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.

Purpose of this paper

Current Population Survey (CPS)

- Primary source for U.S. statistics on unemployment and labor force
- Contains many internal inconsistencies

Our paper

- Documents these problems
- Proposes a reconciliation

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)

Contacts same address again next month to ask same questions

- In any given month, some people are being asked first time, others 2nd, and others an 8th time.

Problem 1: Rotation-group bias

- Bailar (JASA, 1975); Solon (JBES, 1986); Halpern-Manners and Warren (Demography, 2012); Krueger, Mas and Niu (REStat, 2017)
- The average answers change the more times people are asked
- Average unemployment rate (July 2001-April 2018)
  - 6.8 percent in rotation 1
  - 5.9 percent in rotation 8
- Average labor-force participation rate
  - 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 unemployment and out of labor force even in month when measured unemployment rate may be rising

Problem 2: Non-random missing observations

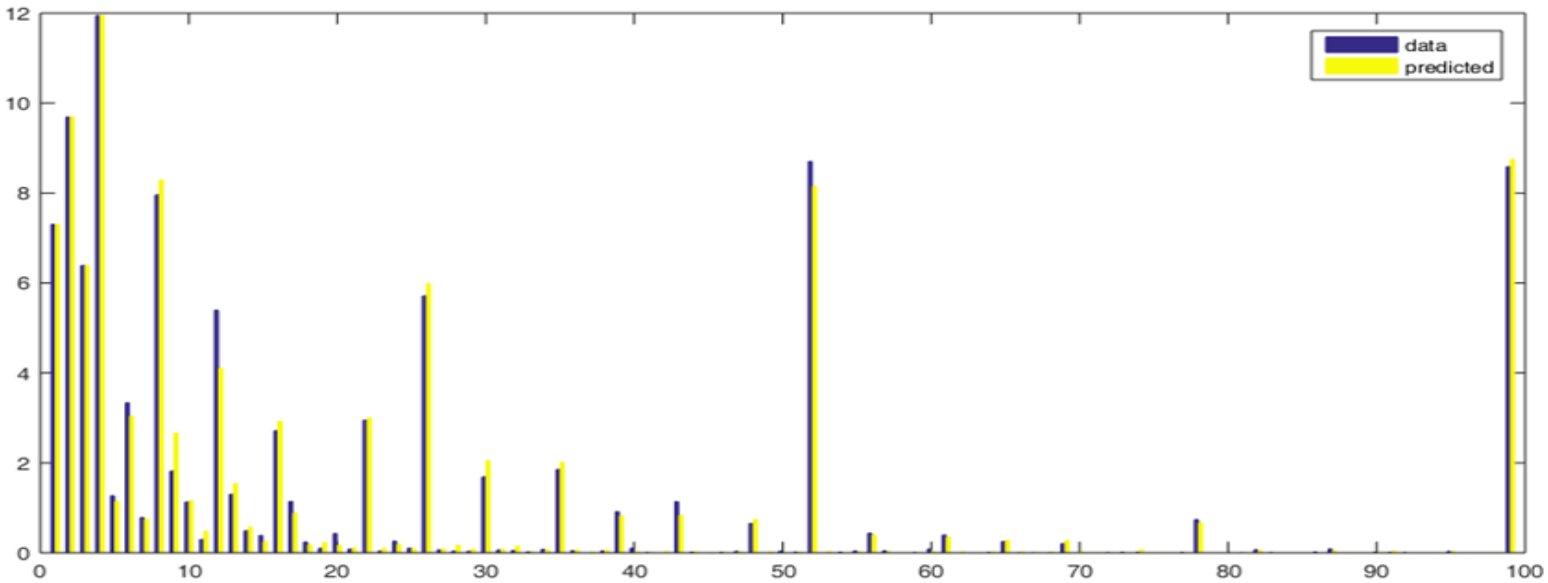
- Abowd and Zellner (JBES, 1985)
- If someone was missing last month but sampled this month, more likely than general population to be U this month
- Missing individuals bias the reported unemployment rate downward

Problem 3: Digit preference

- Preference for even numbers
  - On average more people report unemployment durations of 2 weeks than 1 week
  - More 6 weeks than 5 weeks
- Preference for rounded numbers
  - Many more 24 weeks than 23 weeks

Problem 4: Reported durations of unemployment inconsistent with reported labor-force histories

- Consider reported unemployment durations of people who were *N* in rotation 1 and *U* in rotation 2
- 2/3 say they have been actively looking for work for longer than 4 weeks



Why does this matter?

