# Misclassification Errors and the Underestimation of the U.S. Unemployment Rate *

Shuaizhang Feng[†] Yingyao Hu[‡]

January 28, 2012

**Abstract**

Using recent results in the measurement error literature, we show that the official U.S. unemployment rate substantially underestimates the true level of unemployment, due to misclassification errors in the labor force status in the Current Population Survey. During the period from January 1996 to August 2011, the corrected monthly unemployment rates are between 1 and 4.4 percentage points (2.1 percentage points on average) higher than the official rates, and are more sensitive to changes in business cycles. The labor force participation rates, however, are not affected by this correction.

Keywords: Unemployment Rate, Labor Force Participation Rate, Misclassification, Measurement Error, Current Population Surveys.

JEL Classification: J21, J64, C14.

*We are grateful to Robert Moffitt and Katheryn Russ for suggestions and comments. We also thank Xianqiang Zou and Jingliang Lu for able research assistance. All remaining errors are our own.

[†]School of Economics, Shanghai University of Finance and Economics, 715 Econ Building, 111 Wuchuan Road, Shanghai 200433, China and IZA. E-mail: shuaizhang.feng@gmail.com

[‡]Corresponding author. Department of Economics, Johns Hopkins University, 440 Mergenthaler Hall, 3400 N. Charles Street, Baltimore, MD 21218, USA and IZA. E-mail: yhu@jhu.edu

# 1 Introduction

The unemployment rate is among the most important and carefully-watched economic indicators in modern society, and often takes center stage in discussions of economic policy. However, there is considerable disagreement over the precise definition and measurement of unemployment, hence the other two labor force statuses: "employed" and "not-in-labor-force".[1] In the United States, the Bureau of Labor Statistics (BLS) reports six alternative measures of unemployment (U1-U6), including the official unemployment rate (U3) which is based on the International Labor Organization (ILO)'s definition.[2] Due to the intrinsic difficulties in classifying some groups of people, such as marginally-attached workers and involuntary part-time workers, into three distinct labor force statuses, the U.S. official unemployment rate is potentially subject to measurement error.

In this paper, we take a latent variables approach and view the reported labor force statuses as functions of the underlying unobserved true labor force statuses. We then impose a structure on the misclassification process and the dynamics of the underlying latent LFS. Using recent results in the measurement error literature, we show that the official U.S. unemployment rate substantially underestimates the true level of unemployment. During the period from January 1996 to August 2011, our corrected unemployment rates are higher than the corresponding official figures by 2.1 percentage points on average. In terms of the monthly differences, the corrected rates are from 1 to 4.4 percentage points higher than the official rates, and are more sensitive to changes in the business cycles.

Official unemployment statistics in U.S. are based on the Current Population Survey (CPS) conducted by the Census Bureau. The CPS interviews around 60,000 households each month to collect basic demographic and labor force status information.

[1] For example, using Canadian data, Jones and Riddell (1999) empirically examine labor market transitions of people with different labor force statuses and find substantial heterogeneity within the nonemployed, such that no dichotomy exists between those unemployed and not-in-labor-force among all nonemployed persons.

[2] The ILO defines "unemployed" as those who are currently not working but are willing and able to work for pay, currently available to work, and have actively searched for work.

Based on the answers to survey questions on job-related activities, the CPS records each individual's labor force status as "employed", "unemployed," or "not-in-labor-force." The misclassification among the three possible values of the labor force status has been a substantial issue in the CPS, as clearly demonstrated by the Reinterview Surveys, in which a small sub-sample of the households included in the original CPS are recontacted and asked the same questions. Treating the CPS reconciled Reinterview Surveys sample as reflecting true labor force statuses, researchers have found that there exists considerable error in the original CPS.[3] Of course, the actual misclassification errors in labor force status are likely to be substantially larger than suggested in reconciled CPS reinterviews, as argued by Poterba and Summers (1995), Biemer and Forsman (1992) and Sinclair and Gastwirth (1996).

The misclassification of labor force statuses in the CPS and other similar surveys has received considerable attention in the literature. To identify the misclassification probabilities, early studies typically relied on a particular exogenous sources of "truth", such as the reconciled CPS reinterview surveys (see e.g. Abowd and Zellner 1985, Poterba and Summers 1986, and Magnac and Visser 1999). Nevertheless, the reinterview surveys are usually small in sample size (approximately 3% of the corresponding CPS sample) and not readily available to outside researchers. Reinterview surveys are also subject to misclassification errors due to many practical limitations.[4] Actually, some studies using other methods show that the reconciled CPS reinterview data may contain even more errors than the original CPS sample (Sinclair and Gastwirth, 1996). Other studies rely on two repeated measures of the labor force status of the same individuals in the same period and assume that the error probabilities are the same for different sub-samples.[5] More recent studies, such as Biemer and Bushery

[3]The CPS reinterview sample consists two components, one is "non-reconciled", in which case no attempt is made to determine which answers are correct, the other is "reconciled", in which case the second interviewer would compare the responses from the first survey with the reinterview answers and try to resolve any conflicts (Poterba and Summers, 1984).

[4]The reinterview may not have been independent of the original interview to the extent that respondents remembered and repeated their answers from the original interview. In addition, several factors make it difficult to conduct the reinterview as an exact replication of the original interview, including (1) Only senior interviewers conducted the reinterview, (2) Almost all reinterviews were conducted by telephone, even if the original interview was conducted in person, and (3) The reinterview may not perfectly "anchor" respondents in the original interview's reference period.

[5]See Sinclair and Gastwirth (1996, 1998), which use the H-W model first proposed by Hui and

(2000) and Bassi and Trivellato (2008), explore the panel nature of the surveys and treat the underlying true labor force status as a latent process to be jointly modeled with the misclassification process.

Most existing studies focus on adjusting flows, i.e., the gross labor flows between two consecutive months, not stocks, such as the unemployment rate and the labor market participation rate. While those studies acknowledge that misclassification errors cause serious problems for flows, they somewhat surprisingly assume that errors tend to cancel out for stocks (e.g. Singh and Rao 1995). The only study that has tried to correct for the unemployment rate is Sinclair and Gastwirth (1998). However, their results rely on a key identification assumption that males and females have the same misclassification error probabilities, which we reject in this paper.

This paper uses recent results in the measurement error literature to identify the misclassification probabilities (Hu, 2008). Our method relies only on short panels formed by matching the CPS monthly data sets, thus avoiding the use of auxiliary information such as the reinterview surveys, which are usually small and subject to errors. Our approach is close to the Markov Latent Class Analysis (MLCA) method proposed by Biemer and Bushery (2000), but we use an eigenvalue-eigenvector decomposition to establish a closed-form global identification, while they took a maximum likelihood approach with local identifiability. Generally speaking, parametric GMM or MLE methods typically rely on a local identification argument that the number of unknowns does not exceed that of the restrictions. Given the observed distribution, our identification and estimation procedure directly leads to the unique true values of the unknown probabilities without using the regular optimization algorithms. Therefore, we do not need to be concerned about choosing initial values or obtaining a local maximum in the estimation procedure. In that sense, our estimates are more reliable than those based on local identification, including Biemer and Bushery (2000). Our assumption regarding the dynamics of the underlying true labor force status is also weaker than their first-order Markov chain assumption. In addition, Biemer and Bushery (2000) use group-level data, which are subject to potential biases from within-group heterogeneities. Our identification results enable us to take advantage

Walter (1980).

of the large sample size of the individual-level CPS data, and therefore, to achieve more efficient estimates.

To control for individual heterogeneities, we separately estimate the misclassification probabilities for each demographic group, defined by individual's gender, race and age. Based on those misclassification probabilities, we then estimate the corrected monthly unemployment rates and the labor force participation rates for all demographic groups, and for the US population as a whole. During the period from January 1996 to August 2011, our corrected unemployment rates are higher than the official ones by up to 4.4 percentage points and on average by 2.1 percentage points, with the differences always statistically significant. The most substantial misclassification errors occur when unemployed individuals misreport as either not-in-labor-force or employed. On the other hand, the corrected labor force participation rates and the official ones are rather close and never statistically significantly different.

The rest of the paper is organized as follows. Section 2 provides theoretical results on the identification and estimation of the misclassification probabilities and the marginal distribution of the underlying labor force status. Section 3 presents our main empirical results on the estimated misclassification probabilities and the corrected unemployment rates, along with reported (official) ones. The last section concludes. Additional estimates and simulation results are included in the online appendix of the paper.

# 2 A closed-form identification result

This section presents a closed-form identification and estimation procedure, which uniquely maps the directly estimable distribution of the self-reported labor force status to the misclassification probabilities and the distribution of the underlying true labor force status. We also evaluate the validity and robustness of the assumptions made in order to achieve identification.

## 2.1 Assumptions and identification results

Let $U_t$ denote the self-reported labor force status in month $t$, and $X$ be a vector of demographic variables such as gender, race and age. By matching the monthly CPS samples, we observe the self-reported labor force status in three periods $(t+1, t, t-9)$, together with the demographic variables $X$ for each individual $i$.[6] For example, if $U_t$ stands for the labor force status of an individual in January 2008, then $U_{t+1}$ and $U_{t-9}$ denote his or her labor force status in February 2008 and in April 2007, respectively. We denote the i.i.d. sample as $\{U_{t+1}, U_t, U_{t-9}, X\}_i$ for $i = 1, 2, ..., N$. The self-reported labor force status $U_t$ is defined as follows:

$$U_t = \begin{cases} 1 & \text{employed} \\ 2 & \text{unemployed} \\ 3 & \text{not-in-labor-force} \end{cases}.$$

We denote the latent true labor force status at period $t$ as $U_t^*$, which takes the same possible values as $U_t$. Let $\Pr(\cdot)$ stand for the probability distribution function of its argument, we outline our assumptions as follows.

**Assumption 1.** *The distribution of misclassification errors only depends the true labor force status in the current period, conditional on individual characteristics, i.e.,*

$$\Pr(U_t | U_t^*, X, \mathcal{U}_{\neq t}) = \Pr(U_t | U_t^*, X)$$

*for all $t$ with $\mathcal{U}_{\neq t} = \{(U_\tau, U_\tau^*),\ \text{for } \tau \neq t\}$.*

Assumption 1 still allows the misclassification errors to be correlated with the true labor force status $U_t^*$ and other variables in other periods through $U_t^*$. This is weaker

[6] Our identification strategy requires matching of three CPS monthly data sets in order to identify the misclassification matrix for the month in the middle of the three months. We choose one month later, i.e., $t+1$, and nine month earlier, i.e., $t-9$, for the following reasons: 1) we want the three periods to be close enough to minimize attrition in CPS samples; 2) we want the three months to cover the 8-month recess period in the CPS rotation structrue so that there are enough variations in the labor force status; 3) Assumption 2 on the dynamics of the latent true labor force status is more likely to be satisfied if we use the data reported a while ago, e.g., nine months earlier.

than the classical measurement error assumption, where the error is independent of everything else, including the true values. Assumption 1 is a standard assumption in the literature and allows the misreporting behavior to be summarized by a simple misclassification matrix. Moreover, Meyer (1988) examines this assumption and finds it likely to be valid for CPS data. Assumption 1 implies that the joint probability of the observed labor force status $\Pr(U_{t+1}, U_t, U_{t-9}|X)$ is associated with the unobserved ones as follows:

$$\begin{aligned} & \Pr(U_{t+1}, U_t, U_{t-9}|X) \\ = & \sum_{U^*_{t+1}} \sum_{U^*_t} \sum_{U^*_{t-9}} \Pr\left(U_{t+1}|U^*_{t+1}, X\right) \Pr\left(U_t|U^*_t, X\right) \Pr\left(U_{t-9}|U^*_{t-9}, X\right) \Pr\left(U^*_{t+1}, U^*_t, U^*_{t-9}|X\right). \end{aligned} \tag{1}$$

Having established the conditional independence of the misclassification process, our next assumption deals with the dynamics of the latent true labor force status.

