Accounting for Mismatch Unemployment

Benedikt Herz and Thijs van Rens

May 2015

Appendices (for online publication)

A Effect of Mismatch on Unemployment

Since we are considering a mean-preserving change in the distribution of labor market tightness, we know that $\bar{\theta} = \bar{\theta}^{CF}$. Then, with $f_i^W = B\theta_i^{1-\mu} \Leftrightarrow \theta_i = \left(f_i^W/B\right)^{\frac{1}{1-\mu}}$, we get

$$\bar{\theta} = E\left[\left(\frac{f_i^W}{B}\right)^{\frac{1}{1-\mu}}\right] = E\left[\left(\frac{f_i^{W,CF}}{B}\right)^{\frac{1}{1-\mu}}\right] = \bar{\theta}^{CF} \tag{1}$$

The unemployment rate follows from the aggregate job finding rate by assuming steady state, i.e. $u = \frac{\lambda}{\lambda + \bar{f}^W}$, where $\lambda << \bar{f}^W$ is the aggregate separation rate, which implies $u^{CF}/u = \left(\bar{f}^W + \lambda\right)/\left(\bar{f}^{W,CF} + \lambda\right) \simeq \bar{f}^W/\bar{f}^{W,CF}$. Substituting $\hat{f}_i^W = \left(f_i^W - \bar{f}^W\right)/\bar{f}^W \Leftrightarrow f_i^W = \bar{f}^W\left(1 + \hat{f}_i^W\right)$ and re-arranging gives

$$\frac{u^{CF}}{u} \simeq \frac{\bar{f}^W}{\bar{f}^{W,CF}} = \left(\frac{E\left[\left(1 + \hat{f}_i^{W,CF}\right)^{\frac{1}{1-\mu}}\right]}{E\left[\left(1 + \hat{f}_i^{W}\right)^{\frac{1}{1-\mu}}\right]}\right)^{1-\mu} \propto \frac{V\left[\theta_i^{CF}/\bar{\theta}^{CF}\right]}{V\left[\theta_i/\bar{\theta}\right]} \tag{2}$$

To show that u^{CF}/u is proportional to the ratio of the variances $\theta_i^{CF}/\bar{\theta}^{CF}$ and $\theta_i/\bar{\theta}$, take logs and assume $1 + \hat{f}_i^W$ is log-normally distributed to get

$$\begin{aligned}
\ln\left(\frac{u^{CF}}{u}\right) &\simeq (1-\mu)\left(\ln E\left[\left(1 + \hat{f}_i^{W,CF}\right)^{\frac{1}{1-\mu}}\right] - \ln E\left[\left(1 + \hat{f}_i^{W}\right)^{\frac{1}{1-\mu}}\right]\right) \\
&= (1-\mu)\left(\begin{array}{c} E\left[\ln\left(1 + \hat{f}_i^{W,CF}\right)^{\frac{1}{1-\mu}}\right] + \frac{1}{2}V\left[\ln\left(1 + \hat{f}_i^{W,CF}\right)^{\frac{1}{1-\mu}}\right] \\ -E\left[\ln\left(1 + \hat{f}_i^{W}\right)^{\frac{1}{1-\mu}}\right] - \frac{1}{2}V\left[\ln\left(1 + \hat{f}_i^{W}\right)^{\frac{1}{1-\mu}}\right] \end{array}\right) \\
&= E\left[\ln\left(1 + \hat{f}_i^{W,CF}\right)\right] + \tfrac{1}{2}\frac{1}{1-\mu}V\left[\ln\left(1 + \hat{f}_i^{W,CF}\right)\right] \\
&- E\left[\ln\left(1 + \hat{f}_i^{W}\right)\right] - \tfrac{1}{2}\frac{1}{1-\mu}V\left[\ln\left(1 + \hat{f}_i^{W}\right)\right] \\
&= \tfrac{1}{2}\frac{1}{1-\mu}\left(V\left[\ln\left(1 + \hat{f}_i^{W,CF}\right)\right] - V\left[\ln\left(1 + \hat{f}_i^{W}\right)\right]\right)
\end{aligned} \tag{3}$$

