Online Appendix

A. Measurement issues and seasonal adjustment

The seasonally adjusted numbers of people unemployed for less than 5 weeks, for between 5 and 14 weeks, 15-26 weeks and for longer than 26 weeks are published by the Bureau of Labor Statistics. To further break down the number unemployed for longer than 26 weeks into those with duration between 27 and 52 weeks and with longer than 52 weeks, we used seasonally unadjusted CPS microdata publicly available at the NBER website (http://www.nber.org/data/cps_basic.html). Since the CPS is a probability sample, each individual is assigned a unique weight that is used to produce the aggregate data. From the CPS microdata, we obtain the number of unemployed whose duration of unemployment is between 27 and 52 weeks and the number longer than 52 weeks. We seasonally adjust the two series using X-12-ARIMA,[27] and calculated the ratio of those unemployed 27-52 weeks to the sum. We then multiplied this ratio by the published BLS seasonally adjusted number for individuals who had been unemployed for longer than 26 weeks to obtain our series $U_t^{7.12}$.[28]

An important issue in using these data is the redesign of the CPS in 1994. Before 1994, individuals were always asked how long they had been unemployed. After the redesign, if an individual is reported as unemployed during two consecutive months, then her duration is recorded automatically as the sum of her duration last month and the number of weeks between the two months' survey reference periods. Note that if an individual was unemployed during each of the two weeks surveyed, but worked at a job in between, that individual would likely self-report a duration of unemployment to be less than 5 weeks before the redesign, but the duration would be imputed to be a number greater than 5 weeks after the redesign.

As suggested by Elsby, Michaels and Solon (2009) and Shimer (2012) we can get an idea of the size of this effect by making use of the staggered CPS sample design. A given address is sampled for 4 months (called the first through fourth rotations, respectively), not sampled for the next 8

[27] An earlier version of this paper dealt with seasonality by taking 12-month moving averages and arrived at similar overall results to those presented in this version. As a further check on the approach used here, we compared the published BLS seasonally adjusted number for those unemployed with duration between 15 and 26 weeks to an X-12-ARIMA-adjusted estimate constructed from the CPS microdata, and found the series to be quite close.

[28] This adjustment is necessary because the published number for unemployed with duration longer than 26 weeks is different from that directly computed from the CPS microdata, although the difference is subtle. The difference arises because the BLS imputes the numbers unemployed with different durations to various factors, e.g., correction of missing observations.

months, and then sampled again for another 4 months (the fifth through eighth rotations). After the 1994 redesign, the durations for unemployed individuals in rotations 2-4 and 6-8 are imputed, whereas those in rotations 1 and 5 are self-reported, just as they were before 1994. For those in rotation groups 1 and 5, we can calculate the fraction of individuals who are newly unemployed and compare this with the total fraction of newly unemployed individuals across all rotations. The ratio of these two numbers is reported in Panel A of Figure A1, and averaged 1.15 over the period 1994-2007 as reported in the second row of Table A1. For comparison, the ratio averaged 1.01 over the period 1989-1993, as seen in the first row. This calculation suggests that if we want to compare the value of U_t^1 as calculated under the redesign to the self-reported numbers available before 1994, we should multiply the former by 1.15. This is similar to the adjustment factors of 1.10 used by Hornstein (2012), 1.154 by Elsby, Michaels and Solon (2009), 1.106 by Shimer (2012), and 1.205 by Polivka and Miller (1998).

For our study, unlike most previous researchers, we also need to specify which categories the underreported newly unemployed are coming from. Figure A1 reports the observed ratios of rotation 1 and 5 shares to the total for the various duration groups, with average values summarized in Table A1. One interesting feature is that under the redesign, the fraction of those with 7-12 month duration from rotations 1 and 5 is very similar to that for other rotations, whereas the fraction of those with 13 or more months is much lower.[29] Based on the values in Table A1, we should scale up the estimated values for U_t^1 and scale down the estimated values of $U_t^{2.3}$ and $U_t^{13.+}$ relative to the pre-1994 numbers. The values for $U_t^{4.6}$ and $U_t^{7.12}$ seem not to have been affected much by the redesign. Our preferred adjustment for data subsequent to the 1994 redesign is to multiply U_t^1 by 1.15, $U_t^{2.3}$ by 0.87, $U_t^{13.+}$ by 0.77, and leave $U_t^{4.6}$ and $U_t^{7.12}$ as is. We then multiplied all of our adjusted duration figures by the ratio of total BLS reported unemployment to the sum of our adjusted series in order to match the BLS aggregate exactly.

