JOLTS Overview

Introduction

The Job Openings and Labor Turnover Survey (JOLTS) tells us how many job openings there are each month, how many workers were hired, how many quit their job, how many were laid off, and how many experienced other separations (which includes worker deaths).

The JOLTS survey design is a stratified random sample of 20,700 nonfarm business and government establishments. The sample is stratified by ownership, region, industry sector, and establishment size class. The establishments are drawn from a universe of over 9.4 million establishments compiled by the Quarterly Census of Employment and Wages (QCEW) program which includes all employers subject to state unemployment insurance laws and federal agencies subject to the Unemployment Compensation for Federal Employees program.

Employment estimates are benchmarked, or ratio adjusted, monthly to the strike-adjusted employment estimates of the Current Employment Statistics (CES) survey. A ratio of CES to JOLTS employment is used to adjust the levels for all other JOLTS data elements.

JOLTS data provide information on all pieces that go into the net change in the number of jobs. These components include hires, layoffs, voluntary quits, and other job separations (which includes retirements and worker deaths). Putting those components together reveals the overall (or net) change. JOLTS data provide information about the end of one month to the end of the next, whereas the monthly employment numbers provide information from the middle of one month to the middle of the next.

The JOLTS estimates also are affected by nonsampling error. Nonsampling error can occur for many reasons including: the failure to include a segment of the population; the inability to obtain data from all units in the sample; the inability or unwillingness of respondents to provide data on a timely basis; mistakes made by respondents; errors made in the collection or processing of the data; and errors from the employment benchmark data used in estimation.[1]

Data

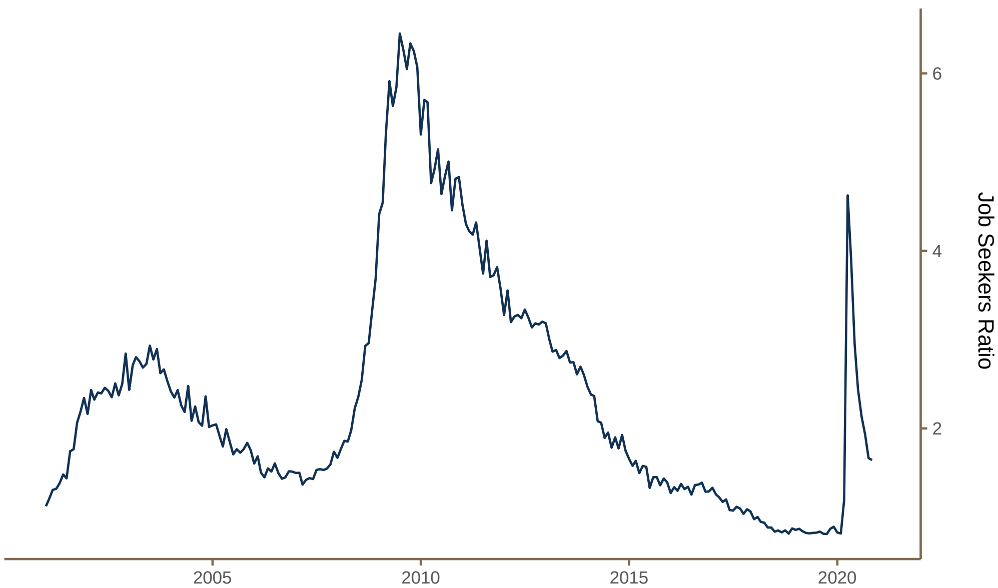
As the United States reels from the COVID-19 pandemic's catastrophic economic damage, the tight labor markets from early 2020 seem like a distant memory. Before the COVID-19 pandemic, the U.S. labor market had been steadily improving for years: the unemployment rate was falling for nearly a decade to rates not seen since the late 1960s, and witnessed the

[1] Nonsampling error occurs when a sample is surveyed rather than the entire population. Which means that there is a chance that the sample estimates may differ from the true population values they represent. The difference, or sampling error, varies depending on the particular sample selected. This variability is measured by the standard error of the estimate. BLS analysis is generally conducted at the 90-percent level of confidence. That means that there is a 90-percent chance, or level of confidence, that an estimate based on a sample will differ by no more than 1.6 standard errors from the true population value because of sampling error. Sampling error estimates are available at the BLS' website

longest streak of private sector job creation on record. Given these impressive headline statistics, it may come as a surprise that unemployed workers had slightly more trouble finding a job than they did at the peak of the last business cycle (in 2006) and have a much lower probability of finding a job than in 2000.

One of the more useful indicators is the job seekers ratio, that is, the ratio of unemployed workers to job openings. On average, there were 10.7 million unemployed workers while there were only 6.5 million job openings. This translates into a job seeker ratio of about 1.6 unemployed workers to every job opening. Another way to think about this: for every 16 workers who were officially counted as unemployed, there were only available jobs for 10 of them.

Figure 1, Job Seeker Ratio



Source: Bureau of Labor Statistics: Job Openings and Labor Turnover Survey

During the late 1990s, roughly 30–35 percent of the unemployed found a job within a month; in 2006, just prior to the Great Recession, the rate peaked at an annual average of 28.3. Over the course of the recession and for many years afterwards, the probability that an unemployed worker found a job would remain relatively low, with an annual average of 17.0 percent in 2010. Only over the last two years has the probability risen back into the 25–30 percent range, and it stood at 27.7 percent in 2018.