where we used that the $E\hat{f}_i^W = 0$. Using $\hat{f}_i^W = \left(f_i^W - \bar{f}^W\right)/\bar{f}^W$ and $f_i^W = \theta_i^{1-\mu}$ we get

$$\begin{aligned}
\ln\left(\frac{u^{CF}}{u}\right) &\simeq \tfrac{1}{2}\frac{1}{1-\mu}\left(V\left[\ln\left(1+\hat{f}_i^{W,CF}\right)\right] - V\left[\ln\left(1+\hat{f}_i^{W}\right)\right]\right) \\
&= \tfrac{1}{2}\frac{1}{1-\mu}\left(V\left[\ln\left(\frac{f_i^{W,CF}}{\bar{f}^{W,CF}}\right)\right] - V\left[\ln\left(\frac{f_i^{W}}{\bar{f}^{W}}\right)\right]\right) \\
&= \tfrac{1}{2}\frac{1}{1-\mu}\left(V\left[\ln f_i^{W,CF}\right] - V\left[\ln f_i^{W}\right]\right) \\
&= \tfrac{1}{2}\left(1-\mu\right)\left(V\left[\ln\theta_i^{CF}\right] - V\left[\ln\theta_i\right]\right) \\
&\simeq \tfrac{1}{2}\left(1-\mu\right)\left(V\left[\theta_i^{CF}/\bar{\theta}^{CF}\right] - V\left[\theta_i/\bar{\theta}\right]\right) \qquad (4)
\end{aligned}$$

where the last equality is simply a first order Taylor approximation saying that $\ln\theta_i \simeq \left(\theta_i - \bar{\theta}\right)/\bar{\theta}$ where $\bar{\theta}$ is the mean of θ_i. Then,

$$\begin{aligned}
\frac{u^{CF}}{u} &\simeq \exp\left(\tfrac{1}{2}\left(1-\mu\right)\left(V\left[\theta_i^{CF}/\bar{\theta}^{CF}\right] - V\left[\theta_i/\bar{\theta}\right]\right)\right) \\
&= \exp\left(\tfrac{1}{2}\left(1-\mu\right)V\left[\theta_i/\bar{\theta}\right]\left(\frac{V\left[\theta_i^{CF}/\bar{\theta}^{CF}\right]}{V\left[\theta_i/\bar{\theta}\right]} - 1\right)\right) \\
&\simeq \exp\left(\tfrac{1}{2}\left(1-\mu\right)V\left[\theta_i/\bar{\theta}\right]\right)\frac{V\left[\theta_i^{CF}/\bar{\theta}^{CF}\right]}{V\left[\theta_i/\bar{\theta}\right]} \qquad (5)
\end{aligned}$$

B Counterfactual Decompositions

Let $u_{XX=YY=0}$ denote the unemployment rate that prevails if we set $\alpha_i^{XX} = \alpha_i^{YY} = 0$. Then, there are two ways to define the contribution of a particular frictions to unemployment.

$$contrib_1^{WM} = u - u_{WM=0} \qquad (6)$$

$$contrib_2^{WM} = u_{JM=WD=0} - u_{WM=JM=WD=0} \qquad (7)$$

By using both estimators, we can disentangle the direct contribution of a friction from its contribution through its correlation with other frictions and thus design a decomposition that is (approximately) additive.

From equation (2) in appendix A, taking a second order Taylor approximation around $\hat{f}_i^W = 0$, we get that

$$u_{XX=0} - u_{XX=YY=0} \simeq \kappa u\left(V\left[\hat{f}_i^W|\hat{\alpha}_i^{XX} = 0\right] - V\left[\hat{f}_i^W|\hat{\alpha}_i^{XX} = \hat{\alpha}_i^{YY} = 0\right]\right) \qquad (8)$$

where κ is some constant of proportionality, which is not of interest here, and u is the actual unemployment rate. Using this approximation, the estimators can be written as

$$\begin{aligned} contrib_1^{WM} &\simeq \kappa (1-\mu)^2 u \left(V\left[\alpha_i^{WM} - \alpha_i^{JM} - \alpha_i^{WD}\right] - V\left[-\alpha_i^{JM} - \alpha_i^{WD}\right]\right) \\ &= \kappa (1-\mu)^2 u \left(V\left[\alpha_i^{WM}\right] - 2Cov\left[\alpha_i^{WM}, \alpha_i^{JM}\right] - 2Cov\left[\alpha_i^{WM}, \alpha_i^{WD}\right]\right) \quad (9) \\ contrib_2^{WM} &\simeq \kappa (1-\mu)^2 u \left(V\left[\alpha_i^{WM}\right] - 0\right) = \kappa (1-\mu)^2 u V\left[\alpha_i^{WM}\right] \quad (10) \end{aligned}$$

So, the difference between the two estimators is that $contrib_1^{WM}$ includes the covariance terms involving α_i^{WM}, whereas $contrib_2^{WM}$ does not.