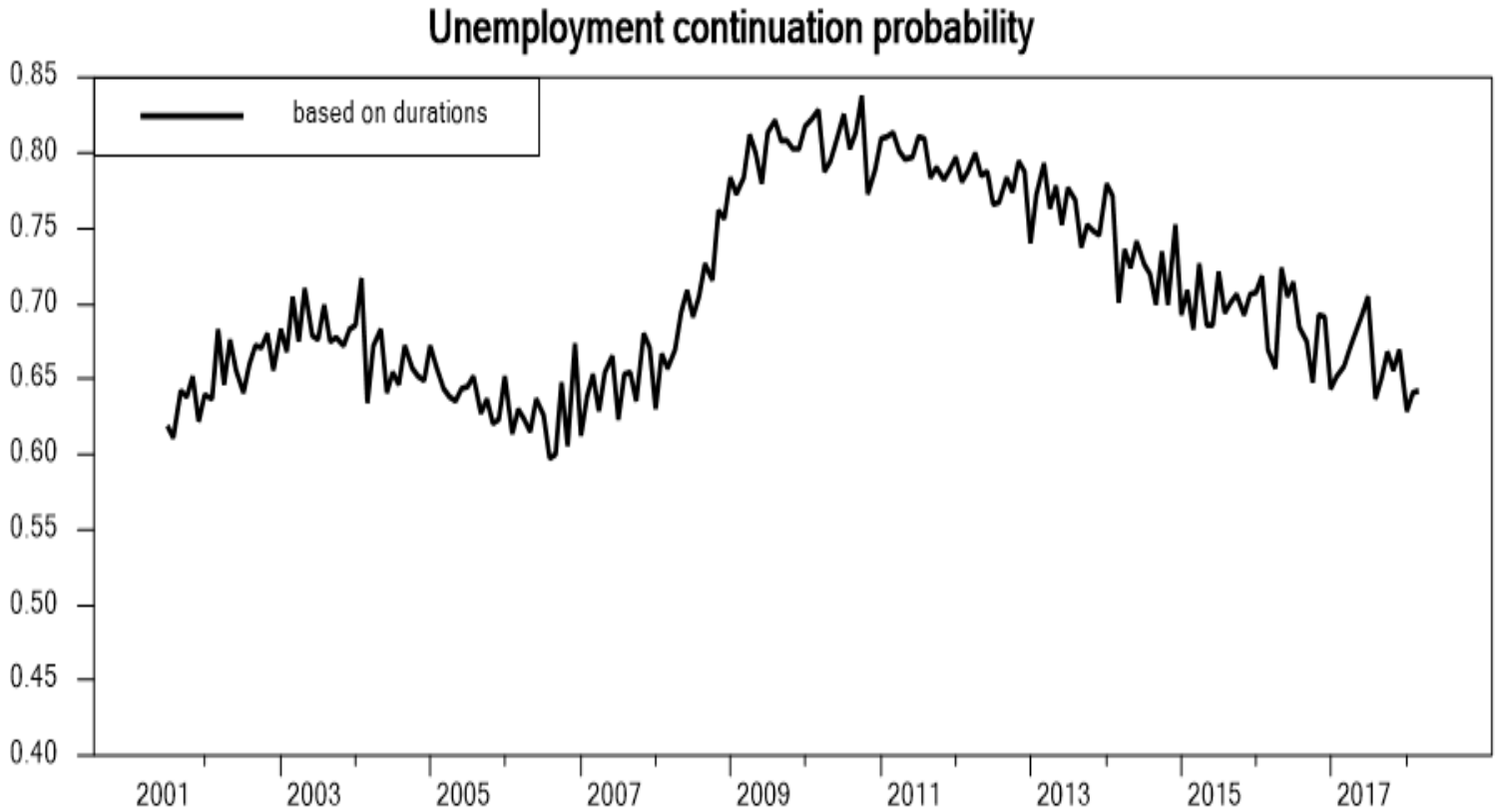
Example 1: What is the probability that someone who is unemployed today will still be unemployed next month?

Duration-based approach

- Calculate ratio of number unemployed in t with duration 5 weeks or greater to number unemployed at t-1
- Variants used by van den Berg and and van der Klaauw (J Econometrics, 2001); Elsby, Michaels and Solon (AEJ Macro, 2001); Shimer (Rev Econ Dyn, 2012)

