's distribution of active matches $\{e_i'\}$ in all sectors. Line $(A4)$ in the problem collects all the exogenous stochastic processes the planner takes as given.

It is easy to see that this is a concave problem where first-order conditions are sufficient for optimality. At an interior solution ($u_i>0$ for all i), the choice of how many unemployed workers u_i to allocate in market i yields the first-order condition

$$Z\Phi\phi_i m_{u_i}\left(\frac{v_i}{u_i}\right)+\beta\mathbb{E}\left[V_{e_i}\left(\mathbf{e}';\mathbf{v}',\phi',Z',\Delta',\Phi'\right)\right]\left(1-\Delta\right)\Phi\phi_i m_{u_i}\left(\frac{v_i}{u_i}\right)=\mu, \qquad \text{(A5)}$$

where μ is the multiplier on constraint $(A1)$. The right-hand side (RHS) of this condition is the shadow value of an additional worker in the unemployment pool available to search. The left-hand side (LHS) is the expected marginal value of an additional unemployed worker allocated to sector i. The derivative of the sector-specific matching function m is written as a function of local market tightness only (with a slight abuse of notation) because of its CRS specification.

The Envelope condition with respect to the state e_i yields:

$$V_{e_i}\left(\mathbf{e};\mathbf{v},\phi,Z,\Delta,\Phi\right) = Z - \mu + \beta(1-\Delta)\mathbb{E}\left[V_{e_i}\left(\mathbf{e}';\mathbf{v}',\phi',Z',\Delta',\Phi'\right)\right], \tag{A6}$$

from which it is immediate to see, by iterating forward, that $\mathbb{E}\left[V_{e_i}\left(\mathbf{e}';\mathbf{v}',\phi',Z',\Delta',\Phi'\right)\right]$ is independent of i, since productivity and the job destruction rate are common across all sectors.[58] Using this result into $(A5)$, the optimal rule for the allocation of unemployed workers across sectors can be written as equation (1) in the main text.

A.2 Heterogenous productivities and destruction rates

We extend the baseline model of Section 2.1 as follows. Individuals (still in measure one) can be either employed in sector i (e_i), or unemployed and searching in sector i (u_i), or out of the labor force. The aggregate labor force is $\ell = \sum_{i=1}^{I}(e_i + u_i) \leq 1$.

Labor productivity in sector i is given by Zz_i, where each idiosyncratic component z_i is strictly positive, i.i.d. across sectors and independent of the aggregate state Z. The non-employed individuals produce output $\zeta Z > 0$ (which can also be interpreted as the value of additional leisure), and the unemployed incur in an extra disutility cost of search $\xi > 0$.

Let the conditional distribution of the vector $\mathbf{z} = \{z_i\}$ be $\Gamma_z(\mathbf{z}', \mathbf{z})$. The idiosyncratic component of the exogenous destruction rate in sector i is δ_i, i.i.d. across sectors and independent of Δ, Z and z_i. Let the conditional distribution of the vector $\delta = \{\delta_i\}$ be $\Gamma_\delta(\delta', \delta)$. The survival probability of a match is then $(1-\Delta)(1-\delta_i)$. The vector $\{Z, \Delta, \Phi, \mathbf{z}, \mathbf{v}, \phi, \delta\}$ takes strictly positive values.

It is convenient to impose additional structure on some conditional distributions: as specified in the text, we assume that $(Z, 1-\Delta, z_i, 1-\delta_i)$ are all positive martingales. The timing of events is exactly as before, with the decision on the size of the labor force for next period taken at the end of the current period. The recursive formulation of the planner's problem has three additional states compared to the problem of Section 2.1: the current number of unemployed workers u, the vector of productive efficiencies $\mathbf{z}$, and the vector of destruction rates δ. The

[58] We are also using the transversality condition $\lim_{t\to\infty}\beta^t(1-\Delta)^t\mathbb{E}\left[V_{e_{it}}\right] = 0$.

planner solves the problem:

$$
\begin{aligned}
V\left(u, \mathbf{e}; \mathbf{z}, \mathbf{v}, \phi, \delta, Z, \Delta, \Phi\right) &= \max_{\{u_i, \ell'\}} \sum_{i=1}^{I} Z z_i \left(e_i + h_i\right) - \xi u + Z\zeta \left[1 - \sum_{i=1}^{I} \left(e_i + h_i\right)\right] \\
&+ \beta \mathbb{E}\left[V\left(u', \mathbf{e}'; \mathbf{z}', \mathbf{v}', \phi', \delta', Z', \Delta', \Phi'\right)\right] && \text{(A7)} \\
s.t. &: \\
\sum_{i=1}^{I} u_i &\leq u && \text{(A8)} \\
h_i &= \Phi \phi_i m\left(u_i, v_i\right) && \text{(A9)} \\
e_i' &= \left(1-\Delta\right)\left(1-\delta_i\right)\left(e_i + h_i\right) && \text{(A10)} \\
u' &= \ell' - \sum_{i=1}^{I} e_i' && \text{(A11)} \\
u_i &\in \left[0, u\right], \ell' \in \left[0, 1\right], && \text{(A12)}
\end{aligned}
$$

$$
\Gamma_{Z,\Delta,\Phi}\left(Z', \Delta', \Phi'; Z, \Delta, \Phi\right),\ \Gamma_{\mathbf{v}}\left(\mathbf{v}'; \mathbf{v}, Z', \Delta', \Phi', \mathbf{z}'\right), \Gamma_{\phi}\left(\phi'; \phi\right), \Gamma_{\mathbf{z}}\left(\mathbf{z}'; \mathbf{z}\right), \Gamma_{\delta}\left(\delta', \delta\right) \quad \text{(A13)}
$$

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

$$
Z\left(z_i - \zeta\right) \Phi \phi_i m_{u_i}\left(\frac{v_i}{u_i}\right) + \beta \mathbb{E}\left[-V_u'\left(\cdot\right) + V_{e_i}'\left(\cdot\right)\right]\left(1-\Delta\right)\left(1-\delta_i\right) \Phi \phi_i m_{u_i}\left(\frac{v_i}{u_i}\right) = \mu, \quad \text{(A14)}
$$

where μ is the multiplier on constraint $(A8)$. The Envelope conditions with respect to the states u and e_i yield:

$$
\begin{aligned}
V_u\left(u, \mathbf{e}; \mathbf{z}, \mathbf{v}, \phi, \delta, Z, \Delta, \Phi\right) &= \mu - \xi && \text{(A15)} \\
V_{e_i}\left(u, \mathbf{e}; \mathbf{z}, \mathbf{v}, \phi, \delta, Z, \Delta, \Phi\right) &= Z\left(z_i - \zeta\right) + \beta(1-\Delta)\left(1-\delta_i\right) \mathbb{E}\left[V_{e_i}' - V_u'\right]. && \text{(A16)}
\end{aligned}
$$

According to the first Envelope 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 this period, net of the foregone output from non-employment, plus its discounted continuation value net of the value of search, conditional on the match not being destroyed.

The optimal decision on the labor force size next period ℓ' requires

$$
\mathbb{E}\left[V_u\left(u', \mathbf{e}'; \mathbf{z}', \mathbf{v}', \phi', \delta', Z', \Delta', \Phi'\right)\right] = 0. \quad \text{(A17)}
$$

By combining $(A17)$ with $(A15)$, we note that the planner will choose the size of the labor force so that the expected shadow value of an unemployed worker $\mathbb{E}\left[\mu'\right]$ equals search disutility ξ (note that ζ does not feature in this equality because both unemployed job-seekers and non-participants produce ζZ).[59]

Using $(A17)$ into the Envelope condition $(A16)$, and exploiting the additional assumption that all the elements of the vector $x = (Z, 1-\Delta, z_i, 1-\delta_i)$ are independent martingales, iterating forward we arrive at:

$$\mathbb{E}\left[V'_{e_i}\right] = \frac{Z\left(z_i - \zeta\right)}{1 - \beta\left(1-\Delta\right)\left(1-\delta_i\right)} \tag{A18}$$

which, substituted into equation $(A14)$ yields

$$Z\left(z_i - \zeta\right)\Phi\phi_i m_{u_i}\left(\frac{v_i}{u_i}\right) + \frac{\beta\left(1-\Delta\right)\left(1-\delta_i\right)}{1-\beta\left(1-\Delta\right)\left(1-\delta_i\right)} Z\left(z_i - \zeta\right)\Phi\phi_i m_{u_i}\left(\frac{v_i}{u_i}\right) = \mu. \tag{A19}$$

Rearranging, we conclude that the planner allocates idle labor to equalize

$$\frac{z_i - \zeta}{1-\beta\left(1-\Delta\right)\left(1-\delta_i\right)}\phi_i m_{u_i}\left(\frac{v_i}{u_i^*}\right) \tag{A20}$$

across sectors, which is expression (2) in Section (2.2) in the main text. Finally, we note that to guarantee and interior solution, i.e., a positive measure of unemployed workers in each sector, we must impose that the lower bound of the distribution of z_i exceeds ζ.

A.3 Endogenous separations

We now allow the planner to move workers employed in sector i into unemployment (or out of the labor force) at the end of the period, before choosing the size of the labor force for next period. There are two changes to the planner's problem of Section $A.2$. First, the law of motion for employment becomes

$$e'_i = \left(1-\Delta\right)\left(1-\delta_i\right)\left(e_i + h_i\right) - \sigma_i. \tag{A21}$$

Second, the planner has another vector of choice variables $\{\sigma_i\}$, with $\sigma_i \in \left[0, \left(1-\Delta\right)\left(1-\delta_i\right)\left(e_i+h_i\right)\right].$

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

$$\mathbb{E}\left[V'_u\left(\cdot\right) - V'_{e_i}\left(\cdot\right)\right]\begin{cases} < 0 & \to \sigma_i = 0 \\ = 0 & \to \sigma_i \in \left(0, \left(1-\Delta\right)\left(1-\delta_i\right)\left(e_i+h_i\right)\right) \\ > 0 & \to \sigma_i = \left(1-\Delta\right)\left(1-\delta_i\right)\left(e_i+h_i\right) \end{cases} \tag{A22}$$

[59]It is clear that our result is robust to allowing ξ to be stochastic and correlated with (Z, Δ, Φ).

depending on whether at the optimum a corner or interior solution arises. If the first-order conditions $(A17)$ hold with equality, then the optimality condition $(A22)$ holds with the " $<$ " inequality and $\sigma_i = 0$. As a result, the planner's allocation rule (2) remains unchanged.

A.4 Heterogeneous sensitivities to the aggregate shock

Let productivity in sector i be Z^{η_i} and let $\log Z$ follow a unit root process with innovation ϵ independent of Δ and distributed as $N(-\sigma_\epsilon/2, \sigma_\epsilon)$. Note that $\mathbb{E}\left[(Z')^{\eta_i}\right] = Z^{\eta_i} \exp\left(\eta_i (\eta_i - 1) \frac{\sigma_\epsilon}{2}\right)$. denote $\Omega_i \equiv \exp\left(\eta_i (\eta_i - 1) \frac{\sigma_\epsilon}{2}\right)$. We maintain that $(1-\Delta, 1-\delta_i)$ follow unit root processes. Using $(A17)$ into the Envelope condition $(A16)$ yields

$$V_{e_i} = Z^{\eta_i} - \zeta Z + \beta(1-\Delta)(1-\delta_i)\mathbb{E}\left[V'_{e_i}\right]. \tag{A23}$$

Solving (A23) forward by using the unit root assumption, we obtain

$$\mathbb{E}\left[V'_{e_i}\right] = \frac{Z^{\eta_i}\Omega_i}{1-\beta(1-\Delta)(1-\delta_i)\Omega_i} - \frac{\zeta Z}{1-\beta(1-\Delta)(1-\delta_i)}.$$

Substituting this expression for $\mathbb{E}\left[V'_{e_i}\right]$ into equation $(A14)$ and rearranging, we conclude that the planner allocates unemployed workers so to equalize

$$\left[\frac{Z^{\eta_i}}{1-\beta(1-\Delta)(1-\delta_i)\Omega_i} - \frac{\zeta Z}{1-\beta(1-\Delta)(1-\delta_i)}\right]\phi_i m_{u_i}\left(\frac{v_i}{u_i^*}\right),$$

across sectors, which is expression (3) in Section (2.3) in the main text. Since η_i could be larger than one, a necessary additional technical condition we must impose is $\beta(1-\Delta)(1-\delta_i)\Omega_i < 1$ for all i.

A.5 Properties of the mismatch index

First, we prove that $0 \leq \mathcal{M}_{\phi t} \leq 1$. Since all the components of the sum in (8) are positive, $\mathcal{M}_{\phi t} \leq 1$. Under maximal mismatch (no markets where unemployment and vacancies coexist), the index is exactly equal to one. To show that $\mathcal{M}_{\phi t} \geq 0$, note that

$$\begin{aligned}
1 - \mathcal{M}_{\phi t} &= \frac{1}{v_t^\alpha u_t^{1-\alpha}} \frac{1}{\left[\sum_{i=1}^{I} \phi_{it}^{\frac{1}{\alpha}}\left(\frac{v_{it}}{v_t}\right)\right]^\alpha} \sum_{i=1}^{I}\left(\phi_{it}^{\frac{1}{\alpha}} v_{it}\right)^\alpha (u_{it})^{1-\alpha} \\
&\leq \frac{1}{v_t^\alpha u_t^{1-\alpha}} \frac{1}{\left[\sum_{i=1}^{I} \phi_{it}^{\frac{1}{\alpha}}\left(\frac{v_{it}}{v_t}\right)\right]^\alpha} \left[\sum_{i=1}^{I}\left(\phi_{it}^{\frac{1}{\alpha}} v_{it}\right)\right]^\alpha \left(\sum_{i=1}^{I} u_{it}\right)^{1-\alpha} \\
&= 1
\end{aligned}$$

where the $\leq$ sign follows from Hölder's inequality. It is easy to show that the index becomes exactly zero in absence of mismatch by substituting the allocation rule (7) into the index.

By inspecting (8), it is also easy to see that the $\mathcal{M}_{\phi t}$ index is invariant to "pure" aggregate shocks that shift the total number of vacancies and unemployed up or down, but leave the vacancy and unemployment shares across markets unchanged.

