# Accounting for Mismatch Unemployment

Benedikt Herz and Thijs van Rens

May 2018

# Appendices
# (for online publication)

## A Model

This appendix provides the details on the derivations for the model. See Section 2.1 in the main text for a description of the model environment. Section A.1 below derives the efficient allocation, as discussed in Section 2.2 in the main text. Section A.2 derives the equilibrium, as used in Section 3.1 in the main text.

### A.1 Efficient Allocation

#### A.1.1 Social Planner Problem

The social planner solves

$$\max_{\{\{u_{it},v_{it}\}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \left( \sum_i \left( f\left(n_{it}; z_{it}\right) + b_{it} u_{it} - g\left(v_{it}; \kappa_{it}\right) \right) \right) \tag{1}$$

subject to

$$n_{it+1} = \left(1-\delta_i\right) n_{it} + m\left(u_{it}, v_{it}; \phi_{it}\right) \tag{2}$$

$$\sum_i u_{it} = 1 - \sum_i n_{it} \tag{3}$$

$$\sum_i v_{it} = 1 - \sum_i n_{it} \tag{4}$$

where constraints (3) and (4) get multipliers $\lambda_t^u$ and $\lambda_t^v$, respectively.

Let $V\left(\{n_{it}\}\right)$ denote the planner's value function in period $t$, which depends on the state variables $n_{it}$ for each segment $i$. The planner's problem can be written in recursive form as the following Bellman equation,

$$V\left(\{n_{it}\}\right) = \max_{\{u_{it},v_{it}\}} \left\{ \sum_i \left( f\left(n_{it}; z_{it}\right) + b_{it} u_{it} - g\left(v_{it}; \kappa_{it}\right) \right) + \beta E_t V\left(\{n_{it+1}\}\right) \right\} \tag{5}$$

where $n_{it+1}$ as in (2) and the maximization is subject to (3) and (4).

### A.1.2 Efficiency Conditions

The first-order conditions for $u_{it}$ and $v_{it}$ are given by

$$m_u\left(u_{it}, v_{it}; \phi_{it}\right) S_{it} = \lambda_t^u - b_{it} \tag{6}$$

$$m_v\left(u_{it}, v_{it}; \phi_{it}\right) S_{it} = \lambda_t^v + g'\left(v_{it}; \kappa_{it}\right) \tag{7}$$

where $S_{it} = \beta E_t V_i\left(\{n_{it+1}\}\right)$ is the discounted expected value of having one more worker employed in segment $i$ next period.

$S_{it}$ is determined by the envelope condition for $n_{it}$ (forwarding one period, taking conditional expectations and multiplying by $\beta$) and satisfies:

$$S_{it} = \beta E_t f'\left(n_{it+1}; z_{it+1}\right) + \beta\left(1-\delta_i\right) E_t S_{it+1} \tag{8}$$

Iterating forward,

$$S_{it} = \beta \sum_{s=0}^{\infty} \beta^s \left(1-\delta_i\right)^s E_t f'\left(n_{it+s+1}; z_{it+s+1}\right) = \frac{z_{it}}{r+\delta_i} \tag{9}$$

where the last equality follows if $f\left(n_{it}; z_{it}\right) = z_{it} n_{it}$ is linear and $z_{it}$ follows a random walk, so that $E_t f'\left(n_{it+s+1}; z_{it+s+1}\right) = z_{it}$.

### A.1.3 Efficient allocation of unemployed workers and vacancies

Dividing the first-order condition for unemployment (6) by that of vacancies (7) and using a Cobb-Douglas matching function $m\left(u_{it}, v_{it}; \phi_{it}\right) = \phi_{it} u_{it}^{\mu} v_{it}^{1-\mu}$, we get an expression for the vacancy-unemployment ratio in each labor market segment,

$$\frac{v_{it}}{u_{it}} = \frac{1-\mu}{\mu} \frac{\lambda_t^u - b_{it}}{\lambda_t^v + g'\left(v_{it}; \kappa_{it}\right)} \tag{10}$$

which is equation (1) in the main text.

Condition (10) reduces to condition (A36) in Şahin, Song, Topa, and Violante (2014) if we set $g\left(v_{it}; \kappa_{it}\right) = \frac{1}{1+\varepsilon} \kappa_{it}^{\varepsilon} v_{it}^{1+\varepsilon}$ and $\lambda_t^v = 0$ (free entry of vacancies). Substituting these assumptions into (10), substituting the result into (6), and solving for $v_{it}$ gives,

$$v_{it} = \frac{1}{\kappa_{it}} \left(\frac{1-\mu}{\mu}\right)^{1/\varepsilon} \left(\frac{1}{\lambda_t^u - b_{it}}\right)^{-\frac{\mu/\varepsilon}{1-\mu}} \left(\mu \phi_{it} S_{it}\right)^{\frac{1/\varepsilon}{1-\mu}} \tag{11}$$

which is equation (A36) in Şahin, Song, Topa, and Violante (2014) if we set $\mu = 1-\alpha$, $\lambda_t^u = \tilde{\mu}$, $b_{it} = 0$, $\phi_{it} = \Phi \phi_i$ and $S_{it} = \frac{z_{it}}{r+\delta_i} = \frac{Z z_i}{1-\beta(1-\Delta)(1-\delta_i)}$ to be consistent with their notation and assumptions.