The job-finding rate moves inversely with the unemployment rate, though the correlation is not perfect. As the unemployment rate initially began to decline at the beginning of the recovery from the Great Recession, the job-finding probability did not increase as quickly. The sustained depression in job-finding was a primary factor that prolonged high U.S. unemployment; today, a relatively high job-finding rate helps keep unemployment low.

However, the job-finding rate is still slightly lower than in 2006 despite an unemployment rate that is nearly a percentage point lower today. To better understand how the job-finding rate has changed over time and why it remains somewhat lower than one might expect, we look at how the probability of finding a job changes with length of unemployment.

Figure 2, Unemployment & Vacany Rate

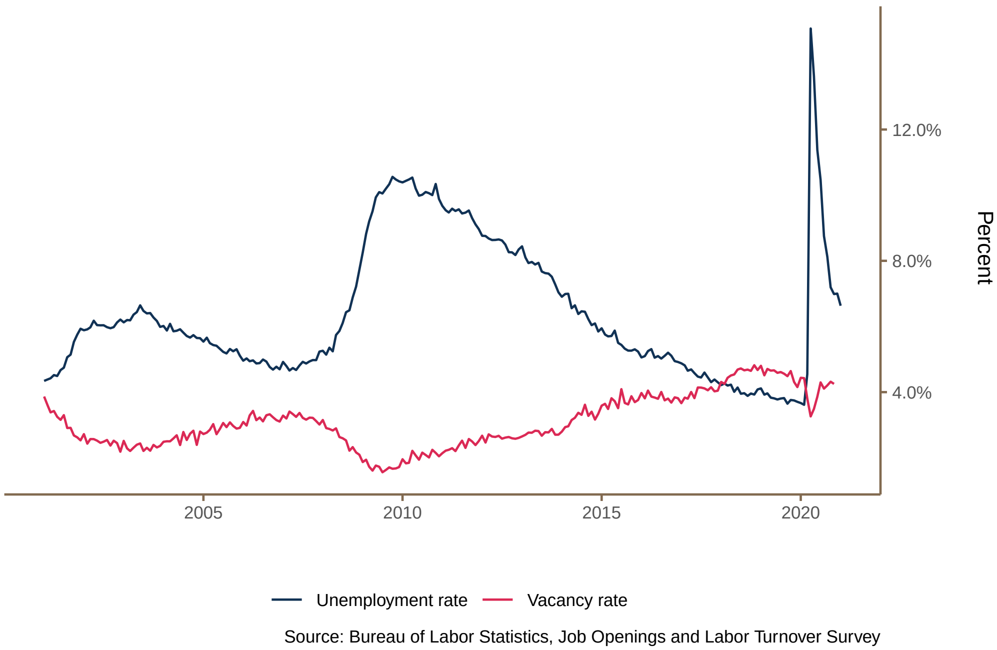
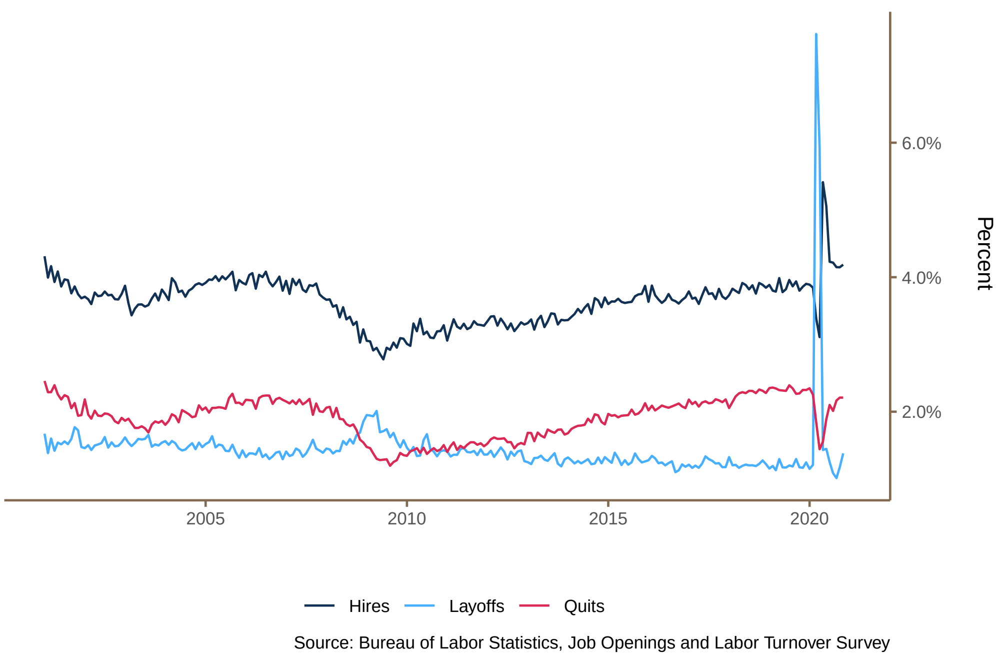
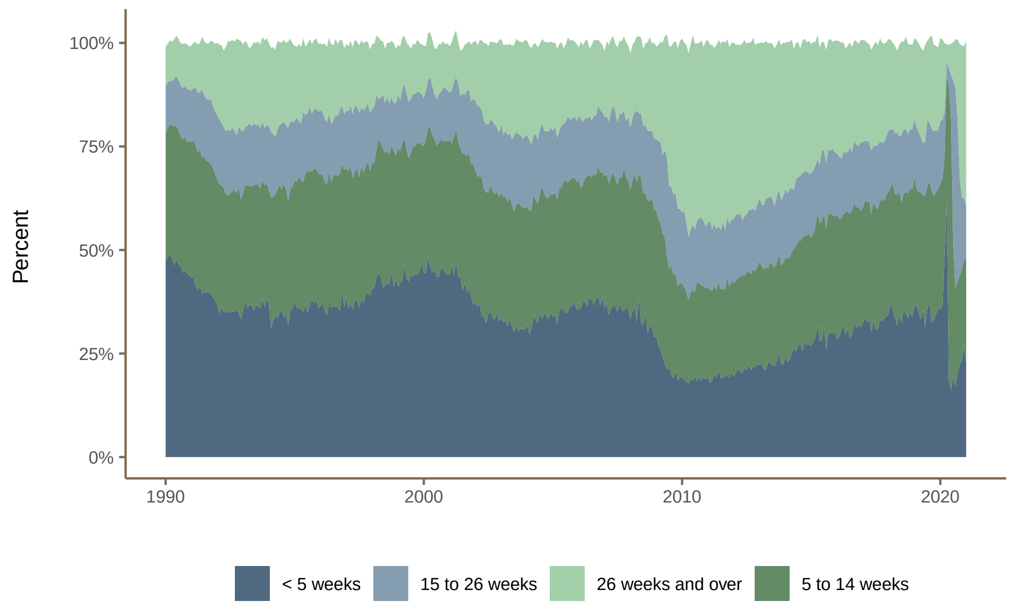