To get an (approximately) additive decomposition, we use

$$contrib\_add^{WM} = \tfrac{1}{2}\left(contrib_1^{WM} + contrib_2^{WM}\right) \quad (11)$$

Because this estimator includes half of the covariance terms and the other half will be attributed to the other frictions, it satisfies

$$contrib\_add^{WM} + contrib\_add^{JM} + contrib\_add^{WD} \simeq contrib^{MMtotal} \quad (12)$$

Figures 6 and 7 show that this approximation is good in the actual data.

C Match Surplus

C.1 Match Surplus in the DMP model

The value of an employed worker in submarket i, W_{it}, and the value of an unemployed worker in that submarket, U_{it}^W, satisfy the following set of Bellman equations,

$$(1+r) W_{it} = w_{it} + \lambda_{it} E_t U_{it+1}^W + (1-\lambda_{it}) E_t W_{it+1} \quad (13)$$

$$(1+r) U_{it}^W = b_{it} + f_{it}^W E_t W_{it+1} + \left(1 - f_{it}^W\right) E_t U_{it+1}^W \quad (14)$$

where λ_{it} is the separation rate, f_{it}^W is the job finding rate, w_{it} is the wage and b_{it} is the flow value of being unemployed, which consists of unemployment benefits and the value of leisure. Worker surplus equals the difference between the payoff from having a job in submarket i minus the payoff of looking for a job in that submarket, $S_{it}^W = W_{it} - U_{it}^W$, so that

$$(1+r) S_{it}^W = w_{it} - b_{it} + \left(1 - \lambda_{it} - f_{it}^W\right) E_t S_{it+1}^W \quad (15)$$

where $w_{it} - b_{it}$ is the worker's flow payoff from having a job net of the payoff from being unemployed, and $\lambda_{it} + f_{it}^W$ is worker turnover.

The value of a filled job in submarket i, J_{it}, and the value of a vacancy in that submarket, U_{it}^J, satisfy the following set of Bellman equations,

$$(1+r) J_{it} = \pi_{it} + \lambda_{it} E_t U_{it+1}^J + (1-\lambda_{it}) E_t J_{it+1} \quad (16)$$

$$(1+r)\,U_{it}^J = -k_{it} + f_{it}^F E_t J_{it+1} + \left(1 - f_{it}^F\right) E_t U_{it+1}^J \tag{17}$$

where f_{it}^F is the worker finding rate, π_{it} are flow profits and k_{it} are vacancy posting costs. Job surplus equals the difference between the payoff from having a filled job in submarket i minus the payoff of having a vacancy in that submarket, $S_{it}^J = J_{it} - U_{it}^J$, so that

$$(1+r)\,S_{it}^J = \pi_{it} + k_{it} + \left(1 - \lambda_{it} - f_{it}^F\right) E_t S_{it+1}^J \tag{18}$$

where $\pi_{it} + k_{it}$ is the firm's flow payoff from having a filled job gross of vacancy posting costs, and $\lambda_{it} + f_{it}^F$ is job turnover.