Hornstein (2012) adopted an alternative adjustment, assuming that all of the imputed newly unemployed came from the $U^{2.3}$ category. He chose to multiply U_t^1 by 1.10 and subtract the added workers solely from the $U_t^{2.3}$ category. As a robustness check we also report results using

[29] One possible explanation is digit preference– an individual is much more likely to report having been unemployed for 12 months than 13 or 14 months. When someone in rotation 5 reports they have been unemployed for 12 months, BLS simply counts them as such, and if they are still unemployed the following month, BLS imputes to them a duration of 13 months. The imputed number of people 13 months and higher is significantly bigger than the self-reported numbers, just as the imputed number of people with 2-3 months appears to be higher than self-reported.

Hornstein's proposed adjustment in Section 5.1, as well as results using no adjustments at all.

An alternative might be to to use the ratios for each t in Figure A1 rather than to use the averages from Table A1. However, as Shimer (2012) and Elsby, Michaels and Solon (2009) mentioned, such an adjustment would be based on only about one quarter of the sample and thus multiplies the sampling variance of the estimate by about four, which implies that noise from the correction procedure could be misleading in understanding the unemployment dynamics.

Table A1. Average ratio of each duration group's share in the first/fifth rotation group to that in total unemployment

	U^1	$U^{2.3}$	$U^{4.6}$	$U^{7.12}$	$U^{13.+}$
1989-1993	1.01	1.01	0.96	1.02	0.97
1994-2007	1.15	0.87	0.95	1.05	0.77

B. Estimation algorithm

The system in Section 2.1 can be written as

$$x_t = Fx_{t-1} + v_t$$

$$y_t = h(x_t) + r_t$$

for $x_t = (\xi_t', \xi_{t-1}', ..., \xi_{t-47}')'$, $E(v_t v_t') = Q$, and $E(r_t r_t') = R$. The function $h(.)$ as well as elements of the variance matrices R and Q depend on the parameter vector $\theta = (\tilde{\delta}_1, \tilde{\delta}_2, \tilde{\delta}_3, R_1, R_{2.3}, R_{4.6}, R_{7.12}, R_{13+}, \sigma_L^w, \sigma_H^w, \sigma_L^x, \sigma_H^x)'$. The extended Kalman filter (e.g., Hamilton, 1994b) can be viewed as an iterative algorithm to calculate a forecast $\hat{x}_{t|t-1}$ of the state vector conditioned on knowledge of θ and observation of $Y_{t-1} = (y_{t-1}', y_{t-2}', ..., y_1')'$ with $P_{t|t-1}$ the MSE of this forecast. With these we can approximate the distribution of y_t conditioned on Y_{t-1} as $N(h(\hat{x}_{t|t-1}), \Omega_t)$ for $\Omega_t = H_t' P_{t|t-1} H_t + R$ and $H_t = \partial h(x_t)/\partial x_t'|_{x_t = \hat{x}_{t|t-1}}$ from which the approximate likelihood function associated with that θ,

$$\ell(\theta) = \textstyle\sum_{t=1}^T \ln f(y_t|Y_{t-1}; \theta)$$

$$\ln f(y_t|Y_{t-1}; \theta) = -(1/2)\ln(2\pi) - (1/2)\ln|\Omega_t| - (1/2)[y_t - h(\hat{x}_{t|t-1})]'\Omega_t^{-1}[y_t - h(\hat{x}_{t|t-1})],$$

can be maximized numerically. The forecast of the state vector can be updated by iterating on

$$K_t = P_{t|t-1} H_t (H_t' P_{t|t-1} H_t + R)^{-1}$$

$$P_{t|t} = P_{t|t-1} - K_t H_t' P_{t|t-1}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t (y_t - h(\hat{x}_{t|t-1}))$$