Figure 3, Hires, Quits, & Layoffs



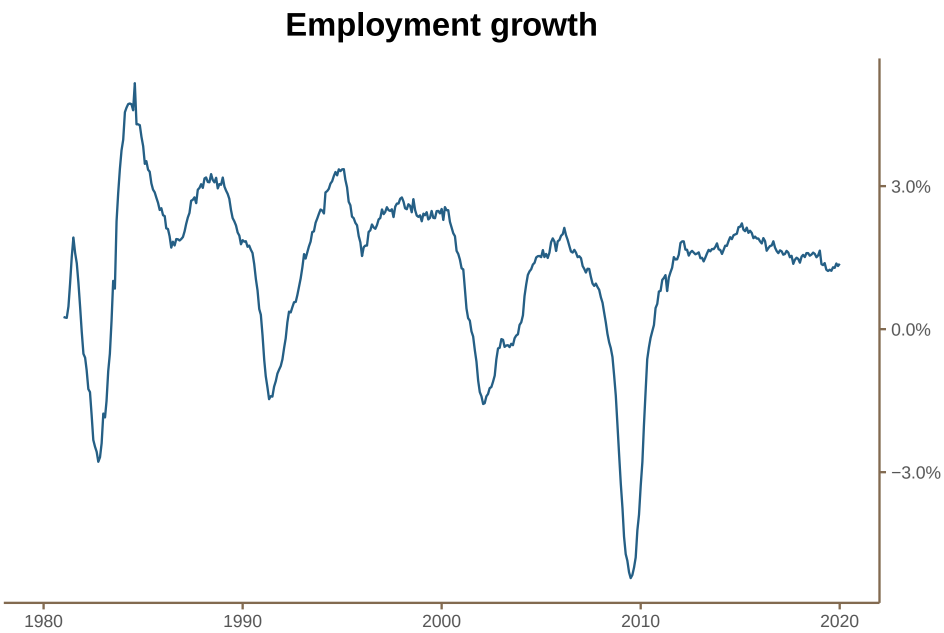
What makes the COVID-19 recession unique is the nature of the initial shock. While past recessions have been primarily caused by economic or financial shocks, the adverse shock to the labor market in 2020 was biological in nature, triggered by a novel virus that forced millions of employees into temporary unemployment by the second quarter of 2020. The record-level rise in temporary unemployment contrasts starkly with past recessions that typically start with an increase in permanent layoffs (Elsby et al., 2010). The path of job vacancies has also been unusual: while vacancies fell throughout the first half of 2020, the drop was much less pronounced than is typical in most recessions. In fact, vacancies at their lowest level were equal to the level that prevailed in 2015, a time typically considered to be a tight labor market. Thus, while the Beveridge curve – the negative relationship between vacancies and unemployment – typically "loops around" during and after a recession, in the early months of the COVID-19 recession, the increase in the unemployment rate was much larger than the corresponding drop in job vacancies.