Why does this matter? (1) U-continuation probability

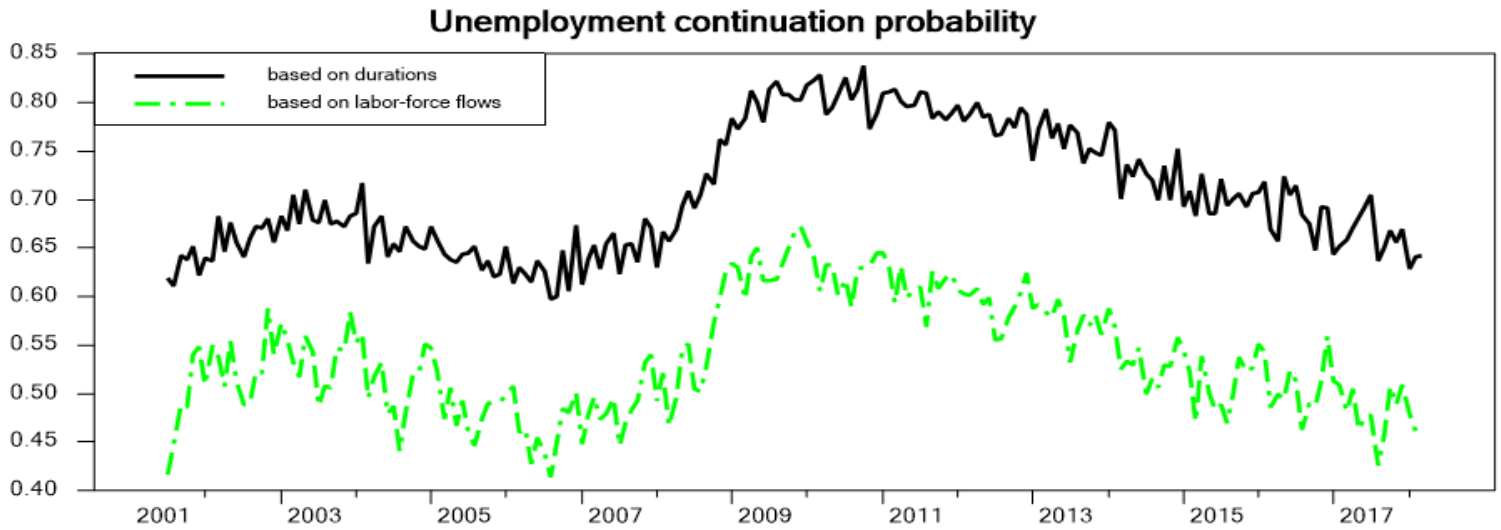
Duration-based measure averages 70.7 %



Why does this matter? (1) U-continuation probability

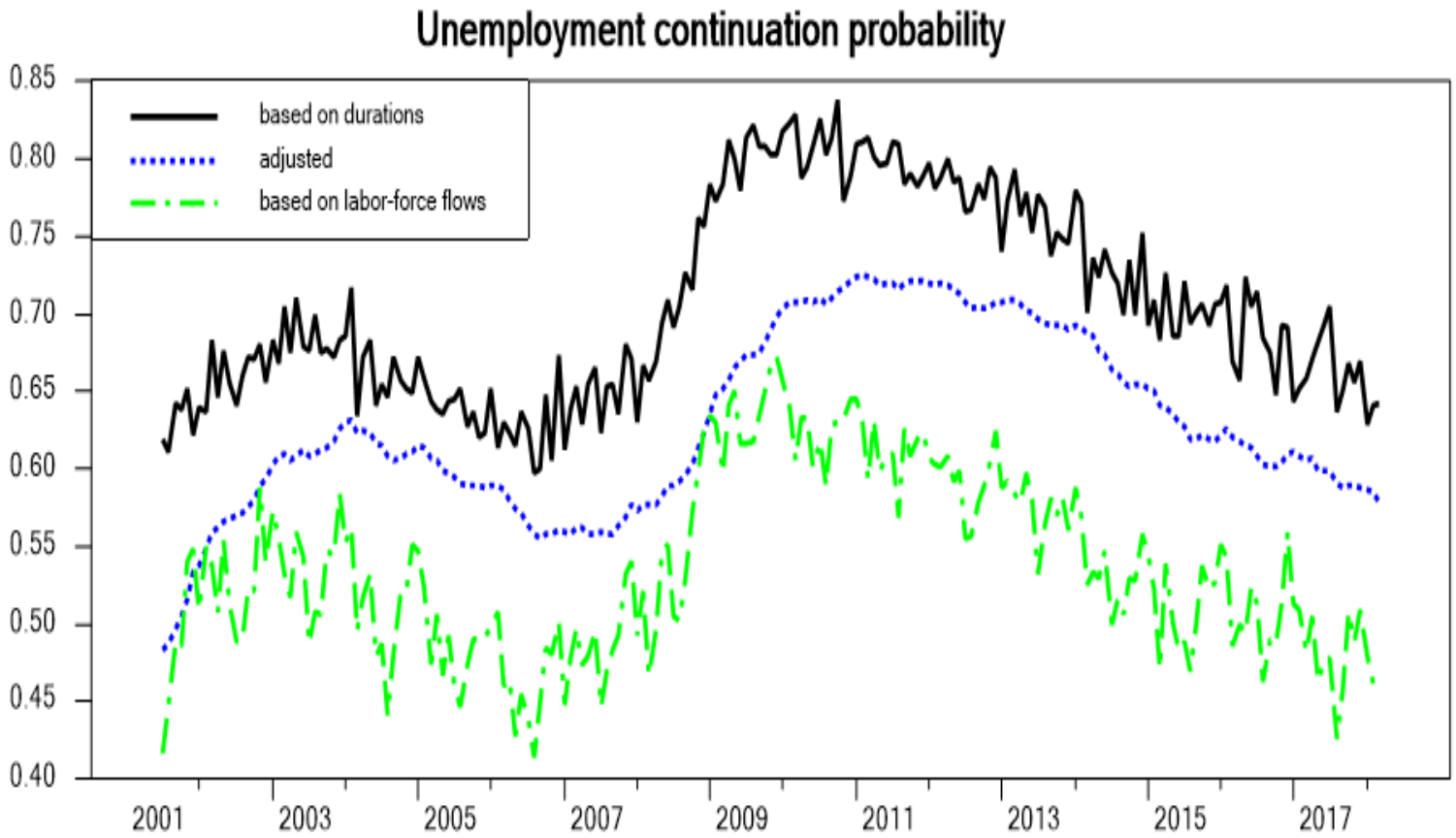
Flows-based approach