$$\hat{x}_{t+1|t} = F \hat{x}_{t|t}$$

$$P_{t+1|t} = F P_{t|t} F' + Q.$$

Prior to the starting date January 1976 for our sample, BLS aggregates are available but not the micro data that we used to construct $U_t^{13.+}$. For the initial value for the extended Kalman filter, we calculated the values that would be implied if pre-sample values had been realizations from an initial steady state, estimating the (4×1) vector $\bar{\xi}_0$ from the average values for $\bar{U}^1, \bar{U}^{2.3}, \bar{U}^{4.6}$, and $\bar{U}^{7.+}$ over February 1972 - January 1976 using the method described in Section 1.1. Our initial guess was then $\hat{x}_{1|0} = \iota_{48} \otimes \bar{\xi}_0$ where ι_{48} denotes a (48×1) vector of ones. Diagonal elements of $P_{1|0}$ determine how much the presample values of ξ_j are allowed to differ from this initial guess $\hat{\xi}_{j|0}$. For this we set $E(\xi_j - \hat{\xi}_{j|0})(\xi_j - \hat{\xi}_{j|0})' = c_0 I_4 + (1-j) c_1 I_4$ with $c_0 = 10$ and $c_1 = 0.1$. The value for c_0 is quite large relative to the range of $\xi_{t|T}$ over the complete observed sample, ensuring that the particular value we specified for $\hat{x}_{1|0}$ has little influence. For $k < j$ we specify the covariance[30] $E(\xi_j - \bar{\xi}_0)(\xi_k - \bar{\xi}_0)' = E(\xi_j - \bar{\xi}_0)(\xi_j - \bar{\xi}_0)'$. The small value for c_1 forces presample ξ_j to be close to ξ_k when j is close to k, again consistent with the observed month-to-month variation in $\hat{\xi}_{t|T}$.

Smoothed inferences about x_t using the full sample of available data, $\hat{x}_{t|T} = E(x_t|Y_T)$ and their variance matrix $P_{t|T} = E[(x_t - \hat{x}_{t|T})(x_t - \hat{x}_{t|T})']$ can be calculated by iterating backwards on the

[30] In other words,

$$P_{1|0} = \begin{bmatrix} c_0 I_4 & c_0 I_4 & c_0 I_4 & \cdots & c_0 I_4 \\ c_0 I_4 & c_0 I_4 + c_1 I_4 & c_0 I_4 + c_1 I_4 & \cdots & c_0 I_4 + c_1 I_4 \\ c_0 I_4 & c_0 I_4 + c_1 I_4 & c_0 I_4 + 2c_1 I_4 & \cdots & c_0 I_4 + 2c_1 I_4 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ c_0 I_4 & c_0 I_4 + c_1 I_4 & c_0 I_4 + 2c_1 I_4 & \cdots & c_0 I_4 + 47 c_1 I_4 \end{bmatrix}.$$

following equations for $t = T-1, T-2, ..., 1$:

$$J_t = P_{t|t} F' P_{t+1|t}^{-1}$$

$$\hat{x}_{t|T} = \hat{x}_{t|t} + J_t(\hat{x}_{t+1|T} - \hat{x}_{t+1|t})$$

$$P_{t|T} = P_{t|t} + J_t(P_{t+1|T} - P_{t+1|t})J_t'.$$

These smoothed inferences $\hat{x}_{t|T}$ and functions of them are plotted in Figures 5-7 and 10.

We calculated standard errors for the estimate $\hat{\theta}$ as in equation (3.13) in Hamilton (1994b):

$$E(\hat{\theta} - \theta)(\hat{\theta} - \theta)' \simeq V = K_1^{-1} K_2 K_1^{-1}$$

$$K_1 = \left. \frac{\partial \ell(\theta)}{\partial \theta \partial \theta'} \right|_{\theta = \hat{\theta}}$$

$$K_2 = \textstyle\sum_{t=1}^{T} \left\{ \left[\left. \frac{\partial \ln f(y_t | Y_{t-1}; \theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} \right] \left[\left. \frac{\partial \ln f(y_t | Y_{t-1}; \theta)}{\partial \theta} \right|_{\theta=\hat{\theta}} \right]' \right\}.$$

To obtain standard errors for the variance decompositions in Figure 7 and Table 3, we generated $J = 1,000$ draws from the asymptotic distribution of $\hat{\theta}$, $\theta^{[j]} \sim N(\hat{\theta}, V)$, $j = 1, ..., J$ and calcuated $q_{s,k}(\theta^{[j]})$ as in equation (21) for each s and each $k = 1, ..., 4$. The standard deviation of $q_{s,k}(\theta^{[j]}) / \sum_{k=1}^{4} q_{s,k}(\theta^{[j]})$ across draws j was used to get the error bands and standard errors in Figure 7 and Table 3.