Figure 4, Unemployment Distribution by Duration



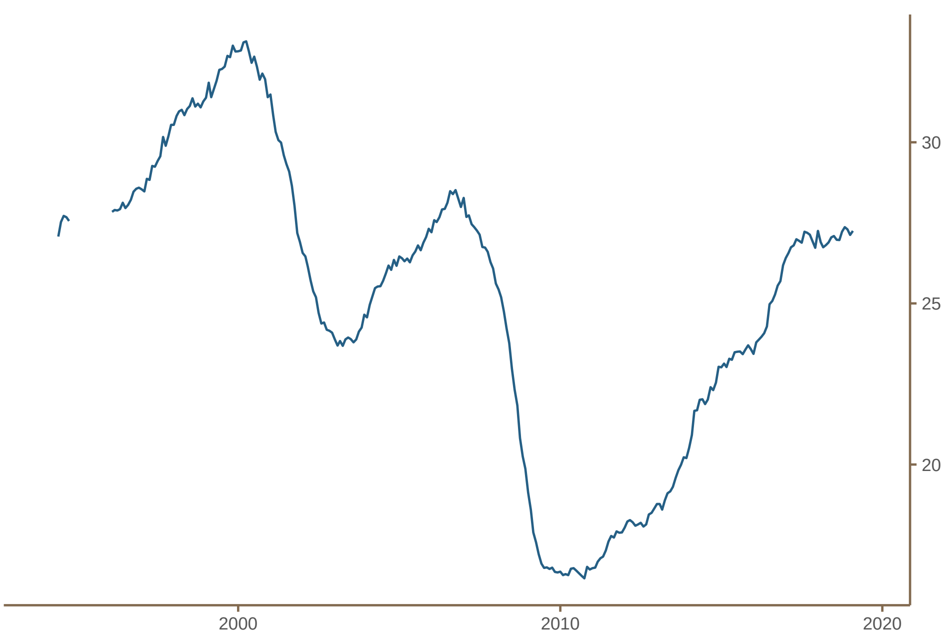
Source: Bureau of Labor Statistics, Job Openings and Labor Turnover Survey

One of the most direct measures of declining labor market dynamism is the rate of job creation. Job creation combines the employment gains at new and growing establishments. While there has been some cyclical fluctuation, job creation as a share of employment has been on a long downward trend since the early 1990s (Davis and Haltiwanger, 2014). At the same time, workers are increasingly less likely to switch jobs. This decline matters for wage growth. First, at least one-third of all hires are made among those already employed, suggesting that job switching is a major part of how workers' careers evolve. Second, part of the decline in hiring comes from the decline in job switching. Indeed, more than 40 percent of the decline in hires and separations can be ascribed to declining job-to-job transitions (Hyatt and Spletzer, 2013). Given that workers generally receive a raise when they transition directly from one job to another, declining job switching has put downward pressure on wage growth.

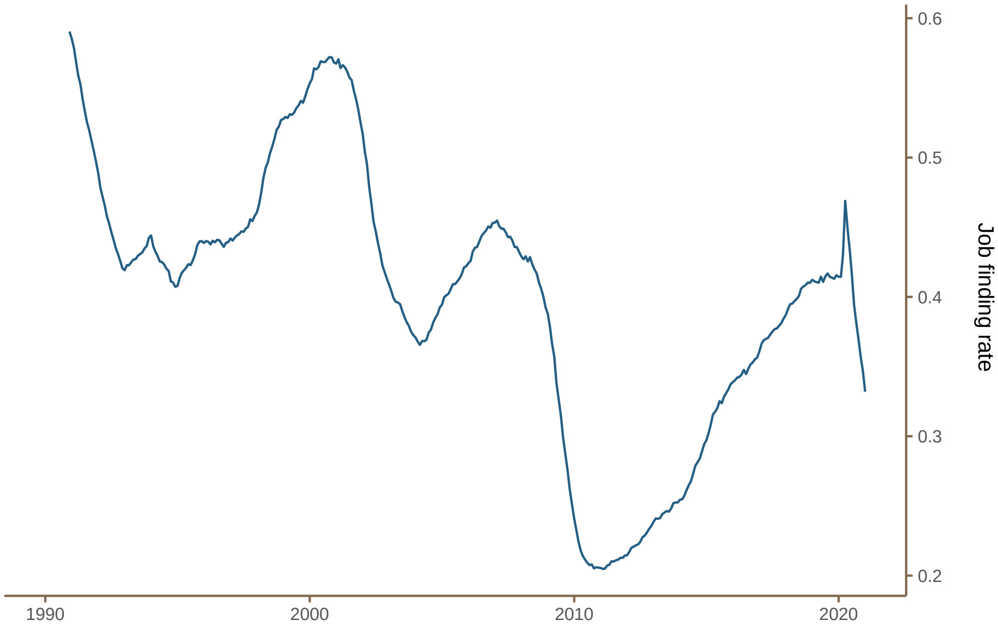


Source: Current Population Survey, Bureau of Labor Statistics

These are not the only statistical measures showing declining flexibility in the U.S. labor market. There have been substantial declines in dynamism—sometimes referred to as labor market fluidity—across a variety of related measures. When job creation, job destruction, job switching, interstate migration, and other indicators of fluidity are combined, Molloy et al. (2016) find that labor market fluidity has been on a downward trend since at least the 1980s, and has fallen by 10 to 15 percent since the 1990s.

