Measuring labor-force participation and the incidence and duration of unemployment

Hie Joo Ahn and James D. Hamilton

Federal Reserve Board of Governors and University of California, San Diego

Opinions expressed herein are those of the authors alone and do not necessarily reflect the views of the Federal Reserve System.

July 16th 2019, NBER SI (CRIW)
Current Population Survey

CPS randomly selects address and seeks to classify each noninstitutionalized individual aged 16 and over:

- **Employed** \((E)\)
  - Worked during reference week for own business or for pay or absent due to vacation, illness, weather
- **Unemployed** \((U)\)
  - Not employed but made specific efforts to find work any time during last 4 weeks
- **Not in labor force** \((N)\)

Rotation group structure (4-8-4)

- 1/8 of households enter the survey for the first time, are surveyed four times, leave the survey for 8 months, return to the survey and are surveyed for 4 months.
Key measurement problems → Propose reconciliation

1. Rotation-group bias
2. Non-random missing observations
3. Number preference
4. Inconsistency between reported duration of unemployment and labor-force status

→ First to provide a unified approach for reconciliation that allows for time-variation in measurement errors

- Unemployment rate ↑ 1.9% (countercyclical)
- Labor-force participation rate ↑ 2.2% (countercyclical, slowly rising trend)
- Mean duration ↓ 9 weeks (countercyclical)
Problem 1: Rotation-group bias

- The UR and LFPR should be the same across rotations.
- Average UR and LFPR (July 2001-April 2018)
  - 6.8 percent in rotation 1, 5.9 percent in rotation 8
  - 66.0 percent in rotation 1, 64.3 percent in rotation 8

**Implication** if track fixed group of individuals over time, in typical month find net flows out of U and out of LF.

**Solution**

- Model statistically the way answers change the more times people have been asked → Rotation-specific interview technology
- Propose a method to calculate measures using any of the 8 interview technologies: first interview technology.
Problem 2: Non-random missing observations

- If someone was sampled last month but missing this month, more likely than general population to have been \( U \) last month
- If someone was missing last month but sampled this month, more likely than general population to be \( U \) this month

Solutions

- Add a fourth observed category \( (M = \text{missing}) \)
- Construct data set in which accounting identities relating stocks and flows hold by construction
- Sum of \( EE, NE, ME, UE \) transitions between rotation 1 and 2 exactly equals number of \( E \) for rotation 2
- Will use observations when individuals are \( E, N, U \) to infer something about status when \( M \)
Problem 3: Number (digit) preference

When reporting the duration of unemployment, more likely to
- round to months or years
- report in even numbers (2, 4, 6, 8, 10 weeks > 1, 3, 5, 7, 9 weeks)
or rounded numbers (3 months, 6 months, 1 year)

Our solution: represent individuals’ perceived duration of unemployment using a parametric monotonic function; model digit-reporting preferences as layer on top of this

Our contribution: our parametric specification allows direct linkage of data on stocks, flows, and durations and includes both digit and interval preference.
Problem 4: Reported durations of unemployment inconsistent with reported labor-force histories

$NU^{5+}$ transitions: Consider $N$ in RG1 $(t)$, $U$ in RG2 $(t + 1)$.

- 2/3 say actively looking for work for longer than 4 weeks
- 8% say searching for 1 year and another 8% say 2 years
Problem 4: Reported durations of unemployment inconsistent with reported labor-force histories

\( NU^{5+} \) transitions: Consider \( N \) in RG1 \((t)\), \( U \) in RG2 \((t + 1)\).

- 2/3 say actively looking for work for longer than 4 weeks
- 8% say searching for 1 year and another 8% say 2 years


- Long-term unemployed had probability 0.25 of exiting U.
- Implied mean duration: no larger than 4 \((1/0.25)\) months.
Problem 4: Reported durations of unemployment inconsistent with reported labor-force histories

**$NU^{5+}$ transitions:** Consider $N$ in RG1 ($t$), $U$ in RG2 ($t + 1$).

- 2/3 say actively looking for work for longer than 4 weeks
- 8% say searching for 1 year and another 8% say 2 years

**Probability of leaving U. VS U. duration:** Consider 2011.

- Long-term unemployed had probability 0.25 of exiting U.
- Implied mean duration: no larger than 4 ($1/0.25$) months.