The standard errors used for Figures 5 and 6 incorporate both filter and parameter uncertainty. The matrix $P_{t|T}$ summarizes uncertainty we would have about x_t even if we knew the true value of the parameters in θ. Given that we also have to estimate θ, the true uncertainty is greater than that represented by $P_{t|T}$. Following Ansley and Kohn (1986) we calculate the total variance as

$$\left. P_{t|T} \right|_{\theta=\hat{\theta}} + Z_t V Z_t'$$

$$\underset{(4 \times 12)}{Z_t} = \left. \frac{\partial \hat{x}_{t|T}}{\partial \theta'} \right|_{\theta=\hat{\theta}}.$$

The values of $\{Z_t\}_{t=1}^{T}$ can be found by numerical differentiation, e.g., replace $\hat{\theta}$ with $\hat{\theta} + \delta e_i$ for $\delta = 10^{-8}$ and e_i the ith column of I_{12} and then redo the iteration to calculate $\hat{x}_{t|T}(\hat{\theta} + \delta e_i)$. The

ith column of Z_t is then $\delta^{-1}[\hat{x}_{t|T}(\hat{\theta}+\delta e_i) - \hat{x}_{t|T}|(\hat{\theta})]$.

C. Derivation of linearized variance and historical decompositions

The state equation $\xi_{t+1} = \xi_t + \varepsilon_{t+1}$ implies

$$\begin{aligned} \xi_{t+s} &= \xi_t + \varepsilon_{t+1} + \varepsilon_{t+2} + \varepsilon_{t+3} + \cdots + \varepsilon_{t+s} \\ &= \xi_t + u_{t+s}. \end{aligned}$$

Letting $y_t = (U_t^1, U_t^{2.3}, U_t^{4.6}, U_t^{7.12}, U_t^{13.+})'$ denote the (5×1) vector of observations for date t, our model implies that in the absence of measurement error y_t would equal $h(\xi_t, \xi_{t-1}, \xi_{t-2}, ..., \xi_{t-47})$ where $h(\cdot)$ is a known nonlinear function. Hence

$$y_{t+s} = h(u_{t+s} + \xi_t, u_{t+s-1} + \xi_t, ..., u_{t+1} + \xi_t, \xi_t, \xi_{t-1}, ..., \xi_{t-47+s}).$$

We can take a first-order Taylor expansion of this function around $u_{t+j} = 0$ for $j = 1, 2, ..., s$,

$$y_{t+s} \simeq h(\xi_t, ..., \xi_t, \xi_t, \xi_{t-1}, ..., \xi_{t-47+s}) + \sum_{j=1}^{s} [H_j(\xi_t, \xi_t, ..., \xi_t, \xi_t, \xi_{t-1}, ..., \xi_{t-47+j})] u_{t+s+1-j}$$

for $H_j(\cdot)$ the (5×4) matrix associated with the derivative of $h(\cdot)$ with respect to its jth argument. Using the definition of u_{t+j}, this can be rewritten as

$$y_{t+s} \simeq c_s(\xi_t, \xi_{t-1}, ..., \xi_{t-47+s}) + \sum_{j=1}^{s} [\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+j})] \varepsilon_{t+j}$$

from which (18) follows immediately.

Similarly, for purposes of a historical decomposition note that the smoothed inferences satisfy

$$\hat{\xi}_{t+s|T} = \hat{\xi}_{t|T} + \hat{\varepsilon}_{t+1|T} + \hat{\varepsilon}_{t+2|T} + \hat{\varepsilon}_{t+3|T} + \cdots + \hat{\varepsilon}_{t+s|T}$$

where $\hat{\varepsilon}_{t+s|T} = \hat{\xi}_{t+s|T} - \hat{\xi}_{t+s-1|T}$. For any date $t+s$ we then have the following model-inferred value for the number of people unemployed:

$$\iota_5{}' h(\hat{\xi}_{t+s|T}, \hat{\xi}_{t+s-1|T}, \hat{\xi}_{t+s-2|T}, ..., \hat{\xi}_{t+s-47|T}).$$