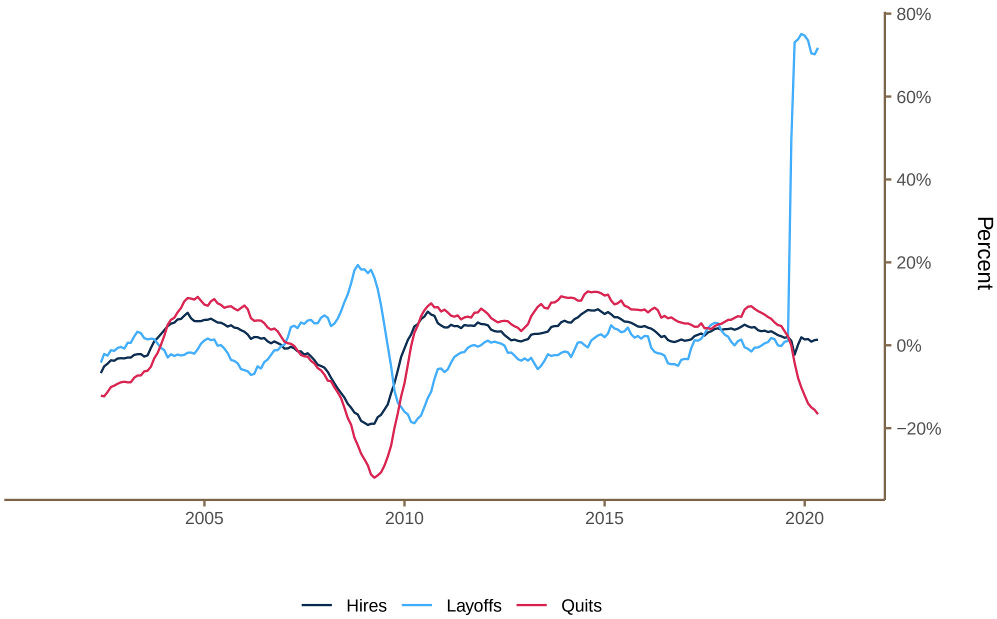


Source: Current Population Survey, Bureau of Labor Statistics



Source: Current Population Survey, Bureau of Labor Statistics

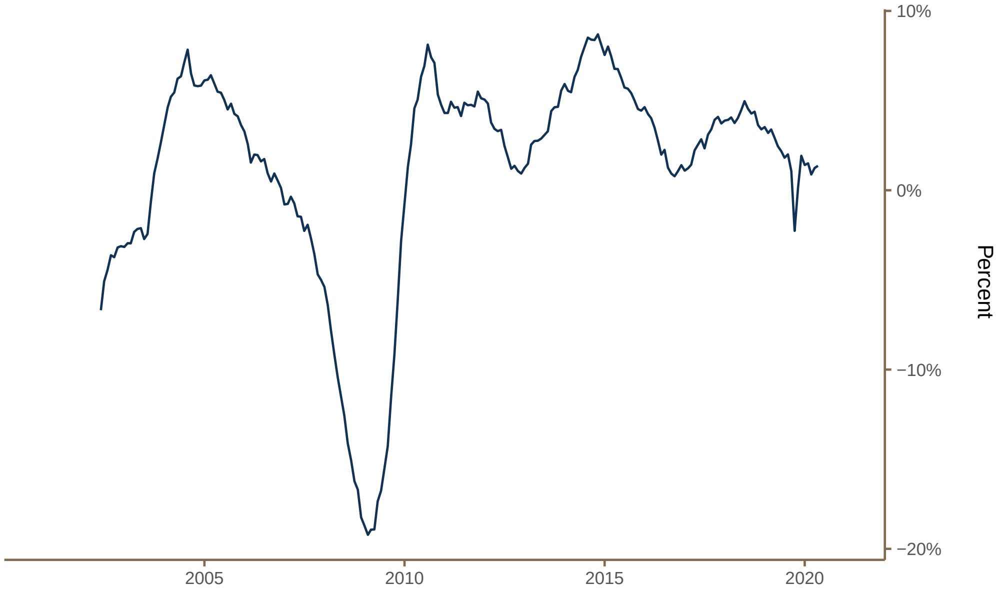
Figure 4, Hires, Quits, & Layoffs (12−month rolling average)



Source: Bureau of Labor Statistics, Job Openings and Labor Turnover Survey

The Great Recession obscures the long-run trend of the data.

Figure 5, Hiring Rate (12−month rolling average)



Source: Bureau of Labor Statistics, Job Openings and Labor Turnover Survey

Adjusting the window to the period after the Great Recession we can see the more recent trend.

Figure 6, Hiring Rate (12−month rolling average)

Source: Bureau of Labor Statistics, Job Openings and Labor Turnover Survey

Smoothing the signal of the series further we can pick out the declining trend in the hiring rate.

Figure 7, Hiring Rate (24−month rolling average)

Source: Bureau of Labor Statistics, Job Openings and Labor Turnover Survey

The unusual nature of the COVID-19 recession makes it difficult to draw on experiences from past recessions to project how the labor market will evolve in the months ahead. For example, during the Great Recession, the initial wave of layoffs was subsequently followed by a prolonged period of lower job finding rates (Elsby, 2009). This led to a significant increase in the long-term unemployment share, which in turn prolonged the recession through negative duration dependence (Krueger, Cramer and Cho, 2014). Currently available data suggest that the dynamics of the COVID-19 recession may play out quite differently. In particular, vacancies continue to be elevated, and job finding rates have not decreased substantially – both the recall rates for the temporary unemployed, as well as the job finding rates for workers who have been permanently laid off.

The pandemic is an urgent reminder of what long-term labor trends have been illustrating for years: low-wage workers need better pathways into decent jobs, and from shrinking occupations to the jobs of tomorrow. Policymakers face a dual imperative: to facilitate safe reemployment as soon as possible, even as COVID-19 continues to surge in many parts of the country, while also helping low-wage workers on the journey to jobs with dignity, stability, and a fair shot at economic mobility. While the risk of mass unemployment has already spurred large public expenditures, more funding and efforts are needed to ensure opportunity reaches those who need it the most.

