Discussion of Hall and Kudlyak

Robert Shimer
University of Chicago

July 11, 2021

In this provocative paper, Hall and Kudlyak show that during a typical recession in the United States from 1949 to 2019, a short-lived burst of job loss explains much of the rapid increase in the unemployment rate. During the subsequent expansion, unemployment declines comparatively slowly and steadily, primarily due to a sharp and persistent decline in the job finding rate, both for the initial group of job losers and for other workers who became unemployed only later in the business cycle. Hall and Kudlyak argue that the elevated jobless rate for the latter group is evidence that unemployment is “contagious” or “infectious.”

For the most part, I will not take issue with their facts, although I will make the (well-known) observation that the 2020 pandemic recession and subsequent expansion featured the fastest increase in unemployment on record, followed by the fastest decrease on record. While this recession was different in many ways from past ones, the fact that unemployment fell so quickly during the early stages of this expansion may be useful for diagnosing why unemployment declined so slowly during prior expansions.

I will focus my attention on the claim that unemployment is contagious or infectious. Hall and Kudlyak are very clear about what they mean by this: “We consider negative feedback from high unemployment to the job finding rate as a key mechanism behind slow unemployment recoveries.” The function $\gamma(u)$ in equation (14) exemplifies this logic. Section 8 of their paper sketches a number of endogenous mechanisms that can generate such feedback. I will not try to critique each — or indeed any — of these mechanisms, since the authors only sketch the mechanisms and a serious critique of all of them would require a book-length treatise. Indeed, I acknowledge that some of these mechanisms may be empirically relevant.

Instead, I will disagree with Hall and Kudlyak’s conclusion that their analysis establishes that there must be a negative feedback loop from unemployment to the job finding rate. They reach this conclusion using the following syllogism: a short-lived initial shock dumps some workers into unemployment but does not directly affect other workers. In the subsequent years, there is elevated unemployment for workers who are initially unaffected by the shock.
Therefore, it must be the case that the shock affected them indirectly, through a contagious effect of high unemployment for the initial job losers on the job finding rate for the remaining workers.

There are two problematic suppositions here. The first is that the initial shock is short-lived. While it is true that the baseline Diamond-Mortensen-Pissarides model generates almost no internal propagation, the shock itself may persist for some time after the onset of the recession. For example, credit markets did not magically revert to their pre-financial crisis state when the NBER declared the recession over in June 2009. I won’t dwell here on that possibility, but of course persistence of shocks is the standard assumption in virtually all models of aggregate fluctuations, including those with unemployment (Merz, 1995; Andolfatto, 1996; Shimer, 2005, 2010).

The second problematic supposition is that the initial shock only affects workers who immediately experience unemployment. While I recognize that this assumption is standard in certain versions of the Diamond-Mortensen-Pissarides framework, it strikes me as useful in some applications, but also as empirically untenable. I will argue that a different type of shock, one that lowers the value of occupation-specific human capital for many workers, is both plausible and consistent with the Hall-Kudlyak facts. Such a shock will do two things. First, it will immediately move some workers into unemployment, with a prolonged search for a new occupation. This is the initial burst of unemployment that the authors emphasize. But second, it will also lead to subsequent elevated job loss for other now-vulnerable workers, even after the economy starts to recover. As these workers also search for a new occupation, the economy may experience a prolonged period with a (slightly) elevated job loss rate and a (significantly) depressed job finding rate. I will illustrate this possibility through a simple quantitative model, deliberately constructed to have no possibility of contagion, and argue that the key mechanism is consistent with a number of other facts that we know about how labor markets function over the business cycle.

The rest of my discussion will proceed in three steps. First, I will use some data which complements Hall and Kudlyak’s to show how movements in job loss and job finding rates influence unemployment over the business cycle, emphasizing not just the counter-cyclicality of the job loss rate, but also the fact that once it is elevated, it stays high for many years. Second, I will describe my model, which features two distinct types of job loss, one of which is associated with a loss of occupation-specific human capital and a significant deterioration in long-run outcomes. Third, I will argue that the mechanism in my model is consistent with evidence on the concentration of secular shifts in industrial composition during recessions (Lilien, 1982; Jaimovich and Siu, 2020), with evidence on the long-run wage consequences of job loss during recessions (Davis and von Wachter, 2011), and with the behavior of the
1 Worker Transitions over the Business Cycle