C.2 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 $\tau_{it} = \tau_i^*$ and $S_{it} = S_i^*$.

$$(1+r)\,S_{it} = y_{it} + E_t\left[(1-\tau_{it+1})\,S_{it+1}\right] \simeq y_{it} + (1-\tau_i^*)\,E_t S_{it+1} + E_t\left[\tau_i^* - \tau_{it+1}\right] S_i^* \tag{19}$$

Now, we can solve forward as if turnover were constant:

$$\begin{aligned} S_{it} &\simeq \frac{1}{1+r}\left\{y_{it} + E_t\left[\tau_i^* - \tau_{it+1}\right] S_i^*\right\} + \frac{1-\tau_i^*}{1+r} E_t S_{it+1} \\ &= \frac{1}{1+r}\sum_{s=0}^{\infty}\left(\frac{1-\tau_i^*}{1+r}\right)^s E_t\left[y_{it+s} + (\tau_i^* - \tau_{it+s+1})\,S_i^*\right] \end{aligned} \tag{20}$$

From the autoregressive processes (which do not need to be independent because of the linearity)

$$E_t y_{it+s} = \bar{y}_t + (1-\delta_y)^s\,(y_{it} - \bar{y}_t) \tag{21}$$

$$E_t\left[\tau_i^* - \tau_{it+s+1}\right] = \tau_i^* - \bar{\tau}_t + (1-\delta_\tau)^{s+1}\,(\bar{\tau}_t - \tau_{it}) \tag{22}$$

Substituting into the expression for surplus

$$\begin{aligned} S_{it} &\simeq \frac{1}{1+r}\sum_{s=0}^{\infty}\left(\frac{1-\tau_i^*}{1+r}\right)^s \left\{\bar{y}_t + (1-\delta_y)^s\,(y_{it}-\bar{y}_t) + (\tau_i^* - \bar{\tau}_t)\,S_i^* + (1-\delta_\tau)^{s+1}\,(\bar{\tau}_t - \tau_{it})\,S_i^*\right\} \\ &= \frac{1}{1+r}\sum_{s=0}^{\infty}\left(\frac{1-\tau_i^*}{1+r}\right)^s \left\{\bar{y}_t + (\tau_i^* - \bar{\tau}_t)\,S_i^*\right\} + \frac{1}{1+r}\sum_{s=0}^{\infty}\left(\frac{(1-\tau_i^*)\,(1-\delta_y)}{1+r}\right)^s (y_{it}-\bar{y}_t) \\ &+ \frac{1-\delta_\tau}{1+r}\sum_{s=0}^{\infty}\left(\frac{(1-\tau_i^*)\,(1-\delta_\tau)}{1+r}\right)^s (\bar{\tau}_t - \tau_{it})\,S_i^* \\ &= \frac{\bar{y}_t + (\tau_i^* - \bar{\tau}_t)\,S_i^*}{r+\tau_i^*} + \frac{y_{it}-\bar{y}_t}{r+\tau_i^*+\delta_y-\delta_y\tau_i^*} + \frac{(1-\delta_\tau)\,(\bar{\tau}_t-\tau_{it})\,S_i^*}{r+\tau_i^*+\delta_\tau-\delta_\tau\tau_i^*} \\ &\simeq \frac{\bar{y}_t + (\tau_i^* - \bar{\tau}_t)\,S_i^*}{r+\tau_i^*} + \frac{y_{it}-\bar{y}_t}{r+\tau_i^*+\delta_y} + \frac{(1-\delta_\tau)\,(\bar{\tau}_t-\tau_{it})\,S_i^*}{r+\tau_i^*+\delta_\tau} \end{aligned} \tag{23}$$

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

$$S_{it} \simeq \frac{(r+\tau_{it})(r+\tau_{it}+\delta_\tau)}{(r+\tau_{it})(r+\tau_{it}+\delta_\tau)+\delta_\tau(1+r+\tau_{it})(\bar{\tau}_t-\tau_{it})}\left(\frac{\bar{y}_t}{r+\tau_{it}}+\frac{y_{it}-\bar{y}_t}{r+\tau_{it}+\delta_y}\right) \quad (24)$$

D Heterogeneity

D.1 Observable Worker Heterogeneity

We implement this approach in two steps. First, we regress the variable of interest on observable worker characteristics using a flexible specification. The variable of interest is either the wage, or an dummy variable indicating whether a worker lost or found a job. Second, we calculate fitted values for 40 worker cells, defined based on 2 gender, 5 education groups (less than high school, high school graduate, some college, college graduate, or more than college), and 4 categories for potential labor market experience (0-10 years, 11-20 years, 21-30 years, 31-40 years after completion of schooling), and calculate worker and job surplus and job finding rates for the average worker in each of these 40 cells.