To show that the mismatch index is increasing in the level of disaggregation, consider an economy where the aggregate labor market is described by two dimensions indexed by (i, j), e.g., I regions $\times$ J occupations. Let $\mathcal{M}_{\phi It}$ be the mismatch index over the I sectors and $\mathcal{M}_{\phi IJt}$ be the one over the $I \times J$ sectors. From the disaggregated matching function, we have $h_{ijt} = \Phi_t \phi_{ijt} v_{ijt}^{\alpha} u_{ijt}^{1-\alpha}$. Summing this expression over j yields

$$h_{it} = \sum_{j=1}^{J} \Phi_t \phi_{ijt} v_{ijt}^{\alpha} u_{ijt}^{1-\alpha} = \Phi_t \left[\sum_{j=1}^{J} \phi_{ijt} \left(\frac{v_{ijt}}{v_{it}} \right)^{\alpha} \left(\frac{u_{ijt}}{u_{it}} \right)^{1-\alpha} \right] v_{it}^{\alpha} u_{it}^{1-\alpha}. \tag{A24}$$

At the aggregated level, we have $h_{it} = \Phi_t \phi_{it} v_{it}^{\alpha} u_{it}^{1-\alpha}$ and therefore (A24) implies that

$$\phi_{it} = \sum_{j=1}^{J} \phi_{ijt} \left(\frac{v_{ijt}}{v_{it}} \right)^{\alpha} \left(\frac{u_{ijt}}{u_{it}} \right)^{1-\alpha}. \tag{A25}$$

Now consider the disaggregated matching index. We have

$$1 - \mathcal{M}_{\phi IJt} = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{\phi_{ijt}}{\bar{\phi}_{IJt}} \left(\frac{v_{ijt}}{v_t} \right)^{\alpha} \left(\frac{u_{ijt}}{u_t} \right)^{1-\alpha} \tag{A26}$$

for

$$\bar{\phi}_{IJt} = \left[\sum_{i=1}^{I} \sum_{j=1}^{J} \phi_{ijt}^{\frac{1}{\alpha}} \left(\frac{v_{ijt}}{v_t} \right) \right]^{\alpha}. \tag{A27}$$

Manipulating the above expression yields

$$\begin{aligned}
1 - \mathcal{M}_{\phi IJt} &= \frac{1}{\bar{\phi}_{IJt} v_t^{\alpha} u_t^{1-\alpha}} \sum_{i=1}^{I} \sum_{j=1}^{J} \phi_{ijt} v_{ijt}^{\alpha} u_{ijt}^{1-\alpha} \\
&= \frac{1}{\bar{\phi}_{IJt} v_t^{\alpha} u_t^{1-\alpha}} \sum_{i=1}^{I} v_{it}^{\alpha} u_{it}^{1-\alpha} \sum_{j=1}^{J} \phi_{ijt} \left(\frac{v_{ijt}}{v_{it}} \right)^{\alpha} \left(\frac{u_{ijt}}{u_{it}} \right)^{1-\alpha} \\
&= \frac{1}{\bar{\phi}_{IJt}} \sum_{i=1}^{I} \phi_{it} \left(\frac{v_{it}}{v_t} \right)^{\alpha} \left(\frac{u_{it}}{u_t} \right)^{1-\alpha}
\end{aligned}$$

where the third step above follows from (A25). Next, manipulating (A27) delivers

$$\begin{aligned}\bar{\phi}_{IJt} &= \left\{\frac{1}{v_t}\sum_{i=1}^{I} v_{it}\left(\left[\sum_{j=1}^{J}\phi_{ijt}^{\frac{1}{\alpha}}\left(\frac{v_{ijt}}{v_{it}}\right)\right]^{\alpha}\right)^{\frac{1}{\alpha}}\right\}^{\alpha} \\ &= \left\{\frac{1}{v_t}\sum_{i=1}^{I} v_{it}\left(\left[\sum_{j=1}^{J}\phi_{ijt}^{\frac{1}{\alpha}}\left(\frac{v_{ijt}}{v_{it}}\right)\right]^{\alpha}\cdot\left[\sum_{j=1}^{J}\frac{u_{ijt}}{u_{it}}\right]^{1-\alpha}\right)^{\frac{1}{\alpha}}\right\}^{\alpha}\end{aligned}$$

where the second step above follows from the identity $\sum_{j=1}^{J} u_{ijt} = u_{it}$. Applying Holder's inequality yields

$$\begin{aligned}\bar{\phi}_{IJt} &\geq \left\{\frac{1}{v_t}\sum_{i=1}^{I} v_{it}\left(\sum_{j=1}^{J}\phi_{ijt}\left(\frac{v_{ijt}}{v_{it}}\right)^{\alpha}\left(\frac{u_{ijt}}{u_{it}}\right)^{1-\alpha}\right)^{\frac{1}{\alpha}}\right\}^{\alpha} \\ &= \left\{\sum_{i=1}^{I}\phi_{it}^{\frac{1}{\alpha}}\left(\frac{v_{it}}{v_t}\right)\right\}^{\alpha} = \bar{\phi}_{It}\end{aligned}$$

where $\bar{\phi}_{It}$ is an expression equivalent to $\bar{\phi}_{IJt}$ in (A27) for the case where the $(I \times J)$ sectors are collapsed into I sectors. Combining results, we have shown that

$$1 - \mathcal{M}_{\phi IJt} \leq \sum_{i=1}^{I}\frac{\phi_{it}}{\bar{\phi}_{It}}\left(\frac{v_{it}}{v_t}\right)^{\alpha}\left(\frac{u_{it}}{u_t}\right)^{1-\alpha} = 1 - \mathcal{M}_{\phi It}$$

and so we must have $\mathcal{M}_{\phi IJt} \geq \mathcal{M}_{\phi It}$.

A.6 Planner's problem with endogenous vacancies

Optimal vacancy creation Consider the planner's problem of Section 2.2 solved in Appendix A.2, the most general of our environments. To simplify the notation, without loss of generality, let z_i denote output in sector i net of the flow output from nonemployment ζ. If we let the creation of vacancies $\{v_i\}$ be under the control of the planner, we have:

$$
\begin{aligned}
V\left(u, \mathbf{e}; \mathbf{z}, \phi, \delta, \kappa, Z, \Delta, \Phi\right) &= \max_{\{u_i, v_i, \ell'\}} \sum_{i=1}^{I} Z z_i \left(e_i + h_i\right) - K_i\left(v_i\right) - \xi u \\
&+ \beta \mathbb{E}\left[V\left(u', \mathbf{e}'; \mathbf{z}', \phi', \delta', \kappa', Z', \Delta', \Phi'\right)\right] &\text{(A28)}\\
s.t. &: \\
\sum_{i=1}^{I} u_i &\leq u &\text{(A29)}\\
h_i &= \Phi \phi_i m\left(u_i, v_i\right) &\text{(A30)}\\
e_i' &= \left(1 - \Delta\right)\left(1 - \delta_i\right)\left(e_i + h_i\right) &\text{(A31)}\\
u' &= \ell' - \sum_{i=1}^{I} e_i' &\text{(A32)}\\
u_i &\in \left[0, u\right], \ell' \in \left[0, 1\right], v_i \geq 0 &\text{(A33)}\\
\Gamma_{Z,\Delta,\Phi}\left(Z', \Delta', \Phi'; Z, \Delta, \Phi\right),\ &\Gamma_{\phi}\left(\phi'; \phi\right), \Gamma_{\mathbf{z}}\left(\mathbf{z}'; \mathbf{z}\right), \Gamma_{\delta}\left(\delta', \delta\right), \Gamma_{\kappa}\left(\kappa', \kappa\right) &\text{(A34)}
\end{aligned}
$$

The optimality condition for vacancy creation is

$$
K_{v_i}\left(v_i^*\right) = \Phi \phi_i m_{v_i}\left(\frac{v_i^*}{u_i^*}\right)\left\{Z z_i + \beta\left(1 - \Delta\right)\left(1 - \delta_i\right)\mathbb{E}\left[V_{e_i}'\left(\cdot\right)\right]\right\}.
$$

Using the expression for $\mathbb{E}\left[V_{e_i}'\left(\cdot\right)\right]$ obtained in $(A18)$ and the functional forms for K_i and m specified in the main text, we obtain expression (16).

Calculation of planner's vacancies We now lay out an algorithm to compute the planner's optimal allocation of vacancies across sectors. Rearranging condition $(A20)$ dictating the optimal allocation of unemployed workers across sectors, given the distribution of vacancies $\{v_i^*\}$, yields

$$
\frac{v_i^*}{u_i^*} = \left[\frac{\mu}{1-\alpha}\frac{1}{\frac{Z z_i \Phi \phi_i}{1 - \beta\left(1-\Delta\right)\left(1-\delta_i\right)}}\right]^{\frac{1}{\alpha}} \qquad \text{(A35)}
$$

where μ is the multiplier on the resource constraint $\sum_{i=1}^{I} u_i \leq u$. Substituting $(A35)$ into (16)

yields an equation for the optimal number of vacancies in sector i which reads

$$v_i^* = \frac{1}{\kappa_i}\left(\frac{\alpha}{1-\alpha}\right)^{1/\varepsilon}\left(\frac{1}{\mu}\right)^{\frac{(1-\alpha)/\varepsilon}{\alpha}}\left[(1-\alpha)\frac{Zz_i\Phi\phi_i}{1-\beta(1-\Delta)(1-\delta_i)}\right]^{\frac{1/\varepsilon}{\alpha}}. \tag{A36}$$

Summing over all $i's$, we arrive at the optimal share of vacancies in sector i

$$\frac{v_i^*}{v_t^*} = \frac{\frac{1}{\kappa_i}\left[\frac{z_i\phi_i}{1-\beta(1-\Delta)(1-\delta_i)}\right]^{\frac{1/\varepsilon}{\alpha}}}{\sum_{i=1}^{I}\frac{1}{\kappa_i}\left[\frac{z_i\phi_i}{1-\beta(1-\Delta)(1-\delta_i)}\right]^{\frac{1/\varepsilon}{\alpha}}} \tag{A37}$$

only as a function of parameters, which is quite intuitive: the higher is productive, matching and job creation efficiency in sector i, relative to the other sectors, the larger its share of vacancies. However, to solve the model, we need to determine the *level* of v_i^* which requires eliminating μ from $(A36)$. Combining again the two first order conditions, and summing across all sectors, we arrive at

$$u^* = \left(\frac{\alpha}{1-\alpha}\right)^{1/\varepsilon}\left[Z\Phi(1-\alpha)\right]^{\frac{1+1/\varepsilon}{\alpha}}\left(\frac{1}{\mu}\right)^{\frac{1+(1-\alpha)/\varepsilon}{\alpha}}\cdot\sum_{i=1}^{I}\frac{1}{\kappa_i}\left[\frac{z_i\phi_i}{1-\beta(1-\Delta)(1-\delta_i)}\right]^{\frac{1+1/\varepsilon}{\alpha}} \tag{A38}$$

which establishes a unique inverse relationship between μ and u^*: the higher the number of idle workers, the lower the shadow value of the constraint.

Equation $(A38)$ suggests an algorithm to solve for v_i^*. At any date, before choosing how to allocate vacancies and unemployed workers, the total number of idle workers is a state variable for the planner, i.e., u^* is known. One can therefore back out μ from $(A38)$, and then v_i^* from $(A36)$ and u_i^* from $(A35)$.

Counterfactual unemployment To perform the counterfactual on unemployment with endogenous vacancies, we use the same iterative procedure described in Section 3.2, with the caveat that the relationship between the planner's job-finding rate and the empirical job-finding rate at date t is now given by

$$f_t^* = \frac{h_t^*}{u_t^*} = \Phi_t\bar{\phi}_{xt}^*\left(\frac{v_t^*}{u_t^*}\right)^{\alpha} = f_t\cdot\frac{1}{1-\mathcal{M}_{xt}}\cdot\left(\frac{u_t}{u_t^*}\right)^{\alpha}\cdot\left[\left(\frac{\bar{\phi}_{xt}^*}{\bar{\phi}_{xt}}\right)\cdot\left(\frac{v_t^*}{v_t}\right)^{\alpha}\right], \tag{A39}$$

where $\bar{\phi}_{xt}$ is given by equation (11), and $\bar{\phi}_{xt}^*$ is the same aggregator with shares (v_{it}^*/v_t^*) instead of (v_{it}/v_t). When $v_{it}^* = v_{it}$ (i.e., $\varepsilon\to\infty$), equation $(A39)$ collapses to the relationship $f^* = [f/(1-\mathcal{M}_{xt})](u_t/u_t^*)^{\alpha}$ that we have used in our baseline counterfactual with exogenous vacancies.

A.7 Model with heterogeneous α

We now extend the model of Section 2.2 and introduce sector-specific matching functions m_i. We retain the constant-return Cobb-Douglas specification, but we allow the vacancy share α (and hence the unemployment share $1-\alpha$) to vary across sectors, i.e., hires in sector i at date t are now given by the matching function

$$h_{it} = \Phi_t \phi_{it} v_{it}^{\alpha_i} u_{it}^{1-\alpha_i}. \tag{A40}$$

By replicating all the steps outlined in Section A.2, we arrive at the set of I first-order conditions (one for each sector i):

$$(1-\alpha_i) \frac{(z_{it}-\zeta)}{1-\beta(1-\Delta_t)(1-\delta_{it})} \left(\frac{v_{it}}{u_{it}^*}\right)^{\alpha_i} = \mu_t \tag{A41}$$

which, together with the adding-up constraint $\sum_{i=1}^{I} u_{it} = u_t$, yields a system of $(I+1)$ equations in $(I+1)$ unknowns $\left\{\{u_{it}^*\}_{i=1}^{I}, \mu_t\right\}$ at every date t, which can be solved numerically.

Since optimal hires are $h_{it}^* = \Phi_t \phi_{it} v_{it}^{\alpha_i} (u_{it}^*)^{1-\alpha_i}$, the mismatch index at t is

$$\mathcal{M}_{xt} = 1 - \frac{h_t}{h_t^*} = 1 - \frac{\sum_{i=1}^{I} \phi_{it} v_{it}^{\alpha_i} u_{it}^{1-\alpha_i}}{\sum_{i=1}^{I} \phi_{it} v_{it}^{\alpha_i} (u_i^*)^{1-\alpha_i}}.$$

Even if this mismatch index has no longer a closed form, it is easy to compute once we have the vector of planner's allocations of unemployed workers across sectors $\{u_{it}^*\}$. The counterfactual unemployment rate is still obtained as described in Section 3.2 of the paper.

B Data Appendix

B.1 Comparison between JOLTS and HWOL vacancies

Vacancies recorded in JOLTS are derived from a sample of about 16,000 business establishments. JOLTS vacancies represent "all unfilled, posted positions available at an establishment on the last day of the month. The vacancy must be for a specific position where work can start within thirty days, and an active recruiting process must be underway for the position." (Faberman, 2009, p. 86). As noted in Section 4, the HWOL database collects ads from job listings posted by employers on thousands of internet job boards and online newspapers. The HWOL program uses a mid-month survey reference period. For example, data for October would be the sum of all posted ads from September 14th through October 13th. This reference period is aligned to the BLS unemployment "job search" time period. The monthly vacancy counts that we use in our calculations are total monthly unduplicated ads appearing in the reference period. This figure therefore includes both newly posted ads and ads reposted from the previous months.