I follow a well-known methodology using matched files from the Current Population Survey to construct a measure of the flow of workers between employment (E), unemployment (U), and not-in-the-labor force (N); see for example Shimer (2012). In each month, this gives me the share of workers in one state transitioning to another state, so for example $\lambda_{EU}^t$ is the share of employed workers in month $t - 1$ who are unemployed in month $t$. I let $\Lambda$ denote the resulting $3 \times 3$ transition matrix,

$$
\Lambda_t = \begin{pmatrix}
1 - \lambda_{EU}^t - \lambda_{EN}^t & \lambda_{EU}^t & \lambda_{EN}^t \\
\lambda_{UE}^t & 1 - \lambda_{UE}^t - \lambda_{UN}^t & \lambda_{UN}^t \\
\lambda_{NE}^t & \lambda_{NU}^t & 1 - \lambda_{NE}^t - \lambda_{NU}^t
\end{pmatrix}.
$$

Note that each row sums to 1. I focus here on data since 2008 in order to avoid over-cluttering the graphs, but the basic points I make here hold in the period when I can construct these flows, since 1967.

I do two exercises with this transition matrix. First, if the economy were in steady state, with constant flows from month to month, $\Lambda_t = \Lambda$, the share of workers in each state would converge to a constant, given by the eigenvector associated with the matrix $\Lambda$’s unit eigenvalue. I compare that implied steady state unemployment rate $u_t^*$ with the actual unemployment rate $u_t$ in each month. Figure 1 shows that the two are very similar, with $u_t^*$ slightly larger than $u_t$ during the two recession periods and slightly smaller during the subsequent expansions. The gaps reflect the fact that unemployment rises (falls) when the flows push it towards a higher (lower) rate. But more importantly for my purposes here, the two lines are similar because the underlying transition rates change relatively slowly over time.

Second, I construct counterfactual unemployment rates to examine the effect of each of the six transition rates in turn. More precisely, I allow just one of the six transition rates in equation (1) to vary over time and hold the other five fixed at their average value. I then look at the eigenvector associated with the unit eigenvalue of the resulting matrix to construct a counterfactual unemployment rate. The green lines in Figure 2 shows that the bulk of the increase in unemployment during and after the 2008–2009 recession is accounted for by the (decline in) the unemployment-to-employment transition probability $\lambda_{UE}^t$, which I refer to also as the job finding rate. Interestingly, however, an early increase in the employment-
Figure 1: The blue line shows the quarterly average unemployment rate from 2008Q1–2021Q2 in the United States. The red line shows the steady state unemployment rate implied by worker flows.

to-unemployment transition probability $\lambda_{EU}^t$, which I call here the job loss rate, accounts for additional upward pressure on the unemployment rate, pushing it up to 6.3 percent by 2009Q1, before a slow and steady decline in job loss pulled it back down to 4.4 percent ten years later. Changes in the other flows, i.e. movements in and out of the labor force, had less of an impact on the steady state unemployment rate.

The story of the pandemic recession in 2020 is dramatically different but equally informative. In April 2020, the job loss rate $\lambda_{EU}^t$ spiked, explaining virtually all of the increase in unemployment. It declined almost as rapidly, so by June 2021, elevated job loss contributed less to the unemployment rate than did the low job finding rate. Still, as I write this comment, it appears that the economy has settled into a fairly normal (and comparatively slow) recovery with unemployment hovering near six percent.

I will argue below that the difference between the pandemic recession and “normal” recessions is that during the pandemic, most workers understood that they could return to their old job, or at least their old occupation, after the worst of the pandemic had passed. In contrast, other recessions are associated with more long-term displacement as workers leave their occupation and gradually retrain for other jobs. This slow process leads to a depressed job finding rate for years after the initial shock. I conjecture that a similar phenomenon will likely arise for those workers lingering in unemployment as the United States emerges from this most recent recession.
Figure 2: Each panel shows the steady state unemployment rate implied by worker flows in red and the part explained by changes in one transition rate in green. United States data, 2008Q1–2021Q2.
2 Two Types of Unemployment

I now describe a mechanical model where workers can be either employed or unemployed and can have either high or low human capital. Idiosyncratic shocks move workers between the two employment statuses and between the two human capital statuses, with a positive correlation between job loss and human capital loss, as in the Ljungqvist and Sargent (1998) model of turbulence. I think of human capital as being something occupation-specific, so a high human capital worker is someone who has found an occupation that he excels at, while a low human capital worker is searching for such an occupation (Kambourov and Manovskii, 2009; Alvarez and Shimer, 2011).

I assume that unemployed workers with high (low) human capital find a job at rate $f_h$ ($f_l$), while employed workers with high (low) human capital lose their jobs at rates $s_h$ ($s_l$), all constant and exogenous. Finally, employed workers with low human capital acquire human capital at rate $\gamma$, while a fraction $\gamma$ of job losses for high human capital workers are associated with a loss of human capital, the Ljungqvist and Sargent (1998) turbulence shock.