**Assumption 2.** *Conditional on individual characteristics, the true labor force status nine months ago has no predictive power over the true labor force status in the next period beyond the current true labor force status, i.e.,*

$$\Pr\left(U^*_{t+1}|U^*_t, U^*_{t-9}, X\right) = \Pr\left(U^*_{t+1}|U^*_t, X\right)$$

*for all $t$.*

Biemer and Bushery (2000) impose a first-order Markov restriction on the dynamics of the latent labor force status, which states $\Pr\left(U^*_{t+1}|U^*_t, U^*_{t-1}, ..., U^*_1\right) = \Pr\left(U^*_{t+1}|U^*_t\right)$. That assumption is likely to be too strong due to the presence of state dependency, serial correlation among idiosyncratic shocks, and unobserved heterogeneity (see e.g. Hyslop 1999). Our assumption 2 is considerably weaker because we use the true labor force status nine month ago. Under Assumption 2, equation (1) may be simplified as follows:

$$\begin{aligned} & \Pr(U_{t+1}, U_t, U_{t-9}|X) \\ = & \sum_{U^*_t} \Pr\left(U_{t+1}|U^*_t, X\right) \Pr\left(U_t|U^*_t, X\right) \Pr\left(U^*_t, U_{t-9}|X\right). \end{aligned} \tag{2}$$

Following the identification results in Hu (2008), we show that all the probabilities containing the latent true labor force status $U_t^*$ on the right-hand-side (RHS) of Equation (2) may be identified under reasonable assumptions. Integrating out $U_{t+1}$ in Equation (2) leads to

$$\Pr\left(U_t, U_{t-9}|X\right) = \sum_{U_t^*} \Pr\left(U_t|U_t^*, X\right) \Pr\left(U_t^*, U_{t-9}|X\right). \tag{3}$$

Following Hu (2008), we introduce our matrix notation. For any given subpopulation with individual characteristics $X = x$, we define the misclassification matrix as follows.

$$\begin{aligned}
& M_{U_t|U_t^*,x} \\
\equiv\ & \begin{bmatrix} \Pr\left(U_t = 1|U_t^* = 1, x\right) & \Pr\left(U_t = 1|U_t^* = 2, x\right) & \Pr\left(U_t = 1|U_t^* = 3, x\right) \\ \Pr\left(U_t = 2|U_t^* = 1, x\right) & \Pr\left(U_t = 2|U_t^* = 2, x\right) & \Pr\left(U_t = 2|U_t^* = 3, x\right) \\ \Pr\left(U_t = 3|U_t^* = 1, x\right) & \Pr\left(U_t = 3|U_t^* = 2, x\right) & \Pr\left(U_t = 3|U_t^* = 3, x\right) \end{bmatrix} \\
\equiv\ & \left[\Pr\left(U_t = i|U_t^* = k, X = x\right)\right]_{i,k}.
\end{aligned}$$

Each column of the matrix $M_{U_t|U_t^*,x}$ describes how an individual (mis)reports his or her labor force status given a possible value of the true labor force status. The matrix $M_{U_t|U_t^*,x}$ contains the same information as the misclassification probabilities $\Pr\left(U_t|U_t^*, x\right)$, which means the identification of $M_{U_t|U_t^*,x}$ implies that of $\Pr\left(U_t|U_t^*, x\right)$. Similarly, we may define

$$\begin{aligned}
M_{U_t,U_{t-9}|x} &\equiv \left[\Pr\left(U_t = i, U_{t-9} = k|x\right)\right]_{i,k}, \\
M_{U_t^*,U_{t-9}|x} &\equiv \left[\Pr\left(U_t^* = i, U_{t-9} = k|x\right)\right]_{i,k}, \\
M_{1,U_t,U_{t-9}|x} &\equiv \left[\Pr\left(U_{t+1} = 1, U_t = i, U_{t-9} = k|x\right)\right]_{i,k}.
\end{aligned}$$

We also define a diagonal matrix as follows:

$$D_{1|U_t^*,x} \equiv \begin{bmatrix} \Pr\left(U_{t+1}=1|U_t^*=1,x\right) & 0 & 0 \\ 0 & \Pr\left(U_{t+1}=1|U_t^*=2,x\right) & 0 \\ 0 & 0 & \Pr\left(U_{t+1}=1|U_t^*=3,x\right) \end{bmatrix}.$$

As shown in Hu (2008), Equations (2) and (3) imply the following two matrix equations:

$$M_{1,U_t,U_{t-9}|x} = M_{U_t|U_t^*,x} D_{1|U_t^*,x} M_{U_t^*,U_{t-9}|x} \tag{4}$$

and

$$M_{U_t,U_{t-9}|x} = M_{U_t|U_t^*,x} M_{U_t^*,U_{t-9}|x}. \tag{5}$$

In order to solve for the unknown matrix $M_{U_t|U_t^*,x}$, we need a technical assumption as follows.

**Assumption 3.** *The distributions of the current self-reported labor force status conditional on different self-reported labor force statuses nine month ago are linearly independent, i.e.,* $\Pr\left(U_t|U_{t-9}=1,x\right)$ *is not equal to a linear combination of* $\Pr\left(U_t|U_{t-9}=2,x\right)$ *and* $\Pr\left(U_t|U_{t-9}=3,x\right)$ *for all* $U_t$ *and* $x$.

This assumption is equivalent to the condition that the matrix $M_{U_t,U_{t-9}|x}$ is invertible. Since it is imposed directly on the observed probabilities, this assumption is directly testable. Under Assumption 3, Equation (5) implies that both $M_{U_t|U_t^*,x}$ and $M_{U_t^*,U_{t-9}|x}$ are invertible. Eliminating matrix $M_{U_t^*,U_{t-9}|x}$ in Equations (4) and (5) leads to

$$M_{1,U_t,U_{t-9}|x} M^{-1}_{U_t,U_{t-9}|x} = M_{U_t|U_t^*,x} D_{1|U_t^*,x} M^{-1}_{U_t|U_t^*,x}. \tag{6}$$

This equation implies that the observed matrix on the left-hand-side (LHS) has an eigenvalue-eigenvector decomposition on the RHS. The three eigenvalues are the three diagonal entries in $D_{1|U_t^*,x}$ and the three corresponding eigenvectors are the three columns in $M_{U_t|U_t^*,x}$. Note that each column of $M_{U_t|U_t^*,x}$ is a distribution so that the column sum is 1, which implies that the eigenvectors are normalized.

In order to make the eigenvector unique for each given eigenvalue, we need the eigenvalues to be distinctive, which is formally stated as follows.

**Assumption 4.** *A different true labor force status leads to a different probability of reporting "employed" in the next period, i.e.,* $\Pr(U_{t+1} = 1|U_t^* = k, x)$ *are different for different* $k \in \{1, 2, 3\}$.

This assumption is also testable from Equation (6). This is because $\Pr(U_{t+1} = 1|U_t^* = k, x)$ for $k \in \{1, 2, 3\}$ are eigenvalues of the observed matrix $M_{1,U_t,U_{t-9}|x} M^{-1}_{U_t,U_{t-9}|x}$. Therefore, Assumption 4 holds if and only if all the eigenvalues of $M_{1,U_t,U_{t-9}|x} M^{-1}_{U_t,U_{t-9}|x}$ in Equation (6) are distinct. Intuitively, this assumption implies that the true labor force status at period $t$ has an impact on the probability of reporting to be employed one period later.

The distinct eigenvalues guarantee the uniqueness of the eigenvectors. Since we do not observe $U_t^*$ in the sample, we need to reveal the value $u_t^*$ for each eigenvector $\Pr(U_t|U_t^* = u_t^*, x)$. In other words, the ordering of the eigenvalues or the eigenvectors is still arbitrary in Equation (6). In order to eliminate this ambiguity, we make the following assumption.

**Assumption 5.** *Each individual is more likely to report the true labor force status than to report any other possible values, i.e.,*

$\Pr(U_t = k|U_t^* = k, x) > \Pr(U_t = j|U_t^* = k, x)$ *for* $j \neq k$.

This assumption does not reveal the value of these misclassification probabilities, nor require the probability of reporting the truth to be larger than 50%. Assumption 5 is consistent with results from CPS reinterviews (see e.g.: Poterba and Summers, 1984) and other validation studies discussed in Bound et al. (2001).

Technically, Assumption 5 implies that the true labor force status is the mode of the conditional distribution of the self-reported labor force status in each column of the eigenvector matrix. Therefore, the ordering of the eigenvectors is fixed and the the eigenvector matrix $M_{U_t|U_t^*,x}$ is uniquely determined from the eigenvalue-eigenvector decomposition of the observed matrix $M_{1,U_t,U_{t-9}|x} M^{-1}_{U_t,U_{t-9}|x}$. In particular, after diagonalizing the directly-estimable matrix $M_{1,U_t,U_{t-9}|x} M^{-1}_{U_t,U_{t-9}|x}$, we rearrange the order of the eigenvectors such that the largest element of each column or each eigenvector, i.e, the mode of the corresponding distribution, is on the diagonal of the eigenvector matrix. Consequently, the misclassification probability $\Pr(U_t|U_t^*, X)$ may be expressed as a closed-form function of the observed probability $\Pr(U_{t+1}, U_t, U_{t-9}|X)$. Such a procedure is constructive because one may estimate the misclassification probability $\Pr(U_t|U_t^*, X)$ by following the identification procedure above.

We summarize the closed-form identification and estimation of the misclassification probability $\Pr(U_t|U_t^*, X)$ as follows.

**Theorem 1.** *Under Assumptions 1, 2, 3, 4, and 5, the misclassification matrix* $\Pr(U_t|U_t^*, X)$ *is uniquely determined by the observed joint probability of the self-reported labor force status in three periods, i.e.,* $\Pr(U_{t+1}, U_t, U_{t-9}|X)$ *, through the unique eigenvalue-eigenvector decomposition in equation (6).*

**Proof:** The results directly follow from Theorem 1 in Hu (2008). A complete proof can be found in the online appendix.

Finally, we may estimate the distribution of the latent true labor force status $\Pr(U_t^*|X)$ using the misclassification probability $\Pr(U_t|U_t^*, X)$ from the following equation:

$$\Pr(U_t|X) = \sum_{U_t^*} \Pr(U_t|U_t^*, X) \Pr(U_t^*|X).$$

This equation implies

$$\begin{bmatrix} \Pr(U_t = 1|x) \\ \Pr(U_t = 2|x) \\ \Pr(U_t = 3|x) \end{bmatrix} = M_{U_t|U_t^*,x} \times \begin{bmatrix} \Pr(U_t^* = 1|x) \\ \Pr(U_t^* = 2|x) \\ \Pr(U_t^* = 3|x) \end{bmatrix}.$$

Since we have identified the misclassification probability $\Pr(U_t|U_t^*, X)$, we may solve for the distribution of the latent true labor force status $\Pr(U_t^*|X)$ from that of the self-reported labor force status $\Pr(U_t|X)$ by inverting the matrix $M_{U_t|U_t^*,x}$. Therefore, the distribution of the latent true labor force status for a given $x$ is identified as follows:

$$\begin{bmatrix} \Pr(U_t^* = 1|x) \\ \Pr(U_t^* = 2|x) \\ \Pr(U_t^* = 3|x) \end{bmatrix} = M_{U_t|U_t^*,x}^{-1} \times \begin{bmatrix} \Pr(U_t = 1|x) \\ \Pr(U_t = 2|x) \\ \Pr(U_t = 3|x) \end{bmatrix}. \tag{7}$$

Given the marginal distribution of the demographic characteristics $X$, $\Pr(X)$, we may identify the marginal distribution of the latent true labor force status $\Pr(U_t^*)$ as follows

$$\Pr(U_t^*) = \sum_X \Pr(U_t^*|X) \Pr(X).$$

This gives the unemployment rate

$$\mu_t^* \equiv \frac{\Pr(U_t^* = 2)}{\Pr(U_t^* = 1) + \Pr(U_t^* = 2)},$$

and the labor force participation rate

$$\rho_t^* \equiv \Pr(U_t^* = 1) + \Pr(U_t^* = 2).$$

Our identification procedure is constructive as it leads directly to an estimator. A nice property of our approach is that if there is no misclassification error in the data, our estimator would produce the same unemployment rate and labor force participation rate as those based on the raw data, under the assumptions above. Our estimator does not require an initial consistent estimate or iterations as in the regular optimization algorithms do.

## 2.2 Evaluation of the assumptions

Before proceeding to empirical work, we evaluate the key assumptions which are essential for our identification results. We perform extensive Monte Carlo simulations to examine the robustness of our estimator to deviations from Assumptions 1 and 2. We also test the validity of Assumptions 3 and 4 directly using CPS data. For

Assumption 5, we argue that it is likely to hold based on previous empirical work in the literature. We summarize the main things we have done here while leaving all detailed results in the online appendix.

Assumption 1 imposes conditional independence of the misreporting process. We have considered three different kinds of deviations to this assumption in our Monte Carlo simulations. In the first case, we allow misreporting errors to be correlated with the latent true labor force status in the previous period, i.e., $\Pr(U_t|U_t^*, \mathcal{U}_{\neq t}) = \Pr(U_t|U_t^*, U_{t-1}^*)$. In the second case, misreporting errors may be correlated with the self-reported labor force status in the previous period, i.e., $\Pr(U_t|U_t^*, \mathcal{U}_{\neq t}) = \Pr(U_t|U_t^*, U_{t-1})$. Lastly, we consider a special case of a general relaxation of Assumption 1, i.e., $\Pr(U_t|U_t^*, \mathcal{U}_{\neq t}) = \Pr(U_t|U_t^*, U_{t-1}^*, U_{t-1})$, where people would report the same value as in the previous period with certain probability if their true labor force status does not change, otherwise, they would report following the baseline misclassification probability $\Pr(U_t|U_t^*)$.[7] In all cases, our simulation results show that our estimator is robust to reasonable deviations from Assumption 1.[8]

Similarly, Assumption 2 imposes conditional independence on the transition of the underlying true labor force status. In the Monte Carlo simulation setup, we relax this assumption to allow the transition of the true LFS to depend on that 9 periods earlier, i.e., $\Pr(U_{t+1}^*|U_t^*, U_{t-9}^*) \neq \Pr(U_{t+1}^*|U_t^*)$. Our simulation results show that the estimator is robust to reasonable deviations to assumption 2.[9]

Assumption 3 requires an observed matrix to be invertible, and therefore, is directly testable from the CPS data. We use bootstrapping to show that the determinant of this matrix is significantly different from zero, which implies that the matrix is invertible.[10]

Under Assumptions 1, 2, and 3, Assumption 4 requires that the eigenvalues of an

[7] We do this in response to a referee's concern that reporting behaviors might be serially-correlated.

