

Appendix for "Visualizing labor market growth with many indicators"

This appendix has four sections; section 1 describes the data series used and their construction, section 2 has technical details for the bubble plots; section 3 has technical details on the heatmap and section 4 provides a decomposition of the change in the unemployment rate into five different components.

Section 1: List Of Labor Market Indicators Used

The following are notes [and Haver Analytics tickers] on data construction for the bubble plots and heatmaps. If a series x_t is available starting in month t_0 and a closely related series y_t is available at and before t_0 , we then define backwards extrapolation of x_t by y_t as $\tilde{x}_t = y_t \frac{x_{t_0}}{y_{t_0}}$ for $t < t_0$. All series are monthly.

-All Employees: Goods-producing Industries (**GoodsEmp**): Haver ticker LAGOODA@USECON.

-Temporary help services employment (**Temps**): The NAICS based series [LAR132A@USECON] is available since 1990. We backwards extrapolate it back to January 1982 with SIC "Help Supply Services" employment [LA7363A@LABOR03]. We backwards extrapolate the resultant series with SIC "Personnel Supply Services" employment [LA7360A@LABOR03] back to January 1972.

-Private services employment excluding temporary workers (**SvcsExTemp2AdjAug83**): Information services employment [LAINFOA@USECON] fell by about 600k in August 1983 and increased about the same the next month. We adjust this outlier by setting the August level equal to the average of the adjacent values. We add this series to private services [LAPSRVA@USECON] ex temporary workers (**Temps**) ex information services [LAINFOA@USECON] to get **SvcsExTemp2AdjAug83**.

-Government employment excluding temporary Decennial Census workers (**GovtExCensus**): Total government employment [LAGOVTA@USECON] less Census workers [LAFGT@USECON]. The latter is available since January 1990, so we set it to 0 before then.

-Total payroll employment ex census workers (**PayrollExCensus**): Adjusted for August 1983 outlier by taking the sum of **GoodsEmp**, **SvcsExTemp2AdjAug83**, **Temps**, and **GovtExCensus**.

-1-month employment diffusion index (**Diffusion1splice**): Proportion of 266 private industries that had increased employment compared to the previous month [ZDLA@LABOR].

The NAICS based series is available since January 1991; before that we used the SIC based analog [ZDLA@LABOR03] available back to 1977.

-Average weekly hours of private production & nonsupervisory workers (**AvgHoursAdjJan82**): This series has a one-time outlier in January 1982 which we eliminate by taking the average of the adjacent values.

-Unemployment rate (**UnempRate**): Defined as the negative of the log of the employment rate; $-\log(\frac{LF_t - U_t}{LF_t})$ where U_t is the total number of unemployed in month t [LTU@USECON] and LF_t is the labor force [LF@USECON]. We use this definition from Fleischman and Roberts (2011) since the unemployment rate, employment population ratio, and labor force participation rate have a multiplicative relationship. Our series will also differ from the published BLS series since the latter is rounded to 1 decimal place and rounding can be important for one-month changes.

-Labor force participation rate (**LFPRNoRound**): Defined as $\frac{LF_t}{POP_t}$ where POP_t is civilian non-institutional population. Differs slightly from published LFPR since the latter is rounded.

-Marginally attached workers (**MarginAttachSA**): Haver series LHWSAN@USECON; "Persons Not in LF: Want a Job, Searched, Available to Work Now". The original series is NSA and is seasonally adjusted with Proc X12 in SAS. Since the series is only available back until January 1994, we backwards extrapolate it with "Discouraged Workers"; its most closely related precursor. A quarterly series for the latter can be constructed from the "U1-U7 range of alternative measures 1948-1993" document available from the BLS. The quarterly series is interpolated to the monthly frequency with Denton interpolation and part-time workers [constructed from the same U1-U7 document].

-Persons who work part-time for economic reasons (**PTER**): Haver series LEPT@USECON. Available back to 1955, but there is a level shift in January 1994 due to the CPS redesign. We eliminate this by treating the 1955-1993 data and 1994-present data as two distinct time series. We then do a simple extrapolation of the 1955-1993 data to January 1994 and use this series to "backwards extrapolate" the 1994-present series.

-Shimer job-finding probability (**JFProbShimer**): Defined as $JFPR_t = 1 - \frac{U_t - U_t^{<5}}{U_{t-1}}$ where U_t is the total number of unemployed in month t [LTU@USECON] and $U_t^{<5}$ is the number of unemployed for less than 5 weeks [LU0@USECON]. We pre-transform $U_t^{<5}$ by multiplying

it by 1.10 before January 1994 to adjust for the break caused by the CPS redesign. This correction was suggested by Shimer (2012).

-Initial Claims for Unemployment Insurance (**InitialClaims**): The DOL series is weekly; Haver aggregates this to a monthly frequency [LICM@USECON].