In order to compare to the baseline in Şahin, Song, Topa, and Violante (2014), in which the vacancy distribution is exogenous, we simply drop the first-order condition for vacancies (7), so that the efficient allocation is described by conditions (6) and (9), which is condition (2) in the main text.

#### A.1.4 Productivity and matching efficiency

If we assume $b_{it} = b_t$ and $g'(v_{it}; \kappa_{it}) = \kappa_t$ and substitute (10) back into (6), we get an efficiency condition that does not depend on the distributions of unemployed worker and vacancies.

$$\phi_{it} S_{it} = \frac{1}{\mu} \left( \frac{\mu}{1-\mu} \right)^{1-\mu} (\lambda_t^u - b_t)^{\mu} (\lambda_t^v + \kappa_t)^{1-\mu} \tag{12}$$

In the efficient allocation, $S_{it} = f'(n_{it}; z_{it}) / (r + \delta_i)$ must be inversely proportional to matching efficiency across labor market segments, which implies that $S_{it}$ must be equalized if matching efficiency is constant across labor market segments, as discussed in Section 2.3 in the main text. Substituting condition (12) into the remaining condition (7) pins down the the multiplier $\lambda_t^u$.

### A.2 Equilibrium Allocation

There are two types of agents in our economy: a large representative household, consisting of a measure 1 of workers, and a large representative firm, consisting of a measure 1 of jobs, which may be filled or unfilled.

#### A.2.1 Households

The household does not have a technology for intertemporal consumption smoothing nor a motive to do so (the utility function is linear in consumption), so that maximizing utility is equivalent to maximizing consumption and maximizing income. Thus, the household chooses to which segments to allocate its unemployed workers in order to solve

$$\max_{\{\{u_{it}\}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \left( \sum_i (w_{it} n_{it} + b_{it} u_{it}) \right) \tag{13}$$

subject to

$$n_{it+1} = (1 - \delta_i) n_{it} + p_{it} u_{it} \tag{14}$$

$$\sum_i u_{it} = 1 - \sum_i n_{it} \tag{15}$$

and taking $w_{it}$ (wages) and $p_{it}$ (job finding probabilities) as given. The endogenous variables $u_{it}$ (number of unemployed workers) and $n_{it}$ (employment), the exogenous variable $b_{it}$ (home production of the unemployed), and the parameters $\delta_i$ (separation probabilities) and $\beta$ (discount factor) are the same as for the social planner problem, and were introduced in the main text.

Writing the problem in recursive form, with $\{n_{it}\}$ as the endogenous state variables, we get the following first-order condition for $u_{it}$,

$$p_{it} S_{it}^W = \lambda_t^u - b_{it} \tag{16}$$

which is worker mobility condition (7) in the main text, where $\lambda_t^u$ is the multiplier on the unemployment constraint (15).

$S_{it}^W$ is the discounted expected value to the household of having one more worker employed in segment $i$ next period, and is found from the envelope condition for $n_{it}$, by forwarding one period, taking conditional expectations and multiplying by $\beta$,

$$S_{it}^W = \beta E_t w_{it+1} + \beta\left(1-\delta_i\right) E_t S_{it+1}^W = \beta \sum_{s=0}^{\infty} \beta^s \left(1-\delta_i\right)^s E_t w_{it+s+1} \tag{17}$$

where the last equality follows from iterating forward. If we further assume that wages follow a random walk, then $S_{it}^W = w_{it}/\left(r+\delta_i\right)$.

#### A.2.2 Firms

The firm chooses in which segment to post its vacancies. To make the problem analogous to the household's problem, we assume that the total number of (filled and unfilled) jobs is constant at measure 1, which constrains the total amount of vacancies. The optimality condition if the amount of vacancies is chosen by the firm (through free entry or any other mechanism) follows as a special case by setting the multiplier on the vacancy constraint to zero. Since the utility function of the household is linear, the firm uses the same discount factor $\beta = 1/\left(1+r\right)$ as the household. Thus, the firm solves

$$\max_{\{\{v_{it}\}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t \left( \sum_i \left( f\left(n_{it}; z_{it}\right) - w_{it} n_{it} - g\left(v_{it}; \kappa_{it}\right) \right) \right) \tag{18}$$

subject to

$$n_{it+1} = \left(1-\delta_i\right) n_{it} + q_{it} v_{it} \tag{19}$$

$$\sum_i v_{it} = 1 - \sum_i n_{it} \tag{20}$$

and taking $w_{it}$ (wages) and $q_{it}$ (job filling probabilities) as given. The endogenous variables $v_{it}$ (number of vacancies) and $n_{it}$ (employment), the exogenous variables $z_{it}$ (production efficiency) and $\kappa_{it}$ (vacancy cost parameter), and the parameters $\delta_i$ (separation probabilities) and $\beta$ (discount factor) are the same as for the social planner problem, and were introduced in the main text.