[8] The detailed Monte Carlo setup can be found at section 3.1.2 in the online appendix and the simulation results can be found at sections 3.2.2-3.2.4 in the online appendix.

[9] The detailed Monte Carlo setup can be found at section 3.1.3 in the online appendix and the simulation results can be found at section 3.2.5 in the online appendix.

[10] Detailed results can be found at section 4 (Table A11) in the online appendix.

observed matrix be distinct. We may also directly test this assumption using the CPS data by estimating the differences between the eigenvalues. Our bootstrapping results show that the absolute differences between the eigenvalues are significantly different from zero, which implies that the eigenvalues are distinctive.[11]

Assumption 5 implies that each individual is more likely to report the true labor force status than any other possible values. We believe this assumption is intuitively reasonable. Also, we are not aware of any studies in the literature (see e.g. previous studies cited in our paper and those reviewed by Bound et al. (2001)) that report anything in violation of this assumption.

# 3 Empirical Results

## 3.1 Matching of monthly CPS data

We use the public-use micro CPS data to estimate the unemployment rate and the labor force participation rate.[12] Each CPS monthly file contains eight rotation groups that differ in month-in-sample. The households in each rotation group are interviewed for four consecutive months after they enter, withdraw temporarily for eight months, then reenter for another four months of interviews before exiting the CPS permanently. Because of the rotational group structure, the CPS can be matched to form longitudinal panels, which enable us to obtain the joint probabilities of the self-reported labor force statuses in three periods.

We follow the algorithm proposed by Madrian and Lefgren (2000) to match adjacent CPS monthly files.[13] There are two main steps in the process of matching. First, the CPS samples are matched based on identifiers. If two individuals in two CPS monthly files (within the corresponding rotational groups) have the same household identifier,

[11] Results can be found at Table A12 of section 4 in the online appendix.

[12] All data are downloaded from www.bls.census.gov/cps_ftp.html. Following BLS practices, we restrict the samples to those aged 16 and over. Sample summary statistics can be found at Table A1 in the online appendix.

[13] See also Feng (2001) and Feng (2008).

household replacement number (which denotes whether this is a replacement of the initial household) and personal identifier (which uniquely identifiers a person within a household), then the two individuals are declared as a "crude match". This step is not perfect and may result in considerable matching errors because there might exist coding errors with respect to those identifiers. Therefore, the second step uses information on sex, age and race to "certify" the crude match. In the matching algorithm we use, if the sex or race reported in the two monthly files corresponding to a crude match are different, or if the age difference is greater than 1 or less than 0, then we discard the match as a false one.

As the previous literature (e.g.: Peracchi and Welch 1995 and Feng 2008) has documented, the matched sample is not representative of the cross-sectional sample in period $t$ due to sample attrition in matching. We use the matching weights to correct for attrition. First, we run a Logit regression for the period $t$ cross-sectional sample, where the dependent variable is either 1 (the observation is matched) or 0 (the observation is not matched), and the independent variables are sex, race, age, schooling, and the labor force status in period $t$. We next calculate the predicted probabilities of being matched for all the observations in the matched sample. The final matched sample is then weighted using the inverse of the predicted match probabilities. This adjustment procedure ensures the cross-sectional sample and the matched sample have the same marginal distributions on the key individual characteristics for period $t$.[14]

## 3.2 Misclassification probabilities

For each demographic group, we pool matched samples to estimate the misclassification probabilities.[15] Table 1 reports results for all the eight groups, including (1)

[14]Under the assumption that attrition is solely based on observables, our correction method using match weights is consistent. To check for robustness of our procedure we have also tried not using matching weights, i.e., not correcting for attrition in matching, and found similar results in terms of corrected unemployment rates. Details can be found at section 5.5 of the online appendix.

[15]To be consistent with the last version of the paper we pool data from January 1996 to December 2009. The estimated misclassification probabilities do not change statistically significantly if we pool all data up to August 2011. Please refer to section 5.3 of the online appendix for details and more

white males aged 40 and younger; (2) white males aged over 40; (3) nonwhite males aged 40 and younger; (4) nonwhite males aged over 40; (5) white females aged 40 and younger; (6) white females aged over 40; (7) nonwhite females aged 40 and younger; (8) nonwhite females aged over 40. There exist some consistent patterns across all the groups. When the actual labor force status is either employed or not-in-labor-force, the probabilities of being misreported to a different labor force status are typically small and never above 6%. The biggest errors come from the unemployed people being misclassified as either not-in-labor-force or employed. Only around 50-70% of unemployed people correctly report their true labor force status. For example, for white males aged 40 and younger, 20% of the unemployed report to be employed, while another 17% of them report as not-in-labor-force. On the other hand, there are considerable heterogeneities among different demographic groups. For example, 10.8% of the unemployed white females aged 40 and younger report as not-in-labor-force, while all other groups have much higher probabilities of reporting to be not-in-labor-force while unemployed.

We also formally test for the differences in the misclassification probabilities between the groups. For example, we consider males vs. females, controlling for race and age categories. We find that employed males are more likely to misreport as unemployed but less likely to misreport as not-in-labor-force than employed females. The differences are always statistically significant at the 5% significance level except for the comparison between nonwhite males aged 40 and younger and nonwhite females aged 40 and younger. When unemployed, the differences are mostly insignificant, with the only exception being that white males aged over 40 are less likely to misreport as being not-in-labor-force compared to white females aged over 40. In addition, when not-in-labor-force, males are more likely to be misclassified as employed.[16]

Some previous studies have made strong assumptions regarding between-group misclassification errors. For example, in order to achieve identification, Sinclair and Gastwirth (1998) assume that males and females have the same misclassification er-

elaborate discussions.

[16] Comparisons between males and females and other demographic characteristics can be found in Table A13 in the online appendix.

ror probabilities (see also Sinclair and Gastwirth 1996), which we can safely reject.[17] In general, our results suggest that the equality assumptions of misclassification probabilities across different demographic groups, which are essential for identification in the H-W models, are unlikely to hold in reality.

The last two rows of Table 1 report misclassification probabilities and associated standard errors for the overall U.S. population. The results are broadly consistent with those in the existing literature. When we compare our estimates of misclassification probabilities with some of those obtained in the existing literature,[18] we see the same general pattern: the biggest misclassification probabilities happen when unemployed individuals misreport their labor force statuses as either not-in-labor-force $(\Pr(U_t = 3|U_t^* = 2))$ or employed $(\Pr(U_t = 1|U_t^* = 2))$, while the other misclassification probabilities are all small. Our point estimates of $\Pr(U_t = 3|U_t^* = 2)$ and $\Pr(U_t = 1|U_t^* = 2)$ are somewhat higher than many of the existing estimates. But our estimates are well within the 95% confidence intervals in many existing studies because of their large standard errors. Due to our methodological advantages and the large sample size we use, we are able to produce much more precise estimates.

### 3.3 The unemployment rate

Given the estimated misclassification matrices, we then calculate distribution of the latent true labor force status for each demographic group based on Equation (7). To estimate $\Pr(U_t|X)$, we use all the eight rotation groups in any given CPS monthly file, which subsequently give us the self-reported unemployment rate and the labor force participation rate. Once we have $\Pr(U_t^*|X)$, we can calculate the corrected unemployment rate and the corrected labor force participation rate. In order to be consistent with officially-announced statistics, all numbers are weighted using final weights provided by CPS.[19]

[17] See the first panel in Table A13 in the online appendix.
[18] These estimates can be found in Table A14 in the online appendix.
[19] The final weights in the CPS micro data have been adjusted for a composite estimation procedure that BLS uses to produce official labor force statistics (Appendix I in BLS, 2000).

Table 2 presents the results for each demographic group. We divide the study period into three sub-periods based on the US business cycles.[20] The first sub-period goes from January 1996 to October 2001, which is roughly the end of the 2001 recession. The second sub-period is from November 2001 to November 2007, corresponding to the expansion period between two recessions (the 2001 recession and the most recent 2007-09 recession). The third sub-period goes from December 2007 to the end of our study period, i.e., Aug 2011, which includes the 2007-2009 recession and its aftermath.

For each demographic group and each sub-period, the corrected unemployment rates are always higher than the reported ones. Note also that for all demographic groups, sub-period 3 posts the highest levels of unemployment, followed by sub-period 2, and then by the first sub-period. This relationship is unchanged using either the reported or the corrected rates. In addition, the degree of underestimation is larger when the level of unemployment is higher. For example, for white males less than 40, in the first sub-period, the corrected unemployment rate is 6.5%, which is higher than the reported unemployment rate by 1.5 percentage points. In the second sub-period, the corrected unemployment rate is 8.2%, which is higher than the reported unemployment rate by 2.1%. The largest differential appears in the latest recession period. In this case, the corrected unemployment rate is 14.5%, which is higher than the reported unemployment rate by 4.4% – a 44% upward adjustment.

We then estimate the unemployment rates and the corresponding standard errors for the US population as a whole, based on the results for all the demographic groups. Based on the last two rows of Table 2, corrected unemployment rates for the US population are 5.9%, 6.9% and 11.5% for the three sub-periods, respectively. Note that the degree of underestimation is substantially larger in the third sub-period, official unemployment rate is 3.4 percentage points lower than the corrected one, while in the first two sub-periods the discrepancies are only 1.5 and 1.8 percentage points, respectively. Figure 1 displays all the monthly values that are seasonally-adjusted. For the whole period, the corrected unemployment rate is always higher than the reported one and the difference is between 1% and 4.4%, and 2.1% on average.

[20]see http://www.nber.org/cycles.html.

The substantial degree of underestimation of the unemployment rate may not be very surprising because most of the misclassification errors are from the unemployed people misreporting their labor force status as either employed or not-in-labor-force. We believe this arises primarily due to the intrinsic difficulties in classifying labor force status for some specific groups of people. Among those not-in-labor-force, marginally-attached workers, especially discouraged workers, could be classified as unemployed because they also desire a job although do not search in the job market. In fact, Jones and Riddell (1999) find that some marginally-attached workers are behaviorally more similar to unemployed than to the rest of those not-in-labor-force. On the other hand, involuntary part-time workers are classified as employed according to the official definition. But many of them could be observationally more similar to unemployed workers.[21] [22]

Table 3 decomposes the underestimation of unemployment rate. For the period of January 1996 to August 2011, the official statistics underestimate the unemployment rate on average by 2.1 percentage points. The degree of underestimation varies, however, by demographic group. On the one hand, the young nonwhite female group posts the largest level of underestimation, at 5 percentage points. On the other hand, the official statistics only underestimate by 1.3 percentage points for white males over 40. In terms of contributions to the total degree of underestimation (last column of Table 3), white females over 40 declare the largest share of the total (27%), followed by white males 40 and younger (21%). Nonwhite groups contributed relatively little as they account for relatively small portions of the US total population.

One particular concern is whether misclassification behaviors and the resulted corrected unemployment rates would depend on labor market conditions. For example, when the labor market is weak and the pool of unemployed people includes a larger share of job losers and others whose status is unambiguous, then the misreporting of unemployment would tend to be less prevalent. In order to test this hypothe-

[21]For example, Farber (1999) examine displaced workers and find temporary and involuntary part-time jobs are part of the transitional process from unemployment to full-time work.

[22]According to the broadest concept of unemployment by BLS, U6, all marginally-attached workers and involuntary part-time workers are counted as unemployed. Our corrected unemployment rate series are substantially lower than U6, as shown by Figure A4 in the online appendix.

sis directly, we have estimated three different misclassification probabilities for each demographic group for the three sub-periods. We do find some evidence that the misclassification probabilities are different in different sub-period corresponding to different labor market conditions. More specifically, sub-period 3 (December 2007 to August 2011), which is characterized by much higher rate of unemployment and presumably much weaker labor market conditions compared to the previous two sub-periods, has lower levels of misclassification in general. Nevertheless, we show that the corrected unemployment series are robust to whether we allow misclassification probabilities to differ in different sub-periods.[23].