Methods

We produce forecasts using a collection of traditional and non-traditional time series methods. This section provides a general overview of the methods used, their benefits, and their limitations. It cannot be emphasized enough that, no matter the strength of a model, it remains exactly that: a model. As such, all statistical models are "wrong." No matter the method used, any model is an attempt to reproduce ("model") the true data generating process of a data series.

Background

For many familiar with the methodology of time series forecasting, this section presents a brief recapitulation of basic concepts which can be found in any standard textbook such as Montgomery, Jennings and Kulachi (2015) or Brockwell and Davis (2002). For a more advanced treatment, one should see Hamilton (1994) or Brockwell and Davis (2006).

Time series analysis is the procedure of using known data values to fit a time series with a suitable model and estimation of the corresponding parameters. It comprises methods that attempt to understand the nature of the time series.

A major assumption that often provides relief in modeling efforts is the linearity assumption. A linear filter, for example, is a linear operation from one time series x_t to another time series y_t.

$$E[y_t] = L(x_t) = A(L^i)\epsilon_t = \left(\sum_{-\infty}^{\infty} \psi_i L^i\right)\epsilon_t = \mu + \sum_{-\infty}^{\infty} \psi_i \epsilon_{t-i},$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$ and L^i is the Lag or Backshift Operator defined as,

$$L^i x_t = x_{t-i} \; \forall \, i \in \mathbb{N}$$

the linear filter can be seen as a process that converts the input, x_t, into an output, y_t with a conversion process that involves all (present, past, and future) values of the input in the form of a summation with different "weights", ψ_i, on each observation x_i. Specifically, y_t can be expressed as

$$y_t = \mu + \sum_{i=o}^{\infty} \psi_i \epsilon_{t-i}$$

where the weights are conventionally some form of moving average M. A moving average M of span N assigns weights $\frac{1}{N}$ to the most recent observations such that the estimate can be written as

$$M_t = \frac{1}{N} \sum_{t=T-N-1}^{N} y_t$$

The covariance between y_t and its value at another point in time y_{t+k} is called the auto-covariance at lag k, and is defined by

$$\gamma_k = \mathrm{Cov}(y_t, y_{t+k}) = E\left[(y_t - \mu)(y_{t+k} - \mu)\right]$$

the autocovariance of the series at lag $k = 0$ is simply the variance of the time series. The autocorrelation coefficient at k is then

$$
\begin{aligned}
\rho_k &= \frac{E[(y_t-\mu)(y_{t+k}-\mu)]}{\sqrt{E[(y_t-\mu)^2]E[(y_{t+k}-\mu)^2]}} \\
&= \frac{\mathrm{Cov}(y_t, y_{t+k})}{\mathrm{Var}(y_t)} \\
&= \frac{\gamma_k}{\gamma_0}
\end{aligned}
$$

The collection of these values ρ_k is called the autocorrelation function. For a finite time series, it is necessary to estimate the autocovariance and autocorrelation functions. A usual estimate of the autorcovariance function is

$$
c_k = \hat{\gamma_k} = \frac{1}{T}\sum_{t=1}^{T-k}(y_t-\bar{y})(y_{t+k}-\bar{y}), \quad k = 1, 2, 3, \ldots, K
$$

and the autocorrelation function is estimated by the sample autocorrelation function (sample ACF)

$$
r_k = \hat{\rho_k} = \frac{c_k}{c_o}
$$

The stationarity of a time series is related to its statistical properties in time. That is, in the more strict sense, a stationary time series exhibits similar "statistical behavior" in time and this is often characterized as a constant probability distribution in time. However, it is usually satisfactory to consider the first two moments of the time series and define stationarity (or weak stationarity) as follows: (1) the expected value of the time series does not depend on time and (2) the autocovariance function defined as $\mathrm{Cov}(y_t, y_{t+k})$ for any lag k is only a function of k and not time: that is, $y(k) = \mathrm{Cov}(y_t, y_{t+k})$. Weak stationarity, then, implies that the mean of the series does not change with time t and that the autocovariance function, $\gamma(k,t)$, depends on k and t only through their difference $k - t$.

For a time-invariant and stable linear filter a stationary input time series $y(t)$ can be written more succinctly as,

$$
\begin{aligned}
y(t) &= \mu + \sum_{i=0}^{\infty}\psi_i\epsilon_t \\
&\mu + \sum_{i=0}^{\infty}\psi_i L^i\epsilon_t \\
&\mu + \Psi(L)\epsilon_t
\end{aligned}
$$

where

$$
\Psi(L)\epsilon_t = \sum_{i=0}^{\infty}\psi_i L^i
$$

The infinite moving average model has a covariance function in the form,

$$
\mathrm{Cov}(y_t, y_{t+k}) = \gamma_y(k,t) = \sum_{i=-\infty}^{\infty}\sum_{j=-\infty}^{\infty}\psi_i\psi_j\gamma_x(i-j+k)
$$

The autocovariance function of y_t is then,

$$\gamma_y(k,t) = \sum_{i=o}^{\infty}\sum_{j=0}^{\infty}\psi_i\psi_j\gamma_\epsilon(i-j+k)$$

$$= \sigma^2\sum_{i=0}^{\infty}\psi_i\psi_{i+k}$$

This representation comes from a theorem by Wold and which essentially states that any nondeterministic weakly stationary time series y_t can be represented as in the above manner. A more intuitive interpretation of this theorem is that a stationary time series can be seen as the weighted sum of the present and past random "disturbances."