Sampled establishments in the JOLTS only report their own direct employees and exclude "employees of temporary help agencies, employee leasing companies, outside contractors, and consultants," which are counted by their employer of record, not by the establishment where they are working.[60] Thus, this approach captures temp-help and leasing workers as long as their employers are sampled in the JOLTS, but does not capture the self-employed contract workforce (these workers typically receive a 1099-MISC form instead of a W-2 form to report payments received for services they provide). On the other hand, the HWOL series includes postings for contract work. In what follows, we often also report HWOL vacancy counts excluding contract work, to make the series more comparable to the JOLTS measure of vacancies, but in our empirical analyses of mismatch with HWOL data we consider all ads, including those for contract work.

We perform two exercises to compare the vacancy counts we get from each data source, one at the regional level and one at the industry level—region and industry are the only dimensions available in both JOLTS and HWOL. First, we compare total vacancies by Census region in Figure C1. The HWOL series tend to be lower than the JOLTS series before 2008 (especially in the South), and higher from 2008 onwards (especially in the Northeast). The two series are closest in the West: here the correlation between the HWOL and JOLTS series is 0.94. In the other three regions the correlation is lower: 0.27 in the Midwest, 0.40 in the South, and 0.54 in the Northeast. Our re-weighing strategy in Section 7 enables us to correct for the possibility that online ads penetration may differ across regions.

[60] See the JOLTS Technical Note at http://www.bls.gov/news.release/jolts.tn.htm.

For about 57% of the job listings, we observe the NAICS code of the employer. Therefore, we are able to directly compare vacancy counts by industry from HWOL to those in the JOLTS. We report in Figure C2 scatterplots of vacancy shares by industry from JOLTS and from HWOL—for the latter, we report both total vacancies, as well as vacancies without contract work. The top panel of the figure reports average vacancy shares over the sample period under consideration. Most data points are close to the 45-degree line, indicating that the vacancy shares by industry in the two series line up fairly well, especially when we omit contract work from HWOL to make it more comparable to the JOLTS. The only two sectors where JOLTS and HWOL show significant differences in vacancy shares are "Public Administration" and "Accommodation and Food Services." The bottom panel reports the change in average vacancy shares between 2006 and the 12 month period around December 2009 for each series. Again, the JOLTS and HWOL series are quite close to each other, with the exception of "Public Administration."

We have investigated whether the missing industry information in HWOL exhibits any systematic patterns that may have skewed our analysis. For robustness, we re-weighted the industry observations in HWOL as follows: first, we dropped observations from individual Job Boards with the highest rates of missing NAICS codes. Then, we re-weighted the remaining observations to correct for any correlation between NAICS missing values and Job Board, occupation or Census region. In other words, if vacancies for specific (Job Board, SOC, Census region) combinations are more likely to have missing NAICS codes, the vacancies that do have NAICS information in those cells are assigned a larger weight in computing total vacancies by industry.[61] The resulting vacancy shares are almost identical to those based on the raw data.

To sum up, the comparison between JOLTS and HWOL vacancy counts suggests that there are some discrepancies in the behavior of two series. The main concerns are (i) the possible over- or under-use of online advertisement in certain sectors (regions and/or industries) and (ii) the presence of an upward trend in the use of online recruitment that could artificially mitigate the drop in job advertisements around the last recession (and inflate the subsequent recovery). We address these issues in Section 7 and show that our quantitative results on mismatch measures are robust.

[61] For example, suppose a (Job Board, SOC, Census region) cell has four observations. Observation one is in NAICS code 11, observations two and three are in NAICS code 13, and observation four has a missing NAICS. Thus, the missing NAICS rate is 0.25. Then, a weight of $1/(1-0.25) = 1.333$ is applied to each observation with non-missing NAICS. So we find 1.333 job vacancies in NAICS code 11, and 2.667 job vacancies in NAICS code 13.

B.2 Matching function estimation

Throughout our analysis we assume matching functions are constant returns to scale. We begin by imposing a Cobb-Douglas specification. At the end of this section we show that, when we allow for a more general CES specification, our results point towards an elasticity of substitution statistically close to one.

To compute market-specific matching efficiency parameters, ϕ_i, and vacancy share α, we use various data sources. At the industry level, we use vacancies and hires from JOLTS, and unemployment counts from the CPS. At the occupation level, we use vacancies from HWOL but do not have a direct measure of hires as in JOLTS. Therefore, we construct hires from the CPS using flows from unemployment into a given occupation i for people who are surveyed in adjacent months. Because these monthly flows are quite noisy, we use a 3-month moving average of the data, and aggregate occupations into five broad occupation groups. For comparison purposes, we replicate the analysis at the industry level using the constructed "CPS hires" as well. At the aggregate level, we perform the estimation using both JOLTS and HWOL vacancies, and both JOLTS and CPS hires.

The estimation of matching functions is subject to an endogeneity problem, as shocks to unobserved matching efficiency may affect the number of vacancies posted by firms—much like TFP shocks affect firm's choice of labor input. To deal with this issue, we follow two strategies suggested by Borowczyk-Martins, Jolivet, and Postel-Vinay (2012). First, they recognize that some of the major movements in matching efficiency inducing a bias in the OLS estimator are low-frequency ones. As a result, modeling explicitly the dynamics of matching efficiency through time-varying polynomials and structural breaks goes a long way towards solving the problem even with the simple OLS estimator. This is the first route we take. At the aggregate level, we estimate:

$$\log\left(\frac{h_t}{u_t}\right) = const + \gamma' QTT_t + \alpha \log\left(\frac{v_t}{u_t}\right) + \epsilon_t, \tag{B1}$$

where QTT_t is a vector of four elements for the quartic time trend which is meant to capture shifts in aggregate matching efficiency (i.e., Φ_t in the model).

At the sectoral level, we are interested in the sector-specific component of matching efficiency orthogonal to common aggregate movements in aggregate matching efficiency. Therefore, at the industry and 2-digit occupation level, we perform the following panel regression:

$$\log\left(\frac{h_{it}}{u_{it}}\right) = \gamma' QTT_t + \chi_{\{t\le 07\}} \log\left(\phi_i^{pre}\right) + \chi_{\{t>07\}} \log\left(\phi_i^{post}\right) + \alpha \log\left(\frac{v_{it}}{u_{it}}\right) + \epsilon_{it}, \tag{B2}$$

where $\chi_{\{t>07\}}$ is an indicator for months after December 2007, the official start of the recession, to absorb sector-specific shifts in matching efficiency.

Borowczyk-Martins, Jolivet and Postel-Vinay (2012) also propose a GMM estimator to take care of the simultaneity bias. This method requires imposing an ARMA(p,q) structure on the matching efficiency process: we follow their model selection protocol and set $p = 3$ and $q = 3$. We use an over-identified GMM estimator implemented with four lags of market tightness and one lag of the job-finding rate as instruments, as they argue it is the one delivering the most precise parameter estimates.

Table C4 displays the full set of estimates of the vacancy share parameter α. In the aggregate regressions, the estimated vacancy share varies between 0.32 and 0.67; in the panel regressions, the estimates are somewhat lower varying between 0.24 and 0.53. To construct our mismatch indices, and in our calculation of mismatch unemployment, we pick a value of $\alpha = 0.5$ for two reasons. First, it is the midpoint of our estimates with aggregate data. Second, our mismatch indices are typically highest for $\alpha = 0.5$; therefore, in the spirit of reporting an upper bound for mismatch unemployment, we use this value.

The estimated quartic time trend (not shown) drops during the recession in all our OLS specifications, consistent with a deterioration of aggregate matching efficiency. With regard to sectoral matching efficiency, in our baseline calculations we use the estimates obtained with JOLTS hires for the industry level mismatch analysis, and those with CPS hires for the occupation level analysis. In all cases, we use the *pre-recession* matching efficiency parameter estimates, and verify the robustness of our findings to this choice. The estimated matching efficiency parameters ϕ_i pre- and post-recession are reported in Tables C6-C8. Beyond movements in the common component Φ_t, the quartic in time, changes over time in sector-specific matching efficiencies are small.

Finally, in order to examine the plausibility of the Cobb-Douglas specification, we generalize (B2) and estimate the following CES specification via minimum distance:

$$\log\left(\frac{h_{it}}{u_{it}}\right) = \gamma' QTT_t + \chi_{\{t\leq 07\}}\log\left(\phi_i^{pre}\right) + \chi_{\{t>07\}}\log\left(\phi_i^{post}\right) + \frac{1}{\sigma}\log\left[\alpha\left(\frac{v_{it}}{u_{it}}\right)^{\sigma} + (1-\alpha)\right] + \epsilon_{it}. \tag{B3}$$

Recall that $\sigma \in (-\infty, 1)$ with $\sigma = 0$ being the Cobb-Douglas case. A simulated annealing algorithm is used to ensure that we attain a global minimum. 95% confidence intervals are computed via bootstrap methods. The estimation results are reported in Table C5. The point estimates of σ range from -0.11 to 0.18 depending on the specification, implying an elasticity between 0.9 and 1.2. In the specification with HWOL vacancies and CPS hires, we cannot reject the null that $\sigma = 0$ at the 5% significance level. In the other specifications with JOLTS data, $\sigma = 0$ lies just outside the 95% confidence interval, but the point estimates are close to zero, implying values close to unity for the elasticity of the matching function.

B.3 Adjustment in sectoral unemployment count

Let u_{it} be the number of unemployed workers at date t whose last job is in sector i, and U_{it} be the true number of unemployed actually searching in sector i at date t. Also let u_{it}^{j} be the number of unemployed whose last job is in sector i and who are searching in sector j. By definition, we have $u_{it} = \sum_{j=1}^{I} u_{it}^{j}$. The key unknown at each date t is the vector $\{U_{it}\}$.

From the panel dimension of CPS we observe h_{it}^{j}, the number of unemployed workers hired in sector j in period t whose last job was in sector i. Let the total number of hires in sector j in period t be h_t^j. Assume that the job-finding rate in sector j is the same for all unemployed, independent of the sector of provenance, with the sole exception if their previous job was in that same sector, in which case their job-finding rate is higher by a factor $\gamma_t \geq 1$, or:

$$\frac{h_{jt}^{j}}{u_{jt}^{j}} = (1+\gamma_t)\,\frac{h_{it}^{j}}{u_{it}^{j}}, \text{ for } i \neq j. \tag{B4}$$

The average hiring rate of sector j is the total number of hires for j divided by the total number of unemployed looking in sector j or:

$$\frac{h_t^j}{U_{jt}} = \sum_{i\neq j}\left(\frac{h_{it}^{j}}{u_{it}^{j}}\right)\left(\frac{u_{it}^{j}}{U_{jt}}\right) + \left(\frac{h_{jt}^{j}}{u_{jt}^{j}}\right)\left(\frac{u_{jt}^{j}}{U_{jt}}\right).$$

Substituting $(B4)$ into the above equation delivers:

$$\frac{h_t^j}{U_{jt}} = \sum_{i\neq j}\left(\frac{h_{it}^{j}}{u_{it}^{j}}\right)\left(\frac{u_{it}^{j}}{U_{jt}}\right) + (1+\gamma_t)\,\frac{h_{it}^{j}}{u_{it}^{j}}\left(\frac{u_{jt}^{j}}{U_{jt}}\right).$$

Because the ratio h_{it}^{j}/u_{it}^{j} is the same across all $i \neq j$, we can pull it out of the sum above and obtain, after rearranging:

$$\frac{h_{it}^{j}}{u_{it}^{j}} = \begin{cases} \left(\frac{h_t^j}{U_{jt}}\right)\left[1+\gamma_t\left(\frac{u_{jt}^{j}}{U_{jt}}\right)\right]^{-1} & \text{if } i \neq j \\ (1+\gamma_t)\left(\frac{h_t^j}{U_{jt}}\right)\left[1+\gamma_t\left(\frac{u_{jt}^{j}}{U_{jt}}\right)\right]^{-1} & \text{if } i = j \end{cases} \tag{B5}$$

Since we do not observe u_{jt}^{j}/U_{jt}, we want to substitute it out. Note that

$$\frac{u_{jt}^{j}}{U_{jt}} = \frac{\frac{h_{jt}^{j}}{h_t^j}\left(\frac{1}{1+\gamma_t}\right)}{1-\frac{h_{jt}^{j}}{h_t^j}\left(\frac{\gamma_t}{1+\gamma_t}\right)}$$

and using this expression in $(B5)$, we arrive at a relationship between the hiring rate from i to j and the average hiring rate in j:

$$\frac{h_{it}^j}{u_{it}^j} = \xi_{it}^j \cdot \frac{h_t^j}{U_{jt}} \tag{B6}$$

where

$$\xi_{it}^j = \begin{cases} 1 - \frac{h_{jt}^j}{h_t^j}\left(\frac{\gamma_t}{1+\gamma_t}\right) & \text{if } i \neq j \\ (1+\gamma_t)\left[1 - \frac{h_{jt}^j}{h_t^j}\left(\frac{\gamma_t}{1+\gamma_t}\right)\right] & \text{if } i = j \end{cases}$$

Rearranging equation $(B6)$ and summing across all j yields, at every t, the I equations:

$$u_{it} = \sum_{j=1}^{N} \frac{1}{\xi_{it}^j}\left(\frac{h_{it}^j}{h_t^j}\right) U_t^j$$

in the $(I+1)$ unknowns $\{U_{jt}\}, \gamma_t$. The last equation needed is the "aggregate consistency" condition

$$\sum_{j=1}^{I} U_{jt} = \sum_{j=1}^{I} u_{jt} \tag{B7}$$

stating that the true distribution of unemployed across sectors must sum to the observed total number of unemployed. We therefore have a system of $(I+1)$ equations in $(I+1)$ unknowns.

In our calculation of unemployment counts, to guarantee a non-negative solution to the linear system, we set to zero all entries in the transition matrices h_{it}^j which account for less than 5% of hires h_t^j in any given sector at any date t. We find that the estimated values of γ_t are all close to one.

B.4 Reweighting of HWOL vacancies

Let v_{irt}^H be the vacancies in the HWOL data for industry $i = 1, ..., I$ and region $r = 1, ..., R$ in month t. Let v_{irt}^J be the corresponding count for JOLTS vacancies. The objective is to reweigh monthly vacancies in HWOL to match those in JOLTS by industry and region (the only two common variables across data sets). We therefore solve, at every t, the following set of $(I \times R)$ equations

$$\begin{aligned} \sum_{i=1}^{I} v_{irt}^H \cdot \omega_{it} \cdot \omega_{rt} &= v_{rt}^J \\ \sum_{r=1}^{R} v_{irt}^H \cdot \omega_{it} \cdot \omega_{rt} &= v_{it}^J \end{aligned}$$

for the $(I \times R)$ vector of weights $\{\omega_{it}, \omega_{rt}\}$. Our solution algorithm imposes that weights must be positive, but this constraint is never binding in practice. Table C13 reports the average estimates of these weights over 2005-2006 and 2010-2011. We then compute reweighed vacancy counts by occupation o in month t as

$$v_{ot}^{H} = \sum_{i=1}^{I} \sum_{r=1}^{R} \omega_{it} \cdot \omega_{rt} \cdot v_{oirt}^{H}.$$

Our reweighed occupational mismatch index of Figure C25 is based on this revised vacancy count.