Let $u_h(t)$ and $u_l(t)$ denote the fraction of workers who are unemployed at time $t$ with high and low human capital, respectively. Let $e_h(t)$ and $e_l(t)$ denote the corresponding employment rates. The worker flows in the previous paragraph imply these evolve according to

$$
\dot{u}_l(t) = s_l e_l(t) - f_l u_l(t) + \gamma s_h e_h(t),
$$
$$
\dot{e}_l(t) = f_l u_l(t) - s_l e_l(t) - \gamma e_l(t),
$$
$$
\dot{u}_h(t) = s_h e_h(t) - f_h u_h(t) - \gamma s_h e_h(t),
$$
$$
\dot{e}_h(t) = f_h u_h(t) - s_h e_h(t) + \gamma e_l(t).
$$

To be clear, this model is entirely and deliberately mechanical, leaving no scope for unemployment to be contagious or infectious.

I want to think quantitatively about what kinds of dynamics this model can generate, but to be clear the numbers I use here are only suggestive. I stress that there are two critical quantitative assumptions: human capital status is much more persistent than employment status, and high human capital workers typically have shorter unemployment durations and longer employment durations than low human capital workers. I think of a time period as a month. I assume that ten percent of low human capital workers switch employment status each month, $s_l = f_l = 0.1$, while only two percent of employed high human capital workers lose their job, $s_h = 0.02$, and sixty percent of unemployed high human capital workers find a job, $f_h = 0.6$. These transition rates ensure that workers with high human capital have
much lower unemployment duration, much higher employment duration, and consequently
much lower unemployment rates. Finally, I also assume $\gamma = 0.02$, so occupation-specific
human capital is highly persistent.

In steady state, 95.9 percent of workers have high human capital, while the remaining 4.1
percent of workers have low human capital. The unemployment rate is 3.2 percent for the
high human capital workers and 54.5 percent for the low human capital workers, so despite
the paucity of low human capital workers, they account for almost half of the 5.3 percent
unemployment rate. The job finding rate for the typical unemployed worker is 38.8 percent
per month, while the job loss rate is 2.2 percent per month, both reflecting the mixture of
human capital in the two pools.

Now suppose the model economy is initially in steady state but is hit by a one-time shock
which adversely impacts workers’ human capital and throws some of them into unemploy-
ment. This might be, for example, a sharp and permanent decline in the industry demanding
the worker’s skills. More precisely, I assume that ten percent of high human capital workers,
employed and unemployed, lose their human capital. In addition, half the employed workers
who lose their human capital also immediately lose their job. This combination of shocks
immediately nearly doubles the unemployment rate to 10.0 percent and, more importantly,
more than triples the share of workers with low human capital to 13.7 percent.

More interesting is the subsequent dynamics, which largely reflect the slow recovery of
the stock of human capital. On impact, the job finding rate falls by 50 log points to 23.6
percent, while the job loss rate increases less dramatically, by 18 log points to 2.6 percent.
Each then recovers monotonically but very slowly. As a result, we get a prolonged period
of high unemployment with a very low job finding rate and a slightly elevated job loss rate.
Figure 3 depicts the resulting counterfactual unemployment rates, showing persistence not
unlike the response of the United States labor market to the financial crisis depicted in
Figure 2.

Figure 4 looks at the same dynamics from a different perspective. The purple line shows
the share of the labor force that is unemployed immediately after the shock at time 0 and
also unemployed at time $t$ (possibly with one or more intervening employment spell). The
blue line shows the share that is employed immediately after the shock but unemployed at
$t$. Notably the two lines cross after six months, so workers who lost their job at the time of
the shock subsequently account for less than half of all unemployment.

If I handed data generated from this model to Hall and Kudlyak, they would note the
high unemployment rate at $t$ of workers who remained employed through the initial shock,
together with the low job finding rate for the typical unemployed worker. Putting those two
facts together, I believe they would reach the same conclusion as in Section 4.3 of their paper,
Figure 3: Model-generated data. The red lines show the steady state unemployment rate implied by worker flows. The green lines show the part accounted for by one transition rate, the job loss rate $\lambda^{EU}$ on the left and the job finding rate $\lambda^{UE}$ on the right.

Figure 4: Model-generated data. The purple line shows the unemployment rate at $t$ of workers who are unemployed at time 0. The blue line shows the unemployment rate at $t$ of workers who are employed at time 0.
“The excess job loss accounts for the magnitude of the initial increase in unemployment, but not its persistence. The persistence is too large to be explained as reflecting only the personal experiences of the extra job-losers dating from the spike.” Their syllogism would then lead them to conclude that unemployment is contagious.