Although very powerful in providing a general representation of any stationary time series, the infinite moving average model given in is not very useful in practice except for certain special cases.

In finite order moving average or MA models, weights that are not set to 0 are represented by the Greek letter θ with a minus sign in front.

$$y_t = \mu + \left(1 - \sum_{i=1}^{q}\theta_i L^i\right)\epsilon_i$$

$$= \mu + \Theta(L)\epsilon_t$$

where $\Theta(L)\epsilon_t = 1 - \sum_{i=1}^{q}\theta_i L^i$

An interpretation of the finite order MA processes is that at any given time, of the infinitely many past disturbances, only a finite number of the disturbances "contribute" to the current value of the time series and that the time window of the contributors "moves" in time, making the "oldest" disturbance obsolete for the next observation. It is indeed not too far fetched to think that some processes might have these intrinsic dynamics. However, for some others, we may be required to consider the "lingering" contributions of the disturbances that happened back in the past. This will of course bring us back to square one in terms of our efforts in estimating infinitely many weights. Another solution to this problem is through autoregressive models in which the infinitely many weights are assumed to follow a distinct pattern and can be successfully represented with only a handful of parameters. We shall now consider some special cases of autoregressive processes.

ARMA

Autoregressive models are based on the idea that current value of the series, x_t can be explained as a linear combination of past values, $x_{t-1}, x_{t-2}, \ldots, x_{t-p}$, together with some random error ϵ. In other words, a series x_t is linearly dependent upon its past values, the degree to which depends on the lag order p.

Let us first consider again the time series

$$y(t) = \mu + \sum_{i=0}^{\infty} \psi_i \epsilon_t$$

$$\mu + \sum_{i=0}^{\infty} \psi_i L^i \epsilon_t$$

$$\mu + \Psi(L)\epsilon_t$$

One approach to modeling this time series is to assume that the contributions of the disturbances that are way in the past should be small compared to the more recent disturbances that the process has experienced. Since the disturbances are independently and identically distributed random variables, we can simply assume a set of infinitely many weights in descending magnitudes reflecting the diminishing magnitudes of contributions of the disturbances in the past.

A simple and yet intuitive set of such weights can be created following an exponential decay pattern. For that we will set $\psi_i = \phi^i$, where $|\,\phi\,| < 1$ to guarantee the exponential "decay." In this notation, the weights on the disturbances starting from the current disturbance and going back in past will be $1, \phi^2, \phi^3, \ldots$

$$y_t = \mu + \epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \cdots$$
$$= \mu + \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i}$$

We also have

$$y_{t-1} = \mu + \epsilon_{t-1} + \phi\epsilon_{t-2} + \phi^2\epsilon_{t-3} + \cdots$$
$$= \mu + \phi \left(\sum_{i=0}^{\infty} \phi^i \epsilon_{t-1-k} + \epsilon_t \right)$$

combining these two equations together we can represent an AR(1) model as a linear process given by

$$\mu + \sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} = \mu + \phi \left(\sum_{i=0}^{\infty} \phi^i \epsilon_{t-i} + \epsilon_t \right)$$

$$\mu + \epsilon_t + \sum_{i=0}^{\infty} \phi^i y_{t-i} + \phi\mu$$

$$\mu - \phi\mu + \sum_{i=0}^{\infty} \phi^i y_{t-i} + \epsilon_t$$

$$\delta + \sum_{i=0}^{\infty} \phi^i y_{t-i} + \epsilon_t$$

where $\delta = (1-\phi)\mu$. A more general way of writing an autoregressive procress is through operator notation as

$$\begin{aligned}\epsilon_t &= \left(1 - \sum_{i=1}^{i} \phi_i L^i\right) y_t \\ &= \Phi(L) y_t\end{aligned}$$

where $\Phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i$

The above process in is called a first-order autoregressive process, AR(l), because it can be seen as a regression of y_t on y_{t-1} and hence the term autoregressive process.

Autoregressive–moving-average (ARMA) models provide a parsimonious description of a (weakly) stationary stochastic process in terms of two polynomials, one for the autoregression (AR) and the second for the moving average (MA). The general ARMA model was described in the 1951 dissertation of Peter Whittle, Hypothesis Testing in Time Series, which generalized Wold's autoregressive representation theorem for univariate stationary processes to multivariate processes, and was popularized in by Box and Jenkins in their classic book Box and Jenkins (1970).

$$y_t = \delta + \sum_{i=0}^{p} \phi^i y_{t-i} + \epsilon_t - \sum_{i=1}^{q} \theta_i \epsilon_{i-1}$$

$$\Psi(L) y_t = \delta + \Theta(L) \epsilon_t$$

which has an infinite moving average process $\mathrm{MA}(q)$

$$y_t = \mu + \sum_{i=o}^{\infty} \psi_i \epsilon_{t-i}$$

ARIMA

The Autoregressive Integrated Moving Average (ARIMA) model is a generalization of an Autoregressive Moving Average (ARMA) model to overcome the possible violation of the assumption that the series x_t is stationary. In order for inferences drawn from the estimates to be reliable when the data are not stationary, the data must be differenced some number of times d to remove seasonality and trend. If a stochastic process has to be differenced d times to reach stationarity, it is said to be integrated of order d or $I(d)$.