- Look at individuals who are U in t and not missing in $t+1$ and calculate fraction who are U in $t+1$ (averages 53.7%)
- Fujita and Ramey (IER, 2009); Elsby, Hobijn and Sahin (BPEA, 2010)



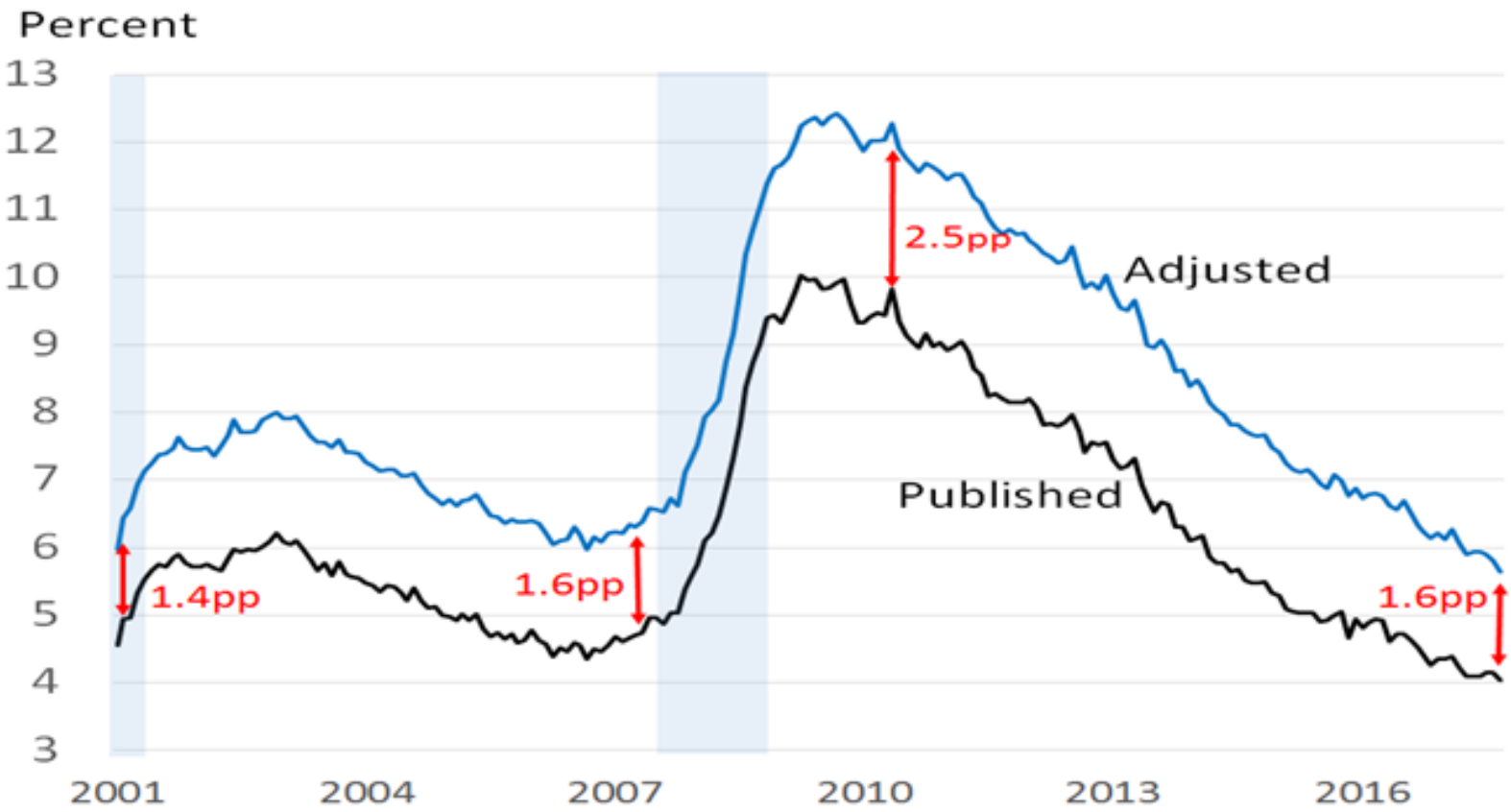
Why does this matter? (1) U-continuation probability