**Solutions to the inconsistency**

- Classify those who make $N_{t-1}U^5_{t-1}$ transitions as $U$ in $t - 1$
- Classify some of those who make $U^{15+}_{t-1}N_t$ transitions as $U$ in $t$
Notation

\[ \pi_{\cdot \cdot \cdot}^{[j]} = (4 \times 1) \text{ vector of observed fractions of each status} \]
\[ X \in \{E, N, M, U\} \text{ in rotation } j \text{ in month } t \]

\[ \Pi_{\cdot \cdot \cdot}^{[j]} = (4 \times 4) \text{ matrix of probabilities that someone who} \]
\[ \text{reports status } X_1 \text{ in rotation } j - 1 \text{ in month } t - 1 \text{ will report} \]
\[ \text{status } X_2 \text{ in month } t \text{ for rotation } j \in J = \{2, 3, 4\} \cup \{6, 7, 8\} \]

\[ \text{Our constructed data exactly satisfy } \pi_{\cdot \cdot \cdot t}^{[j]} = \Pi_{\cdot \cdot \cdot}^{[j]} \pi_{\cdot \cdot \cdot}^{[j-1]} \text{ for } j \in J \]
1. Rotation group bias: Fix based on RG1 technology

Halpern-Manners and Warren (2012)

- Saying that you looked for a job hard but failed carries stigma
- Individuals may believe that follow-up questions for U are onerous
  - ”What have you done to look for work in the last 4 weeks?”
  - ”How long have you been looking for a job?”
- The unemployment rates in rotations 2-8 are likely to 
  **understate** the truth. → fixed based on RG1 technology
Statistical description of the Rotation Group Bias

\( R_t[j] \): a [4 by 4] matrix that contains the probabilities that an individual would have answered LFS \( X^1 \) using technology 1 given answered LFS \( X^j \) in technology \( j \)

\[
R_t[j] \pi_t[j] = \pi_t[1] = \pi_t^* 
\]

\( \Pi_t^* \): LFS transition probabilities if interviewed with RG1 technology

\[
\Pi_t^* = R_t[j] \Pi_t[j] (R_t^{-1}[j])^{-1} 
\]

\( R_t[j] \) can be characterized by
- More missing in RG.1 and 5
- Fewer E/N in in RG.1 and 5 (\( \theta_E^{EM}, \theta_N^{NM} \))
- Rising N, falling U, as surveyed more (\( \theta_{NU}^{[j]} \)): increased over time!
2. Non-random missing observations

After correcting for the rotation group bias

**Goal**: recover the fraction of E, U, and N from missing individuals

- True LFS of missing individuals might have been $E$, $U$, or $N$
- Probability of becoming $E$ in $t$ of those missing in $t - 1 = \text{weighted average of probabilities of becoming } E \text{ in } t \text{ of those who were } E, U, \text{ and } N \text{ in } t - 1$.
  - $\rightarrow ME$ transitions as mixtures of $EE$, $NE$, $UE$
- Same for $MN$ and $MU$
- Calculate the mixing weights: share of $E$, $U$, and $N$ within $M$
- Allows to count size of missing observations
Fractions of $M$ interpreted as $E$, $N$, or $U$ each month

- **Rising trend** in missing individuals who might have been $E/N$.
- **Countercyclical** behavior of $m_U$: unemployed individuals are more likely to be missed during a weak labor market.
3: Number (Digit) preference

Distribution of unemployment duration

Preference for even numbers (red circle), rounded numbers (rounding to month: blue; half year: pink)
Distribution of U. duration free from number preference

Step 1: Model underlying latent distribution of perceived durations using a parametric monotonic distribution

- Baseline specification: mixture of exponentials
  \[ \pi_U^\dagger(\tau) = w_1(1 - p_1)p_1^\tau + w_2(1 - p_2)p_2^\tau \]
- \( w_i \) = fraction of unemployed of type \( i \) \( (w_1 + w_2 = 1) \)
- A fraction of the population \( \pi_U w_i (1 - p_i) \) lose their job each week and have unemployment-continuation prob \( p_i \) \( (i = 1, 2) \)
Distribution of U. duration free from number preference

Step 1: Model underlying latent distribution of perceived durations using a parametric monotonic distribution

- Baseline specification: mixture of exponentials
  \[
  \pi_U^\dagger(\tau) = w_1(1 - p_1)p_1^\tau + w_2(1 - p_2)p_2^\tau
  \]

- \(w_i\) = fraction of unemployed of type \(i\) \((w_1 + w_2 = 1)\)