We have also examined the effect of misclassification on the labor force participation (LFP) rates. For each demographic group for the three sub-periods: January 1996 to October 2001, November 2001 to November 2007, and December 2007 to August 2011, the corrected labor force participation rates are always higher than the reported ones, but the differences are small and not statistically significant. For the US population as a whole, average difference between corrected and official LFP rates is less than 2%, and not statistically significant. For the three sub-periods, the corrected labor force participation rates are 68.1%, 67.3% and 66.8%, respectively. The reported rates are only slightly lower, at 67.1%, 66.2% and 65.2%, respectively. [24] Therefore, misclassification errors cause little change to the labor force participation rate. Compared with the number of unemployed people, the total number of people who are in labor force is much larger. Hence any corrections due to misclassification errors will have a relatively small effect.

# 4 Conclusion

This paper examines misclassification errors in labor force status using CPS data. Similar to previous studies, we show that there exist considerable misclassifications from unemployed to not-in-labor-force and from unemployed to employed. The results at least partly reflect the intrinsic difficulties in classifying labor force statuses of

[23] Detailed results can be found at section 5.4 in the online appendix.
[24] Detailed results are shown in section 7 of the online appendix.

certain groups of people, such as marginally attached workers (especially discouraged workers) and part-time workers for economic reasons, into three distinct categories. We correct for such errors and show that the official U.S. unemployment rate significantly underestimates the true level of unemployment in the United States. For the period from January 1996 to August 2011, our corrected unemployment rates are higher than the reported ones by 2.1 percentage points on average, with differences ranging from 1 to 4.4 percentage points and always statistically significant. In addition, our estimates suggest that unemployment might be much more sensitive to business cycles than previously thought, as the degree of underestimation is larger in magnitude when unemployment rate is higher.

# References


J.M. Abowd and A. Zellner. Estimating gross labor-force flows. *Journal of Business and Economic Statistics*, 3(3):254–283, 1985.

F. Bassi and U. Trivellato. A latent class approach for estimating gross flows in the presence of correlated classification errors. In P. Lynn, editor, *Methodology of Longitudinal Studies*. Chichester, Wiley, 2008.

P.P. Biemer and J.M. Bushery. On the validity of markov latent class analysis for estimating classification error in labor force data. *Survey Methodology*, 26(2):139–152, 2000.

P.P. Biemer and G. Forsman. On the quality of reinterview data with application to the Current Population Survey. *Journal of the American Statistical Association*, 87(420):915–923, 1992.

BLS. *Current Population Survey: Design and Methodology.* Bureau of Labor Statistics, 2000. Technical Paper 63RV.

J. Bound, C. Brown, and N. Mathiowetz. Measurement error in survey data. In *Handbook of Econometrics*, volume 5, pages 3705–3843. 2001.

H.S. Farber. Alternative and part-time employment arrangements as a response to job loss. *Journal of Labor Economics*, 17(S4):142–169, 1999.

S. Feng. The longitudinal matching of Current Population Surveys: A proposed algorithm. *Journal of Economic and Social Measurement*, 27(1-2):71–91, 2001.

S. Feng. Longitudinal matching of recent Current Population Surveys: Methods, non-matches and mismatches. *Journal of Economic and Social Measurement*, 33 (4):241–252, 2008.

Y. Hu. Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution. *Journal of Econometrics*, 144: 27–61, 2008.

S.L. Hui and S.D. Walter. Estimating the error rates of diagnostic tests. *Biometrics*, 36:167–171, 1980.

D.R. Hyslop. State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women. *Econometrica*, 67(6):1255–1294, 1999.

S.R.G. Jones and C.W. Riddell. The measurement of unemployment: an empirical approach. *Econometrica*, 67(1):147–162, 1999.

B. Madrian and L. Lefgren. An approach to longitudinally matching Current Population Survey (CPS) respondents. *Journal of Economic and Social Measurement*, 26:31–62, 2000.

T. Magnac and M. Visser. Transition models with measurement errors. *Review of Economics and Statistics*, 81(3):466–474, 1999.

B.D. Meyer. Classification-error models and labor-market dynamics. *Journal of Business and Economic Statistics*, 6(3):385–390, 1988.

F. Peracchi and F. Welch. How representative are matched cross sections? evidence from the Current Population Survey. *Journal of Econometrics*, 68(1):153–179, 1995.

J.M. Poterba and L.H. Summers. Response variation in the CPS: caveats for unemployment analysts. *Monthly Labor Review*, pages 37–43, March 1984.

J.M. Poterba and L.H. Summers. Reporting errors and labor market dynamics. *Econometrica*, 54(6):1319–1338, 1986.

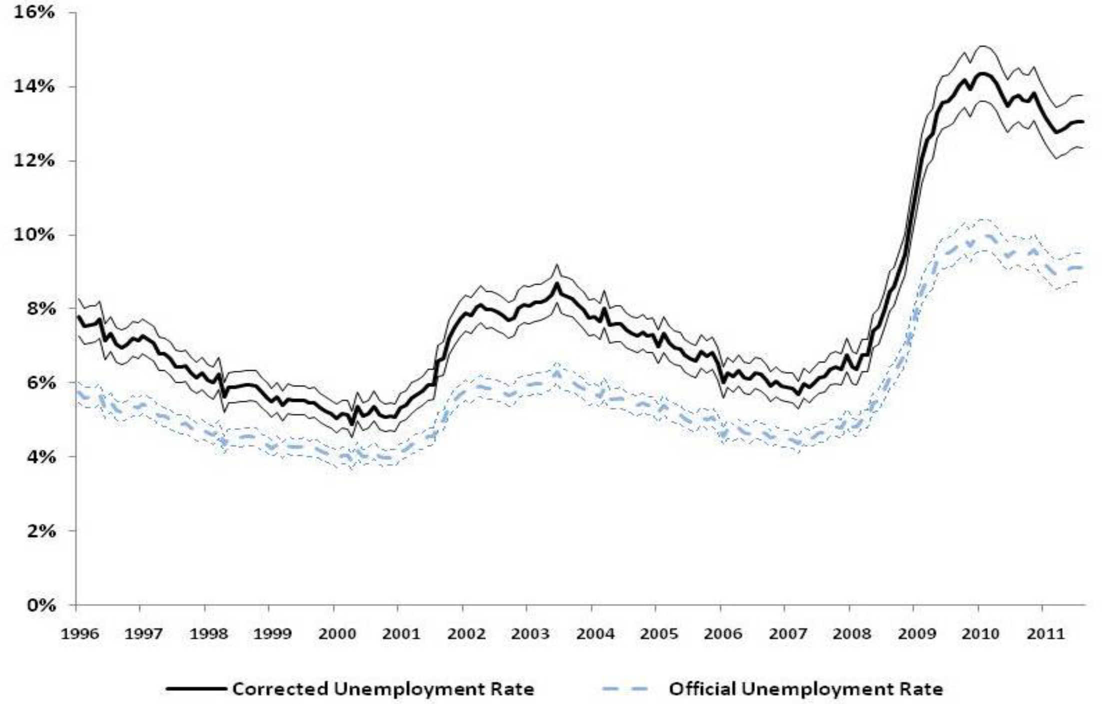
J.M. Poterba and L.H. Summers. Unemployment benefits and labor market transitions: a multinomial logit model with errors in classification. *Review of Economics and Statistics*, 77(2):207–216, 1995.

M.D. Sinclair and J.L. Gastwirth. On procedures for evaluating the effectiveness of reinterview survey methods: application to labor force data. *Journal of the American Statistical Association*, 91(435):961–969, 1996.

M.D. Sinclair and J.L. Gastwirth. Estimates of the errors in classification in the labour force survey and their effect on the reported unemployment rate. *Survey methodology*, 24(2):157–169, 1998.

A.C. Singh and J.N.K. Rao. On the adjustment of gross flow estimates for classification error with application to data from the Canadian labour force survey. *Journal of the American Statistical Association*, 90(430):478–488, 1995.

Figure 1: Corrected and official (reported) unemployment rates




Note: Figure displays seasonally-adjusted corrected unemployment rates (in solid line) and official unemployment rates (in dashed line) for the whole population from Jan 1996 to Aug 2011. The corresponding thin lines signify 95% upper and lower confidence bounds. For seasonally adjustment, we use Census Bureau's WinX12 software.

Table 1: Misclassification probabilities (%) for different demographic groups

| | $\Pr(i\|j) \equiv \Pr(U_t = i\|U_t^* = j)$ | | | | | |
|---|---|---|---|---|---|---|
| Demographic group | Pr(2\|1) | Pr(3\|1) | Pr(1\|2) | Pr(3\|2) | Pr(1\|3) | Pr(2\|3) |
| (1) Male/White/age≤40 | 0.9 | 1.3 | 20.1 | 17.2 | 6.0 | 0.0 |
| | (0.06) | (0.07) | (1.28) | (2.69) | (0.42) | (0.39) |
| (2) Male/White/age > 40 | 0.4 | 0.9 | 16.5 | 18.8 | 1.4 | 0.1 |
| | (0.03) | (0.05) | (1.14) | (2.34) | (0.07) | (0.07) |
| (3) Male/Nonwhite/age ≤ 40 | 1.1 | 2.2 | 13.4 | 18.1 | 5.0 | 4.3 |
| | (0.10) | (0.13) | (1.21) | (3.91) | (0.36) | (1.26) |
| (4) Male/Nonwhite/age > 40 | 0.7 | 1.5 | 15.5 | 22.0 | 1.2 | 0.0 |
| | (0.08) | (0.10) | (1.81) | (5.55) | (0.16) | (0.12) |
| (5) Female/White/age≤40 | 0.6 | 2.1 | 18.6 | 10.8 | 4.4 | 0.0 |
| | (0.05) | (0.10) | (1.59) | (4.10) | (0.27) | (0.08) |
| (6) Female/White/age>40 | 0.3 | 1.4 | 17.9 | 28.2 | 1.0 | 0.0 |
| | (0.03) | (0.07) | (1.46) | (3.16) | (0.06) | (0.01) |
| (7) Female/Nonwhite/age≤40 | 1.1 | 2.6 | 11.8 | 29.4 | 2.2 | 0.0 |
| | (0.09) | (0.16) | (1.54) | (8.24) | (0.70) | (0.01) |
| (8) Female/Nonwhite/age>40 | 0.4 | 1.8 | 13.9 | 25.0 | 1.2 | 0.7 |
| | (0.07) | (0.11) | (1.89) | (5.82) | (0.09) | (0.17) |
| Overall | 0.6 | 1.5 | 17.3 | 20.2 | 2.9 | 0.2 |
| | (0.02) | (0.03) | (0.59) | (1.39) | (0.10) | (0.09) |

Note: Bootstrap standard errors based on 500 repetitions are reported in parentheses.

Table 2: Unemployment rates (%) averaged over three sub-periods for different demographic groups

| Demographic group | Sub-period 1 (1996/01-2001/10) | | Sub-period 2 (2001/11-2007/11) | | Sub-period 3 (2007/12-2011/8) | |
|---|---|---|---|---|---|---|
| | reported | corrected | reported | corrected | reported | corrected |
| (1) Male/White/age≤40 | 5.0 | 6.5 | 6.1 | 8.2 | 10.1 | 14.5 |
| | (0.2) | (0.4) | (0.3) | (0.5) | (0.5) | (0.8) |
| (2) Male/White/age>40 | 2.7 | 3.4 | 3.4 | 4.5 | 6.3 | 8.9 |
| | (0.1) | (0.2) | (0.2) | (0.2) | (0.3) | (0.5) |
| (3) Male/Nonwhite/age≤40 | 10.1 | 11.1 | 10.8 | 12.0 | 16.0 | 19.3 |
| | (0.5) | (0.9) | (0.5) | (1.1) | (0.7) | (1.4) |
| (4) Male/Nonwhite/age>40 | 4.8 | 6.5 | 5.8 | 8.0 | 9.6 | 13.9 |
| | (0.2) | (0.5) | (0.3) | (0.6) | (0.4) | (1.0) |
| (5) Female/White/age≤40 | 5.1 | 6.4 | 5.8 | 7.3 | 8.3 | 10.9 |
| | (0.2) | (0.4) | (0.3) | (0.5) | (0.4) | (0.7) |
| (6) Female/White/age>40 | 2.7 | 4.4 | 3.2 | 5.3 | 5.4 | 9.1 |
| | (0.1) | (0.3) | (0.1) | (0.3) | (0.2) | (0.5) |
| (7) Female/Nonwhite/age≤40 | 10.0 | 14.5 | 10.3 | 14.9 | 13.4 | 19.8 |
| | (0.5) | (1.5) | (0.5) | (1.6) | (0.6) | (2.0) |
| (8) Female/Nonwhite/age>40 | 4.2 | 5.1 | 5.2 | 6.8 | 7.2 | 10.0 |
| | (0.2) | (0.5) | (0.2) | (0.6) | (0.3) | (0.9) |
| Overall | 4.4 | 5.9 | 5.1 | 6.9 | 8.1 | 11.5 |
| | (0.1) | (0.2) | (0.1) | (0.2) | (0.1) | (0.3) |

Note: Numbers reported in parentheses are bootstrap standard errors based on 500 repetitions.