The reasons for the first step are threefold. First, it allows us to control for observable characteristics, race and marital status, which are not used to define worker cells because doing so would result in too few observations per cell. When we calculate fitted values, we set these variables equal to a reference category, effectively calculating hypothetical wages and worker flows as if all workers were white, non-hispanic and married. Second, the regression allows us to control for differences in education and experience within cells. Third, using fitted values makes sure that there are no missing values: if there are no workers in a given cell, we generate a virtual worker with gender, education and experience equal to the cell average.

The regression specification we use must be flexible enough to not change the features of the data, but restrictive enough so that we can identify fitted values for all cells. We include fourth order polynomials in all controls, plus interactions of the first order effects of all controls with each other as well as with state or industry dummies, so that we get the following specification for worker w in state or industry i,

$$\begin{aligned} y_{wi} = {} & D_i'\beta_0 + \beta_1 f_{wi} + \beta_2 b_{wi} + \beta_3 m_{wi} + \beta_4 s_{wi} + \beta_5 x_{wi} \\ & + \beta_6 s_{wi}^2 + \beta_7 s_{wi}^3 + \beta_8 s_{wi}^4 + \beta_9 x_{wi}^2 + \beta_{10} x_{wi}^3 + \beta_{11} x_{wi}^4 \\ & + \beta_{12} f_{wi} * s_{wi} + \beta_{13} f_{wi} * s_{wi}^2 + \beta_{14} f_{wi} * x_{wi} + \beta_{15} f_{wi} * x_{wi}^2 \\ & + s_{wi} * D_i'\beta_{16} + x_{wi} * D_i'\beta_{17} + x_{wi} * S_{wi}'\beta_{18} + x_{wi}^2 * S_{wi}'\beta_{19} + \varepsilon_{wi} \end{aligned} \quad (25)$$

where D_i is a vector of dummies for states or industries, f_{wi} is a dummy variable for female workers, b_{wi} a dummy for African-American workers, m_{wi} a dummy for married workers, s_{wi} is schooling in years, x_{wi} is potential labor market experience (age minus schooling minus 6) and S_{wi} is a vector of dummies for the five education categories.

The dependent variable y_{wit} is either the logarithm of the wage or a dummy variable indicating whether that worker lost or found a job. If y_{wit} is a dummy variable, we use a probit model to guarantee that the fitted values lie between 0 and 1. For wages we use a log-linear specification, as is common in the literature, see Card (1999).[30] In order to get fitted values for wages, we use the fitted values for log wages and apply the correction factor suggested by Carmeron and Trivedi (2010).[31] For the regressions of the probability to find or loose a job we use the sample weights from the basic monthly files. For regressions of wages we use the earnings weights, because wages are only available in the outgoing rotation groups.

The second step controls for differences in gender, education and experience across cells in a fully non-parametric manner. We first take relative deviations from the average across submarkets and only then take the (weighted) average over worker groups. Therefore, any differences in dispersion because of differences in the composition of the work force over the 40 cells are controlled for. For the weighted average in this step, we use invariant weights, equal to the average over years of the sample weights of each group, in order to avoid aggregation issues.

Controlling for worker heterogeneity in profits is more difficult, because we do not observe profits at the worker level. We attempt to still control for heterogeneity, by assuming that worker heterogeneity affects profits in the same way it affects wages. Then, we can control for heterogeneity by multiplying profits by the ratio of wages controlled for worker heterogeneity w_{it}^* over raw wages w_{it}, $\log \pi_{it}^* = \log \pi_{it}^{\text{NIPA}} - \log w_{it}^{\text{CPS}} + \log w_{it}^{*\text{CPS}}$ or $\log \pi_{it}^* = \log \pi_{it}^{\text{NIPA}} - \log w_{it}^{\text{NIPA}} + \log w_{it}^{*\text{CPS}}$. We explore the robustness of our results if we do not control profits and wages for worker heterogeneity.