The first-order condition for $v_{it}$,

$$q_{it} S_{it}^F = \lambda_t^v + g'\left(v_{it}; \kappa_{it}\right) \tag{21}$$

is job mobility condition (8) in the main text, where $\lambda_t^v$ is the multiplier on the vacancy constraint (20) with $\lambda_t^v = 0$ if we assume free entry of vacancies, and $S_{it}^F$ is the discounted expected value to the firm of having one more worker employed in segment $i$ next period.

$$\begin{aligned} S_{it}^{F} &= \beta E_t \left[ f\left(n_{it+1}; z_{it+1}\right) - w_{it+1} \right] + \beta \left(1-\delta_i\right) E_t S_{it+1}^{F} \\ &= \beta \sum_{s=0}^{\infty} \beta^s \left(1-\delta_i\right)^s E_t \left[ f\left(n_{it+s+1}; z_{it+s+1}\right) - w_{it+s+1} \right] \end{aligned} \tag{22}$$

If we further assume that profits follow a random walk, then $S_{it}^{F} = \left(f'\left(n_{it}; z_{it}\right) - w_{it}\right) / \left(r + \delta_i\right)$.

### A.3 Condition Efficiency of Equilibrium

We get the equilibrium vacancy-unemployment ratio from equilibrium condition (15) and the definitions of $\alpha_{it}^{WM}$, $\alpha_{it}^{JM}$, $\alpha_{it}^{MT}$and $\alpha^{WD}$,

$$\theta_{it}^{\mathrm{eqm}} = \frac{\exp\left(\alpha_i^{MT}\right)}{\exp\left(\alpha_i^{WD}\right)} \frac{\exp\left(\alpha_i^{WM}\right)}{\exp\left(\alpha_i^{JM}\right)} = \frac{1-\mu}{\mu} \frac{\lambda_t^u - b_{it}}{\lambda_t^v + g'\left(v_{it}; \kappa_{it}\right)} \tag{23}$$

where the multipliers $\lambda_t^u$ and $\lambda_t^v$ are the same as in the efficient allocation, because the planner maximizes household's income plus firm's profits, so that the shadow price of an additional unemployed worker and an additional vacancy for the planner equal the shadowprice for the household and firm, respectively. Comparing equilibrium condition (23) to efficiency condition (1), it is immediate that the equilibrium is efficient.

## B Counterfactual Decompositions

Equation (5) in the main text expresses the relative contribution of mismatch to the aggregate job finding rate in terms of the deviations from the four no-mismatch equilibrium conditions.

$$\log \bar{p}_t^* - \log \bar{p}_t = \tfrac{1}{2}\mu\left(1-\mu\right) V\left[\alpha_{it}^{WM} - \alpha_{it}^{JM} + \alpha_{it}^{MT} - \alpha_{it}^{WD}\right] \tag{24}$$

We use this expression for counterfactual analysis, where we 'shut down' a friction by setting the corresponding $\alpha$ equal to a constant, e.g. to evaluate the job finding rate in the absence of worker mobility frictions we set $\alpha_{it}^{WM} = 0$.

There are two ways to define the contribution of a particular friction to unemployment. First, we can shut down the friction, leaving all other frictions in place, and compare the resulting counterfactual aggregate job finding rate to the actual job finding rate.

$$\Delta \log \bar{p}_t^{WM,1} = \tfrac{1}{2}\mu\left(1-\mu\right)\left(V\left[\alpha_{it}^{WM} - \alpha_{it}^{JM} + \alpha_{it}^{MT} - \alpha_{it}^{WD}\right] - V\left[0 - \alpha_{it}^{JM} + \alpha_{it}^{MT} - \alpha_{it}^{WD}\right]\right) \tag{25}$$

Alternatively, we can shut down all other frictions, leaving only the friction we are considering in place, and compare the resulting counterfactual job finding rate to the job finding rate that would prevail in the absence of all sources of mismatch.

$$\Delta \log \bar{p}_t^{WM,2} = \tfrac{1}{2}\mu\left(1-\mu\right)\left(V\left[\alpha_{it}^{WM}\right] - 0\right) \tag{26}$$

The difference between the two estimators is that $\Delta \log \bar{p}_t^{WM,1}$ includes the covariance terms of $\alpha_i^{WM}$ with the other alphas, whereas $\Delta \log \bar{p}_t^{WM,2}$ does not. The contribution of all frictions adds up to more than the total amount of mismatch by the first estimator, and to less than the total by the second estimator.

By combining both estimators, we can disentangle the direct contribution of a friction from its contribution through its correlation with other frictions and thus design an additive decomposition:

$$\Delta \log \bar{p}_t^{WM,1} = \tfrac{1}{2}\mu\left(1-\mu\right)\left(V\left[\alpha_{it}^{WM}\right] - 2Cov\left[\alpha_{it}^{WM}, \alpha_{it}^{JM}\right] + 2Cov\left[\alpha_{it}^{WM}, \alpha_{it}^{MT}\right] - 2Cov\left[\alpha_{it}^{WM}, \alpha_{it}^{WD}\right]\right) \tag{27}$$

so that

$$\begin{aligned}\Delta \log \bar{p}_t^{WM} &= \tfrac{1}{2}\left(\Delta \log \bar{p}_t^{WM,1} + \Delta \log \bar{p}_t^{WM,2}\right) \\ &= V\left[\alpha_{it}^{WM}\right] - 2Cov\left[\alpha_{it}^{WM}, \alpha_{it}^{JM}\right] + 2Cov\left[\alpha_{it}^{WM}, \alpha_{it}^{MT}\right] - 2Cov\left[\alpha_{it}^{WM}, \alpha_{it}^{WD}\right]\end{aligned} \tag{28}$$

and similarly for the other frictions. Because this estimator includes half of the covariance terms of $\alpha_{it}^{WM}$ with the other alphas, with the remaining half being attributed to the other frictions, it satisfies

$$\Delta \log \bar{p}_t^{WM} + \Delta \log \bar{p}_t^{JM} + \Delta \log \bar{p}_t^{MT} + \Delta \log \bar{p}_t^{WD} = \log \bar{p}_t^{*} - \log \bar{p}_t \tag{29}$$

The contribution of all frictions adds up to overall mismatch.