C Additional figures and tables

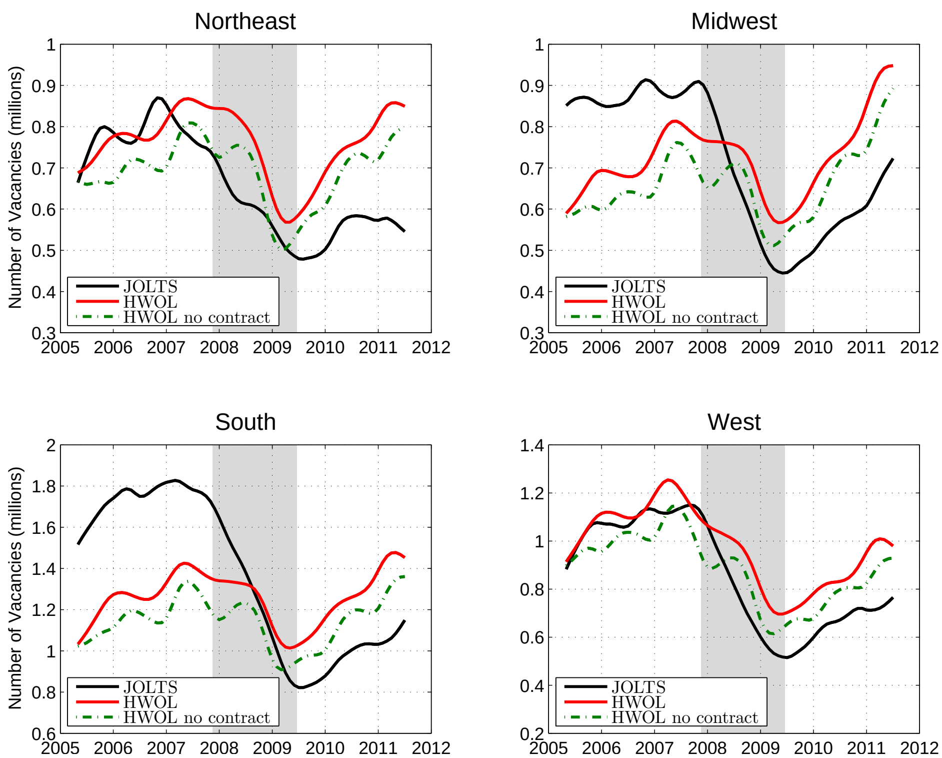
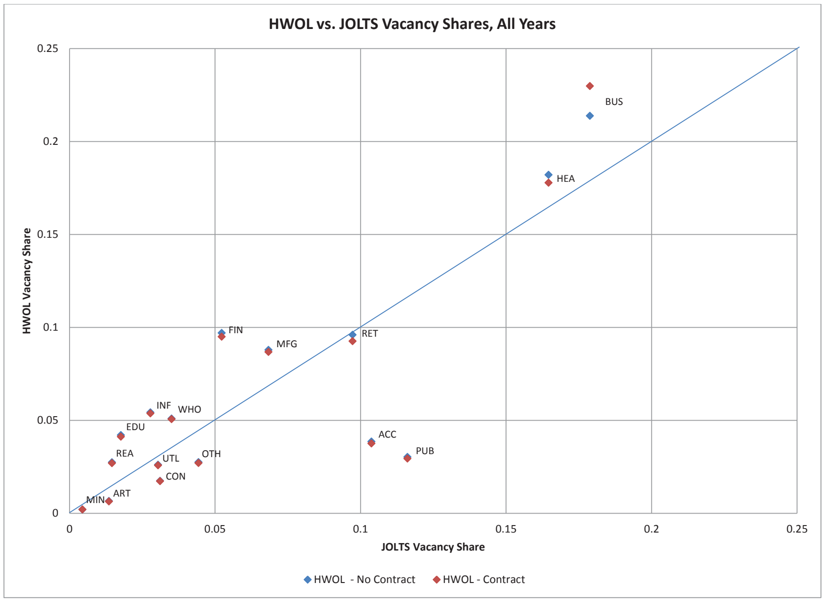


Figure C1: Comparison between the JOLTS and the HWOL (The Conference Board Help Wanted OnLine Data Series). Top-left panel: Northeast, Top-right panel: Midwest, Bottom-left panel: South, Bottom-right panel: West.



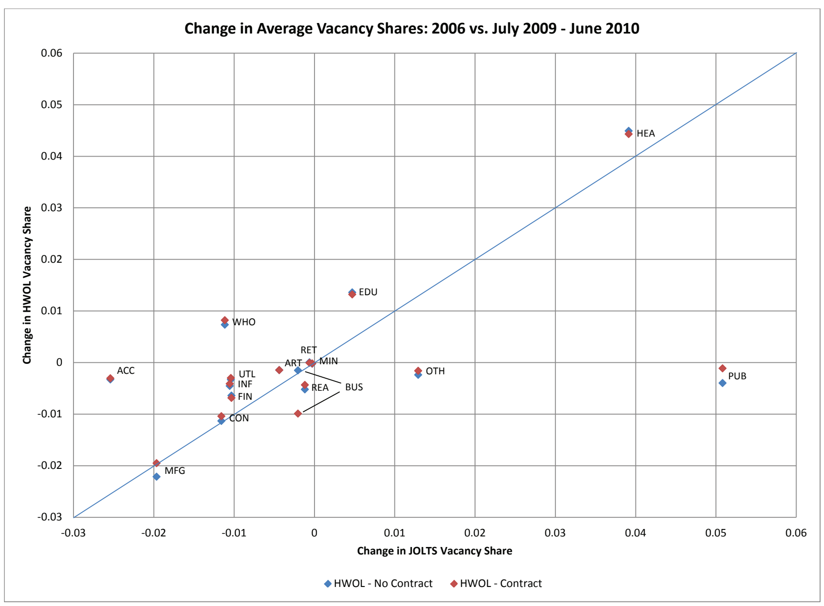


Figure C2: Top panel: comparison between vacancy shares in the JOLTS and HWOL (The Conference Board Help Wanted OnLine Data Series) for the May 2005 to June 2011 period. Bottom panel: change in average vacancy shares from 2006 to July 2009-June 2010 in the JOLTS and the HWOL. See Table C1 for an explanation of industry labels.

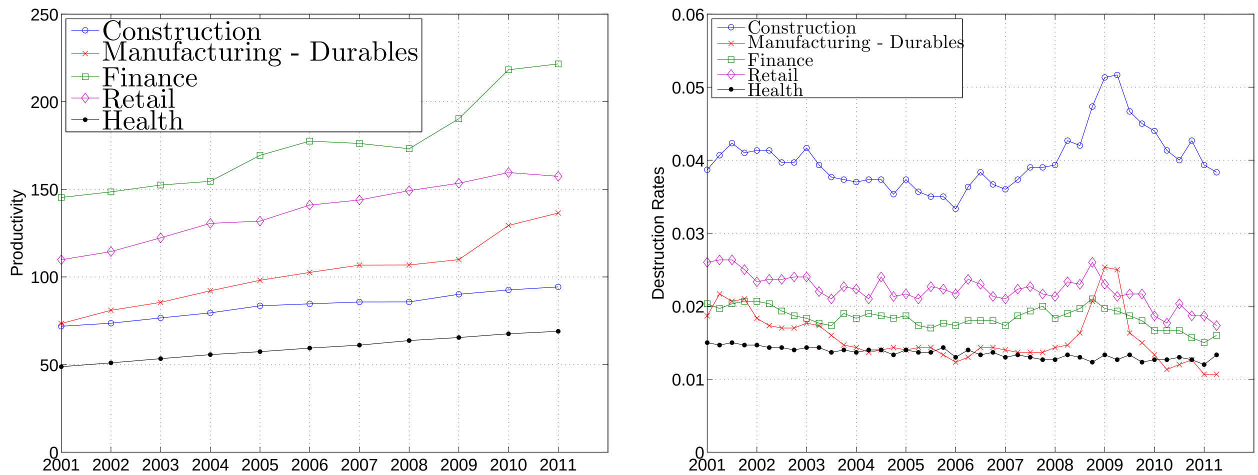


Figure C3: Productivity levels (left panel) and job destruction rates (right panel) for selected industries. Source: BEA and BLS for productivity levels and BED for job destruction rates.

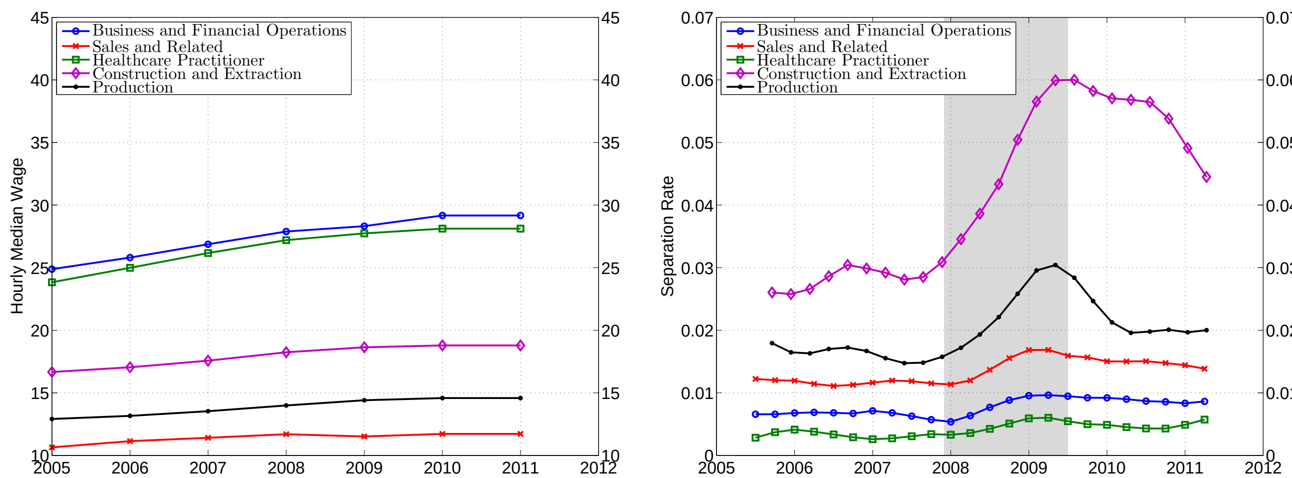


Figure C4: Wages (left panel) and job separation rates (right panel) for selected occupations. Source: OES for wages and CPS for job separation rates.

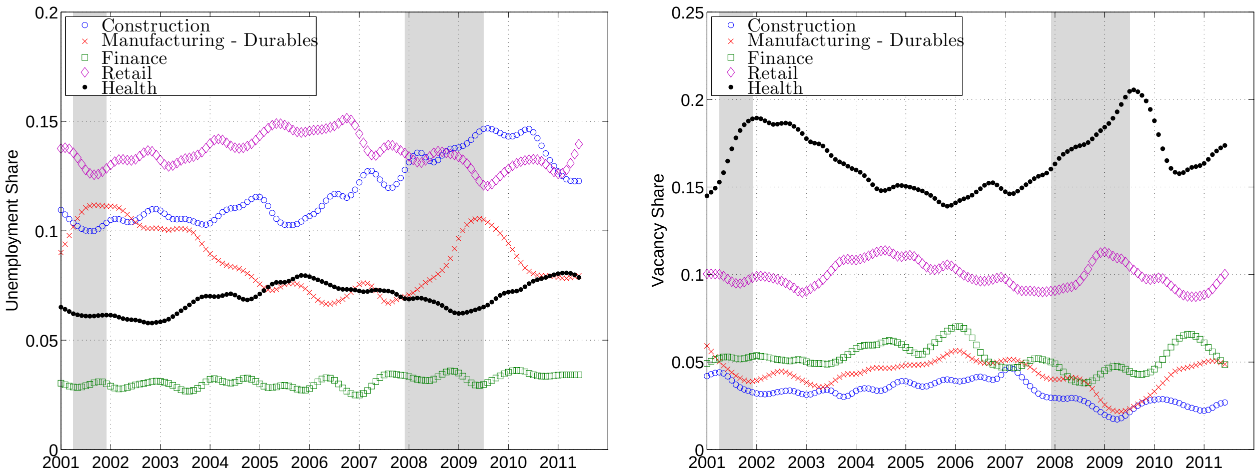


Figure C5: Unemployment and vacancy shares by selected industry.

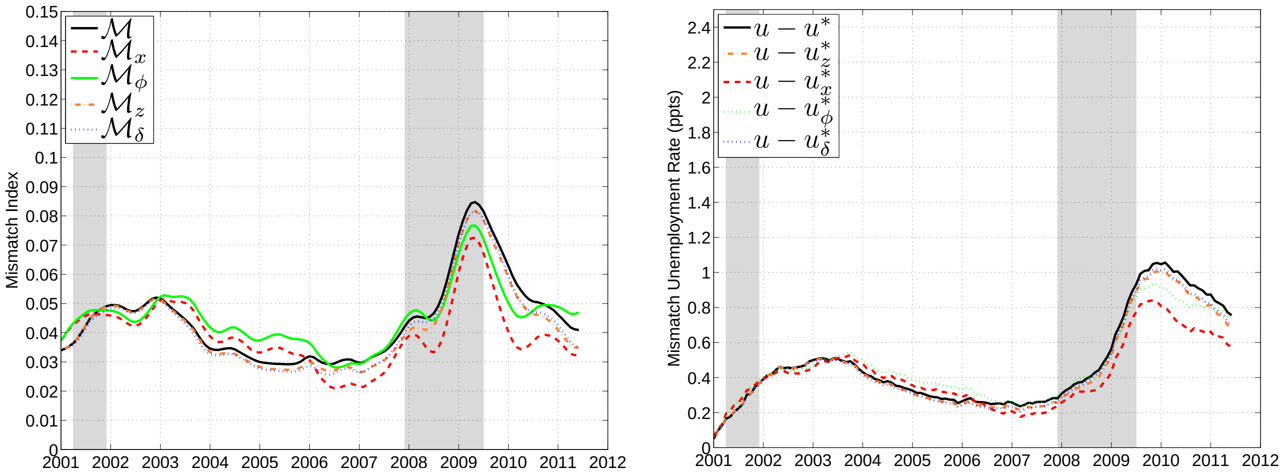


Figure C6: Mismatch indexes $\mathcal{M}_t$, $\mathcal{M}_{xt}$, $\mathcal{M}_{\phi t}$, $\mathcal{M}_{zt}$, and $\mathcal{M}_{\delta t}$ by industry (left panel) and the corresponding mismatch unemployment rates (right panel).