D.2 Unobservable Heterogeneity

If job amenities are constant over time, true worker surplus is given by $\hat{S}_{it}^W + c_i^W$ and the true job surplus equals $\hat{S}_{it}^F + c_i^F$. Then, we can control for compensating differentials by using $\hat{S}_{it}^W$, $\hat{S}_{it}^F$, $\hat{f}_{it}^W$ and $\hat{f}_{it}^F$ in deviations from their time series averages. To see how this works, note that equations (2), (3) and (4) hold in each year, so that,

$$\hat{f}_{it}^W + \hat{S}_{it}^W + c_i^W = \alpha_{it}^{WM} \Rightarrow \widehat{\hat{f}}_{it}^W + \widehat{\hat{S}}_{it}^W = \hat{\alpha}_{it}^{WM} \tag{26}$$

$$\hat{f}_{it}^F + \hat{S}_{it}^F + c_i^F = \alpha_{it}^{JM} \Rightarrow \widehat{\hat{f}}_{it}^F + \widehat{\hat{S}}_{it}^F = \hat{\alpha}_{it}^{JM} \tag{27}$$

$$\hat{S}_{it}^W + c_i^W - \hat{S}_{it}^F - c_i^F = \alpha_{it}^{WD} \Rightarrow \widehat{\hat{S}}_{it}^W - \widehat{\hat{S}}_{it}^F = \hat{\alpha}_{it}^{WD} \tag{28}$$

where $\widehat{\hat{x}}_{it}$ denotes a variable in deviation from its time series average, where the variable itself is in deviation from its average across submarkets, $\widehat{\hat{x}}_{it} = \hat{x}_{it} - \overline{\hat{x}}_i$ and $\hat{x}_{it} = (x_{it} - \bar{x}_t)/\bar{x}_t$ and $\hat{\alpha}_{it} = \alpha_{it} - \bar{\alpha}_i$ denotes the adjustment costs in deviations from their

[30] Card, D. (1999). The Causal Effect of Education on Earnings. In: O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics,* Volume 3A, pp. 1801-1863. Elsevier.

[31] Cameron, A. C. and P. K. Trivedi (2010). *Microeconometrics Using Stata* (Revised ed.). Stata Press books. StataCorp LP.

time series average. For the industry data, we calculate deviations from the time series average separately for the SIC sample 1979-1997 and the NAICS sample 1998-2009. Taking deviations from the time series averages is like including state or industry-specific fixed effects and controls for time-invariant compensating differentials.

E Disaggregation and the Level of Mismatch

In appendix A, we showed that

$$\frac{u^{CF}}{u} \simeq \exp\left(\tfrac{1}{2}\left(1-\mu\right)V\left[\theta_i/\bar{\theta}\right]\right)\frac{V\left[\theta_i^{CF}/\bar{\theta}^{CF}\right]}{V\left[\theta_i/\bar{\theta}\right]} \tag{29}$$

which is observable except the variance ratio, which we get from Barnichon and Figura. Notice that $\exp\left(\frac{1}{2}\left(1-\mu\right)V\left[\theta_i/\theta\right]\right) \simeq 1$ so that we can safely ignore this part of the correction factor.

Barnichon and Figura show that

$$\ln\left(V_n\left[\frac{\theta_i}{\bar{\theta}}\right]\right) \simeq \ln a_0 + a_{geo}\ln n_{geo} + a_{occ}\ln n_{occ} \tag{30}$$

where V_n is the variance of θ_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 estimates the parameters of this relation using UK data to and find $a_{geo} = 0.13$ and $a_{occ} = 0.67$. This implies

$$\ln\left(\frac{V\left[\theta_i^{CF}/\bar{\theta}^{CF}\right]}{V\left[\theta_i/\bar{\theta}\right]}\right) = a_{geo}\ln\left(\frac{1}{n_{geo}}\right) + a_{occ}\ln\left(\frac{1}{n_{occ}}\right) \tag{31}$$

because by assumption θ_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,

$$\frac{V\left[\theta_i^{CF}/\bar{\theta}^{CF}\right]}{V\left[\theta_{geo*ind,i}/\theta_{geo*ind}\right]} = \exp\left(0.13\ln\left(4.64\right) + 0.67\ln\left(10.7\right)\right) = 6.0 \tag{32}$$

which is the correction factor for aggregation that we use in Section 4.2.1 in the main text.