# C Match Surplus with Time-Varying Payoffs and Turnover

In order to be able to solve forward for match surplus, take a linear approximation of the Bellman equation around $\delta_{it} = \delta_i^*$ and $S_{it} = S_i^*$.

$$\left(1+r\right) S_{it} = y_{it} + E_t\left[\left(1-\delta_{it+1}\right) S_{it+1}\right] \simeq y_{it} + \left(1-\delta_i^*\right) E_t S_{it+1} + E_t\left[\delta_i^* - \delta_{it+1}\right] S_i^* \tag{30}$$

Now, we can solve forward as if the separation probability were constant:

$$\begin{aligned}S_{it} &\simeq \frac{1}{1+r}\left\{y_{it} + E_t\left[\delta_i^* - \delta_{it+1}\right] S_i^*\right\} + \frac{1-\delta_i^*}{1+r} E_t S_{it+1} \\ &= \frac{1}{1+r}\sum_{s=0}^{\infty}\left(\frac{1-\delta_i^*}{1+r}\right)^s E_t\left[y_{it+s} + \left(\delta_i^* - \delta_{it+s+1}\right) S_i^*\right]\end{aligned} \tag{31}$$

From the autoregressive processes for payoffs and separation rates,

$$y_{it+1}^k = \left(1-\rho_y^k\right) y_{it}^k + \rho_y^k \bar{y}_t^k + \varepsilon_{y,it+1}^k \Rightarrow E_t y_{it+s}^k = \bar{y}_t^k + \left(1-\rho_y^k\right)^s \left(y_{it}^k - \bar{y}_t^k\right) \tag{32}$$

$$\delta_{it+1} = \left(1-\rho_\delta\right)\delta_{it} + \rho_\delta \bar{\delta}_t + \varepsilon_{\tau,it+1}^k \Rightarrow E_t \delta_{it+s} = \bar{\delta}_t + \left(1-\rho_\delta\right)^s \left(\delta_{it} - \bar{\delta}_t\right) \tag{33}$$

we get (dropping the $k$ superscripts for simplicity)

$$E_t y_{it+s} = \bar{y}_t + \left(1-\rho_y\right)^s \left(y_{it} - \bar{y}_t\right) \tag{34}$$

$$E_t\left[\delta_i^* - \delta_{it+s+1}\right] = \delta_i^* - \bar{\delta}_t + \left(1-\rho_\delta\right)^{s+1}\left(\bar{\delta}_t - \delta_{it}\right) \tag{35}$$

Substituting into the expression for surplus

$$\begin{aligned}
S_{it} &\simeq \frac{1}{1+r}\sum_{s=0}^{\infty}\left(\frac{1-\delta_i^*}{1+r}\right)^s \left\{\bar{y}_t + \left(1-\rho_y\right)^s \left(y_{it}-\bar{y}_t\right) + \left(\delta_i^* - \bar{\delta}_t\right) S_i^* + \left(1-\rho_\delta\right)^{s+1}\left(\bar{\tau}_t - \tau_{it}\right) S_i^*\right\} \\
&= \frac{1}{1+r}\sum_{s=0}^{\infty}\left(\frac{1-\delta_i^*}{1+r}\right)^s \left\{\bar{y}_t + \left(\delta_i^* - \bar{\delta}_t\right) S_i^*\right\} + \frac{1}{1+r}\sum_{s=0}^{\infty}\left(\frac{\left(1-\delta_i^*\right)\left(1-\rho_y\right)}{1+r}\right)^s \left(y_{it}-\bar{y}_t\right) \\
&+ \frac{1-\rho_\delta}{1+r}\sum_{s=0}^{\infty}\left(\frac{\left(1-\delta_i^*\right)\left(1-\rho_\delta\right)}{1+r}\right)^s \left(\bar{\delta}_t - \delta_{it}\right) S_i^* \\
&= \frac{\bar{y}_t + \left(\delta_i^* - \bar{\delta}_t\right) S_i^*}{r+\delta_i^*} + \frac{y_{it}-\bar{y}_t}{r+\delta_i^* + \rho_y - \rho_y \delta_i^*} + \frac{\left(1-\rho_\tau\right)\left(\bar{\delta}_t - \delta_{it}\right) S_i^*}{r+\delta_i^* + \rho_\delta - \rho_\delta \delta_i^*} \\
&\simeq \frac{\bar{y}_t + \left(\delta_i^* - \bar{\delta}_t\right) S_i^*}{r+\delta_i^*} + \frac{y_{it}-\bar{y}_t}{r+\delta_i^* + \rho_y} + \frac{\left(1-\rho_\delta\right)\left(\bar{\delta}_t - \delta_{it}\right) S_i^*}{r+\delta_i^* + \rho_\delta}
\end{aligned} \tag{36}$$