But of course there is no contagion in this simple mechanical model. The persistent dynamics reflect the poor labor market experience of the ten percent of workers who lost their human capital due to the initial shock, only some of whom immediately experienced unemployment. Unemployment persistence is a manifestation of the compositional shift in the skill distribution caused by the initial adverse shock.

3 Relationship to Other Labor Market Facts

My simple model with two types of unemployment builds on a long tradition in macroeconomics which emphasizes that sectoral shifts are concentrated in recessions (Lilien, 1982). One can think of two versions of such a model. In the first, a new sector of the economy opens up, pulling workers out of old sectors and causing a brief recession due to frictional unemployment. This version of the sectoral shifts model is empirically implausible because we do not see evidence of that growth sector during any post-World War II recession, at least in the United States (Abraham and Katz, 1986). In the second version of the model, an adverse shock, say a financial crisis, leads to a recession. The pace of reallocation then accelerates during the recession, possibly due to the low opportunity cost of time (Caballero and Hammour, 1994). This is the model of sectoral shifts that I have in mind.

I’ll offer three pieces of evidence supporting this view of recessions. First, Jaimovich and Siu (2020) show that since 1990, employment in routine tasks has looked like a step function, flat during expansions and dropping precipitously during recessions. If a worker employed in a routine task wants to find a new stable job, he typically has to retrain and look for work in a non-routine (manual or cognitive) occupation. In a similar vein, key manufacturing industries such as steel and automobile fabrication declined sharply during the deep recessions in 1974 and 1982 with little recovery during the subsequent expansion. The road back to stable employment was arduous for those workers.

Second, Davis and von Wachter (2011) look at the subsequent earnings of men who lose a job where they had three or more years of tenure. There are two important conclusions in that paper. First, on average these men lose 1.4 years of discounted pre-displacement earnings during the twenty years after the initial job loss. As Davis and von Wachter (2011) emphasize, this is far more than can be explained by a Diamond-Mortensen-Pissarides model where unemployed workers suffer no loss in human capital and quickly recover their employment.
status. It is consistent with the model I’ve described here, where job loss is correlated with human capital loss, since human capital and wages are positively correlated. The second conclusion is that losing a job when the unemployment rate is above 8 percent is twice as bad as losing a job at an average point in the business cycle. It leads to 2.8 years of lost pre-displacement earnings. This is consistent with the idea that these turbulence shocks, linking job loss to skill loss, are concentrated during periods with elevated unemployment. Most job loss during expansions is not associated with human capital loss, while the reverse is true during recessions.

The third piece of evidence is the “dog that didn’t bark,” the 2020 pandemic recession. Why did unemployment recover so quickly from its peak? This recession featured a temporary stop in activity, not a period of cleansing and rapid adjustment. Many restaurant workers were laid off during the peak of the pandemic, but no one expected the restaurant industry to disappear. Indeed, many of the workers went on temporary layoff and have now returned to the same or similar establishments. Without the loss in human capital, rapid recovery was possible and has occurred. Still, even in response to this unusual shock, some jobs may be lost forever. For example, the pandemic likely accelerated a preexisting decline in brick-and-mortar retail, forcing those workers to look for new stable employment, an arduous process. My expectation is that it will take years for the employment rate to return to its pre-pandemic level, as in past recoveries.

I want to close by recognizing that the model I’ve described here probably does not account for all fluctuations in unemployment. For example, it would have a hard time explaining why the number of unemployed new entrants to the labor force gradually rises during recessions and the early stages of recoveries (Hall and Kudlyak’s Figure 15). A more satisfactory model would probably have a persistent underlying shock. It would probably also have uncertainty about which growing sectors of the economy will absorb the workers leaving the declining sectors, which may propagate the shock and delay the recovery. And it may have a somewhat inelastic supply of job vacancies, which would slow expansions (even if it would not explain the observed decline in vacancies during a recession). But a proper accounting will give a large role to the importance of human capital losses during downturns, and may not cede any role to contagious or infectious unemployment.

Whether unemployment is contagious matters for economic policy. According to the theories mentioned by Hall and Kudlyak, high unemployment for some workers causes high unemployment for other workers, an externality. This may create a role for programs that moderate the cyclicality of unemployment, e.g. by subsidizing work sharing during recessions. In contrast, if efficient sectoral reallocation is concentrated during downturns due to the low opportunity cost of time, work-sharing programs would hinder that necessary reallocation.
and hence create long-run productive inefficiencies. Hall and Kudlyak have identified an interesting possibility. Future research, with better data and more serious models, will help us understand whether the possibility is correct.

References