- A fraction of the population \(\pi_U w_i(1 - p_i)\) lose their job each week and have unemployment-continuation prob \(p_i\) \((i = 1, 2)\)

Step 2: Model reported durations as probabilistic transformation of perceived durations

- Preference for even numbers, digits and rounded numbers
Distribution of U. duration free from number preference

Step 1: Model underlying latent distribution of perceived durations using a parametric monotonic distribution

- Baseline specification: mixture of exponentials
  \[ \pi_U(\tau) = w_1 (1 - p_1)p_1^\tau + w_2 (1 - p_2)p_2^\tau \]
  - \( w_i \) = fraction of unemployed of type \( i \) (\( w_1 + w_2 = 1 \))
  - A fraction of the population \( \pi_U w_i (1 - p_i) \) lose their job each week and have unemployment-continuation prob \( p_i \) (\( i = 1, 2 \))

Step 2: Model reported durations as probabilistic transformation of perceived durations

- Preference for even numbers, digits and rounded numbers
  \[ \hat{p}_1 = 0.83, \hat{p}_2 = 0.97, \hat{w}_1 = 0.42, \hat{w}_2 = 0.58 \]
  \( \rightarrow \) Also use these parameters for the NU adjustment

Hie Joo Ahn and James D. Hamilton
4: Inconsistency bw U. duration and LFS: $NU^{5+}$

Distribution of unemployment duration reported for people who were $N$ in rotation 1 and $U$ in rotation 2

- Those with records $N_{t-1} U^{5+}_t$ perceive their status at $t - 1$ to have been looking for a job ($U_t$), though they reported $N_{t-1}$
- Propose to reclassify $N_{t-1} U^{5+}_{t-1} \rightarrow U_{t-1} U^{5+}_t$: countercyclical!
Corroboration based on LFS history

Preceding $N$ carries information similar to $U$ in predicting future $E$. The pattern of duration dependence is similar b/w $UUU$ and $UNU$.

### Re-employment probability in month $t$ by LFS history

<table>
<thead>
<tr>
<th>$UUU$</th>
<th>Probability</th>
<th>$UNU$</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U_{t-2}, U_{t-1}^{5+}$</td>
<td>0.15</td>
<td>$N_{t-2}, U_{t-1}^{5+}$</td>
<td>0.13</td>
</tr>
<tr>
<td>$U_{t-3}^{5.14}, U_{t-2}^{5.14}, U_{t-1}^{15.26}$</td>
<td>0.16</td>
<td>$U_{t-3}^{5.14}, N_{t-2}, U_{t-1}^{15.26}$</td>
<td>0.14</td>
</tr>
<tr>
<td>$U_{t-3}^{15.26}, U_{t-2}^{15.26}, U_{t-1}^{15.26}$</td>
<td>0.14</td>
<td>$U_{t-3}^{15.26}, N_{t-2}, U_{t-1}^{15.26}$</td>
<td>0.15</td>
</tr>
<tr>
<td>$U_{t-3}^{15.26}, U_{t-2}^{27+}, U_{t-1}^{27+}$</td>
<td>0.11</td>
<td>$U_{t-3}^{15.26}, N_{t-2}, U_{t-1}^{27+}$</td>
<td>0.10</td>
</tr>
<tr>
<td>$U_{t-3}^{27+}, U_{t-2}^{27+}, U_{t-1}^{27+}$</td>
<td>0.08</td>
<td>$U_{t-3}^{27+}, N_{t-2}, U_{t-1}^{27+}$</td>
<td>0.07</td>
</tr>
</tbody>
</table>
Corroboration based on LFS history

Preceding $N$ carries information similar to $U$ in predicting future $E$. The pattern of duration dependence is similar b/w $UUU$ and $UNU$.