For an episode starting at some date t, we can then calculate

$$\iota_5{}' h(\hat{\xi}_{t|T}, \hat{\xi}_{t|T}, \hat{\xi}_{t|T}, ..., \hat{\xi}_{t|T}, \hat{\xi}_{t-1|T}, ..., \hat{\xi}_{t+s-47|T}).$$

This represents the path that unemployment would have been expected to follow between t and $t+s$ as a result of initial conditions at time t if there were no new shocks between t and $t+s$. Given this path for unemployment that is implied by initial conditions, we can then isolate the contribution of each separate shock between t and $t+s$. Using the linearization in equation (18) allows us to represent the realized deviation from this path in terms of the contribution of individual historical shocks as in (22).

D. Alternative estimates of unemployment-continuation probabilities

There is an unresolved controversy in the literature about how to measure outflows from unemployment. Our measure described in footnote 1 follows van den Berg and van Ours (1996), van den Berg and van der Klaauw (2001), Elsby, Michaels and Solon (2009), Shimer (2012), and Elsby, Hobijn and Şahin (2013) in deriving flow estimates from the observed change in the number of unemployed by duration. An alternative approach, employed by Fujita and Ramey (2009) and Elsby, Hobijn and Şahin (2010), is to look at only those individuals for whom there is a matched observation of unemployment in month $t-1$ and a status of employment or out of the labor force in month t. In the absence of measurement error, the two estimates should be the same, but in practice they turn out to be quite different. One reason for the discrepancy is misclassification. For example, an individual who goes from long-term unemployed to out of the labor force to back to long-term unemployed in three successive months counts as a successful "graduate" from long-term unemployment using matched flows but is contributing to the stubborn persistence of long-term unemployment when using the stock data. A follow-up paper to Elsby, Hobijn and Şahin (2010) by Elsby et al. (2011) documented that of the individuals who were employed or out of the labor force in month $t-1$ and who were recorded as unemployed in month t, more than half reported their duration of unemployment to be 5 weeks or longer. Another important reason is that individuals for whom two consecutive observations are available differ in important ways from those for whom some observations are missing. Abowd and Zellner (1985) and Frazis et al. (2005) acknowledged that these measurement errors are more likely to bias the matched flow data than the stock data and suggested methods to correct the bias.

Since our goal is to understand how the reported stock of long-term unemployed came to be so high and why it falls so slowly, we feel that our approach, which is consistent with the observed stock data by construction, is preferable for our application.

E. Details of robustness tests

The standard errors in Table 3 were calculated as follows. For each model, we generated 500 draws for the k-dimensional parameter vector (where k is reported in the first row of the table) from a $N(\hat{\theta}, \hat{V})$ distribution where $\hat{\theta}$ is the MLE and $\hat{V}$ is the $(k \times k)$ variance matrix from inversed hessian of the likelihood function. For each draw of $\theta^{(\ell)}$ we calculated the values implied by that $\theta^{(\ell)}$ and then calculated the standard error of that inference across the draws $\theta^{(1)}, ..., \theta^{(500)}$.

Time-varying genuine duration dependence. Vishwanath (1989) and Blanchard and Diamond (1994) developed theoretical models in which genuine duration dependence could be linked to market tightness. See Whittaker and Isaacs (2014) for a detailed discussion of the conditions that can trigger extended unemployment benefits.

Shimer (2012) argued that this time-aggregation bias would result in underestimating the importance of outflows in accounting for cyclical variation in unemployment, and Fujita and Ramey (2009), Shimer (2012) and Hornstein (2012) all formulated their models in continuous time.

Allowing for structural shocks. For the factor model, the variance decomposition (19) becomes

$$
\begin{aligned}
& E(y_{t+s} - \hat{y}_{t+s|t})(y_{t+s} - \hat{y}_{t+s|t})' \\
= & \sum_{j=1}^{s} [\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+j})](\lambda\lambda' + Q)[\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+j})]' \\
= & \sum_{j=1}^{s} [\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+j})]\lambda\lambda'[\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+j})]' \\
& + \sum_{j=1}^{s} \sum_{m=1}^{4} Q_m [\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+j})e_m][\Psi_{s,j}(\xi_t, \xi_{t-1}, ..., \xi_{t-47+j})e_m]'
\end{aligned}
$$

for Q_m the row m, column m element of Q.