Finally, setting $\delta_i^* = \delta_{it}$ and $S_i^* = S_{it}$ and rearranging we get the expression in the main text.

$$S_{it} \simeq \frac{\left(r+\delta_{it}\right)\left(r+\delta_{it}+\rho_\delta\right)}{\left(r+\delta_{it}\right)\left(r+\delta_{it}+\rho_\delta\right) + \rho_\delta\left(1+r+\delta_{it}\right)\left(\bar{\delta}_t - \delta_{it}\right)}\left(\frac{\bar{y}_t}{r+\delta_{it}} + \frac{y_{it}-\bar{y}_t}{r+\delta_{it}+\rho_y}\right) \tag{37}$$

# D Disaggregation and the Level of Mismatch

Unemployment due to mismatch across states and 2-digit industries is about an order of magnitude smaller than unemployment due to mismatch across 3-digit occupations. As mentioned in the main text, we believe this is because states and 2-digit industries are not sufficiently disaggregated, and most of the mismatch is within a state or an industry. In this appendix, we provide some suggestive evidence for this claim.

We address the aggregation issue in two ways. First, we disaggregate further. For the purposes of this appendix only, we use data that are disaggregated by both state and industry. Instead of 50 states or 33 industries, this gives us $50 * 33 = 1650$ labor market segments. Although 1650 submarkets is probably a more realistic segmentation of the US labor market, it is in all likelihood still to coarse. Therefore, the second part of our solution is to find a correction factor that relates the observed amount of mismatch in our data to the amount of mismatch we would observe if we were to disaggregate to the right level.

Disaggregation by both states and industries, while alleviating the aggregation problem, gives rise to a different bias because of sampling error. Barnichon and Figura (2015) use a very large dataset consisting of the universe of job seekers in the UK.[34] The US data, however, are survey-based and in our dataset we have only about $23,000$ unemployed workers per year, which means that the 1650 labor market segments on average contain only 14 observations and because not all states and industries are equally large, some cells are even much smaller than that. As a result, our estimates for the job finding rate in each segment will be very imprecise. This sampling error will translate into dispersion across segments and bias our estimate for the amount of mismatch unemployment. We address this issue by estimating the variance of the sampling error in each segment and correcting the estimated variance of the job finding rates by subtracting the average variance of the sampling error.[35]

## D.1 Correction factor

An ideal labor market segment would consist of very similar jobs within a geographic area that allows workers to commute to these jobs without moving house. Using UK data, Barnichon and Figura (2015) estimate the correct level of disaggregation would be to use 232 so-called travel-to-work areas and 353 detailed occupational groups. They then aggregate these data to a level that is comparable to US states and major occupational categories and find that the observed amount of mismatch decreases by a factor 6. We argue that a similar correction factor is appropriate for our estimate of mismatch across 1650 state-industry segments.

From equation (5), we know that mismatch is approximately proportional to the variance of log vacancy-unemployment ratios, $V\left[\hat{\theta}_{it}\right]$. Barnichon and Figura show that

$$\ln\left(V_n\left[\hat{\theta}_i\right]\right) \simeq \ln a_0 + a_{geo}\ln n_{geo} + a_{occ}\ln n_{occ} \tag{38}$$

where $V_n$ is the variance of $\hat{\theta}_i$ based on a higher level of aggregation and $n = N/N^{CF}$ is the ratio of the observed versus the correct number of labor market segments. They also estimate the parameters of this relation using UK data to and find $a_{geo} = 0.13$ and

[34] The exercise we describe here is not present in the published article, but may be found in the December 2011 working paper version, available at https://www.bde.es/f/webbde/GAP/Secciones/SalaPrensa/Agenda/Eventos/12/May/barnichon_figura.pdf.

[35] Workers in each segment find a job with probability $f_i^W$. The variance of the realization of this Bernoulli process equals $f_i^W\left(1-f_i^W\right)$, so that the variance of the observed mean probability is equal to $f_i^W\left(1-f_i^W\right)/N_i$, where $N_i$ is the number of observations in segment $i$. The variance of the signal in $f_i^W$ across segments, by the ANOVA formula, is then given by the observed variance $var\left(f_i^W\right)$ minus the average variance of the sampling error $E\left[f_i^W\left(1-f_i^W\right)/N_i\right]$. We do not use segments with less than 5 observations because these would contribute more noise than signal.

$a_{occ} = 0.67$. This implies

$$\ln\left(\frac{V\left[\hat{\theta}_i^{CF}\right]}{V\left[\hat{\theta}_i\right]}\right) = a_{geo}\ln\left(\frac{1}{n_{geo}}\right) + a_{occ}\ln\left(\frac{1}{n_{occ}}\right) \tag{39}$$

because by assumption $\hat{\theta}_i^{CF}$ are the finding rates for the right level of disaggregation so that $n_{geo}^{CF} = n_{occ}^{CF} = 1$.