-Total vacancies (**JOLTSOpens**): The JOLTS based series [LJJTLA@USECON] starts in December 2000. For the "dots" chart, this is the only data we use. For the heatmap, we backwards extrapolate the series with a Composite Help Wanted Index constructed using the methods described in Barnichon (2010). The composite index combines the discontinued newspaper help-wanted index with the online help-wanted index (both from the Conference Board).

-Private hires and private quits (**JOLTSPrivHires**, **JOLTSPrivQuits**): The JOLTS based series [LJHTPLA@USECON and LJQTPLA@USECON] start in December 2000. For the "dots" chart, this is the only data we use. For the heatmaps, we backwards extrapolate these series to 1990:Q2 with synthetic quarterly data on hires and quits from Davis, Faberman and Haltiwanger (2013). The synthetic quarterly data are available at Steven Davis's website. We interpolate their data to the monthly frequency using Chow-Lin interpolation and monthly data on initial claims and published CPS-based data on labor force status flows.

-NFIB: Percent of Small Businesses With Positions Not Able to Fill Right Now (**NFIBCantFill**): Haver series is NFIB2@SURVEYS. Quarterly data from 1973:Q4 - 1985:Q4 has been interpolated to monthly frequency.

-NFIB: Percent Planning to Increase Employment, Net (SA, %) (**NFIBHiringPlans**): Haver series is NFIB1@SURVEYS. Quarterly data from 1973:Q4 - 1985:Q4 has been interpolated to monthly frequency.

-Conference Board Survey Question; Appraisal Present Sit: Employment, Jobs Plentiful (% Respondents; SA) (**ConfBoardJobAvail**): Haver ticker is EPN@CBDB.

Section 2: Treatment of Data For Bubble Plots

UnempRate, **NFIBCantFill**, **NFIBHiringPlans**, **ConfBoardJobAvail** are differenced and **Diffusion1splice** is entered in levels. The remaining series are all log-differenced. Let x_t^i denote the transformed value of indicator i in period t , and let $x_t^{i,n} = \frac{1}{h} \sum_{h=0}^{n-1} x_{t-h}^i$. For the bubble plots of the n -month changes, we calculate the mean and standard deviation

of $x_t^{i,n}$ over the sample starting $n - 1$ months after January 1989. Denote these moments by $\bar{x}^{i,n}$ and $\sigma_x^{i,n}$. We convert $x_t^{i,n}$ into a z-score by $z_t^{i,n} = \frac{x_t^{i,n} - \bar{x}^{i,n}}{\sigma_x^{i,n}}$. Let Pay_t denote the level of **PayrollExCensus** in month t . We convert $z_t^{i,n}$ into a monthly payroll equivalent with the formula

$$(1) \tilde{z}_t^{i,n} = \frac{1}{n} Pay_{t-n} [\exp(n\{z_t^{i,n} \sigma_x^{Pay,n} + \bar{x}^{Pay,n}\}) - 1]$$

By simple substitution we see that $\tilde{z}_t^{Pay,n}$ is the average monthly change in payroll employment from month $t - n$ to month t . To get the tick mark labels for the x-axis in the bubble plot, we define a payroll equivalent unit as $\frac{\bar{x}^{Pay,n}}{\sigma_x^{Pay,n}}$ z-score units and plug each of the points $\dots, \frac{-2\bar{x}^{Pay,n}}{\sigma_x^{Pay,n}}, \frac{-\bar{x}^{Pay,n}}{\sigma_x^{Pay,n}}, 0, \frac{\bar{x}^{Pay,n}}{\sigma_x^{Pay,n}}, \frac{2\bar{x}^{Pay,n}}{\sigma_x^{Pay,n}}, \frac{3\bar{x}^{Pay,n}}{\sigma_x^{Pay,n}}, \dots$ into $\tilde{z}_t^{i,n}$ in equation (1). The y-axis labels are defined similarly by replacing " t " with " $t - n$ " everywhere above starting with equation (1). The origin in the plot [ordered pair (0,0) in z-score space] corresponds to the ordered pair $(\frac{1}{n} Pay_{t-n} \exp(\bar{x}^{Pay,n}), \frac{1}{n} Pay_{t-2n} \exp(\bar{x}^{Pay,n}))$ in "payroll equivalent" space. In nonrecession periods, the x-origin coordinate will be slightly larger than the y-coordinate origin.

Section 3: Determining time-varying trends and color shading for heatmap

Again, we let x_t^i denote a transformed indicator for month t and we let ΔUR_t denote the first difference of **UnempRate** [as defined above]. We estimate the time varying parameter regression model.

$$(2) x_t^i = \alpha_t^i + \beta_t^i \Delta UR_t + \varepsilon_t^i$$

where α_t^i and β_t^i are the time-varying intercept and slope. We identify α_t^i as the trend of x_t^i at time t since that would be its average value with a constant unemployment rate. We set $\alpha_t^{UR} = 0$ so that there is no time-variation in the natural unemployment rate¹. We estimate (2) with a constant gain least squares (CGLS) algorithm based on Branch and Evans (2005).