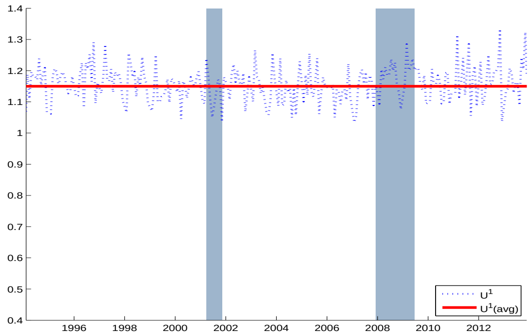
Time aggregation. Elsby, Michaels and Solon (2009) questioned the theoretical suitability of a continuous-time conception of unemployment dynamics, asking if it makes any sense to count a worker who loses a job at 5:00 p.m. one day and starts a new job at 9:00 a.m. the next as if they had been unemployed at all. We agree, and think that defining the central object of interest to be the fraction of those newly unemployed in month t who are still unemployed in month $t + k$, as in our baseline model, is the most useful way to pose questions about unemployment dynamics. Nevertheless, and following Kaitz (1970), Perry (1972), Sider (1985), Baker (1992), and Elsby, Michaels and Solon (2009) we also estimated a version of our model formulated in terms of weekly frequencies as an additional check for robustness.

We can do so relatively easily if we make a few simplifying assumptions. We view each month t as consisting of 4 equally-spaced weeks and assume that in each of these weeks there is an inflow of w_{it} workers of type i, each of whom has a probability $p_{it}(0) = \exp[-\exp(x_{it})]$ of exiting unemployment the following week. This means that for those type i individuals who were newly unemployed during the first week of month t, $w_{it}[p_{it}(0)]^3$ are still unemployed as of the end of the month. Thus for the model interpreted in terms of weekly transitions, equations (10) and (11) would be replaced by

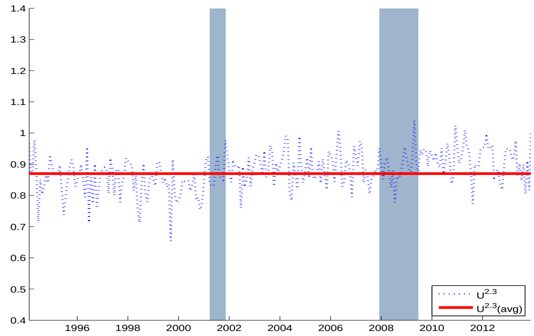
$$U_t^1 = \sum_{i=H,L} \{w_{it} + w_{it}[p_{it}(0)] + w_{it}[p_{it}(0)]^2 + w_{it}[p_{it}(0)]^3\} + r_t^1$$

$$U_t^{2.3} = \sum_{i=H,L} \sum_{s=1}^{4} \left\{w_{i,t-1}[p_{i,t-1}(1)]^{8-s} + w_{i,t-2}[p_{i,t-2}(2)]^{12-s}\right\} + r_t^{2.3}$$

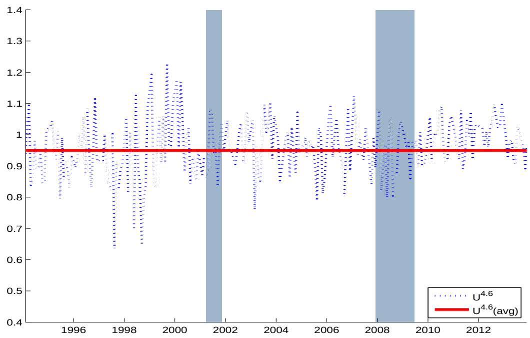
for $p_{it}(\tau)$ given by (5) and (8). Note that although this formulation is conceptualized in terms of weekly inflow and outflows w_i and p_i, the observed data y_t are the same monthly series used in our other formulations, and the number of parameters is the same as for our baseline formulation.