In the UK data that Barnichon and Figura use, the correct number of geographic areas is about 232 (travel to work areas). The US population is larger than the UK population, but the land area is larger as well. Therefore, Barnichon and Figura assume the number of geographic units is the same in the same in the two countries. Since we work with 50 states, $1/n_{geo} = 232/50 = 4.64$. The same UK data have 353 detailed occupational groups, which should be similar in the US. We use 33 broad industries. Assuming these broad industry categories are comparable to broad occupations categories, we get $1/n_{occ} = 353/33 = 10.7$. This implies a correction factor for the variance of labor market tightness of,

$$\frac{V\left[\hat{\theta}_i^{CF}\right]}{V\left[\hat{\theta}_i^{geo*ind}\right]} = \exp\left(0.13\ln\left(4.64\right) + 0.67\ln\left(10.7\right)\right) = 6.0 \tag{40}$$

which is the same correction factor that Barnichon and Figura used.

## D.2 Results

Mismatch across state*industry segments contributes 15% to unemployment, comparable to mismatch across occupation-state segments and substantially more than mismatch across states or industries only. The bias because of sampling error is fairly small, bringing the contribution of mismatch down to 14%, indicating the dispersion in job finding rates across segments is large compared to the sampling error. After muliplying by 6 to correct for aggregation, these estimates suggest that mismatch is responsible for 84% of unemployment. It is important to note that a good amount of guesswork was needed for the aggregation correction and the estimate is therefore rather imprecise. Nevertheless, these estimates indicate that it is quite possible that mismatch across states and industries is of the same order of magnitude as mismatch across detailed occupations, and that mismatch is an important contributor to unemployment.

# E Additional Tables

Table 2A
State-level data, cell sizes 1979-2015

| | | job finding rate | | | wage | | |
|---|---|---|---|---|---|---|---|
| | | min | mean | max | min | mean | max |
| Alabama | AL | 292 | 630 | 12656 | 1484 | 2020 | 2808 |
| Alaska | AK | 409 | 780 | 11703 | 1600 | 2005 | 2400 |
| Arizona | AZ | 323 | 528 | 12178 | 1584 | 1943 | 2650 |
| Arkansas | AR | 322 | 543 | 11940 | 1480 | 1867 | 2433 |
| California | CA | 2079 | 4069 | 67582 | 8996 | 12626 | 14951 |
| Colorado | CO | 296 | 701 | 16924 | 1776 | 2777 | 3637 |
| Connecticut | CT | 159 | 602 | 16889 | 1590 | 2514 | 3781 |
| Delaware | DE | 210 | 441 | 11684 | 1078 | 1974 | 2706 |
| District of Columbia | DC | 196 | 549 | 10318 | 671 | 1542 | 2380 |
| Florida | FL | 960 | 1663 | 35367 | 4633 | 6318 | 8201 |
| Georgia | GA | 374 | 721 | 16769 | 1999 | 2780 | 3846 |
| Hawaii | HI | 179 | 390 | 10919 | 1301 | 1862 | 2375 |
| Idaho | ID | 252 | 569 | 11443 | 1522 | 1939 | 2321 |
| Illinois | IL | 1002 | 1839 | 33016 | 4506 | 6317 | 7992 |
| Indiana | IN | 256 | 737 | 17025 | 2100 | 2626 | 3853 |
| Iowa | IA | 218 | 576 | 15461 | 2011 | 2651 | 3438 |
| Kansas | KS | 266 | 497 | 13196 | 1950 | 2287 | 2922 |
| Kentucky | KY | 338 | 632 | 13105 | 1785 | 2070 | 2779 |
| Louisiana | LA | 259 | 571 | 13233 | 1310 | 1758 | 2983 |
| Maine | ME | 254 | 593 | 14862 | 1418 | 2176 | 3200 |
| Maryland | MD | 239 | 613 | 16907 | 1647 | 2608 | 3839 |
| Massachusetts | MA | 400 | 1028 | 33071 | 2330 | 4532 | 7689 |
| Michigan | MI | 707 | 1831 | 32804 | 3258 | 5498 | 7984 |
| Minnesota | MN | 274 | 695 | 19757 | 2001 | 3076 | 4367 |
| Mississippi | MS | 297 | 584 | 12266 | 1159 | 1773 | 2633 |
| Missouri | MO | 253 | 674 | 14410 | 1870 | 2449 | 3000 |
| Montana | MT | 265 | 546 | 12356 | 1265 | 1816 | 2518 |
| Nebraska | NE | 193 | 393 | 13611 | 1483 | 2396 | 2977 |
| Nevada | NV | 359 | 659 | 15305 | 1608 | 2244 | 3456 |
| New Hampshire | NH | 177 | 472 | 17533 | 1429 | 2416 | 3875 |
| New Jersey | NJ | 613 | 1291 | 33084 | 3055 | 5123 | 7987 |
| New Mexico | NM | 249 | 514 | 11069 | 1050 | 1662 | 2330 |
| New York | NY | 1282 | 2401 | 53144 | 5950 | 8921 | 12941 |
| North Carolina | NC | 488 | 1052 | 35075 | 2759 | 4638 | 8276 |
| North Dakota | ND | 240 | 400 | 13214 | 1644 | 2154 | 2466 |
| Ohio | OH | 767 | 1740 | 35137 | 3964 | 6143 | 8497 |
| Oklahoma | OK | 214 | 486 | 12383 | 1414 | 1971 | 2538 |
| Oregon | OR | 398 | 713 | 13963 | 1539 | 2010 | 2933 |
| Pennsylvania | PA | 888 | 1700 | 33274 | 4347 | 6327 | 8188 |
| Rhode Island | RI | 219 | 610 | 13968 | 1142 | 2092 | 3170 |
| South Carolina | SC | 238 | 564 | 11190 | 1607 | 2017 | 2730 |
| South Dakota | SD | 233 | 440 | 13238 | 1733 | 2346 | 2825 |
| Tennessee | TN | 324 | 603 | 11655 | 1912 | 2121 | 2523 |
| Texas | TX | 1316 | 2090 | 38910 | 6873 | 7870 | 8576 |
| Utah | UT | 249 | 480 | 14645 | 1798 | 2157 | 3158 |
| Vermont | VT | 184 | 424 | 12303 | 1289 | 1931 | 2611 |
| Virginia | VA | 216 | 598 | 15987 | 2338 | 2931 | 3631 |
| Washington | WA | 363 | 746 | 13372 | 1668 | 2337 | 2944 |
| West Virginia | WV | 320 | 648 | 11352 | 1368 | 1784 | 2495 |
| Wisconsin | WI | 288 | 723 | 16871 | 2291 | 2952 | 3773 |
| Wyoming | WY | 231 | 442 | 11638 | 1304 | 1867 | 2449 |