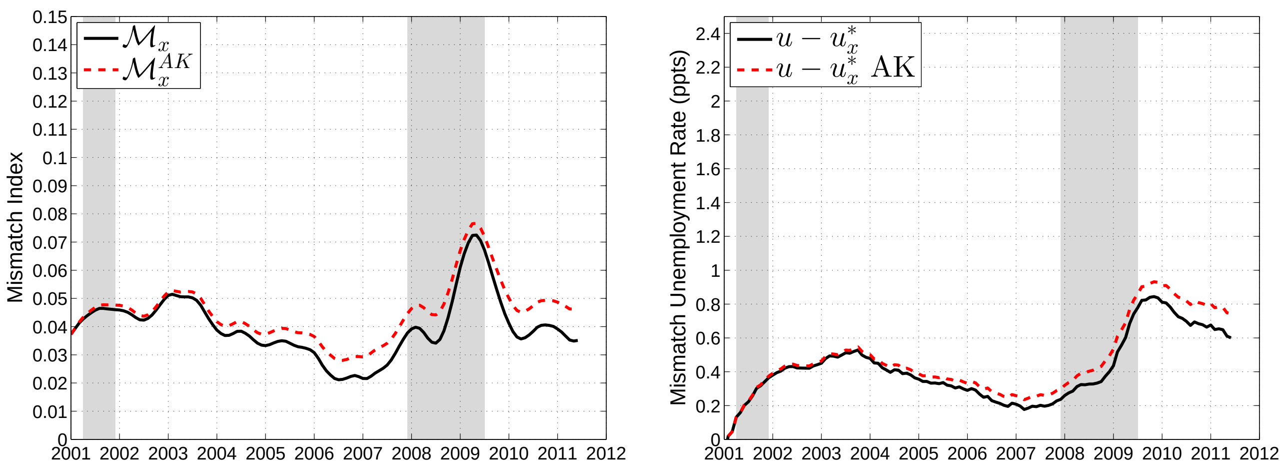


Figure C7: Mismatch indexes $\mathcal{M}_{xt}$ by industry (left panel) and corresponding mismatch unemployment rates (right panel) for the baseline specification and with the Abraham-Katz (AK) specification with heterogenous sensitivities to aggregate shocks.

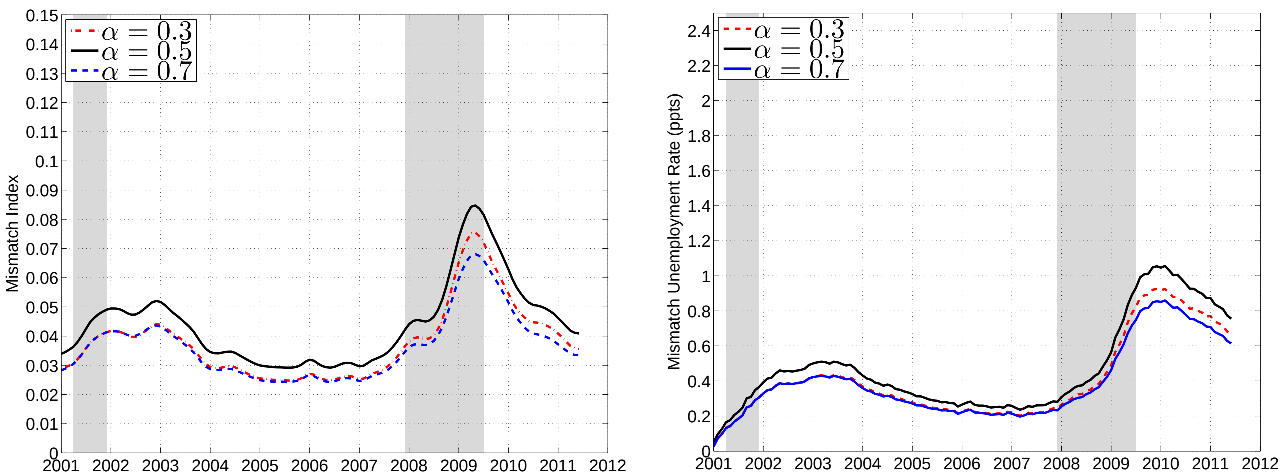


Figure C8: Mismatch index $\mathcal{M}_t$ by industry (left panel) and the corresponding mismatch unemployment rates (right panel) for various values of α, the vacancy share parameter in the matching function

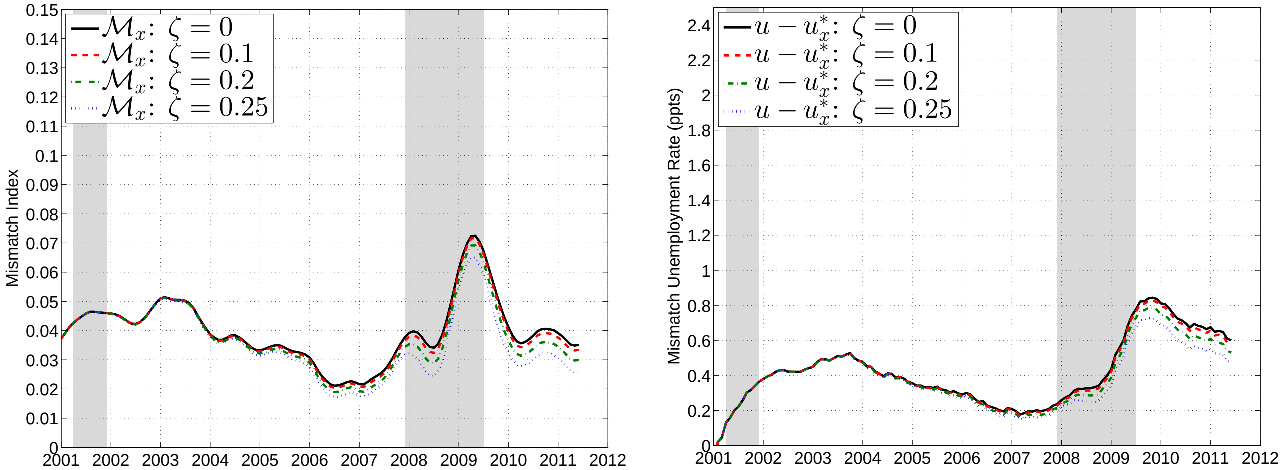


Figure C9: Mismatch index $\mathcal{M}_t$ by industry for different values of the utility flow from non-employment ζ (left panel) and the corresponding mismatch unemployment rates (right panel).

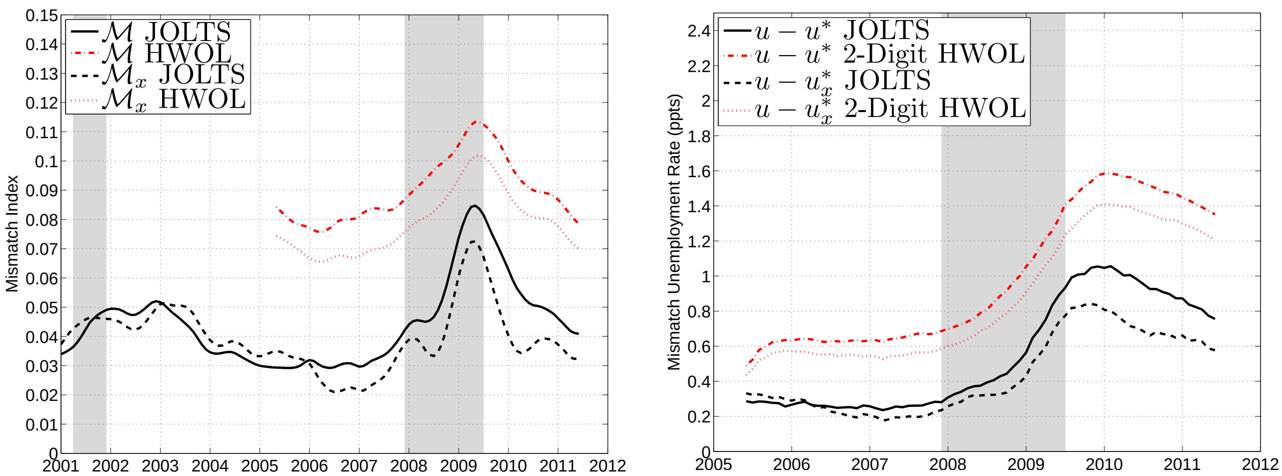


Figure C10: Mismatch indexes $\mathcal{M}_t$ (left panel) and the corresponding mismatch unemployment rates (right panel) across industries using the industry classification in JOLTS and the 2-digit industry classification in HWOL (The Conference Board Help Wanted OnLine Data Series).

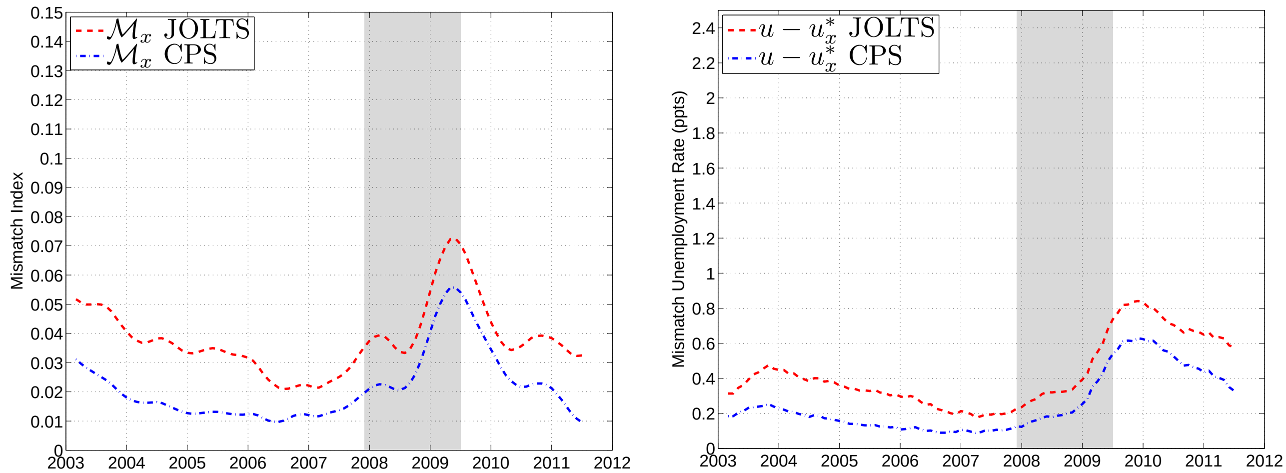


Figure C11: Mismatch index $\mathcal{M}_{xt}$ by industry (left panel) and the corresponding mismatch unemployment rates (right panel) using the JOLTS measure of hires and an estimate of of hires from unemployment based on the CPS.

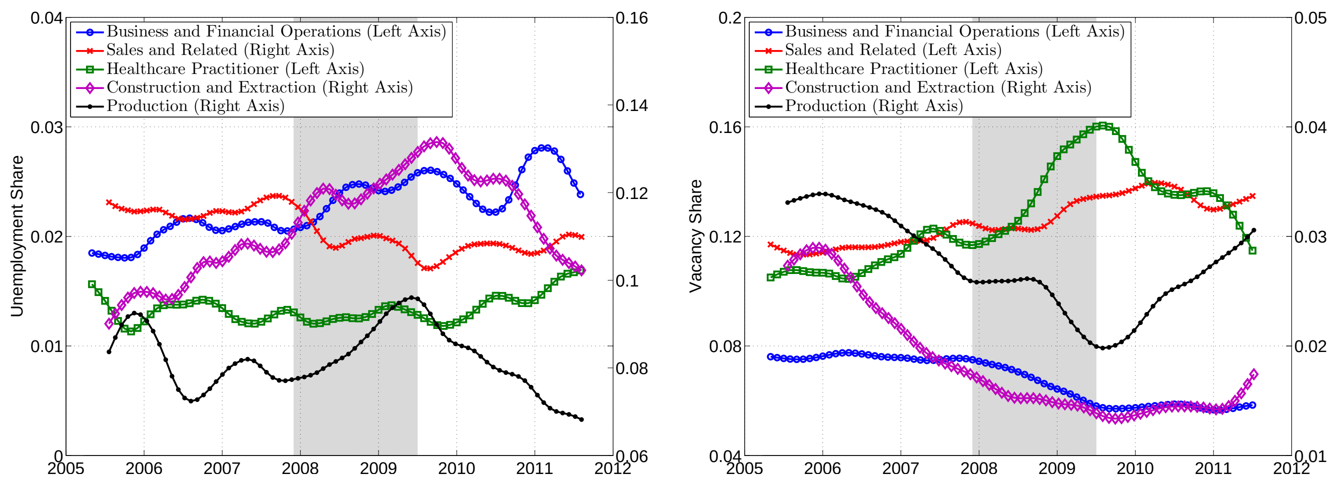


Figure C12: Unemployment and vacancy shares by selected occupation.

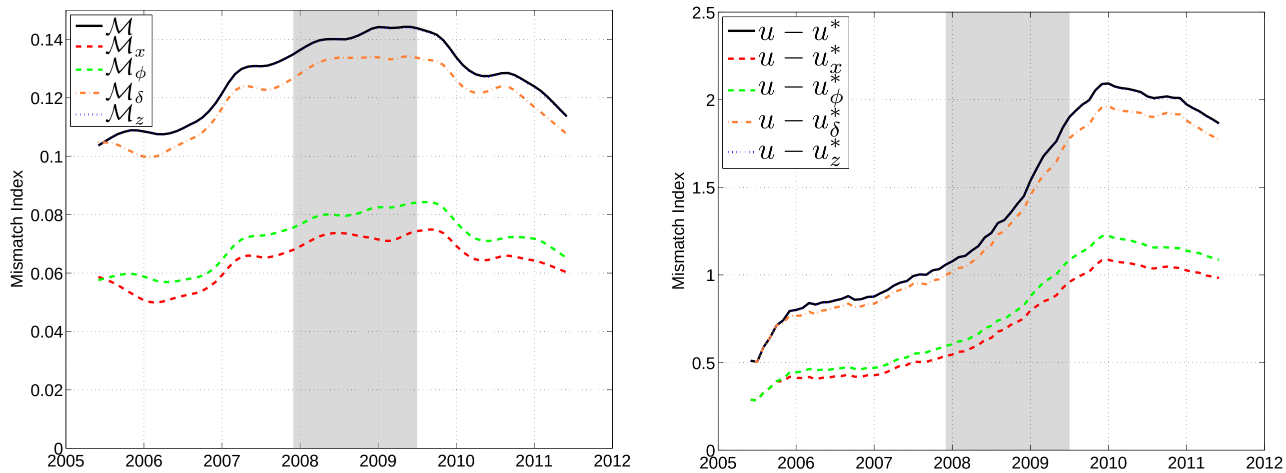


Figure C13: Mismatch indexes $\mathcal{M}_t$, $\mathcal{M}_{xt}$, $\mathcal{M}_{\phi t}$, $\mathcal{M}_{zt}$, and $\mathcal{M}_{\delta t}$ by occupation (left panel) and the corresponding mismatch unemployment rates (right panel).