Table 3: Decomposition of underestimation in unemployment rates by demographic groups

| | Underestimation in unemployment rate | Group share in US population | Contribution to underestimation | Relative contribution |
|---|---|---|---|---|
| Demographic group | $(a) = \widehat{\mu_t^*} - \mu_t$ | $(b)$ | $(c) = (a) \times (b)$ | $(d) = \frac{(c)}{\sum(c)}$ |
| (1) Male/White/age≤40 | 2.41 | 18.24 | 0.44 | 20.57 |
| (2) Male/White/age>40 | 1.34 | 21.80 | 0.29 | 13.65 |
| (3) Male/Nonwhite/age≤40 | 1.72 | 4.46 | 0.08 | 3.59 |
| (4) Male/Nonwhite/age>40 | 2.65 | 3.73 | 0.10 | 4.63 |
| (5) Female/White/age≤40 | 1.68 | 17.91 | 0.30 | 14.08 |
| (6) Female/White/age>40 | 2.37 | 24.20 | 0.57 | 26.82 |
| (7) Female/Nonwhite/age≤40 | 5.05 | 4.99 | 0.25 | 11.79 |
| (8) Female/Nonwhite/age>40 | 1.76 | 4.68 | 0.08 | 3.86 |
| Total | | 100.00 | 2.14 | 100.00 |

Note: Table reports averages over the January 1996 to August 2011 period. All numbers are rounded. (a) Underestimation in the unemployment rate (%), which equals the average corrected unemployment rate $\widehat{\mu_t^*}$ minus the average official unemployment rate $\mu_t$; (b) Population share of the demographic group; (c) Contribution to the total US underestimation in the unemployment rate (%), which equals (a) times (b); (d) Relative contribution to the total underestimation, which equals (c) divided by its column sum.

# Online Appendix: Misclassification Errors and the Underestimation of the U.S. Unemployment Rate

Shuaizhang Feng* Yingyao Hu†

January 28, 2012

**Abstract**

This online appendix accompanies the paper "Misclassification Errors and the Underestimation of the U.S. Unemployment Rate" by Shuaizhang Feng and Yingyao Hu. Section 1 of the appendix lists summary statistics of the CPS sample used in the paper. Section 2 of the appendix provides a detailed proof of theorem 1 in the paper. Section 3 evaluates assumptions 1 and 2 in the paper through detailed monte carlo simulations. Section 4 tests assumptions 3 and 4 in the paper directly using CPS data. Additional empirical results including robustness checks are presented in sections 5, 6 and 7.

*School of Economics, Shanghai University of Finance and Economics, 715 Econ Building, 111 Wuchuan Road, Shanghai 200433, China and IZA. E-mail: shuaizhang.feng@gmail.com

†Corresponding author. Department of Economics, Johns Hopkins University, 440 Mergenthaler Hall, 3400 N. Charles Street, Baltimore, MD 21218, USA and IZA. E-mail: yhu@jhu.edu

## Outline

# 1 Summary statistics

We use monthly basic CPS data from January 1996 to August 2011. The whole study period is divided into three sub-periods based on the US business cycles.[1] The first sub-period goes from the beginning of our study period to October 2001, which is roughly the end of the 2001 recession. The second sub-period goes from November 2001 to November 2007, corresponding to the expansion period between two recessions (the 2001 recession and the most recent 2007-09 recession). The third sub-period goes from December 2007 to the end of our study period (Aug 2011), which includes the 2007-2009 recession and its aftermath.

Table A1 presents simple summary statistics. Because all variables are 0/1 dummies, we only report sample means. The sample includes all persons aged 16 years and over in CPS monthly files. The first three rows show labor force statuses. For the whole study period, around 63% are employed, and 4% are unemployed and the rest are not in the labor force. Compared with the first two sub-periods, sub-period 3 is characterized by much higher levels of unemployment and presumably reflecting considerably weaker labor market conditions.

The next three rows summarize some demographic characteristics in the sample which we use to divide the whole sample into different demographic groups. The percentages of females, nonwhites and those aged below 40 are 53%, 16% and 43%, respectively. The last row shows sample sizes. For the whole period we have a sample size of over 19 million.

[1]See http://www.nber.org/cycles.html.

Table A1: Summary statistics: sample means

| | Sub-period 1 (1996/01-2001/10) | Sub-period 2 (2001/11-2007/11) | Sub-period 3 (2007/12-2011/8) | All Periods (1996/01-2011/08) |
|---|---|---|---|---|
| Employed | 0.64 | 0.63 | 0.60 | 0.63 |
| Unemployed | 0.03 | 0.03 | 0.05 | 0.04 |
| Not in labor force | 0.33 | 0.33 | 0.35 | 0.34 |
| Female | 0.53 | 0.53 | 0.52 | 0.53 |
| Nonwhite | 0.15 | 0.16 | 0.17 | 0.16 |
| 40 or younger | 0.46 | 0.42 | 0.41 | 0.43 |
| Sample size | 6,540,589 | 7,790,199 | 4,736,019 | 19,066,807 |

Note: Sample restricted to those aged 16 and above. Standard deviations are not reported as all variables are 0/1 dummies.

## 2 Proof of Theorem 1

This section provides a formal proof of Theorem 1, which states that Under Assumptions 1, 2, 3, 4, and 5, the misclassification probability of the labor force status, i.e., $\Pr(U_t|U_t^*, X)$, is uniquely determined by the observed joint probability of the self-reported labor force status in three periods, i.e., $\Pr(U_{t+1}, U_t, U_{t-9}|X)$, through a unique eigenvalue-eigenvector decomposition.

Assumptions 1 implies that the observed joint probability of the self-reported labor force status equals

$$\begin{aligned}
& \Pr(U_{t+1}, U_t, U_{t-9}|X) \\
= & \sum_{U_{t+1}^*} \sum_{U_t^*} \sum_{U_{t-9}^*} \Pr\left(U_{t+1}, U_t, U_{t-9}, U_{t+1}^*, U_t^*, U_{t-9}^*|X\right) \\
= & \sum_{U_{t+1}^*} \sum_{U_t^*} \sum_{U_{t-9}^*} \Pr\left(U_{t+1}|U_{t+1}^*, U_t^*, U_{t-9}^*, U_t, U_{t-9}, X\right) \Pr\left(U_t|U_{t+1}^*, U_t^*, U_{t-9}^*, U_{t-9}, X\right) \\
& \times \Pr\left(U_{t-9}|U_{t+1}^*, U_t^*, U_{t-9}^*, X\right) \Pr\left(U_{t+1}^*, U_t^*, U_{t-9}^*|X\right) \\
= & \sum_{U_{t+1}^*} \sum_{U_t^*} \sum_{U_{t-9}^*} \Pr\left(U_{t+1}|U_{t+1}^*, X\right) \Pr(U_t|U_t^*, X) \Pr\left(U_{t-9}|U_{t-9}^*, X\right) \Pr\left(U_{t+1}^*, U_t^*, U_{t-9}^*|X\right).
\end{aligned} \tag{1}$$

Furthermroe, Assumption 2 implies that

$$\begin{aligned}
& \Pr(U_{t+1}, U_t, U_{t-9}|X) \\
= & \sum_{U_{t+1}^*} \sum_{U_t^*} \sum_{U_{t-9}^*} \Pr\left(U_{t+1}|U_{t+1}^*, X\right) \Pr(U_t|U_t^*, X) \Pr\left(U_{t-9}|U_{t-9}^*, X\right) \\
& \times \Pr\left(U_{t+1}^*|U_t^*, X\right) \Pr\left(U_t^*, U_{t-9}^*|X\right) \\
= & \sum_{U_t^*} \left( \sum_{U_{t+1}^*} \Pr\left(U_{t+1}|U_{t+1}^*, X\right) \Pr\left(U_{t+1}^*|U_t^*, X\right) \right) \Pr(U_t|U_t^*, X) \\
& \times \left( \sum_{U_{t-9}^*} \Pr\left(U_{t-9}|U_{t-9}^*, X\right) \Pr\left(U_t^*, U_{t-9}^*|X\right) \right) \\
= & \sum_{U_t^*} \Pr(U_{t+1}|U_t^*, X) \Pr(U_t|U_t^*, X) \Pr(U_t^*, U_{t-9}|X).
\end{aligned} \tag{2}$$

Integrating out $U_{t+1}$ leads to

$$\Pr\left(U_t, U_{t-9}|X\right) = \sum_{U_t^*} \Pr\left(U_t|U_t^*, X\right) \Pr\left(U_t^*, U_{t-9}|X\right). \tag{3}$$

We then define

$$\begin{aligned}
M_{U_t|U_t^*,x} &= \left[\Pr\left(U_t = i|U_t^* = k, X = x\right)\right]_{i,k} \\
M_{U_t,U_{t-9}|x} &= \left[\Pr\left(U_t = i, U_{t-9} = k|X = x\right)\right]_{i,k}, \\
M_{U_t^*,U_{t-9}|x} &= \left[\Pr\left(U_t^* = i, U_{t-9} = k|X = x\right)\right]_{i,k}, \\
M_{1,U_t,U_{t-9}|x} &= \left[\Pr\left(U_{t+1} = 1, U_t = i, U_{t-9} = k|X = x\right)\right]_{i,k}
\end{aligned}$$

and a diagonal matrix

$$D_{1|U_t^*,x} = diag\left[\Pr\left(U_{t+1} = 1|U_t^* = k, X = x\right)\right]_k.$$

As shown in Hu (2008), Equations (2) and (3) are equvalent to the following two matrix equations

$$M_{1,U_t,U_{t-9}|x} = M_{U_t|U_t^*,x} D_{1|U_t^*,x} M_{U_t^*,U_{t-9}|x} \tag{4}$$

and

$$M_{U_t,U_{t-9}|x} = M_{U_t|U_t^*,x} M_{U_t^*,U_{t-9}|x}. \tag{5}$$

Assumption 3 implies that the matrix $M_{U_t,U_{t-9}|x}$ is invertible. We may then consider

$$\begin{aligned}
M_{1,U_t,U_{t-9}|x} M_{U_t,U_{t-9}|x}^{-1} &= M_{U_t|U_t^*,x} D_{1|U_t^*,x} M_{U_t^*,U_{t-9}|x} \left(M_{U_t|U_t^*,x} M_{U_t^*,U_{t-9}|x}\right)^{-1} \\
&= M_{U_t|U_t^*,x} D_{1|U_t^*,x} \left(M_{U_t^*,U_{t-9}|x} M_{U_t^*,U_{t-9}|x}^{-1}\right) M_{U_t|U_t^*,x}^{-1} \\
&= M_{U_t|U_t^*,x} D_{1|U_t^*,x} M_{U_t|U_t^*,x}^{-1}. \qquad (6)
\end{aligned}$$

This equation implies that the observed matrix on the left-hand-side (LHS) has an eigenvalue-eigenvector decomposition on the RHS. The three eigenvalues are the three diagonal entries in $D_{1|U_t^*,x}$ and the three eigenvectors are the three columns in $M_{U_t|U_t^*,x}$. Note that each column of $M_{U_t|U_t^*,x}$ is a distribution so that the column sum is 1, which implies that the eigenvectors are normalized. Assumption 4 implies that the eigenvalues are distinctive, and therefore, the three eigenvectors are linearly independent.

Assumption 5 implies that the true labor force status is the mode of the conditional distribution of the self-reported labor force status in each column of the eigenvector matrix. After diagonalizing the directly-estimable matrix $M_{1,U_t,U_{t-9}|x} M_{U_t,U_{t-9}|x}^{-1}$, we rearrange the order of the eigenvectors such that the largest element of each column or each eigenvector, i.e, the mode of the corresponding distribution, is on the diagonal of the eigenvector matrix. Therefore, the ordering of the eigenvectors is fixed and the the eigenvector matrix $M_{U_t|U_t^*,x}$ is uniquely determined from the eigenvalue-eigenvector decomposition of the observed matrix $M_{1,U_t,U_{t-9}|x}M^{-1}_{U_t,U_{t-9}|x}$. *QED.*

# 3 Evaluation of Assumptions 1 &2 by Monte Carlo simulations

In this section, we use simulated data to show the robustness of the estimator in Feng and Hu (2011). First, we present a baseline data generating process (DGP) which satisfies all the maintained assumptions, and show the consistency of our estimator. Second, we let the DGP deviate from the baseline case to check the robustness of our estimator when each of the assumptions is violated.