Entries in the table are the number of observations used to calculate the job finding rate and the average wage in a state-year cell.

Table 2B
Industry-level data (SIC), cell sizes 1979-1997

| | | job finding rate | | | | wage | |
|---|---|---|---|---|---|---|---|
| | | min | mean | max | min | mean | max |
| Mining | MIN | 179 | 642 | 1770 | 1053 | 1698 | 2896 |
| Construction | CON | 3721 | 6106 | 9114 | 8410 | 9342 | 10647 |
| Lumber & wood prods, ex furniture | LUM | 282 | 550 | 1058 | 1068 | 1218 | 1520 |
| Furniture & fixtures | FUR | 157 | 327 | 576 | 786 | 998 | 1231 |
| Stone, clay, concrete, glass prods | MNR | 126 | 321 | 614 | 766 | 995 | 1317 |
| Primary metals | PMT | 140 | 521 | 1566 | 985 | 1461 | 2353 |
| Fabricated metals | FMT | 223 | 754 | 1639 | 1693 | 2226 | 3334 |
| Machinery, ex electrical | MAC | 370 | 1005 | 2350 | 3237 | 4264 | 5682 |
| Electrical machinery, equip supplies | ELC | 292 | 890 | 1789 | 2482 | 3527 | 4735 |
| Motor vehicles & equip | MVH | 241 | 699 | 1789 | 1434 | 1952 | 2215 |
| Other transportation equip | OVH | 129 | 432 | 842 | 1316 | 1965 | 2333 |
| Professional & photo equip, watches | PHO | 88 | 219 | 397 | 969 | 1158 | 1397 |
| Misc mfg industries | MMA | 227 | 395 | 700 | 807 | 917 | 1092 |
| Food & kindred prods | FOO | 662 | 1173 | 1874 | 2401 | 3094 | 3960 |
| Textile mill prods | TEX | 133 | 393 | 751 | 779 | 1258 | 1581 |
| Apparel & other finished textil prods | APP | 447 | 870 | 1398 | 1199 | 1853 | 2505 |
| Paper & allied prods | PAP | 96 | 237 | 426 | 943 | 1257 | 1528 |
| Printing, publishing & allied inds | PUB | 372 | 605 | 830 | 2346 | 2831 | 3186 |
| Chemicals & allied prods | CHE | 178 | 381 | 645 | 1801 | 2251 | 2734 |
| Petroleum & coal prods | OIL | 15 | 56 | 106 | 237 | 336 | 482 |
| Rubber & misc plastic prods | RUB | 188 | 389 | 699 | 1123 | 1277 | 1412 |
| Leather & leather prods | LEA | 51 | 195 | 474 | 176 | 368 | 741 |
| Transportation | TRA | 1172 | 1811 | 2688 | 6460 | 7673 | 8682 |
| Communications | COM | 243 | 323 | 437 | 2235 | 2767 | 3327 |
| Utilities & sanitary services | UTI | 155 | 298 | 507 | 2135 | 2724 | 3122 |
| Wholesale trade | WHO | 943 | 1532 | 2322 | 5982 | 6757 | 7388 |
| Retail trade | RET | 7763 | 10259 | 13961 | 26903 | 29533 | 32618 |
| Banking & other finance | FIN | 421 | 688 | 927 | 4550 | 5569 | 6394 |
| Insurance & real estate | INS | 722 | 1014 | 1436 | 5276 | 5855 | 6814 |
| Business services | BSV | 1157 | 2290 | 3101 | 3400 | 5835 | 7560 |
| Automobile & repair services | ASV | 581 | 871 | 1281 | 1692 | 2128 | 2481 |
| Personal serv ex private hhs | PSV | 1031 | 1674 | 2360 | 3488 | 4274 | 5036 |
| Entertainment & recreation | ENT | 726 | 1051 | 1391 | 1658 | 2285 | 2995 |
| Health services | HEA | 1669 | 2274 | 3129 | 12566 | 15434 | 17922 |
| Educational services | EDU | 1243 | 1838 | 2855 | 14875 | 16391 | 18584 |
| Social services | SOC | 584 | 858 | 1072 | 2535 | 3389 | 4327 |
| Misc professional services | MSV | 644 | 954 | 1410 | 3697 | 5976 | 7755 |

Entries in the table are the number of observations used to calculate the job finding rate and the average wage in an industry-year cell. Industries are defined according to the 2-digit Standard Industrial Classification (SIC).