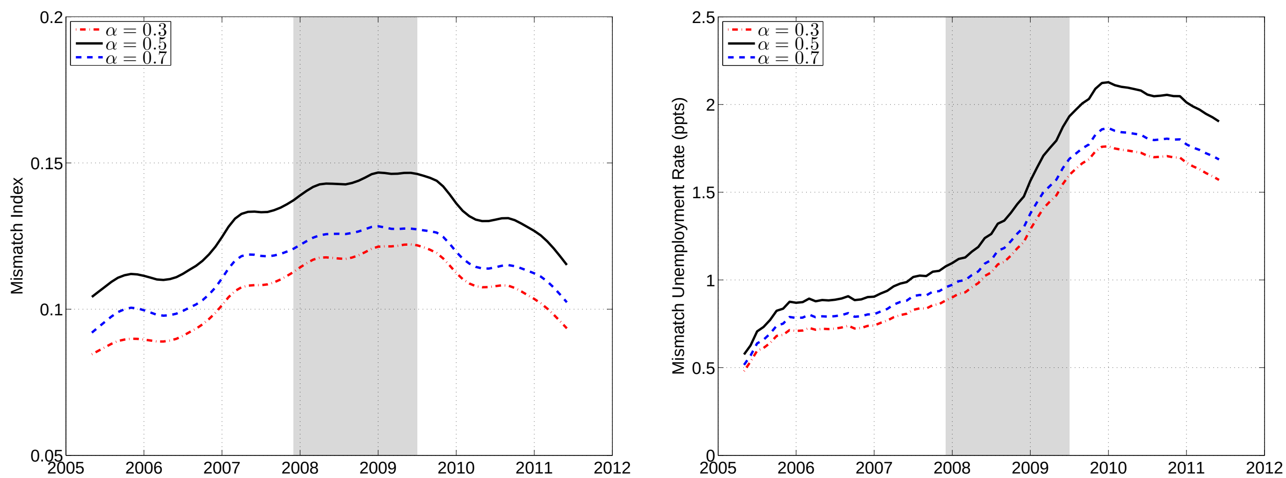


Figure C14: Mismatch index $\mathcal{M}_t$ by occupation (left panel) and the corresponding mismatch unemployment rates (right panel) for various values of α, the vacancy share parameter of the matching function.

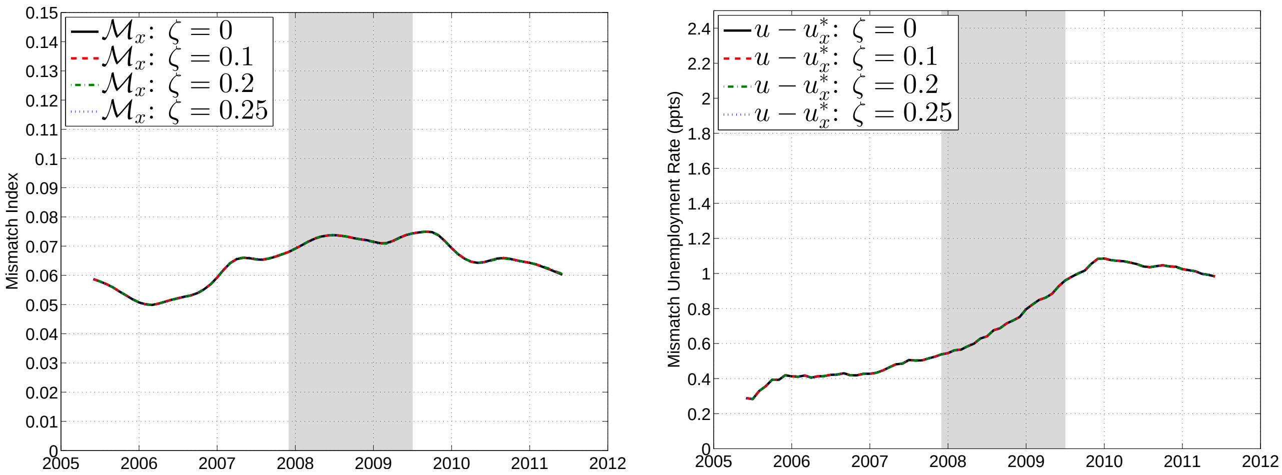


Figure C15: Mismatch index $\mathcal{M}_t$ by occupation for different values of the flow utility from nonemployment ζ (left panel), and the corresponding mismatch unemployment rates (right panel).

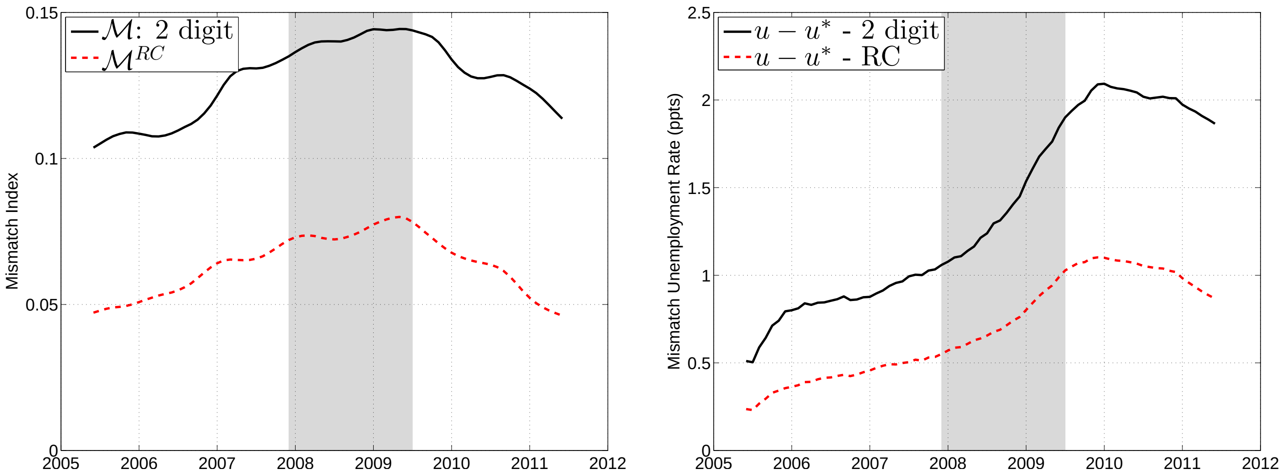


Figure C16: Mismatch indexes $\mathcal{M}$ across four occupations groups (routine/cognitive, manual/non-manual, and across 2-digit occupations (left panel). Corresponding mismatch unemployment rates (right panel).

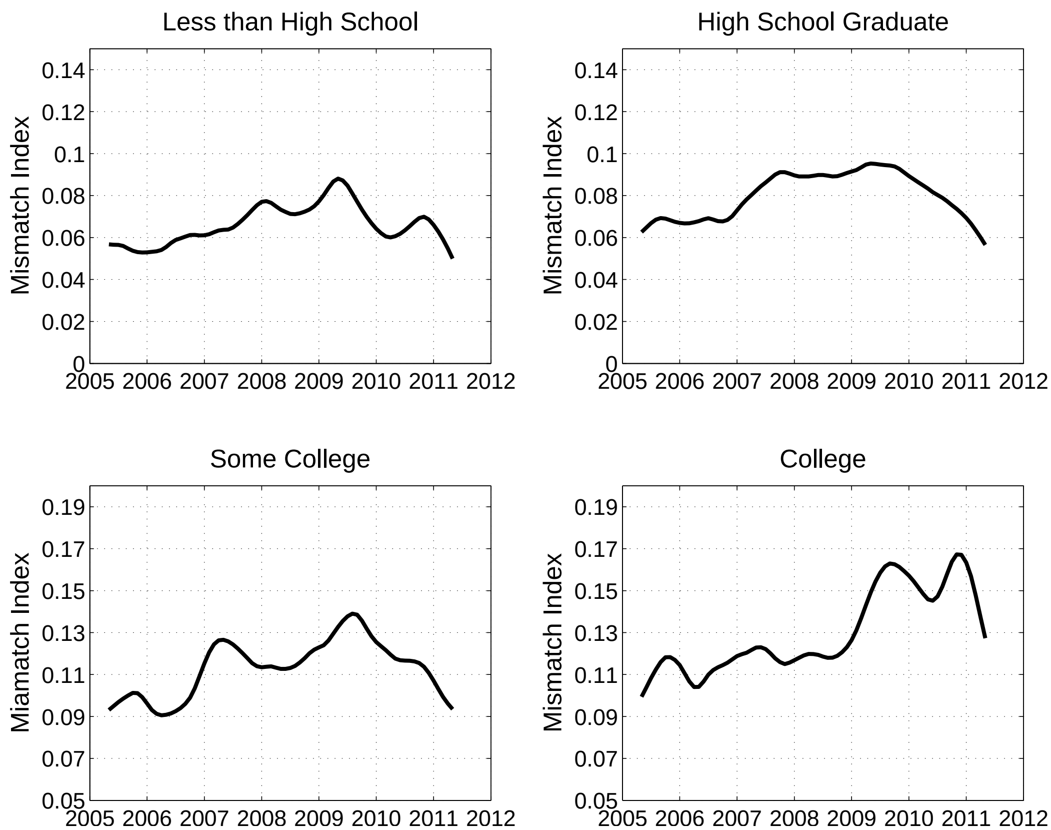


Figure C17: Mismatch indexes ($\mathcal{M}_t$) by occupation for different education groups.

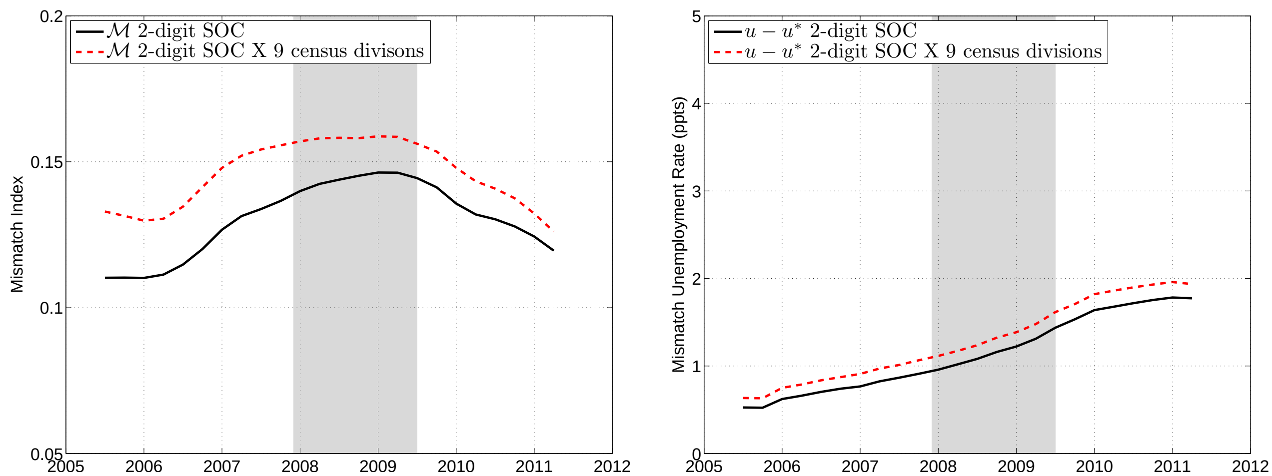


Figure C18: Mismatch index $\mathcal{M}_t$ by occupation and location (left panel) and the corresponding mismatch unemployment rates (right panel).

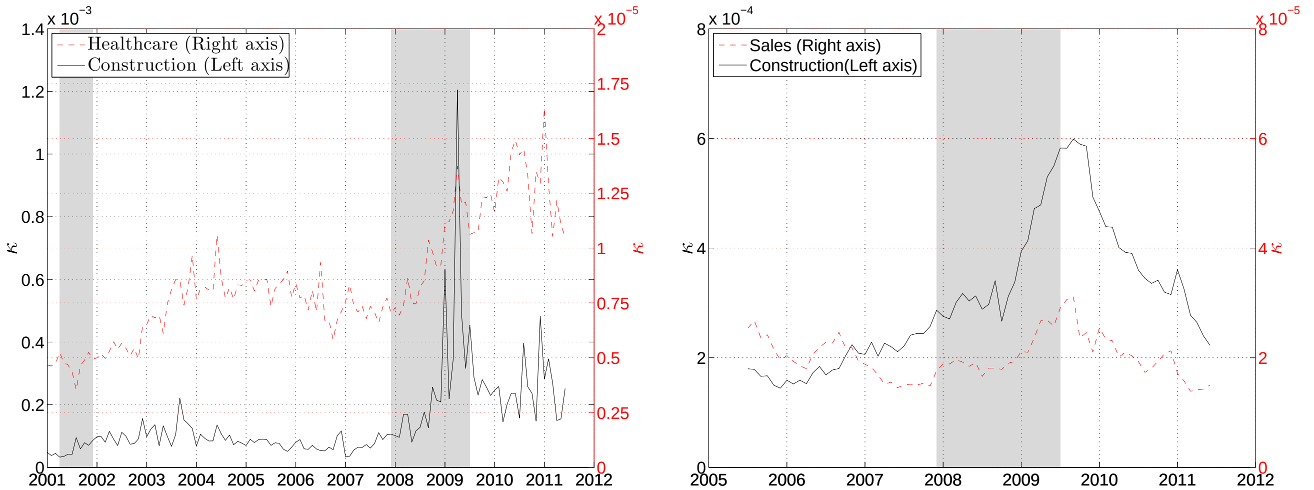


Figure C19: Time series of κ estimated with $\varepsilon = 1$ for two selected industries: construction and health care (left panel) and two selected occupations: construction and extraction occupations, and sales and related occupations (right panel). The cost is normalized by average annual labor productivity of the industry (annual wage for the occupation).

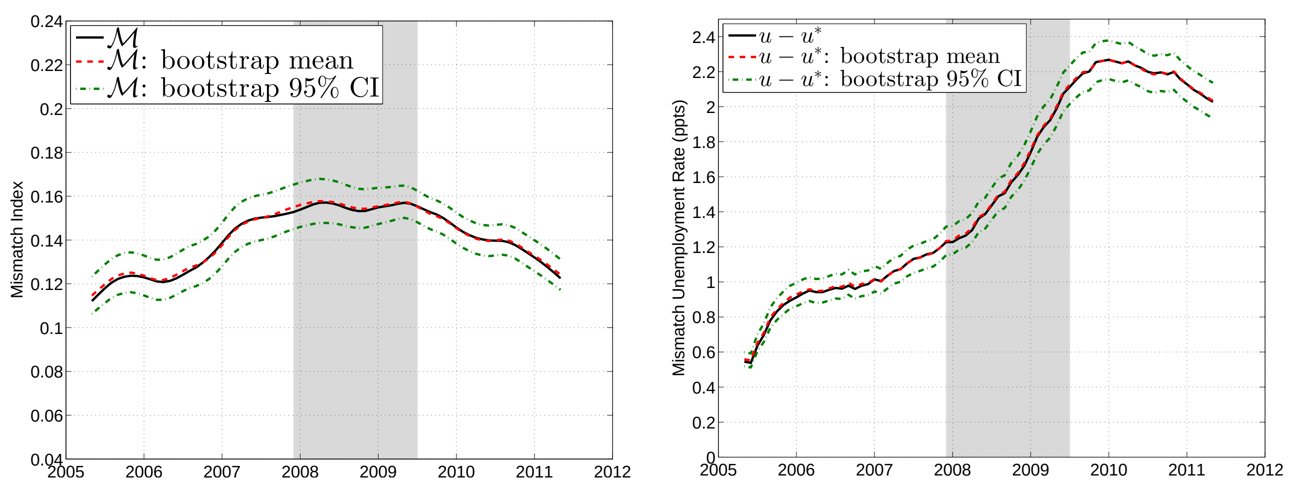


Figure C20: Mismatch index $\mathcal{M}_t$ by occupation (left panel), mismatch unemployment rate (right panel) and the corresponding 95% confidence intervals.