## 3.1 Simulation setups

We start with the definition of notations. Let $U_t^*$ and $U_t$ denote the true and self-reported labor force status (LFS) in period $t$, respectively. The marginal distribution of the true LFS is denoted as:

$$\Pr(U_t^*) = \left[\Pr(U_t^* = 1), \Pr(U_t^* = 2), \Pr(U_t^* = 3)\right]^T,$$

where $\Pr(U_t^* = k)$ is the probability that the true LFS is $k$ (1:employed, 2:unemployed, 3:not-in-labor-force). Given the marginal distribution, we may generate the true LFS for each observation in our simulated sample. We let the underlying true LFS follow a first-order Markov process, defined by the following Markovian transition matrix:

$$M_{U_t^*|U_{t-1}^*} = \left[\Pr(U_t^* = i|U_{t-1}^* = j)\right]_{i,j},$$

where $\Pr(U_t^* = i|U_{t-1}^* = j)$ is the conditional probability $\Pr(U_t^* = i|U_{t-1}^* = j)$. We assume the Markov kernal is time-invariant. Therefore, the two-period Markov transition matrix is

$$M_{U_t^*|U_{t-2}^*} = M_{U_t^*|U_{t-1}^*} M_{U_{t-1}^*|U_{t-2}^*} = M^2_{U_t^*|U_{t-1}^*}.$$

In general, a k-period transition matrix is $M_{U_t^*|U_{t-k}^*} = M^k_{U_t^*|U_{t-1}^*}$. We generate the series of LFS according to these conditional probabilities. The self-reported LFS $U_t$ is generated according to the true LFS $U_t^*$ and the misclassification probability

$$M_{U_t|U_t^*} = \left[\Pr(U_t = i|U_t^* = j)\right]_{i,j}.$$

#### 3.1.1 Setup #1: consistency under maintained assumptions

First, we present the baseline DGP which satisfies all the maintained assumptions to show the consistency of our estimator.

We start by choosing the marginal distribution of the true LFS at period $t-10$, $\Pr(U^*_{t-10})$ and the Markov transition matrix $M_{U^*_t|U^*_{t-1}}$, where parameter values are chosen to mimic real CPS data. Each observation contains $U^*_t$ and $U_t$ in several periods, which are genreated as follows:

Step 1: draw the true LFS $U^*_{t-10}$ at period $t-10$ according to the distribution $\Pr(U^*_{t-10})$;

Step 2: draw $U^*_{t-9}$ according to the Markovian transition matrix $M_{U^*_t|U^*_{t-1}}$ and $U^*_{t-10}$. That means if $U^*_{t-10}$ in step 1 equals 1, we use the distribution in the first column of $M_{U^*_t|U^*_{t-1}}$ to generate $U^*_{t-9}$; if $U^*_{t-10}$ in step 1 equals 2, we use the second column of $M_{U^*_t|U^*_{t-1}}$; if $U^*_{t-10}$ in step 1 equals 3, we use the third column of $M_{U^*_t|U^*_{t-1}}$;

Step 3, draw $U^*_{t-8}$ using $U^*_{t-9}$ in step 2 and $M_{U^*_t|U^*_{t-1}}$;

Step 4, draw $U^*_{t-1}$ using $U^*_{t-8}$ in step 3 and $M_{U^*_t|U^*_{t-7}} = M^7_{U^*_t|U^*_{t-1}}$;

Step 5, draw $U^*_t$ using $U^*_{t-1}$ in step 4 and $M_{U^*_t|U^*_{t-1}}$;

Step 6, draw $U^*_{t+1}$ using $U^*_t$ in step 5 and $M_{U^*_t|U^*_{t-1}}$;

After we have generated the true LFS in six periods $\left(U^*_{t+1}, U^*_t, U^*_{t-1}, U^*_{t-8}, U^*_{t-9}, U^*_{t-10}\right)$. we then generate the observed (misreported) LFS.

Step 7, draw $U_s$ using $U^*_s$ and $M_{U_t|U^*_t}$, respectively for $s = t+1, t, t-1, t-8, t-9, t-10$. The observed LFS $U_{t+1}, U_t, U_{t-1}, U_{t-8}, U_{t-9}, U_{t-10}$ are generated independently one by one.

In order to show consistency, we only need $(U_{t+1}, U_t, U_{t-9})$ in each observation. We then repeat the steps 1-7 $N$ times to obtain an i.i.d. sample containing $\{U_{i,t+1}, U_{i,t}, U_{i,t-9}\}$ for $i = 1, 2, ..., N$.

We choose the following parameter values to mimic the observed CPS data:

$$M_{U_t|U^*_t} = \begin{bmatrix} 0.98 & 0.2 & 0.01 \\ 0.01 & 0.6 & 0.01 \\ 0.01 & 0.2 & 0.98 \end{bmatrix}, \Pr(U^*_{t-10}) = \begin{bmatrix} 0.6256 \\ 0.0544 \\ 0.32 \end{bmatrix}$$

$$M_{U^*_t|U^*_{t-1}} = \begin{bmatrix} 0.98 & 0.1 & 0.015 \\ 0.01 & 0.85 & 0.005 \\ 0.01 & 0.05 & 0.98 \end{bmatrix}.$$

Note that all simulation results are qualitatively robust to perturbations of parameter values within reasonable ranges. For brevity, we only report results for the parameter values as chosen.

We then apply our estimator to this simulated sample of $\{U_{i,t+1}, U_{i,t}, U_{i,t-9}\}$ to estimate $M_{U_t|U_t^*}$ as well as $U_t^*$ to check the consistency of our estimator.

#### 3.1.2 Setup #2: checking robustness of assumption 1

In this subsection, we relax assumption 1 in the DGP by allowing the misclassification probability matrix $M_{U_t|U_t^*}$ to vary with the self-reported LFS $U_{t-1}$ or the true LFS $U_{t-1}^*$.

**Case 1: reported LFS depends on last period true LFS** In this case, We relax our assumption 1 to allow the misreporting probability to depend on the true labor force status in the previous period. That is,

$$\begin{aligned}\Pr(U_t|U_t^*, U_{<t}, U_{<t}^*) &= \Pr(U_t|U_t^*, U_{t-1}^*) \\ &\neq \Pr(U_t|U_t^*).\end{aligned}$$

Under this relaxed version, we only need to change step 7 in the baseline setup #1 as follows:

Step 7, draw $U_s$ using $U_s^*, U_{s-1}^*$ and $\Pr(U_t|U_t^*, U_{t-1}^*)$ respectively for $s = t+1, t, t-9$.

The conditional probability $\Pr(U_t|U_t^*, U_{t-1}^*)$ may be expressed as three misclassification probabilities: $\Pr(U_t|U_t^*, U_{t-1}^* = 1)$ , $\Pr(U_t|U_t^*, U_{t-1}^* = 2)$, and $\Pr(U_t|U_t^*, U_{t-1}^* = 3)$. In matrix notation, we may have

$$M_{U_t|U_t^*,U_{t-1}^*} \equiv \left[\begin{array}{ccc} M_{U_t|U_t^*,U_{t-1}^*=1} & M_{U_t|U_t^*,U_{t-1}^*=2} & M_{U_t|U_t^*,U_{t-1}^*=3} \end{array}\right].$$

We may allow $M_{U_t|U_t^*,U_{t-1}^*=k}$ for $k = 1, 2, 3$ to deviate from the baseline misclassification probability $M_{U_t|U_t^*}$. Obviously, there are many ways to deviate from $M_{U_t|U_t^*}$ or $\Pr(U_t|U_t^*)$ to $M_{U_t|U_t^*,U_{t-1}^*=k}$ or $\Pr(U_t|U_t^*, U_{t-1}^* = k)$. In our simulation, the matrices $M_{U_t|U_t^*,U_{t-1}^*=k}$ are generated by letting the entries in $M_{U_t|U_t^*}$ to deviate according to the confidence intervals in the baseline case.

Let the original

$$M_{U_t|U_t^*} = \begin{bmatrix} m_{1|1} & m_{1|2} & m_{1|3} \\ m_{2|1} & m_{2|2} & m_{2|3} \\ m_{3|1} & m_{3|2} & m_{3|3} \end{bmatrix}.$$

Under assumption 1, we have $M_{U_t|U_t^*,U_{t-1}^*} = \left[ \begin{array}{ccc} M_{U_t|U_t^*} & M_{U_t|U_t^*} & M_{U_t|U_t^*} \end{array} \right]$. This misclassification matrix transforms joint distribution of true LFS in periods $t$ and $t-1$ into observed LFS at period t, i.e., $\Pr(U_t) = M_{U_t|U_t^*,U_{t-1}^*} \Pr(U_t^*, U_{t-1}^*)$, where

$$\Pr(U_t) = \left[ \begin{array}{c} p(U_t = 1) \\ p(U_t = 2) \\ p(U_t = 3) \end{array} \right]$$

and

$$\Pr(U_t^*, U_{t-1}^*) = \left[ \begin{array}{c} p(U_t^* = 1, U_{t-1}^* = 1) \\ p(U_t^* = 2, U_{t-1}^* = 1) \\ p(U_t^* = 3, U_{t-1}^* = 1) \\ p(U_t^* = 1, U_{t-1}^* = 2) \\ p(U_t^* = 2, U_{t-1}^* = 2) \\ p(U_t^* = 3, U_{t-1}^* = 2) \\ p(U_t^* = 1, U_{t-1}^* = 3) \\ p(U_t^* = 2, U_{t-1}^* = 3) \\ p(U_t^* = 3, U_{t-1}^* = 3) \end{array} \right].$$

In order to relax assumption 1, we allow the misclassification probabilities $m_{i|j} = \Pr(U_t = i|U_t^* = j)$ to vary according to their confidence intervals. Let the estimated standard error of $m_{i|j}$ be $s_{i|j}$, then obtain the 95% confidence interval (CI) $\left[\underline{m}_{i|j}, \overline{m}_{i|j}\right]$. Define:

$$\underline{M_{U_t|U_t^*}} = \left[ \begin{array}{lll} 1 - \underline{m}_{2|1} - \underline{m}_{3|1} & \underline{m}_{1|2} & \underline{m}_{1|3} \\ \underline{m}_{2|1} & 1 - \underline{m}_{1|2} - \underline{m}_{3|2} & \underline{m}_{2|3} \\ \underline{m}_{3|1} & \underline{m}_{3|2} & 1 - \underline{m}_{1|3} - \underline{m}_{2|3} \end{array} \right],$$

$$\overline{M_{U_t|U_t^*}} = \left[ \begin{array}{lll} 1 - \overline{m}_{2|1} - \overline{m}_{3|1} & \overline{m}_{1|2} & \overline{m}_{1|3} \\ \overline{m}_{2|1} & 1 - \overline{m}_{1|2} - \overline{m}_{3|2} & \overline{m}_{2|3} \\ \overline{m}_{3|1} & \overline{m}_{3|2} & 1 - \overline{m}_{1|3} - \overline{m}_{2|3} \end{array} \right],$$

which are the deviated misclassification probability matrices generated by allowing the off-diagonal entries to deviate to the upper or lower bounds of their confidence intervals. Note that the off-diagonal elements are misclassification error probabilities while the diagonal elements are probabilities that the LFS is corrected reported.

In general, we can consider the following deviations:

$$M_{U_t|U_t^*,U_{t-1}^*} = \left[ \begin{array}{ccc} M_{U_t|U_t^*,U_{t-1}^*=1} & M_{U_t|U_t^*,U_{t-1}^*=2} & M_{U_t|U_t^*,U_{t-1}^*=3} \end{array} \right],$$

where

$$M_{U_t|U_t^*,U_{t-1}^*=k} = (1-\lambda_k)\,\underline{M_{U_t|U_t^*}} + \lambda_k \overline{M_{U_t|U_t^*}}$$

with different combinations of the constants $(\lambda_1, \lambda_2, \lambda_3)$. Note that there are large numbers of possible combinations. For example, when we consider three possible values, $0, 0.5$,and 1, for $\lambda_k$, which corresponds to the 95% lower bound, the baseline value (no deviation), and the 95% upper bound of the misclassification errors, there are $3^3 - 3 = 24$ cases in which the $\lambda_1, \lambda_2, \lambda_3$ are not the same. (if $\lambda_1 = \lambda_2 = \lambda_3$ then assumption 1 holds).

We choose the parameters to mimic the observed CPS sample as follows:

$$M_{U_t|U_t^*} = \begin{bmatrix} 0.98 & 0.2 & 0.01 \\ 0.01 & 0.6 & 0.01 \\ 0.01 & 0.2 & 0.98 \end{bmatrix}, M_{U_t^*|U_{t-1}^*} = \begin{bmatrix} 0.98 & 0.1 & 0.015 \\ 0.01 & 0.85 & 0.005 \\ 0.01 & 0.05 & 0.98 \end{bmatrix}$$

$$\underline{M_{U_t|U_t^*}} = \begin{bmatrix} 0.984 & 0.16 & 0.008 \\ 0.008 & 0.68 & 0.008 \\ 0.008 & 0.16 & 0.984 \end{bmatrix}, \overline{M_{U_t|U_t^*}} = \begin{bmatrix} 0.976 & 0.24 & 0.012 \\ 0.012 & 0.52 & 0.012 \\ 0.012 & 0.24 & 0.976 \end{bmatrix} \text{ and } \Pr(U_{t-10}^*) = \begin{bmatrix} 0.6256 \\ 0.0544 \\ 0.32 \end{bmatrix}.$$

**Case 2: reported LFS depends on last period reported LFS** In case 2, we relax Assumption 1 to

$$\begin{aligned} \Pr(U_t|U_t^*, U_{<t}, U_{<t}^*) &= \Pr(U_t|U_t^*, U_{t-1}) \\ &\neq \Pr(U_t|U_t^*), \end{aligned}$$

where the misclassification probabilities may be different for different self-reported LFS $U_{t-1}$ in the previous period. Under this relaxed version, we only need to change step 7 in the baseline setup #1 as follows:

Step 7, draw $U_s$ using $U_s^*, U_{s-1}$ and $\Pr(U_t|U_t^*, U_{t-1})$ respectively for $s = t+1, t, t-9$.