F Additional Tables

Table 2A
State-level data, cell sizes 1979-2009

		job finding rate			wage		
		min	mean	max	min	mean	max
Alabama	AL	246	463	1050	5572	7405	10385
Arizona	AZ	204	361	615	5723	7194	10136
Arkansas	AR	205	390	573	5394	6955	9695
California	CA	1193	2288	4367	33428	46222	55085
Colorado	CO	186	448	929	6735	10056	13724
Connecticut	CT	176	391	1020	5125	8475	13579
Delaware	DE	188	316	603	3915	6839	9680
District of Columbia	DC	221	362	616	3031	6014	8470
Florida	FL	613	979	1878	17616	23416	28878
Georgia	GA	279	448	1034	6686	9797	12993
Hawaii	HI	166	294	532	4980	6792	9031
Idaho	ID	222	402	633	5929	7672	9340
Illinois	IL	602	1172	2117	17472	22925	26513
Indiana	IN	217	508	1125	7001	9360	13773
Iowa	IA	186	414	727	7269	9593	12266
Kansas	KS	234	359	539	7041	8412	10860
Kentucky	KY	231	446	806	6219	7505	10646
Louisiana	LA	219	409	828	4993	6408	9327
Maine	ME	212	405	763	5563	8023	12070
Maryland	MD	223	409	866	6225	9627	14703
Massachusetts	MA	350	717	1712	8758	16852	26276
Michigan	MI	541	1237	2477	12401	20266	26395
Minnesota	MN	193	488	1019	7525	10782	15800
Mississippi	MS	205	424	825	4585	6568	9960
Missouri	MO	209	467	936	6185	8852	11739
Montana	MT	229	394	600	5053	7570	9731
Nebraska	NE	158	301	464	6045	8830	10966
Nevada	NV	214	413	967	5925	8014	12561
New Hampshire	NH	175	336	709	5198	8183	14088
New Jersey	NJ	395	833	1701	12034	18992	26772
New Mexico	NM	200	372	583	4399	6622	9359
New York	NY	779	1479	2426	23758	33195	43087
North Carolina	NC	329	674	1134	10670	17115	28086
North Dakota	ND	215	324	483	7093	8385	10462
Ohio	OH	562	1163	2461	15462	22371	27942
Oklahoma	OK	191	341	648	5588	7574	10028
Oregon	OR	220	473	827	5980	7626	11589
Pennsylvania	PA	605	1128	2357	16717	23033	27018
Rhode Island	RI	207	415	985	4127	7269	11423
South Carolina	SC	190	390	660	5600	7197	8994
South Dakota	SD	198	322	468	7450	9320	10964
Tennessee	TN	206	415	860	6452	7630	9445
Texas	TX	629	1200	1802	24797	28094	31313
Utah	UT	191	359	634	6780	8127	12329
Vermont	VT	186	322	539	5121	7141	10211
Virginia	VA	218	421	819	8204	10689	13362
Washington	WA	264	497	874	6501	8557	10981
West Virginia	WV	270	476	1040	5267	6407	8814
Wisconsin	WI	242	509	970	8277	10397	13490
Wyoming	WY	191	311	501	5153	7110	9524

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
Construction	CON	1792	3383	5301	26427	35743	42343
Lumber & wood prods, excl. furniture	LUM	202	373	689	2831	3985	5405
Furniture & fixtures	FUR	122	249	382	2188	2935	3526
Stone, clay, concrete, glass prods	MNR	134	258	460	1939	2978	4095
Primary metals	PMT	116	377	1071	2564	4180	7121
Fabricated metals	FMT	195	494	1021	4368	6495	10061
Machinery, ex electrical	MAC	249	647	1606	8025	12245	16888
Electrical machinery, equip supplies	ELC	196	542	1004	5850	9929	13868
Motor vehicles & equip	MVH	167	487	1150	4477	5552	6305
Other transportation equip	OVH	161	321	588	3057	5551	6788
Professional & photo equip, watches	PHO	123	197	301	2390	3284	3901
Misc mfg industries	MMA	189	280	460	2321	2927	3557
Food & kindred prods	FOO	289	676	1173	5759	8881	11817
Textile mill prods	TEX	153	282	532	2231	3626	4698
Apparel & other finished textile prods	APP	257	506	780	3555	5451	7740
Paper & allied prods	PAP	126	205	313	2379	3550	4578
Printing, publishing & allied inds	PUB	235	385	543	6343	8496	9520
Chemicals & allied prods	CHE	156	281	483	4564	6301	7940
Rubber & misc plastic prods	RUB	163	274	464	2730	3623	4255
Leather & leather prods	LEA	102	181	327	545	1113	2134
Transportation	TRA	641	1014	1705	18969	23240	25327
Communications	COM	180	243	326	5445	7701	9537
Wholesale trade	WHO	489	874	1413	16195	21548	24065
Retail trade	RET	2677	4753	7064	71748	92571	104841
Banking & other finance	FIN	283	424	580	12118	15675	17562
Business services	BSV	546	1181	1667	10913	19947	25720
Automobile & repair services	ASV	281	505	796	7332	9202	10389
Personal serv ex private hhs	PSV	474	836	1250	11642	16321	18720
Entertainment & recreation	ENT	430	582	783	5233	7476	9724
Health services	HEA	696	1116	1637	35425	44720	50893
Educational services	EDU	508	946	1481	36027	45360	54301
Social services	SOC	297	469	616	7694	10604	14509
Misc professional services	MSV	340	563	864	12861	20610	25762