Our reconciled flows-based measure averages 62.9%



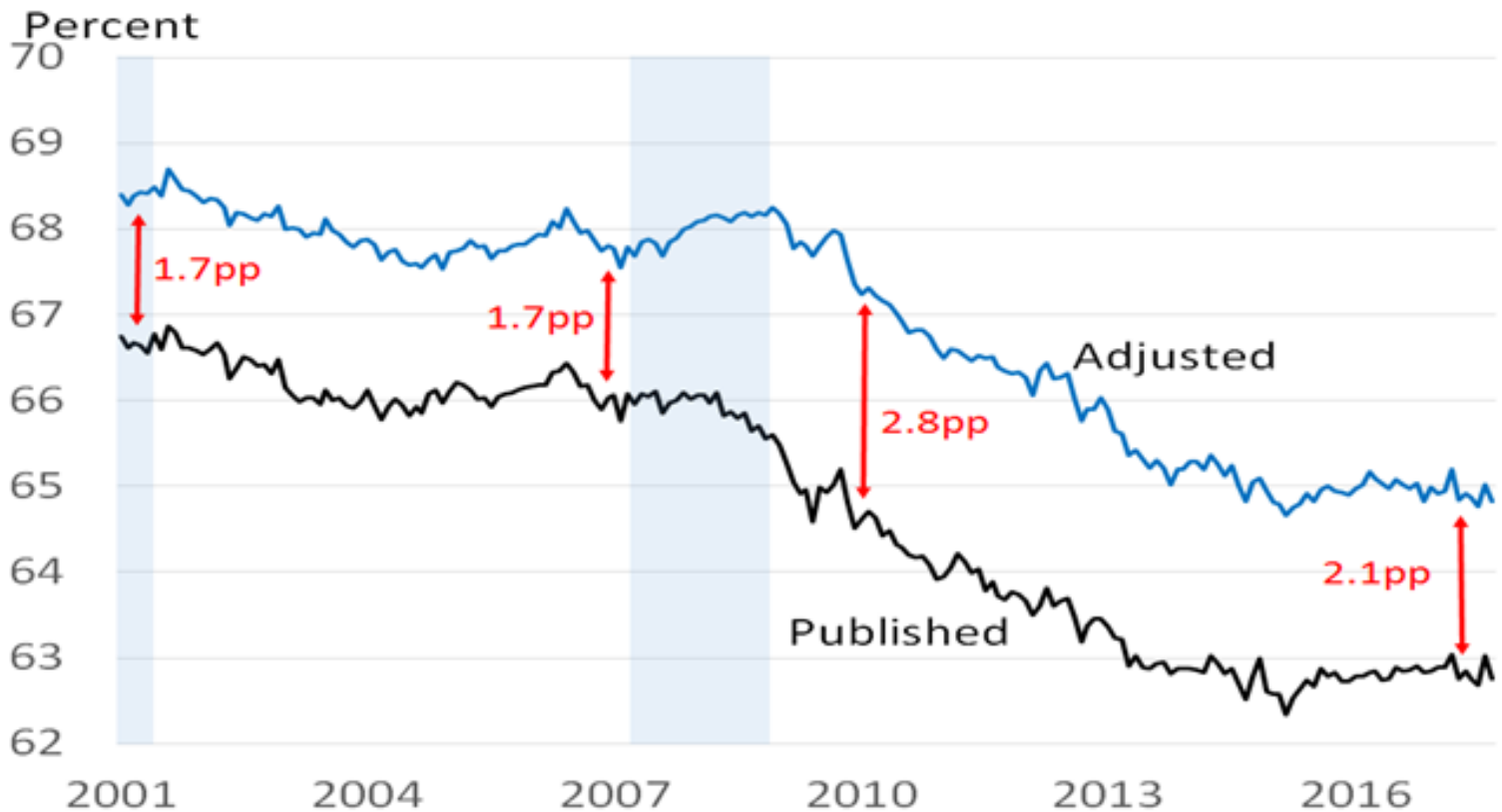
Why does this matter? (2) Unemployment rate

On average, our corrections add 1.9% to the unemployment rate



Why does this matter? (3) Labor-force participation rate

We conclude BLS underestimates labor-force participation rate by 2.2% on average and the gap has increased.