CGLS Algorithm: Get initial estimates of $[\begin{smallmatrix} \alpha_{t_0}^i \\ \beta_{t_0}^i \end{smallmatrix}]$ and $R_{t_0}^i = Var(\beta_{t_0}^i)$ by estimating OLS on a training sample or, if x_t^i is not available for a long period [e.g. the JOLTS data], a training sample on a related series. Set the decay factor κ to a preset constant [we use $\kappa = 0.0075$]. The routine then resembles weighted least squares with geometrically declining

¹Incorporating a time varying NAIRU by using $\Delta(UR_t - (-\log(1 - NAIRU_t^{CBO})))$ [where $NAIRU_t^{CBO}$ is the Congressional Budget Office's estimate of NAIRU] in place of ΔUR_t in equation (2) had only a small impact on the estimates of α_t^i .

weights that have a half-life of 92 months. α_t^i and β_t^i are then recursively computed using the equations:

$$(3) R_t^i = R_{t-1}^i + \kappa \left(\begin{bmatrix} 1 \\ \Delta U R_t \end{bmatrix} \begin{bmatrix} 1 & \Delta U R_t \end{bmatrix} - R_{t-1}^i \right)$$

$$(4) \begin{bmatrix} \alpha_t^i \\ \beta_t^i \end{bmatrix} = \begin{bmatrix} \alpha_{t-1}^i \\ \beta_{t-1}^i \end{bmatrix} + \kappa (x_t^i - \begin{bmatrix} 1 & \Delta U R_t \end{bmatrix} \begin{bmatrix} \alpha_{t-1}^i \\ \beta_{t-1}^i \end{bmatrix}) (R_t^i)^{-1} \begin{bmatrix} 1 \\ \Delta U R_t \end{bmatrix}$$

In the next step, x_t^i is normalized as $z_t^i = \frac{x_t^i - \alpha_t^i}{\sqrt{\frac{1}{t_1 - t_0 + 1} \sum_{h=t_0}^{t_1} (x_h^i - \alpha_h^i)^2}}$ where t_0 is the first period of the sample used for x_t^i [no earlier than January 1981] and t_1 the last. Unlike the HP filter, the CGLS based estimate is a one-sided estimate of trend. A disadvantage of this approach is that α_t^i is not the best estimate of trend for the middle of the sample; an advantage is that past values of the trend are not continuously updated except for revisions to the source data. We compute the first principal component (PC) \hat{F}_t of the normalized data series over the January 1981-present sample using the approach described by Stock and Watson (2002) to handle missing values of the short series. All of the series shown in the heatmap are used to estimate the first PC except for payroll employment, since we are using payroll employment's subcomponents and we don't want to "double count". For each series, we run the regression

$$(5) x_t^i - \alpha_t^i = \gamma^i \hat{F}_t + \epsilon_t^i$$

and compute the sample second moment of ϵ_t^i as $\mu_2^{\epsilon^i} = \frac{1}{t_1 - t_0 + 1} \sum_{h=t_0}^{t_1} (\epsilon_h^i)^2$. The scaled-value used for the i th one-month change heatmap is

$$(6) hmap_t^i = \frac{x_t^i - \alpha_t^i}{\sqrt{\mu_2^{\epsilon^i}}}$$

When "labor market growth" is exactly average in the sense that $\hat{F}_t = 0$; normality of ϵ_t^i implies $hmap_t^i$ is approximately standard normal. We assign heatmap colors based on 90%, 70%, and 50% coverage bands of the standard normal distribution: $|hmap_t^i| > 1.645 \rightarrow$ green/blue; $1.645 \geq |hmap_t^i| > 1.036 \rightarrow$ light color; $1.036 \geq |hmap_t^i| > 0.674 \rightarrow$ faint color; $0.674 \geq |hmap_t^i| \rightarrow$ white. In a constant unemployment rate environment, about 10% of the cells will be the dark red or dark green. For n -month average changes, we denote the average value of $(\frac{1}{n} \sum_{k=0}^{n-1} \epsilon_{t-k}^i)^2$ as $\mu_2^{\epsilon^i, n}$. The standardized heatmap value for the n -month average change is then:

$$(7) hmap_t^{i,n} = \frac{\frac{1}{n} \sum_{k=0}^{n-1} [x_{t-k}^i - \alpha_{t-k}^i]}{\sqrt{\mu_2^{\epsilon^i, n}}}$$

Again, $hmap_t^{i,n}$ will be approximately standard normal if ϵ_t^i is normally distributed and we use the same cutoff values to assign colors.