Similar to Case 1, the conditional probability $\Pr(U_t|U_t^*, U_{t-1})$ may be considered as the joint of three distinct misclassification probabilities, i.e., $\Pr(U_t|U_t^*, U_{t-1} = 1)$ , $\Pr(U_t|U_t^*, U_{t-1} = 2)$, and $\Pr(U_t|U_t^*, U_{t-1} = 3)$. In matrix notation, we have

$$M_{U_t|U_t^*,U_{t-1}} \equiv \begin{bmatrix} M_{U_t|U_t^*,U_{t-1}=1} & M_{U_t|U_t^*,U_{t-1}=2} & M_{U_t|U_t^*,U_{t-1}=3} \end{bmatrix},$$

with

$$M_{U_t|U_t^*,U_{t-1}=k} = (1-\lambda_k)\,\underline{M_{U_t|U_t^*}} + \lambda_k \overline{M_{U_t|U_t^*}}.$$

The rest of the simulation setup and the parameters chosen are the same as in Case 1.

**Case 3: reported LFS depends on both last period true LFS and last period reported LFS** In case 3, we consider the more general relaxation of Assumption 1 to

$$\begin{aligned}\Pr(U_t|U_t^*, U_{<t}, U_{<t}^*) &= \Pr(U_t|U_t^*, U_{t-1}^*, U_{t-1}) \\ &\neq \Pr(U_t|U_t^*),\end{aligned}$$

where the misclassification probabilities may be different for different self-reported LFS $U_{t-1}$ as well as true LFS $U_{t-1}^*$ in the previous period. Because possible deviations are too large to be tractable, following a suggestion by a referee, we consider a special case in which:

$$\Pr(U_t|U_t^*, U_{<t}, U_{<t}^*) = \begin{cases} p \times I\left(U_t = U_{t-1}\right) + (1-p)\Pr(U_t|U_t^*) & \text{if } U_t^* = U_{t-1}^* \\ \Pr(U_t|U_t^*) & \text{otherwise} \end{cases},$$

where $I(\cdot)$ is a 0-1 indicator function.

This case allows us to directly evaluate how correlated reporting behavior affects our results. The idea is that people who misreport in one period and have the same true LFS in the next period are likely to report the same way as in the previous period.

Under this relaxed version, we only need to add one step to the baseline setup #1 as follows:

Step 8, replace $U_t$ as $U_{t-1}$ with probability $p$ if $U_t^* = U_{t-1}^*$ .

All parameters are the same as in the baseline setup #1, except $p$, which is allowed to vary from 0 to 1 to assess the robustness of our estimator.

### 3.1.3 Setup #3: checking robustness of Assumption 2

In this section, we relax Assumption 2 to

$$\Pr(U_{t+1}^*|U_t^*, U_{t-9}^*) \neq \Pr(U_{t+1}^*|U_t^*).$$

Under this relaxed assumption, we only need to change steps 5 and 6 for simplicity in setup #1 as follows:

Step 5, draw $U_t^*$ using $U_{t-1}^*$, $U_{t-10}^*$ and $\Pr(U_{t+1}^*|U_t^*, U_{t-9}^*)$.

Step 6, draw $U_{t+1}^*$ using $U_t^*$, $U_{t-9}^*$ and $\Pr(U_{t+1}^*|U_t^*, U_{t-9}^*)$.

We start with the original Markovian transition kernel $M_{U_{t+1}^*|U_t^*} = M_{U_t^*|U_{t-1}^*} = \begin{bmatrix} m_{1|1} & m_{1|2} & m_{1|3} \\ m_{2|1} & m_{2|2} & m_{2|3} \\ m_{3|1} & m_{3|2} & m_{3|3} \end{bmatrix}$.
Under assumption 2, we have $M_{U_{t+1}^*|U_t^*,U_{t-9}^*} = \begin{bmatrix} M_{U_t^*|U_{t-1}^*} & M_{U_t^*|U_{t-1}^*} & M_{U_t^*|U_{t-1}^*} \end{bmatrix}$. We may consider $M_{U_{t+1}^*|U_t^*,U_{t-9}^*=k}$ for $k = 1, 2, 3$ as deviations from the baseline Markovian transition probability $M_{U_{t+1}^*|U_t^*}$. Similar to what was discussed in the previous subsection, there are many ways to deviate from $M_{U_{t+1}^*|U_t^*}$ or $\Pr(U_{t+1}^*|U_t^*)$ to $M_{U_{t+1}^*|U_t^*,U_{t-9}^*=k}$ or $\Pr(U_{t+1}^*|U_t^*, U_{t-9}^* = k)$. In our simulation, the matrices $M_{U_{t+1}^*|U_t^*,U_{t-9}^*=k}$ are generated by letting the entries in $M_{U_{t+1}^*|U_t^*}$ to deviate according to their confidence intervals in the baseline case.

In order to relax assumption 2, we allow the Markov transition probabilities $m_{i|j} = \Pr(U_{t+1}^* = i|U_t^* = j)$ to vary according to their confidence intervals. Let the estimated standard error of $m_{i|j}$ be $s_{i|j}$, then obtain the 95% confidence interval (CI) $\left[\underline{m}_{i|j}, \overline{m}_{i|j}\right]$. We define

$$\underline{M_{U_t^*|U_{t-1}^*}} = \begin{bmatrix} 1 - \underline{m}_{2|1} - \underline{m}_{3|1} & \underline{m}_{1|2} & \underline{m}_{1|3} \\ \underline{m}_{2|1} & 1 - \underline{m}_{1|2} - \underline{m}_{3|2} & \underline{m}_{2|3} \\ \underline{m}_{3|1} & \underline{m}_{3|2} & 1 - \underline{m}_{1|3} - \underline{m}_{2|3} \end{bmatrix}$$

$$\overline{M_{U_t^*|U_{t-1}^*}} = \begin{bmatrix} 1 - \overline{m}_{2|1} - \overline{m}_{3|1} & \overline{m}_{1|2} & \overline{m}_{1|3} \\ \overline{m}_{2|1} & 1 - \overline{m}_{1|2} - \overline{m}_{3|2} & \overline{m}_{2|3} \\ \overline{m}_{3|1} & \overline{m}_{3|2} & 1 - \overline{m}_{1|3} - \overline{m}_{2|3} \end{bmatrix},$$

which are the deviated Markov transition matrices generated by allowing the off-diagonal entries (error probabilities) to deviate to the upper or lower bounds of their confidence intervals.

In general, we can consider the following deviations:

$$M_{U_{t+1}^*|U_t^*,U_{t-9}^*} \equiv \begin{bmatrix} M_{U_{t+1}^*|U_t^*,U_{t-9}^*=1} & M_{U_{t+1}^*|U_t^*,U_{t-9}^*=2} & M_{U_{t+1}^*|U_t^*,U_{t-9}^*=3} \end{bmatrix},$$

with

$$M_{U_{t+1}^*|U_t^*,U_{t-9}^*=k} = (1 - \lambda_k)\, \underline{M_{U_t^*|U_{t-1}^*}} + \lambda_k \overline{M_{U_t^*|U_{t-1}^*}}.$$

with different combinations of the constants $(\lambda_1, \lambda_2, \lambda_3)$. Again, there are large numbers of possible combinations. For example, when we consider three possible values, $0, 0.5$,and 1, for $\lambda_k$, which corresponds to the 95% lower bound, the baseline value (no deviation), and the 95% upper bound of the misclassification errors, there are $3^3 - 3 = 24$ cases in which the $\lambda_1, \lambda_2, \lambda_3$ are not the same. (if $\lambda_1 = \lambda_2 = \lambda_3$ then assumption 2 holds).

We choose the following parameter values:

$$\underline{M_{U_t^*|U_{t-1}^*}} = \begin{bmatrix} 0.984 & 0.08 & 0.013 \\ 0.008 & 0.89 & 0.003 \\ 0.008 & 0.03 & 0.984 \end{bmatrix}, \overline{M_{U_t^*|U_{t-1}^*}} = \begin{bmatrix} 0.976 & 0.12 & 0.017 \\ 0.012 & 0.81 & 0.007 \\ 0.012 & 0.07 & 0.976 \end{bmatrix}.$$

Other parameter values are the same as in the baseline setup #1.

## 3.2 Simulation results

This section reports all simulation results. Based on the generated data, we produce the following estimates:

*Unemp_C*: corrected unemployment rate, which was estimated from observed data using the proposed estimator.

*Unemp_R*: reported unemployment rate, which are uncorrected and subject to misclassification error.

*Unemp_T*: true unemployment rate implied by the generated sample.

*LFP_C*: corrected labor force participation rate.

*LFP_R*: reported labor force participation rate.

*LFP_T*: true labor force participation rate.

In all cases we report both the mean value of the statistic as well as the 95% confidence lower and upper bounds from 500 repetitions. In all tables we set samples size to be 100,000 in order to match the width of confidence intervals of the estimates based on the observed CPS data. The only exception is Table 1 where we vary sample size to show consistency.

### 3.2.1 Setup #1: consistency of our estimator

Table A2 shows the results when we maintain all assumptions in the paper. One can see that as sample size increases, the estimated *Unemp_C* become closer to the true underlying unemployment rate $Unemp\_T$, which is 7.98%. Even with sample size at 10,000, the mean of our estimates from 500 repetitions are quite close to the true value and the 95% confidence intervals always cover the true value. When sample size is increased to 100,000, mean estimate of $Unemp\_C$ is 8.02%, which is very close to $Unemp\_T$ and the width of the confidence interval become relatively small (2.4%). In contrast, the reported $Unemp\_R$ has a mean value of 6.32% which severely underestimate the true level of unemployment. In addition, its 95% confidence intervals do not cover $Unemp\_T$ in all cases.

Results for Labor Force Participation (LFP) rate are quite similar. With sample size of 10,000, the corrected mean of $LFP\_C$ is 65.8%, which is exactly the true LFP. On the other hand, reported LFP has a relatively large and statistically significant bias.

We have also reported $\Pr(U_t = U_{t-1}|U_t^* = U_{t-1}^*)$, which measures whether people tend to report the same LFS if their true LFS does not change. Note that although we only condition $U_t$ on $U_t^*$, because the misclassification probabilities are the same for all time periods and because the rate of transition among different underlying true LFS is relatively slow, (mis)reporting behaviors across time are correlated. In the generated data, for those who has the same true LFS in periods t-1 and t, about 94% would report the same LFS in both periods. Thus our model setup is able to capture the observed correlations in (mis)reporting across time.

### 3.2.2 Setup #2 case 1: relaxing Assumption 1 to allow observed LFS to depend on last period true LFS

We then report simulation results when we relax assumption 1 to allow observed LFS to also depends on last period true LFS. As described in the previous section, the degree of deviation is controlled by $\lambda$. In Table A3.1, we consider the 24 combination where $\lambda$ can take values of 0, 0.5 and 1. Note that in this case we have:

$$
\begin{aligned}
&M_{U_t|U_t^*,U_{t-1}^*=k} = (1-\lambda_k)\,\underline{M_{U_t|U_t^*}} + \lambda_k\overline{M_{U_t|U_t^*}} \\
&= \underline{M_{U_t|U_t^*}} = \begin{bmatrix} 0.984 & 0.16 & 0.008 \\ 0.008 & 0.68 & 0.008 \\ 0.008 & 0.16 & 0.984 \end{bmatrix} if\ \lambda_k = 0 \\
&= M_{U_t|U_t^*} = \begin{bmatrix} 0.98 & 0.2 & 0.01 \\ 0.01 & 0.6 & 0.01 \\ 0.01 & 0.2 & 0.98 \end{bmatrix} if\ \lambda_k = 0.5 \\
&= \overline{M_{U_t|U_t^*}} = \begin{bmatrix} 0.976 & 0.24 & 0.012 \\ 0.012 & 0.52 & 0.012 \\ 0.012 & 0.24 & 0.976 \end{bmatrix} if\ \lambda_k = 1
\end{aligned}
$$

The corrected unemployment rate using our proposed method ($Unemp\_C$) are relatively close to the true unemployment rate ($Unemp\_T$) and the 95% confidence interval for $Unemp\_C$ always cover the true value of unemployment, which is 7.98%. Also, the width of the 95% confidence interval is relatively small and close to the case with no deviations (see column 3 of Table A2). On the other hand, the reported unemployment rate ($Unemp\_R$) consistently underestimate the true level of unemployment rate, and its 95% upper intervals are always lower than $Unemp\_T$. Thus our proposed estimator consistently outperforms the reported (uncorrected) even when assumption 1 is violated to some extent.