Solution step 1: keep track of missing individuals

- Add a fourth observed category (M = missing) for every individual
- 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
- $\pi_t^{[j]}$ = (4 x 1) vector of fractions of population in rotation j in month t who are measured E, N, M, or U
- $\Pi_t^{[j]}$ = (4 x 4) matrix of probabilities that someone who reports status X_1 in rotation $j-1$ in month $t-1$ will report status X_2 in rotation j in month t
- Our constructed data exactly satisfy $\Pi_t^{[j]}\pi_{t-1}^{[j-1]} = \pi_t^{]j]}$

Step 2: Parameterize how average answers across rotations change gradually over time

- Choose a rotation on which to normalize, e.g. $\pi_t^* = \pi_t^{[1]}$
- Summarize average differences between rotation j answers and rotation 1 answers in month t in terms of (4 x 4) matrix $\bar{R}_t^{[j]}$
- Find (4 x 4) matrix Π_t^* satisfying $\Pi_t^* \pi_{t-1}^* = \pi_t^*$ that best fits data

$$\Pi_t^{[j]} \approx (\bar{R}_t^{[j]})^{-1} \Pi_t^* \bar{R}_t^{[j-1]} \text{ for } j \in J = \{2,3,4\} \cup \{6,7,8\}$$
$$\pi_t^{[1]} \approx \Pi_t^* \pi_{t-1}^*$$
$$\pi_t^{[5]} \approx (\bar{R}_t^{[5]})^{-1} \Pi_t^* \pi_{t-1}^*$$

Where does rotation bias come from?

- Stigma from repeatedly admitting can't find job
  - Little rotation bias when status is reported by someone else
- Desire to end interview quickly
  - Increasing numbers of people claim retired or disabled as j increases
  - But retired-disabled in $j > 1$ are more likely to become E or U than retired-disabled in $j = 1$
- Normalizing on $j = 1$ minimizes inconsistency between reported durations and recorded unemployment-continuation probabilities
  - If $j = 1$ some UN become UU
  - If $j > 1$ some UN become NN
- Conclusion: our preferred normalization is $j = 1$

Step 3: Adjustments for nonrandom missing observations

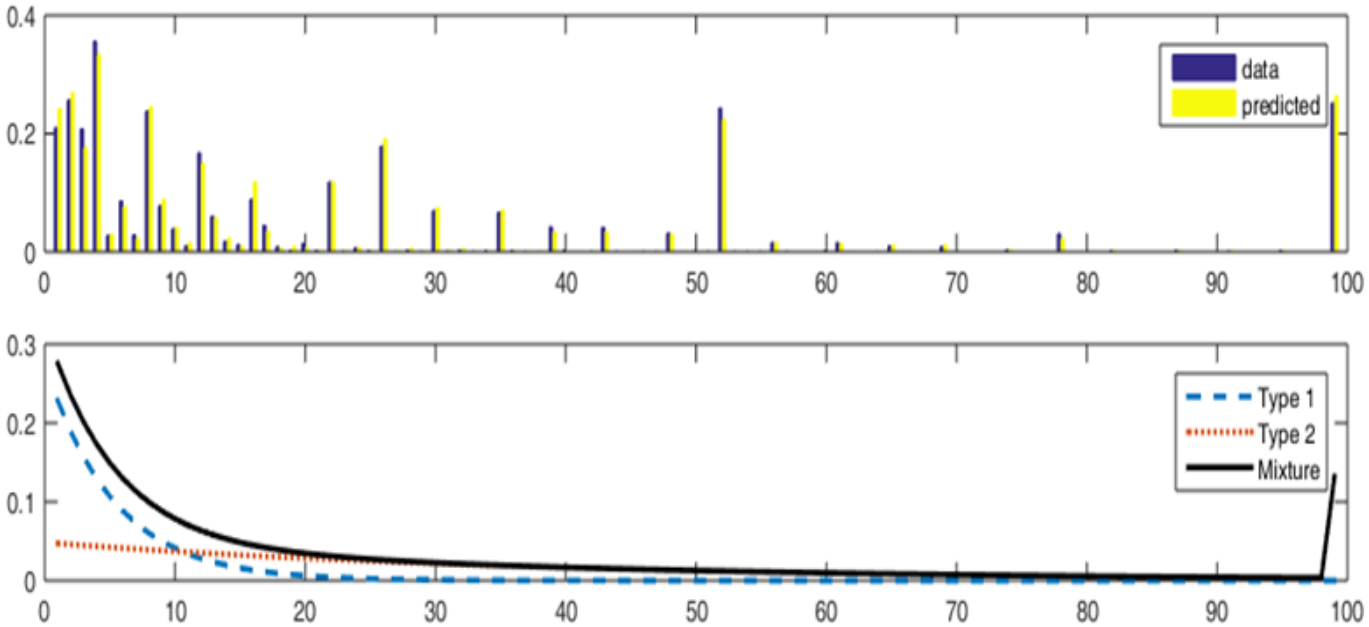
- Interpret ME transitions as mixtures of EE, NE, UE transitions
- Same for MN and MU