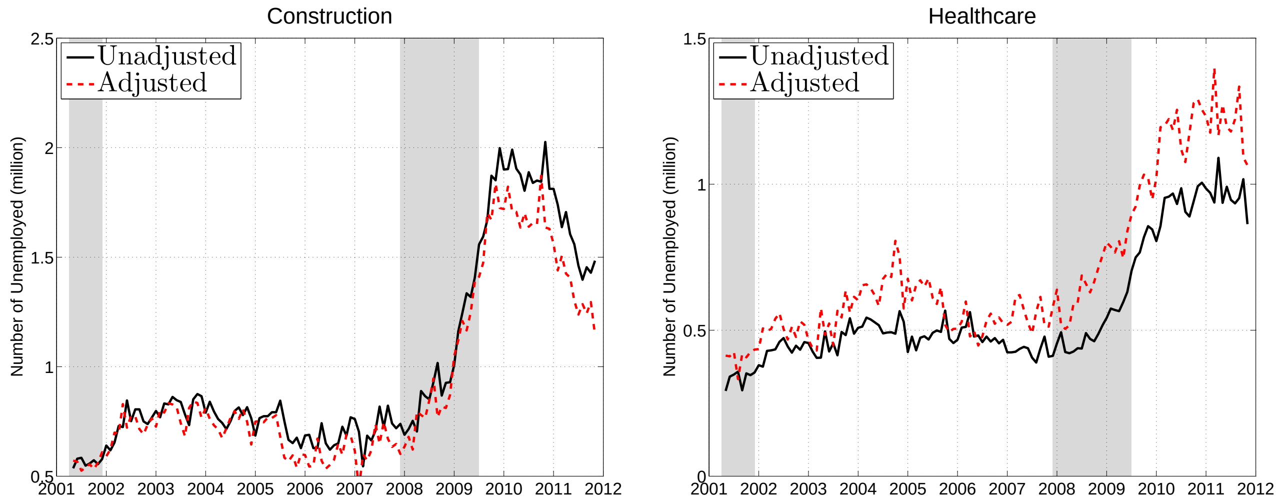


Figure C21: Adjusted unemployment counts for selected industries.

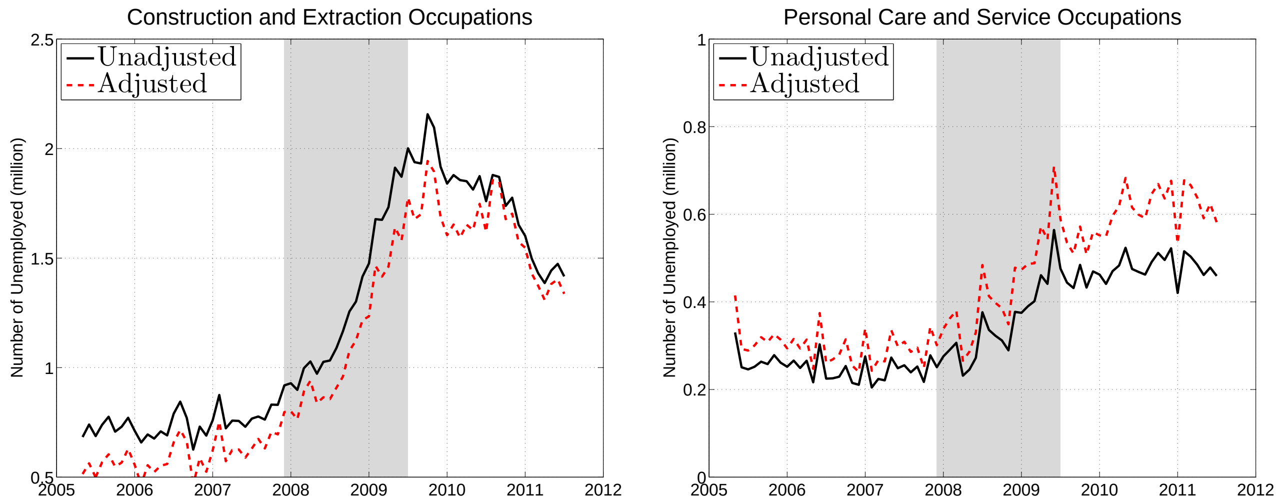


Figure C22: Adjusted unemployment counts for selected occupations.

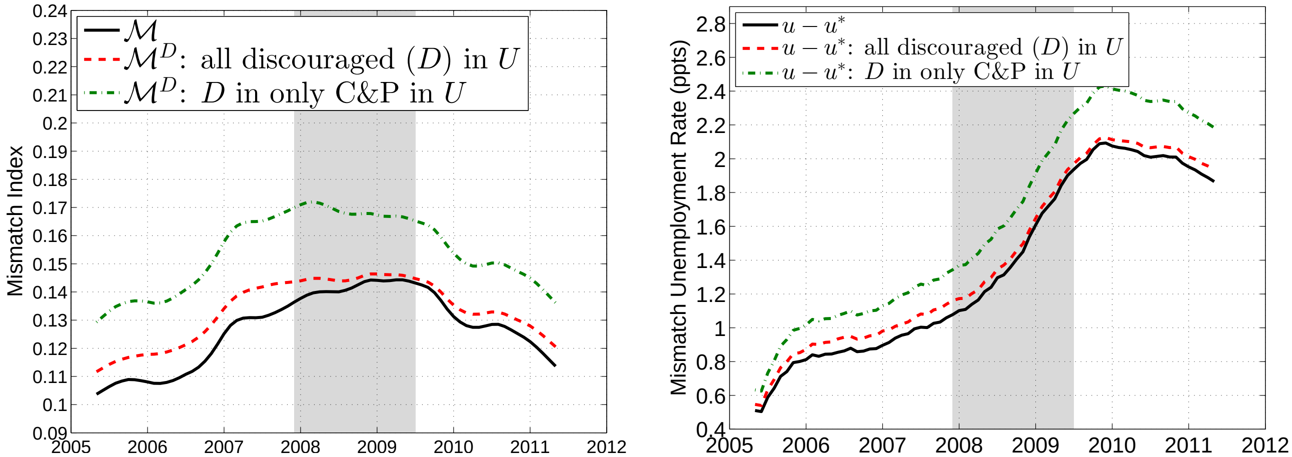


Figure C23: Mismatch indexes $\mathcal{M}_t$ by occupation (left panel) and the corresponding mismatch unemployment rates (right panel) including discouraged workers. The first correction includes all discouraged workers (D) among the unemployed (U). The second is a correction only for Construction (C) and Production-related (P) occupations.

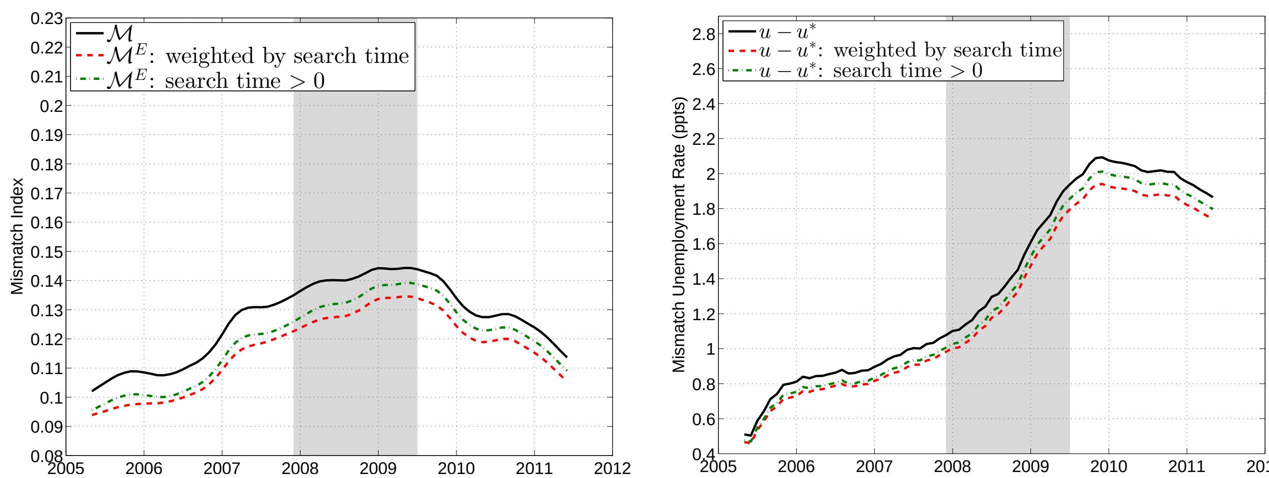


Figure C24: Mismatch indexes $\mathcal{M}_t$ by occupation (left panel) and the corresponding mismatch unemployment rates (right panel) including employed job seekers. The first correction weights employed workers by their reported search time in ATUS (relative to the search time of the unemployed), the second does not.

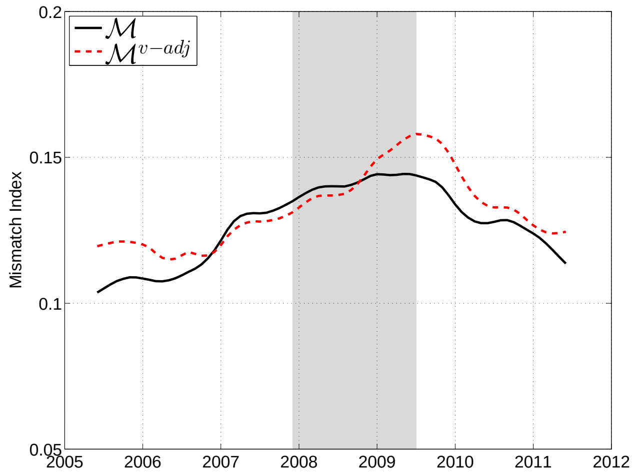


Figure C25: Mismatch index by 2-digit occupation: unadjusted index and index computed with reweighted HWOL vacancies

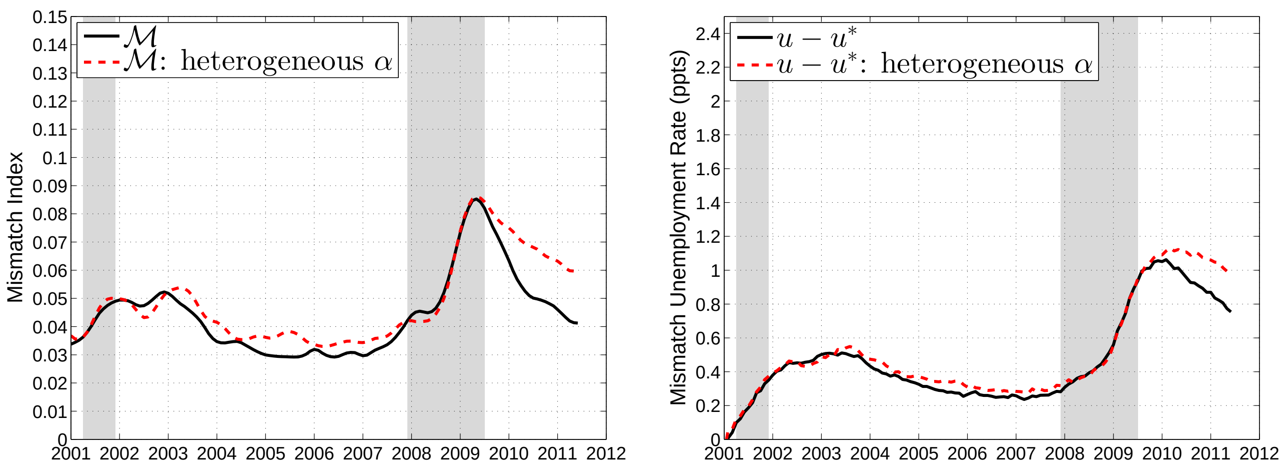


Figure C26: Mismatch indexes $\mathcal{M}_t$ by industry (left panel) and the corresponding mismatch unemployment rates (right panel) in the model with heterogenous α.

Code	Industry
ACC	Accomodation and Food Services
ART	Arts, Entertainment and Recreation
CON	Construction
EDU	Education Services
FIN	Finance and Insurance
PUB	Government
HEA	Health Care and Social Assistance
INF	Information
MFG	Manufacturing-Durable Goods
MFG	Manufacturing-Nondurable Goods
MIN	Mining
OTH	Other Services
BUS	Professional and Business Services
REA	Real Estate and Rental and Leasing
RET	Retail Trade
UTL	Transportation, Warehousing and Utilities
WHO	Wholesale Trade

Table C1: Industry classification in the JOLTS. The codes in the left column are those used in Figure C2.

Code	Occupation	Classification
110000	Management Occupations	Cognitive/Non-routine
130000	Business and Financial Operations Occupations	Cognitive/Non-routine
150000	Computer and Mathematical Occupations	Cognitive/Non-routine
170000	Architecture and Engineering Occupations	Cognitive/Non-routine
190000	Life, Physical, and Social Science Occupations	Cognitive/Non-routine
210000	Community and Social Service Occupations	Cognitive/Non-routine
230000	Legal Occupations	Cognitive/Non-routine
250000	Education, Training, and Library Occupations	Cognitive/Non-routine
270000	Arts, Design, Entertainment, Sports, and Media Occupations	Cognitive/Non-routine
290000	Healthcare Practitioners and Technical Occupations	Cognitive/Non-routine
310000	Healthcare Support Occupations	Manual/Non-routine
330000	Protective Service Occupations	Manual/Non-routine
350000	Food Preparation and Serving Related Occupations	Manual/Non-routine
370000	Building and Grounds Cleaning and Maintenance Occupations	Manual/Non-routine
390000	Personal Care and Service Occupations	Manual/Non-routine
410000	Sales and Related Occupations	Cognitive/Routine
430000	Office and Administrative Support Occupations	Cognitive/Routine
470000	Construction and Extraction Occupations	Manual/Routine
490000	Installation, Maintenance, and Repair Occupations	Manual/Routine
510000	Production Occupations	Manual/Routine
530000	Transportation and Material Moving Occupations	Manual/Routine

Table C2: 2-digit SOC Codes used in our empirical analysis. The classification in the right column is that used in Figure C16.