Table 2C
Industry-level data (NAICS), cell sizes 1998-2015

| | | job finding rate | | | wage | | |
|---|---|---|---|---|---|---|---|
| | | min | mean | max | min | mean | max |
| Mining | MIN | 124 | 272 | 563 | 908 | 1206 | 1612 |
| Construction | CON | 2977 | 5179 | 9942 | 8249 | 9798 | 11839 |
| Nonmetallic mineral product manufacturing | MNR | 81 | 158 | 328 | 454 | 621 | 788 |
| Primary metals and fabricated metal products | MET | 322 | 553 | 1081 | 1790 | 2248 | 2726 |
| Machinery manufacturing | MAC | 203 | 396 | 783 | 1361 | 1856 | 2759 |
| Computer and electronic product manufacturing | CEM | 144 | 377 | 763 | 1085 | 1633 | 2145 |
| Electrical equipment, appliance manufacturing | ELC | 86 | 206 | 622 | 427 | 948 | 2092 |
| Transportation equipment manufacturing | VEH | 340 | 606 | 1457 | 2297 | 2715 | 3137 |
| Wood products | LUM | 56 | 196 | 357 | 249 | 637 | 960 |
| Furniture and fixtures manufacturing | FUR | 102 | 196 | 425 | 418 | 656 | 821 |
| Miscellaneous and not specified manufacturing | MMA | 236 | 439 | 784 | 1354 | 1525 | 1650 |
| Food manufacturing, Beverage, and tobacco products | FOO | 433 | 666 | 980 | 2376 | 2535 | 2778 |
| Textile, apparel, and leather manufacturing | TEX | 192 | 374 | 554 | 563 | 991 | 1962 |
| Paper and printing | PAP | 167 | 347 | 565 | 958 | 1604 | 2491 |
| Petroleum and coal products manufacturing | OIL | 22 | 40 | 74 | 186 | 224 | 265 |
| Chemical manufacturing | CHE | 136 | 267 | 579 | 1447 | 1589 | 1804 |
| Plastics and rubber products | RUB | 92 | 204 | 334 | 543 | 824 | 1221 |
| Wholesale trade | WHO | 595 | 977 | 1395 | 3461 | 5071 | 6565 |
| Retail trade | RET | 3653 | 5440 | 8383 | 16701 | 19049 | 20716 |
| Transportation and warehousing | TRA | 980 | 1568 | 2481 | 6516 | 7233 | 8248 |
| Utilities | UTI | 98 | 166 | 245 | 1555 | 1787 | 2221 |
| Publishing industries (except internet) | PUB | 80 | 182 | 389 | 599 | 853 | 1154 |
| Broadcasting and Telecommunications | COM | 235 | 476 | 801 | 1793 | 2296 | 2828 |
| Information and data processing services | INF | 52 | 201 | 907 | 381 | 1115 | 3040 |
| Finance | FIN | 448 | 821 | 1462 | 4891 | 5507 | 6104 |
| Insurance | INS | 227 | 394 | 745 | 2806 | 3102 | 3405 |
| Real estate | RES | 289 | 483 | 823 | 1955 | 2192 | 2494 |
| Rental and leasing services | REN | 28 | 126 | 298 | 204 | 434 | 651 |
| Professional and technical services | PSV | 1133 | 1881 | 3082 | 7180 | 8934 | 9919 |
| Administrative and support services | ASV | 1275 | 2964 | 5016 | 2705 | 4742 | 5769 |
| Educational services | EDU | 1113 | 2021 | 2952 | 14799 | 16710 | 17871 |
| Hospitals | HOS | 420 | 745 | 1282 | 7842 | 8930 | 12006 |
| Health care services, except hospitals | HEA | 510 | 1685 | 3045 | 5504 | 9676 | 11607 |
| Social assistance | SOC | 583 | 971 | 1561 | 2863 | 3294 | 3646 |
| Arts, entertainment, and recreation | ENT | 901 | 1254 | 1756 | 2991 | 3176 | 3688 |
| Accommodation | ACC | 501 | 740 | 1098 | 1875 | 2198 | 2590 |
| Food services and drinking places | FSV | 2647 | 3942 | 5614 | 8317 | 9660 | 10118 |
| Other services (excl. government) | MSV | 1330 | 1867 | 2735 | 6271 | 6969 | 7553 |

Entries in the table are the number of observations used to calculate the job finding rate and the average wage in an industry-year cell. Industries are defined according to the 2-digit North American Industrial Classification System (NAICS).