| Re-employment probability in month $t$ by LFS history |
|---------------------------------|-----------------|-----------------|-----------------|
| $UUU$                           | Probability     | $UNU$           | Probability     |
| $U_{t-2}, U_{t-1}^{5+}$         | 0.15            | $N_{t-2}, U_{t-1}^{5+}$ | 0.13            |
| $U_{t-3}^{5.14}, U_{t-2}^{5.14}, U_{t-1}^{15.26}$ | 0.16            | $U_{t-3}^{5.14}, N_{t-2}, U_{t-1}^{15.26}$ | 0.14            |
| $U_{t-3}^{15.26}, U_{t-2}^{15.26}, U_{t-1}^{15.26}$ | 0.14            | $U_{t-3}^{15.26}, N_{t-2}, U_{t-1}^{15.26}$ | 0.15            |
| $U_{t-3}^{15.26}, U_{t-2}^{27+}, U_{t-1}^{27+}$ | 0.11            | $U_{t-3}^{15.26}, N_{t-2}, U_{t-1}^{27+}$ | 0.10            |
| $U_{t-3}^{27+}, U_{t-2}^{27+}, U_{t-1}^{27+}$ | 0.08            | $U_{t-3}^{27+}, N_{t-2}, U_{t-1}^{27+}$ | 0.07            |

Marginally attached workers in $N$ category

- 40% of $NU^{5+}$ transitions: they account for only 2.2% of $N$.
- spend as much time for job search as $U$.(ATUS: 154 vs 143)
5. Inconsistency bw U.duration and UU probability

Probability that someone who is U. in RG 1 with duration $\tau$ weeks will still be U in RG 2. $\rightarrow$ Duration dependence in UU continuation

$$\hat{\pi}_{UU}(\tau) = \eta_1(\tau)\gamma_{1,uu} + \eta_2(\tau)\gamma_{2,uu} \quad (\gamma_{1,uu} = 0.37, \gamma_{2,uu} = 0.58)$$

Are $\gamma_{UU}^1$ and $\gamma_{UU}^2$ consistent with $\hat{p}^1$ and $\hat{p}^2$? Without errors, we would also expect

$$\frac{\gamma_{2,uu}}{1 - \gamma_{2,um}} \approx \hat{p}^2_{4.33}$$

Type 2: Perceived durations inconsistent with matched flows.

To reconcile U. duration and UU continuations from the matched flows data, type 2 transitions should be adjusted.
Adjust some \textit{UN} transitions to \textit{UU} continuations

- Similarity b/w \textit{UUU} and \textit{UNU} $\rightarrow$ Some of \textit{UN} $\rightarrow$ \textit{UU}
- Discrepancy b/w flows and stock duration data in Type 2 unemployed’s UU $\rightarrow$ Adjust type 2’s UU $\rightarrow$ majority of \textit{U}_t^{15.\text{+}}
Adjust some UN transitions to UU continuations

- Similarity b/w UUU and UNU → Some of UN → UU
- Discrepancy b/w flows and stock duration data in Type 2 unemployed’s UU → Adjust type 2’s UU → majority of $U^{15.\,+}_t$

Reconciliation: Find hidden $U^{15.\,+}_{t-1} U_t$ from $U^{15.\,+}_{t-1} N_t$

- Fraction $\xi_{UN}$: $U^{15.\,+}_{t-1} U_t$
- Fraction $1 - \xi_{UN}$: $U^{15.\,+}_{t-1} N^*_t \approx N_{t-1}, N_t$

Classify 63% of those reported $U^{15.\,+}_{t-1} N_t$ transitions as $U_t$. SS flows

$$P(E_{t+1} | N_t, U^{15.\,+}_{t-1}) = \underbrace{\xi_{UN} P(E_{t+1} | U_t, U^{15.\,+}_t)}_{=0.633} + (1 - \xi_{UN}) P(E_{t+1} | N_t, N_{t-1}) \approx P(E_{t+1} | N^*_t, U^{15.\,+}_{t-1})$$
Gap between reconciled and published: weakly countercyclical (← NU adjustment)

- RGB(+0.5%p), Missing (+0.3%p), NU adjustment (1.1%p)
Gap between reconciled and published: weakly counter-cyclical ($\leftarrow$ NU adjustment) with slowly rising trend ($\leftarrow$ RGB, missing)

- RGB(+1.2%p), Missing (+0.2%p), NU adjustment (0.8%p)

Contributions of different adjustments to labor-force participation rate
Published: 25 weeks VS Adjusted: 16 weeks

Average duration of unemployment

- Inflows $\leftarrow NU^{5+}$ adjustment
- Adjusted UU continuation from matched flows data $< UU$ continuation implied by the durations $\leftarrow EU^{5+}$ (on-the-job search spells), perceived durations reflecting discouragement
Conclusion

CPS contains multiple internal inconsistencies.

- Rotation group bias
- Non-random missing observations
- Number preference
- Inconsistency between reported durations and LFS histories.