$$\begin{bmatrix} \pi^*_{ME,t} \\ \pi^*_{MN,t} \\ \pi^*_{MU,t} \end{bmatrix} = \begin{bmatrix} \pi^*_{EE,t} & \pi^*_{NE,t} & \pi^*_{UE,t} \\ \pi^*_{EN,t} & \pi^*_{NN,t} & \pi^*_{UN,t} \\ \pi^*_{EU,t} & \pi^*_{NU,t} & \pi^*_{UU,t} \end{bmatrix} \begin{bmatrix} m_{E,t-1} \\ m_{N,t-1} \\ m_{U,t-1} \end{bmatrix}$$

- This adjustment allocates about 15% of M to E, N, or U

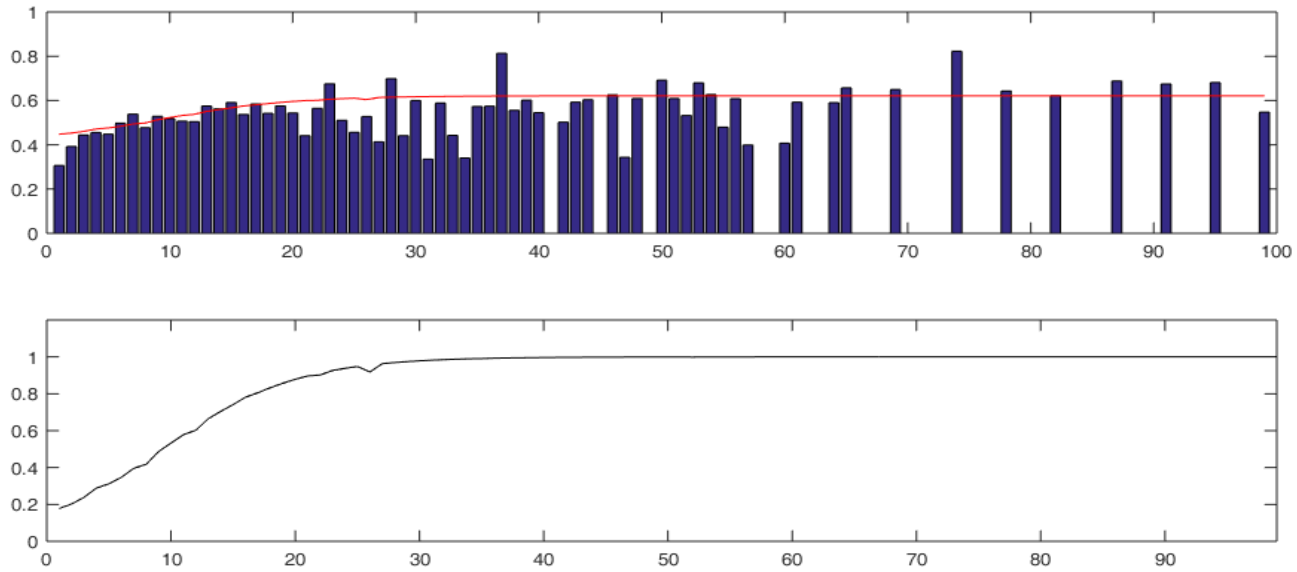
Step 4: Model preference for reporting certain numbers

- Model perceived durations using a parametric monotonically decreasing function
- Best fit to data: mixture of 2 exponentials with implied weekly unemployment-continuation probabilities of $p_1 = 0.8$ and $p_2 = 0.97$



Modeling observed UU continuations

- Can calculate probability that an individual with reported duration τ weeks is type 1 or 2
- Can calculate observed probability $\gamma_{i,UU}$ that an individual of type i will remain unemployed
- Fits data very well, but observed unemployment continuation probability $\gamma_{2,UU}$ is much lower than perceived probability p_2