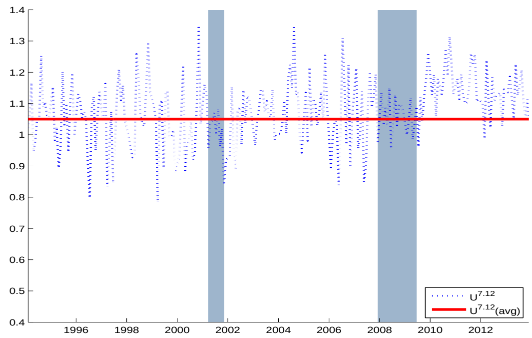
Panel A: U^1



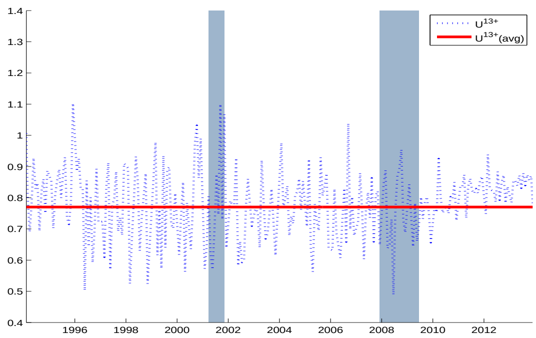
Panel B: $U^{2.3}$



Panel C: $U^{4.6}$



Panel D: $U^{7.12}$



Panel E: $U^{13.+}$

Figure A1. Ratio of each duration group's share in the first and fifth rotation groups to that in all rotation groups

Appendix References

Abowd, John M., and Arnold Zellner (1985). "Estimating Gross Labor-Force Flows," *Journal of Business and Economic Statistics*, 3(3): 254-283.

Ansley, F. Craig, and Robert Kohn. (1986). "Prediction Mean Squared Error for State Space Models with Estimated Parameters," *Biometrika*, 73(2):467-473.

Blanchard, Olivier J. and Peter Diamond (1994). "Ranking, Unemployment Duration, and Wages," *The Review of Economic Studies,* 61(3):417-434.

Darby, Michael R., John C. Haltiwanger, and Mark W. Plant (1986). "The Ins and Outs of Unemployment: The Ins Win," NBER Working Paper # 1997, National Bureau of Economic Research, Inc.

Elsby, Michael W. L., Bart Hobijn, and Ayşegül Şahin (2010). "The Labor Market in the Great Recession," *Brookings Papers on Economic Activity,* Spring 2010: 1-56.

Elsby, Michael W. L., Bart Hobijn, and Ayşegül Şahin (2013). "Unemployment Dynamics in the OECD," *Review of Economics and Statistics*, 95(2): 530-548.

Elsby, Michael W. L., Bart Hobijn, Ayşegül Şahin, and Robert G. Valletta (2011). "The Labor Market in the Great Recession: An Update to September 2011," *Brookings Papers on Economic Activity,* Fall 2011: 353-371.

Frazis, Harley J., Edwin L. Robison, Thomas D. Evans, and Martha A. Duff (2005). "Estimating Gross Flows Consistent with Stocks in the CPS," *Monthly Labor Review* September 2005: 3-9.

Hamilton, James D. (1994b). "State-Space Models," in Robert F. Engle and Daniel L. McFadden, eds., Handbook of Econometrics, Volume IV, Chapter 50, pp. 3039-3080. Amsterdam: Elsevier.

Kaitz, Hyman B. (1970). "Analyzing the Length of Spells of Unemployment," *Monthly Labor Review*, November 1970, pp.11-20, Bureau of Labor Statistics.

Perry, George L. (1972). "Unemployment Flows in the U.S. Labor Market," *Brookings Papers on Economic Activity,* 3(2):245–292.

Polivka, Anne E. and Stephen M. Miller (1998). "The CPS after the Redesign: Refocusing the Economic Lens," in John Haltiwanger, Marilyn E. Manser, and Robert Topel, eds., Labor Statistics Measurement Issues, pp.249-86.

Sider, Hal (1985). "Unemployment Duration and Incidence: 1968-1982," *American Economic Review*, 75(3):461-72.

Vishwanath, Tara. (1989). “Job Search, Stigma Effect, and Escape Rate from Unemployment,” *Journal of Labor Economics*, 7(4):487-502.

Whittaker, Julie M. and Katelin P. Isaacs (2014). "Unemployment Insurance: Programs and Benefits," *Congressional Research Service Report for Congress*, 7-5700.