Code	Occupation
111000	Top Executives
113000	Operations Specialties Managers
119000	Other Management Occupations
131000	Business Operations Specialists
132000	Financial Specialists
151000	Computer Occupations
211000	Counselors, Social Workers, and Other Community and Social Service Specialists
252000	Preschool, Primary, Secondary, and Special Education School Teachers
272000	Entertainers and Performers, Sports and Related Workers
291000	Health Diagnosing and Treating Practitioners
311000	Nursing, Psychiatric, and Home Health Aides
339000	Other Protective Service Workers
352000	Cooks and Food Preparation Workers
353000	Food and Beverage Serving Workers
359000	Other Food Preparation and Serving Related Workers
372000	Building Cleaning and Pest Control Workers
373000	Grounds Maintenance Workers
399000	Other Personal Care and Service Workers
411000	Supervisors of Sales Workers
412000	Retail Sales Workers
413000	Sales Representatives, Services
419000	Other Sales and Related Workers
433000	Financial Clerks
434000	Information and Record Clerks
435000	Material Recording, Scheduling, Dispatching, and Distributing Workers
436000	Secretaries and Administrative Assistants
439000	Other Office and Administrative Support Workers
452000	Agricultural Workers
472000	Construction Trades Workers
493000	Vehicle and Mobile Equipment Mechanics, Installers, and Repairers
499000	Other Installation, Maintenance, and Repair Occupations
512000	Assemblers and Fabricators
514000	Metal Workers and Plastic Workers
519000	Other Production Occupations
533000	Motor Vehicle Operators
537000	Material Moving Workers

Table C3: 3-digit SOC Codes used in our empirical analysis.

	Aggregate regressions				Panel regressions	
	JOLTS		HWOL		Industry (JOLTS)	Occupation (HWOL)
	OLS	GMM	OLS	GMM	OLS	OLS
JOLTS Hires	0.654	0.661	–	–	0.532	–
	(0.010)	(0.037)	–	–	(0.013)	–
Sample Size	126	126	–	–	2,142	–
CPS Hires	0.318	0.298	0.332	0.536	0.241	0.279
	(0.017)	(0.136)	(0.038)	(0.059)	(0.014)	(0.016)
Sample Size	126	126	72	72	404	370

Table C4: OLS and GMM estimates of the vacancy share α using the JOLTS and HWOL datasets. S.E. in parenthesis. See Section B.2 for details.

	JOLTS		HWOL	
	α	σ	α	σ
JOLTS Hires	0.576	0.152	-	-
	[0.542,0.603]	[0.051,0.242]		
CPS Hires	0.301	0.18	0.239	-0.108
	[0.267,0.350]	[0.08,0.303]	[0.194,0.291]	[-0.226,0.004]

Table C5: Estimates of the vacancy share α and CES substitutability parameter σ, using industry and occupation level data. 95-5 confidence intervals computed via bootstrap. Sample sizes are the same as in Table C4.

Industry	ϕ^{pre}	ϕ^{post}
Mining	1.71	1.36
Arts	1.69	1.87
Construction	1.66	1.73
Accommodations	1.53	1.60
Retail	1.47	1.46
Professional and Business Services	1.43	1.45
Real Estate	1.41	1.22
Wholesale	1.21	1.35
Other	1.14	1.16
Transportation and Utilities	1.14	1.16
Manufacturing - Nondurables	0.96	1.00
Education	0.94	1.02
Health	0.93	1.05
Government	0.87	0.89
Finance	0.85	0.73
Manufacturing - Durables	0.84	0.78
Information	0.76	0.70

Table C6: Estimates of industry-specific match efficiencies using hires from the JOLTS.

Industry Groups	Industry	ϕ^{pre}	ϕ^{post}
Group 1	Construction Mining	0.50	0.55
Group 2	Manufacturing Other Transportation and Utilities	0.42	0.44
Group 3	Accommodations Arts Professional and Business Services Retail Wholesale	0.38	0.39
Group 4	Education Finance Government Health Information Real Estate	0.33	0.33

Table C7: Estimates of industry-specific match efficiencies using hires from the CPS.

Occupation Groups	Occupation	ϕ^{pre}	ϕ^{post}
Service	Protective Service Occupations Food Preparation and Serving Related Occupations Building and Grounds Cleaning and Maintenance Occupations Personal Care and Service Occupations	0.58	0.63
Natural Resources, Construction and Maintenance	Construction and Extraction Occupations Installation, Maintenance, and Repair Occupations	0.56	0.63
Production, Transportation and Material Moving	Production Occupations Transportation and Material Moving Occupations	0.48	0.52
Sales and Office	Sales and Related Occupations Office and Administrative Support Occupations	0.37	0.35
Management, Professional and Related	Management Occupations Business and Financial Operations Occupations Computer and Mathematical Occupations Architecture and Engineering Occupations Life, Physical, and Social Science Occupations Community and Social Service Occupations Legal Occupations Education, Training, and Library Occupations Arts, Design, Entertainment, Sports, and Media Occupations Healthcare Practitioners and Technical Occupations Healthcare Support Occupations	0.32	0.33

Table C8: Estimates of occupation-specific match efficiencies using hires from the CPS.

	Index	$u_{06} - u^*_{06}$	$u_{10.09} - u^*_{10.09}$	$\Delta(u - u^*)$	$\Delta(u - u^*)/\Delta u$
JOLTS Hires	$\mathcal{M}$	0.26	1.01	0.75	13.9%
	$\mathcal{M}_x$	0.24	0.84	0.59	11.0%
	$\mathcal{M}_x^{AK}$	0.28	0.89	0.61	11.2%
	$\mathcal{M}_x^{v^*}(\varepsilon = 0.5)$	0.67	1.90	1.22	22.5%
	$\mathcal{M}_x^{v^*}(\varepsilon = 1.0)$	0.35	1.24	0.90	16.6%
	$\mathcal{M}_x^{v^*}(\varepsilon = 2.0)$	0.27	0.95	0.69	12.7%
	$\mathcal{M}_\phi$	0.29	0.92	0.63	11.7%
	$\mathcal{M}_z$	0.24	0.96	0.72	13.4%
	$\mathcal{M}_\delta$	0.23	0.98	0.74	13.7%
	$\mathcal{M}^{u-adj}$	0.25	0.89	0.65	11.9%
	$\mathcal{M}(\alpha = 0.3)$	0.22	0.89	0.67	12.4%
	$\mathcal{M}(\alpha = 0.5)$	0.26	1.01	0.75	13.9%
	$\mathcal{M}(\alpha = 0.7)$	0.22	0.82	0.60	11.1%
	$\mathcal{M}_x^{break}$	0.25	0.92	0.67	12.4%
	$\mathcal{M}_x(\zeta = 0.10)$	0.24	0.82	0.59	10.8%
	$\mathcal{M}_x(\zeta = 0.20)$	0.23	0.79	0.56	10.3%
	$\mathcal{M}_x(\zeta = 0.25)$	0.22	0.73	0.51	9.4%
CPS Hires	$\mathcal{M}$	0.27	1.03	0.77	12.4%
	$\mathcal{M}_x$	0.10	0.61	0.51	9.4%
HWOL	$\mathcal{M}$	0.63	1.51	0.88	16.3%
	$\mathcal{M}_x$	0.56	1.35	0.79	14.7%

Table C9: Changes in mismatch unemployment at the industry level. All the changes are calculated as the difference between October 2009 and the average of 2006. Note that $\Delta u = 5.4$ percentage points.

	Index	$u_{06}-u^*_{06}$	$u_{10.09}-u^*_{10.09}$	$\Delta(u-u^*)$	$\Delta(u-u^*)/\Delta u$
2-digit	$\mathcal{M}$	0.85	2.00	1.15	21.3%
	$\mathcal{M}_x$	0.42	1.02	0.60	11.1%
	$\mathcal{M}_x^{v^*}(\varepsilon=0.5)$	1.08	2.60	1.52	28.1%
	$\mathcal{M}_x^{v^*}(\varepsilon=1.0)$	0.75	1.81	1.07	19.7%
	$\mathcal{M}_x^{v^*}(\varepsilon=2.0)$	0.58	1.41	0.83	15.3%
	$\mathcal{M}^{u-adj}$	0.84	2.00	1.16	21.4%
	$\mathcal{M}^{v-adj}$	0.92	2.12	1.19	22.1%
	$\mathcal{M}^D$ (all discouraged in U)	0.92	2.03	1.11	20.6%
	$\mathcal{M}^D$ (D in C&P in U)	1.06	2.33	1.27	23.4%
	$\mathcal{M}^E$ (E: weighted by search time)	0.78	1.90	1.13	20.9%
	$\mathcal{M}^E$ (E: fraction searching)	0.79	1.97	1.18	21.8%
	$\mathcal{M}_\phi$	0.46	1.15	0.69	12.8%
	$\mathcal{M}_z$	0.85	2.00	1.15	21.2%
	$\mathcal{M}_\delta$	0.80	1.86	1.05	19.5%
	$\mathcal{M}(\alpha=0.3)$	0.72	1.69	0.96	17.8%
	$\mathcal{M}(\alpha=0.5)$	0.85	2.00	1.14	21.3%
	$\mathcal{M}(\alpha=0.7)$	0.79	1.77	0.98	18.1%
	$\mathcal{M}_x^{break}$	0.42	0.98	0.56	10.4%
	$\mathcal{M}(\zeta=0.10)$	0.42	1.02	0.60	11.1%
	$\mathcal{M}(\zeta=0.20)$	0.42	1.02	0.60	11.1%
	$\mathcal{M}(\zeta=0.25)$	0.42	1.02	0.60	11.1%
3-digit	$\mathcal{M}$	1.33	2.91	1.58	29.3%
	$\mathcal{M}_x$	0.79	1.73	0.94	17.4%
	$\mathcal{M}_\phi$	0.83	1.85	1.02	18.8%
	$\mathcal{M}_z$	1.33	2.91	1.58	29.2%
	$\mathcal{M}_\delta$	1.29	2.80	1.50	27.8%

Table C10: Changes in mismatch unemployment at the occupation level. All the changes are calculated as the difference between October 2009 and the average of 2006. Note that $\Delta u = 5.4$ percentage points.

Index	$u_{Q1.01} - u^*_{Q1.01}$	$u_{06.03} - u^*_{06.03}$	$\Delta(u - u^*)$	$\Delta(u - u^*)/\Delta u$
$\mathcal{M}$	0.09	0.50	0.41	22.8%
$\mathcal{M}_x$	0.10	0.50	0.41	21.7%
$\mathcal{M}^{u-adj}$	0.11	0.43	0.32	17.8%
$\mathcal{M}_x^{v^*}(\varepsilon = 1.0)$	0.20	0.70	0.50	26.8%

Table C11: Changes in mismatch unemployment at the industry level for the 2001 recession. All the changes are calculated as the difference between June 2003 (month in which the unemployment rate peaked for the 2001 recession) and the average of 2001Q1. Note that $\Delta u = 1.8$ percentage points.

		2005-2007		2008-2011	
	Occupation	D	U	D	U
11	Management	3.86	4.44	4.24	5.47
13	Business and Financial	2.23	2.26	2.24	2.70
15	Computer and Math	0.90	1.22	1.17	1.36
17	Architecture and Engineering	0.72	0.77	0.84	1.30
19	Life, Physical, and Social Science	0.58	0.45	0.40	0.46
21	Community and Social Service	0.79	0.80	0.79	0.84
23	Legal	0.43	0.45	0.81	0.46
25	Education, Training, and Library	4.85	3.22	5.31	2.84
27	Arts, Design, Entertainment, Sports, and Media	1.94	1.95	2.81	1.92
29	Healthcare Practitioners and Technical	1.90	1.48	2.26	1.45
31	Healthcare Support	2.29	2.34	1.85	1.94
33	Protective Service	1.47	1.76	1.96	1.44
35	Food Preparation and Serving Related	10.62	9.47	9.99	8.19
37	Building and Grounds Cleaning and Maintenance	6.78	6.18	6.22	5.71
39	Personal Care and Service	6.06	3.85	6.27	3.50
41	Sales and Related	15.01	12.94	12.62	11.75
43	Office and Administrative Support	12.91	13.18	12.79	12.60
45	Fishing and Farming	1.63	1.48	1.93	1.41
47	Construction and Extraction	8.40	11.04	8.93	13.34
49	Installation, Maintenance, and Repair	2.48	3.04	2.64	3.42
51	Production	6.13	8.94	5.80	9.23
53	Transportation and Material Moving	8.02	8.75	8.11	8.69

Table C12: Distribution of discouraged and unemployed workers across occupations; percent of D and U in each occupation.

	Weight 2005-2006	Weight 2010-2011
Industry		
Accomodation and Food Services	2.25	2.43
Arts, Entertainment and Recreation	1.07	1.03
Construction	1.42	1.32
Education Services	0.44	0.55
Finance and Insurance	0.49	0.56
Government	2.94	2.35
Health Care and Social Assistance	0.79	0.83
Information	0.49	0.58
Manufacturing-Durable Goods	0.81	0.64
Manufacturing-Nondurable Goods	0.75	0.63
Mining	0.82	1.23
Other Services	1.34	1.14
Professional and Business Services	0.34	0.35
Real Estate and Rental and Leasing	0.56	0.52
Retail Trade	0.92	1.04
Transportation, Warehousing and Utilities	1.00	1.07
Wholesale Trade	0.61	0.73
Region		
Northeast	0.90	0.99
West	1.18	0.97
Southwest	0.68	0.92
South	1.17	1.23

Table C13: Estimated weights which equalize monthly JOLTS and HWOL (The Conference Board Help Wanted OnLine Data Series) vacancy counts by industry and region (average weight is normalized to one each month).

	α	ϕ
Mining	0.5549	1.4503
	(0.056)	(0.110)
Construction	0.3999	1.1542
	(0.040)	(0.083)
Durable goods manufacturing	0.5757	0.7565
	(0.026)	(0.026)
Nondurable goods manufacturing	0.5381	0.8250
	(0.030)	(0.033)
Wholesale trade	0.5126	1.0329
	(0.029)	(0.020)
Retail trade	0.6488	1.3904
	(0.042)	(0.051)
Transportation and warehousing	0.4174	0.8851
	(0.037)	(0.030)
Information	0.5103	0.6210
	(0.030)	(0.018)
Financial activities	0.6485	0.6936
	(0.053)	(0.014)
Real estate	0.3528	1.0877
	(0.055)	(0.044)
Professional & business services	0.5922	1.2406
	(0.028)	(0.018)
Education	0.401	0.7213
	(0.056)	(0.036)
Healthcare	0.6932	0.7459
	(0.026)	(0.011)
Arts, entertaiment, and recreation	0.3511	1.2342
	(0.051)	(0.068)
Accommodation & food services	0.5543	1.3247
	(0.024)	(0.025)
Other	0.3836	0.9120
	(0.044)	(0.029)
Government	0.7891	0.7454
	(0.042)	(0.012)

Table C14: Estimates of α and ϕ by industry. Standard errors are in parentheses.