Step 5. Addressing ambiguity between N and U

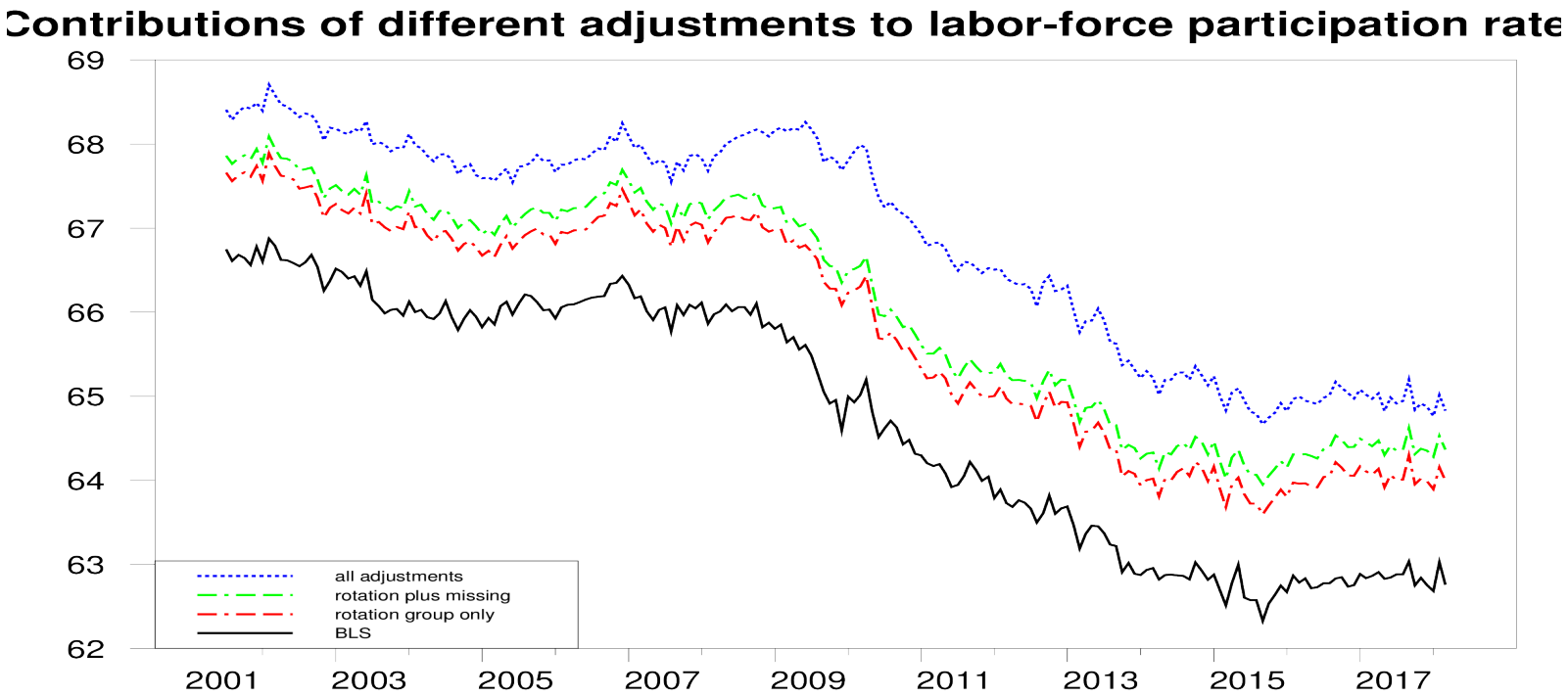
- Two-thirds of observed NU transitions say they have been looking for a job for longer than 4 weeks
- Distribution of reported durations of job search is very similar to those who were counted as U
  - $\hat{p}_2 = 0.9738$ for U
  - $\hat{p}_2 = 0.9746$ for $NU^{5,+}$
- Probabilities of $NU^{5,+}$ getting a job similar to U with similar dependence on reported length of job search
- Many of the $NU^{5,+}$ told the interviewer the previous month that they wanted a job

Proposed solution

- We propose to classify $NU^{5.+}$ as U in $t-1$
- This can explain about half the discrepancy between reported p_2 and objective $\gamma_{2,UU}$
- We attribute remaining discrepancy to misreported durations
- Our approach is similar to common solution of reclassifying UNU as UUU (Rothstein, BPEA 2011; Elsby et al., BPEA 2011, JME 2015; Farber and Valletta, J. Human Resources 2015)
- Our approach: reclassify $UNU^{5.+}$ as UUU

Effects of corrections

On average, correcting for rotation bias adds 1.2%, missing observations 0.2%, and *NU* 0.8% to labor-force-participation rate and gap between our estimate and BLS has increased over time.



Conclusion

CPS contains multiple internal inconsistencies.

- People's answers change the more times they have been asked the questions.
- Missing observations are not random.
- People prefer to report some numbers over others.
- Reported durations are inconsistent with reported labor-force histories.

Our paper is the first unified reconciliation, and concludes that BLS:

- underestimates the unemployment rate and labor-force participation rate
- underestimates new inflows into unemployment
- overestimates average duration of unemployment and the share of long-term unemployment

Researchers who use unadjusted CPS data will:

- overestimate unemployment-continuation probabilities and underestimate new inflows into unemployment if relying on stock duration data
- underestimate number of newly unemployed and number of *UU* continuations if relying on matched flows data