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 1998-2009

		job finding rate			wage		
		min	mean	max	min	mean	max
Construction	CON	1651	2644	5491	33309	40992	46426
Nonmetallic mineral product manufacturing	MNR	119	166	269	2013	2160	2330
Primary metals and fabricated metal products	MET	251	363	673	6150	7448	8054
Machinery manufacturing	MAC	200	314	554	4869	6145	7784
Computer and electronic product manufacturing	CEM	170	294	524	3906	5172	6279
Electrical equipment, appliance manufacturing	ELC	135	220	486	1698	3519	5812
Transportation equipment manufacturing	VEH	260	400	882	7423	8366	8780
Wood products	LUM	146	197	304	1810	2480	2946
Furniture and fixtures manufacturing	FUR	132	189	340	1860	2470	2922
Miscellaneous and not specified manufacturing	MMA	206	290	507	4318	4842	5249
Food manufacturing, beverage and tobacco	FOO	272	389	523	7418	7670	8533
Textile, apparel, and leather manufacturing	TEX	205	308	402	2252	3761	5636
Paper and printing	PAP	180	272	388	3795	5814	7661
Chemical manufacturing	CHE	148	219	431	4564	4857	5573
Plastics and rubber products	RUB	143	198	288	1768	2839	3486
Wholesale trade	WHO	457	602	882	14396	18216	21820
Retail trade	RET	1909	2531	4094	58587	62809	67348
Transportation and warehousing	TRA	616	817	1392	22394	23884	26988
Publishing industries (except internet)	PUB	112	166	271	1968	2899	3811
Broadcasting and Telecommunications	COM	193	344	621	6729	7425	9733
Information and data processing services	INF	107	237	647	1509	4810	9828
Finance	FIN	297	488	897	14925	17440	18960
Professional and technical services	PSV	651	983	1813	27312	32028	35463
Administrative and support services	ASV	708	1341	2485	8745	16145	21859
Educational services	EDU	643	991	1580	42900	49378	53350
Hospitals	HOS	295	452	724	23158	26792	35410
Health care services, except hospitals	HEA	328	707	1336	18172	28243	35478
Social assistance	SOC	359	488	756	10284	12040	14479
Arts, entertainment, and recreation	ENT	554	656	950	10297	11475	14963
Accommodation	ACC	329	433	637	6188	7093	8229
Food services and drinking places	FSV	1317	1724	2677	25960	29458	31395
Other services (excl. government)	MSV	682	894	1360	23891	25986	27923

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).

Table 3
Disaggregation and the Level of Mismatch

	N	Est MMU $\bar{u}^{MM}/\bar{u}$	Corr sampl error	Implied MMU (corr for agg)
States	50	2.3%	2.2%	
Industries	33*	2.1%	2.0%	
States*Industries	1650*	15.0%	14.0%	84%

Procedure to correct for sampling error and aggregation is explained in section 4.2.1 and appendix E.
*We use 33 broad industries for the SIC classification 1979-1997, and 32 for the NAICS classification 1998-2009. As a result, we have 1650 state*industry cells before until 1997 and 1600 cells from 1998 onwards.