Section 4: Decomposing change in unemployment rate into payroll employment growth and other components.

Define the following variables, with t subscripts denoting month t .

- PAY_t : Nonfarm payroll employment (SA).

$$-P\dot{A}Y_t = \frac{PAY_t}{PAY_{t-1}} - 1.$$

- H_t : Civilian household survey employment (SA).

- H_t^{Smooth} : Population break adjusted civilian household survey employment from BLS (SA).

$$-\dot{H}_t^{Smooth} = \frac{H_t^{Smooth}}{H_{t-1}^{Smooth}} - 1.$$

- U_t : Number of unemployed (SA).

- L_t : Civilian labor force (SA).

- $UR_t^* = \frac{100U_t}{LF_t}$: Unrounded unemployment rate in percentage points.

- L_t^{Smooth} : Population break adjusted civilian labor force from BLS.

- POP_t : Civilian noninstitutional population.

$$-LFP_t = \frac{LF_t}{POP_t}$$

$$-L\dot{F}P_t = \frac{LFP_t}{LFP_{t-1}} - 1$$

$$-POP_t^{Smooth} = \frac{L_t^{Smooth}}{LFPR_t}$$

$$-EMPPOP_t = \frac{H_t}{POP_t}$$

We define the two scaling factors $\alpha_t = \frac{100EMPPOP_t}{LFPR_{t-1}} \approx 100(1 - \frac{UR_t^*}{100})$ and $\beta_t = \frac{H_{t-1}^{Smooth}}{POP_t^{Smooth} EMPPOP_{t-1}} \approx 1$.

Then, by defining the following contribution terms:

- $LFP_t^{Cont} = \alpha_t L\dot{F}P_t$: Labor force participation rate growth contribution.

- $DISCREP_t^{Cont} = \alpha_t \beta_t (P\dot{A}Y_t - \dot{H}_t^{Smooth})$: Contribution from discrepancy between payroll growth and household growth

- $PAY_t^{Cont} = -\alpha_t \beta_t PAY_t$: Contribution from payroll employment growth.

- $POP_t^{Cont} = \alpha_t \beta_t P\dot{O}H_t^{Smooth}$: Contribution from population growth.

we can approximate the one-month change in the unemployment rate very closely with the decomposition

$$(8)\Delta UR_t^* \approx PAY_t^{Cont} + POP_t^{Cont} + DISCREP_t^{Cont} + LFP_t^{Cont} = \Delta \hat{U}R_t^*$$

If we define the approximation error $\epsilon_t^{\Delta UR} = \Delta UR_t^* - \Delta \hat{U}R_t^*$, we have the exact decomposition

$$(9)\Delta UR_t^* = PAY_t^{Cont} + POP_t^{Cont} + DISCREP_t^{Cont} + LFP_t^{Cont} + \epsilon_t^{\Delta UR}$$

For an h -month change in the unemployment rate, we have

$$(10)UR_{t+h}^* - UR_t^* = \sum_{k=1}^h PAY_{t+k}^{Cont} + \sum_{k=1}^h POP_{t+k}^{Cont} + \sum_{k=1}^h DISCREP_{t+k}^{Cont} + \sum_{k=1}^h LFP_{t+k}^{Cont} + \sum_{k=1}^h \epsilon_{t+k}^{\Delta UR}$$

and again $\sum_{k=1}^h \epsilon_t^{\Delta UR}$ is very close to 0 for, say, $h \leq 36$. For $h = 36$, for example the maximum value of $|\sum_{k=1}^{36} \epsilon_t^{\Delta UR}|$ since January 1997 is 0.035, much less than one-tenth of a percentage point.

References:

Barnichon, Regis (2010). "Building a composite Help-Wanted Index." *Economics Letters* 91 (2010) 175–178.

Branch, William A. and George W. Evans (2006). "A simple recursive forecasting model." *Economics Letters* 109 (2006) 159–166.

Davis, Steven J., Faberman, Jason, and John Haltiwanger (2013). "The Establishment-Level Behavior of Vacancies and Hiring". *Quarterly Journal of Economics*, 01 May 2013.

Fleischman, Charles A. and John M. Roberts (2011). "From Many Series, One Cycle: Improved Estimates of the Business Cycle from a Multivariate Unobserved Components Model." Federal Reserve Finance and Economics Discussion Series (FEDS) Working Paper No. 2011-46.

Stock, James H. and Mark W. Watson (2002) "Macroeconomic Forecasting Using Diffusion Indexes." *Journal of Business & Economic Statistics*, April 2002, Vol. 20, No. 2.

Shimer, Robert (2012). "Reassessing the ins and outs of unemployment." *Review of Economic Dynamics*, 2012, volume 15, pp. 127-148.