Our paper is the first unified reconciliation, and concludes

- The published unemployment rate and labor-force participation rate are underestimated.
- The new inflows into unemployment are underestimated.
- The mean duration of unemployment is overestimated.
Duration: the ratio of number unemployed for 5 weeks and over in $t$ to number unemployed in $t - 1$

Flows: the fraction of those who continue to be $U$ in $t + 1$ out of those who are $U$ in $t$
Adjustments to the unemployment rate

On average, correcting for rotation bias adds 0.5%, missing observations 0.3%, and $NU$ adjustment adds 1.1% to UR.
Adjustments to the labor-force participation rate

Correcting for rotation bias adds 1.2%p, missing observations 0.2%p, and NU 0.8%p to LFPR.
Adjustment to the mean duration of unemployment

Published: 25 weeks VS Adjusted: 16 weeks
Evolution of measurement errors in UR and LFPR

Unemployment rate

Labor-force participation rate

Hie Joo Ahn and James D. Hamilton
Maximum likelihood estimates

Postulate that probability that someone who is unemployed with duration $\tau$ in RG 1 will be $X \in \{E, N, M, U\}$ in RG 2 is $\dot{\pi}_{UX}(\tau) = \eta_1(\tau)\gamma_{1,UX} + \eta_2(\tau)\gamma_{2,UX}$ for $\gamma_{i,UX}$ unrestricted parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>[1]</th>
<th>[2]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_{1,UE}$</td>
<td>0.3183</td>
<td>0.0053</td>
</tr>
<tr>
<td>$\gamma_{1,UN}$</td>
<td>0.2179</td>
<td>0.0032</td>
</tr>
<tr>
<td>$\gamma_{1,UM}$</td>
<td>0.0909</td>
<td>0.0025</td>
</tr>
<tr>
<td>$\gamma_{1,UU}$</td>
<td>0.3729</td>
<td>...</td>
</tr>
<tr>
<td>$\gamma_{2,UE}$</td>
<td>0.1153</td>
<td>0.0092</td>
</tr>
<tr>
<td>$\gamma_{2,UN}$</td>
<td>0.2353</td>
<td>0.0087</td>
</tr>
<tr>
<td>$\gamma_{2,UM}$</td>
<td>0.0735</td>
<td>0.0028</td>
</tr>
<tr>
<td>$\gamma_{2,UU}$</td>
<td>0.5759</td>
<td>...</td>
</tr>
</tbody>
</table>
A simple summary of how to use reported duration

$$\eta_2(\tau) = \text{probability someone who reports duration } \tau \text{ is of type 2}$$
Is this a reasonable estimate? Additional corroboration

- Average fraction of population with reported $NU$ that are really type 2 $UU$ continuations: 0.0028
- Average fraction of population with reported $UN$ that are interpreted as type 2 $UU$ continuations: 0.0026
Adjusting monthly estimates

**Estimation of time-varying parameters** Filtering idea similar to Kalman filter

We use exponential smoothing to calculate a weighted average of recent observations through date $t$ to infer how the adjustment parameters $\theta_t$ are changing over time. If $\theta_t$ denotes an estimate using observations from month $t$ alone, we calculate

$$\bar{\theta}_t = (1 - \lambda)\theta_t + \lambda\bar{\theta}_{t-1}.$$ 

$$\lambda = 0.98$$
Alternative measures of unemployment rate

Alternative measures of new inflows into unemployment
## Features of measurement errors in UR and LFPR

<table>
<thead>
<tr>
<th></th>
<th>Unemployment rate</th>
<th>Labor force participation rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Trend</td>
</tr>
<tr>
<td>Total</td>
<td>1.9</td>
<td>No</td>
</tr>
<tr>
<td>Rotation group bias</td>
<td>0.5</td>
<td>Slowly rising</td>
</tr>
<tr>
<td>Missing observations</td>
<td>0.3</td>
<td>No</td>
</tr>
<tr>
<td>N/U misclassification</td>
<td>1.1</td>
<td>No</td>
</tr>
</tbody>
</table>
Features of measurement errors in UR and LFPR

Unemployment rate

Labor-force participation rate
Size of $NU^{5.+}$: Countercyclical

$m_{Nt}^\#$: Percent of population who report $NU^{5.+}$
Size of $UN$ whom we interpret as $UU$: Countercyclical

Percent of population $UN \rightarrow UU$

$m_{Nt}^b$