A standard convention for expressing the ARIMA model is $\mathrm{ARIMA}(p, d, q)$, where:

- p denotes the number of lag observations included in the model, also called the lag order.
- d denotes the number of times that the series is differenced, also called the degree of differencing.
- q denotes the size of the moving average window.

We will call a time series homogeneous, non-stationary if it is not stationary but its first difference, that is, $w_t = y_t - Y_{t-1} = (1-L)y_t$,or higher-order differences, $w_t = (I-L)^d y_t$ produce a stationary time series. We will further call y_t an autoregressive integrated moving average (ARIMA) process of orders p, d, and q $\mathrm{ARIMA}(p,d,q)$ if its d-th difference, denoted by $w_t = (I-L)^d y_t$ produces a stationary $\mathrm{ARMA}(p,q)$ process. The term integrated is used since, $d=1$, for example, we can write y_t as the sum (or "integral") of the w_t process as

$$y_t = \delta + \sum_{i=0}^{p} \phi^i y_{t-i} + \epsilon_t - \sum_{i=1}^{q} \theta_i \epsilon_{i-1}$$

$$\left(1 - \sum_{i=0}^{p} \phi^i L^i\right) y_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \epsilon_t$$

$$\left(1 - \sum_{i=0}^{p} \phi^i L^i\right) = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) (1-L)^d$$

$$\left(1 - \sum_{i=0}^{p} \phi^i L^i\right) (1-L)^d \, y_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \epsilon_t$$

$$\Phi(L)(1-L)^d y_t = \delta + \Theta(L)\epsilon_t$$

Selection of the hyperparameters (p,d,q) are often chosen through an inspection of the ACF and PACF, but can also be selected through AIC (Akaike Information Criterion), AICc (corrected AIC) and BIC (Bayesian Information Criterion). But note that the selection of the hyperparameters is not unique.

SVM

RNN

Random Forest

Unit Root

A unit root is a feature of some stochastic processes (such as random walks) that can cause problems in statistical inference involving time series models. A linear stochastic process has a unit root if 1 is a root of the process's characteristic equation. Shocks to a unit root process have permanent effects which do not decay as they would if the process were stationary. The characteristic roots (roots of the characteristic equation) also provide qualitative information about the behavior of the variable whose evolution is described by the dynamic equation. For a differential equation parameterized on time, the variable's evolution is stable if and only if the real part of each root is negative. For difference equations, such as a standard time series, there is stability if and only if the absolute value of each root is less than 1.

An augmented Dickey–Fuller test (ADF) tests the null hypothesis that a unit root is present in a time series sample. The alternative hypothesis is different depending on which version of the test is used, but is usually stationarity or trend-stationarity.

The intuition behind the test is that if the series is characterised by a unit root process then the lagged level of the series (y_{t-1}) will provide no relevant information in predicting the change in y_t besides the one obtained in the lagged changes. The OLS estimate (based on an n-observation time series) of the autocorrelation parameter ρ is given by

$$\hat{\rho}(n) = \frac{\sum\limits_{t=1}^{n} y_{t-1}y_t}{\sum\limits_{t=1}^{n} y_t^2}$$

if $|\rho| < 1$ then

$$\sqrt{n}\,(\hat{\rho} - \rho) \implies \sim N(0, 1 - \rho^2)$$

To compute the test statistics, we fit the augmented Dickey–Fuller regression

$$\Delta y_t = \alpha + \beta y_{t-1} + \delta + \sum_{j=1}^{k} \psi_i \Delta y_{t-i} + \epsilon_t$$

Depending on the specifications, the constant term α or time trend δ is omitted and k is the number of lags specified.

References

Box, George, and Gwilym Jenkins. 1970. *Time Series Analysis: Forecasting and Control*. San Francisco, CA:Holden-Day.

Brockwell, Peter J., and Richard A. Davis. 2002. *Introduction to Time Series and Forecasting*. . 2 ed., New York, NY:Springer.

Brockwell, Peter J., and Richard A. Davis. 2006. *Time Series: Theory and Methods*. . 2 ed., New York, NY:Springer.

Davis, Steven, and John Haltiwanger. 2014. "Labor Market Fluidity and Economic Performance." National Bureau of Economic Research Working Paper 20479.

Elsby, Michael. 2009. "Evaluating the Economic Significance of Downward Nominal Wage Rigidty." *Journal of Monetary Economics*, 56(2): 154–169.

Elsby, Michael, Bart Hobijn, , and Aysegul Sahin. 2010. "The Labor Market in the Great Recession." *Brookings Papers on Economic Activity*, 2010: 1–70.

Hamilton, James D. 1994. *Time Series Analysis*. Princeton, NJ:Princeton University Press.

Hyatt, Henry, and James Spletzer. 2013. "The recent decline in employment dynamics." *IZA Journal of Labor Economics*, 2(5): 1–24.

Krueger, Alan, Judd Cramer, and David Cho. 2014. "Are the Long-Term Unemployed on the Margins of the Labor Market?" *Brookings Papers on Economic Activity*, 2014(Spring).

Molloy, Raven, Christopher L. Smith, Riccardo Trezzi, and Abigail Wozniak. 2016. "Understanding Declining Fluidity in the U.S. Labor Market." *Brookings Papers on Economic Activity*, 2016(Spring).

Montgomery, Douglas, Cheryl Jennings, and Murat Kulachi. 2015. *Introduction to Time Series Analysis and Forecasting.* . 2 ed., Cambridge, MA:Wiley.