Table A3.2 consider the case with slightly more deviations, and presents results for the 24 possible combinations of $\lambda$ taking values of -0.5, 0.5 and 1.5. In this case, we have:

$$
\begin{aligned}
&M_{U_t|U_t^*,U_{t-1}^*=k} = (1-\lambda_k)\,\underline{M_{U_t|U_t^*}} + \lambda_k \overline{M_{U_t|U_t^*}} \\
&= \begin{bmatrix} 0.988 & 0.12 & 0.006 \\ 0.006 & 0.76 & 0.006 \\ 0.006 & 0.12 & 0.988 \end{bmatrix} if\ \lambda_k = -0.5 \\
&= M_{U_t|U_t^*} = \begin{bmatrix} 0.98 & 0.2 & 0.01 \\ 0.01 & 0.6 & 0.01 \\ 0.01 & 0.2 & 0.98 \end{bmatrix} if\ \lambda_k = 0.5 \\
&= \begin{bmatrix} 0.972 & 0.28 & 0.014 \\ 0.014 & 0.44 & 0.014 \\ 0.014 & 0.28 & 0.972 \end{bmatrix} if\ \lambda_k = 1.5
\end{aligned}
$$

As expected, results shown in Table A3.2 are somewhat worse than those in Table A3.1. For example, when $\{\lambda_1, \lambda_2, \lambda_3\} = \{0.5, -0.5, 0.5\}$, the mean of $Unemp\_C$ is 9.19%, implying a upward bias of 1.21%, and also with the width of the confidence interval being 3.4%. Nevertheless, to some extent results in Table A3.2 are still acceptable because in all cases the 95% confidence intervals contain the ture value of unemployment, which is 7.98%. In contrast, none of the confidence intervals for $Unemp\_R$ cover the true value of unemployment rate.

It is tempting to test whether our estimator could "fail" if there are "too much" deviations in the misclassification probabilities. In Table A3.3 we allow $\lambda$ to take values between -3 and 4. When $\lambda$ are too far away out of the [0,1] range, the implied misclassification matrix may contain elements smaller than 0 or larger than 1. To deal with this case, we apply a normalization procedure, which first transfer any elements greater than 1 to 1, and transfer any elements smaller than 0 to 0, and then divide each elements by its column sum to make sure each column sum to 1. After the normalization, we have:

$M_{U_t|U_t^*,U_{t-1}^*=k} =$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} if\ \lambda_k = -3$$

$$= \begin{bmatrix} 0.952 & 0.48 & 0.024 \\ 0.024 & 0.04 & 0.024 \\ 0.024 & 0.48 & 0.952 \end{bmatrix} if\ \lambda_k = 4$$

Table A3.3 reports the results. In general the biases are quite large and the confidence intervals are too wide to be informative. For example, in the last row we have the mean value of corrected unemployment rate of 28.5% while the true value is 7.98%. The 95% confidence lower bound is 0.65% and the upper bound is 81%. Thus it is possible for the estimator to "fail" if the deviations from assumption 1 are too large.

Corresponding results for labor force participation rates (LFP) are presented in Tables A4.1-4.3. Table A4.1 and A4.2 report results when $\lambda$s are relatively small. The corrected LFP has mean values close to the true LFP, which is 65.8%, and the 95% confidence interval is relatively tight. Table A4.3 report results when $\lambda$s are relatively large. Even in this case the biases are not big compared to results reported in Table A3.3 for unemployment rates. But again in the last row, the mean of corrected LFP is 73.7%, representing an upward bias of around 8%.

### 3.2.3 Setup #2 case 2: relaxing Assumption 1 to allow observed LFS to depend on last period observed LFS

We then consider a different type of deviation to assumption 1, and allow observed LFS $U_t$ to depend on both $U_t^*$ and last period observed LFS $U_{t-1}$. The setup and parameters chosen are otherwise similar to the previous case.

In Table A5.1, we consider the 24 combination where $\lambda$ can take values of 0, 0.5 and 1. Similar to Table A3.1, the corrected unemployment rate using our proposed method ($Unemp\_C$) are relatively close to the true unemployment rate ($Unemp\_T$) and the 95% confidence interval for $Unemp\_C$ always cover the true value of unemployment, which is 7.98%. Also, the width of the 95% confidence interval is relatively small and close to the case with no deviations (see column 3 of Table A2). On the other hand, the reported unemployment rate ($Unemp\_R$) consistently underestimate the true level of unemployment rate, and its 95% upper intervals are always lower than $Unemp\_T$.

Table A5.2 consider the case with slightly more deviations, and presents results for the 24 possible combinations of $\lambda$ taking values of -0.5, 0.5 and 1.5. As expected, results shown in Table A5.2 are somewhat worse than those in Table A5.1 but still acceptable.

Table A5.3 presents results when $\lambda$s are allowed to vary between -3 and 4. Note here we have rather large biases and wide confidence bounds. For example, in the last row of table A5.3, mean value of $Unemp\_C$ is as large as 71.9%.

Table A6.1-6.3 display results for labor force participation rates (LFP). Once again, the corrected LFP using our proposed method ($LFP\_C$) are very close to the true LFP ($LFP\_T$) and the 95% confidence interval for $LFP\_C$ always cover the true value of LFP, which is 65.8%. This is even the case when $\lambda$s are relatively large as in Table A6.3. On the other hand, the reported LFP ($LFP\_R$) consistently underestimate the true level of LFP and its 95% confidence intervals do not cover the true values of LFP.

#### 3.2.4 Setup #2 case 3: reported LFS depends on both last period true LFS and last period reported LFS

We then report results for setup #2 case 3, allowing $U_t$ to depends on $U_t^*$ as well as $U_{t-1}^*$ and $U_{t-1}$. But we only focus on a very special case in which people consistently (mis)report. When $U_t^* = U_{t-1}^*$, we replace the value of $U_t$ to be equal to $U_{t-1}$ with probability $p$, which ranges from 0 to 1 at the interval of 0.1.

Table A7 reports results for unemployment rates and Table A8 reports results for LFP rates. The results are very similar to the no deviation case and our corrected estimator works very well. The results are kind of expected because in the baseline case $\Pr(U_t = U_{t-1}|U_t^* = U_{t-1}^*)$ is already around 94%. But the results here more formally show that our estimator is robust to consistent (mis)reporting behavior.

#### 3.2.5 Setup #3: relaxing Assumption 2

Lastly, we report simulation results when we relax assumption 2 to allow $U_{t+1}^*$ depends on both $U_t^*$ and $U_{t-9}^*$. As described in the previous section, the degree of deviation is controlled by $\lambda$. In Table A9.1, we consider the 24 combination where $\lambda$ can take values of 0, 0.5 and 1. Note that in this case we have:

$$M_{U^*_{t+1}|U^*_t,U^*_{t-9}=k} = (1-\lambda_k)\,\underline{M_{U^*_{t+1}|U^*_t}} + \lambda_k \overline{M_{U^*_{t+1}|U^*_t}}$$
$$= \underline{M_{U^*_{t+1}|U^*_t}} = \begin{bmatrix} 0.984 & 0.08 & 0.013 \\ 0.008 & 0.89 & 0.003 \\ 0.008 & 0.03 & 0.984 \end{bmatrix} \; if\ \lambda_k = 0$$
$$= M_{U^*_{t+1}|U^*_t} = \begin{bmatrix} 0.98 & 0.1 & 0.015 \\ 0.01 & 0.85 & 0.005 \\ 0.01 & 0.05 & 0.98 \end{bmatrix} \; if\ \lambda_k = 0.5$$
$$= \overline{M_{U^*_{t+1}|U^*_t}} = \begin{bmatrix} 0.976 & 0.12 & 0.017 \\ 0.012 & 0.81 & 0.007 \\ 0.012 & 0.07 & 0.976 \end{bmatrix} \; if\ \lambda_k = 1$$

The corrected unemployment rate using our proposed method (*Unemp_C*) are relatively close to the true unemployment rate (*Unemp_T*) and the 95% confidence interval for *Unemp_C* always cover the true value of unemployment. Also, the width of the 95% confidence interval is relatively small and close to the case with no deviations (see column 3 of Table A2). On the other hand, the reported unemployment rate (*Unemp_R*) consistently underestimate the true level of unemployment rate, and its 95% upper intervals are always lower than *Unemp_T*. Thus our proposed estimator consistently outperforms the reported (uncorrected) even when assumption 2 is violated to some extent.

Table A9.2 considers the case with slightly more deviations, and presents results for the 24 possible combinations of $\lambda$ taking values of -0.5, 0.5 and 1.5. In this case, we have:

$$M_{U^*_{t+1}|U^*_t,U^*_{t-9}=k} = (1-\lambda_k)\,\underline{M_{U^*_{t+1}|U^*_t}} + \lambda_k \overline{M_{U^*_{t+1}|U^*_t}}$$
$$= \begin{bmatrix} 0.988 & 0.06 & 0.011 \\ 0.006 & 0.93 & 0.001 \\ 0.006 & 0.01 & 0.988 \end{bmatrix} \; if\ \lambda_k = -0.5$$
$$= M_{U^*_{t+1}|U^*_t} = \begin{bmatrix} 0.98 & 0.1 & 0.015 \\ 0.01 & 0.85 & 0.005 \\ 0.01 & 0.05 & 0.98 \end{bmatrix} \; if\ \lambda_k = 0.5$$
$$= \begin{bmatrix} 0.972 & 0.14 & 0.019 \\ 0.014 & 0.77 & 0.009 \\ 0.014 & 0.09 & 0.972 \end{bmatrix} \; if\ \lambda_k = 1.5$$

As expected, results shown in Table A9.2 are somewhat worse than those in Table A9.1. For example, when $\{\lambda_1, \lambda_2, \lambda_3\} = \{1.5, -0.5, -0.5\}$, the mean of *Unemp_C* is 8.82%, implying a upward bias of about 1%. Nevertheless, to some extent results in Table A9.2 are still acceptable because in all cases the 95% confidence intervals contain the true value of unemployment, which varies by the combinations of $\lambda$s.

In contrast, none of the confidence intervals for $Unemp\_R$ cover the true value of unemployment rate.

We also test whether our estimator could "fail" if there are "too much" deviations in the Markovian assumption. In Table A9.3 we allow $\lambda$ to take values between -3 and 4. When $\lambda$ are too far away out of the [0,1] range, the implied transition matrix may contain elements smaller than 0 or larger than 1. To deal with this case, we apply a normalization procedure, which first transfer any elements greater than 1 to 1, and transfer any elements smaller than 0 to 0, and then divide each elements by its column sum to make sure each column sum to 1. After the normalization, we have:

$$
\begin{aligned}
&M_{U^*_{t+1}|U^*_t,U^*_{t-9}=k} = \\
&= \begin{bmatrix} 1 & 0 & 0.001 \\ 0 & 1 & 0 \\ 0 & 0 & 0.999 \end{bmatrix} \; if\ \lambda_k = -3 \\
&= \begin{bmatrix} 0.952 & 0.24 & 0.029 \\ 0.024 & 0.57 & 0.019 \\ 0.024 & 0.19 & 0.952 \end{bmatrix} \; if\ \lambda_k = 4
\end{aligned}
$$

Table A9.3 reports the results. In some cases the biases are large and the 95% confidence intervals do not cover the true unemployment rate. For example, in the first row of Table A9.3 we see a statistically significant downward bias, while the last row shows a statistically significant upward bias.

Corresponding results for labor force participation rates (LFP) are presented in Tables A10.1-10.3. Table A10.1 and A10.2 report results when $\lambda$s are relatively small. The corrected LFP has mean values close to the true LFP, and the 95% confidence interval is relatively tight. Table A10.3 report results when $\lambda$s are relatively large. Results are better than reported in Table A9.3 because the 95% confidence intervals always cover true LFP. Thus violation of assumption 2 is a more severe issue for unemployment rates than for LFP.

Table A2: Simulation results under maintained assumptions

| Sample size | $10,000$ | $100,000$ | $1,000,000$ |
|---|---|---|---|
| $Unemp\_C$ (%) | 8.60 $(5.27, 14.38)$ | 8.02 $(6.86, 9.28)$ | 7.98 $(7.60, 8.36)$ |
| $Unemp\_R$ (%) | 6.32 $(5.68, 6.97)$ | 6.32 $(6.13, 6.50)$ | 6.32 $(6.27, 6.38)$ |
| $Unemp\_T$ (%) | 7.98 $(7.30, 8.62)$ | 7.98 $(7.76, 8.18)$ | 7.98 $(7.91, 8.04)$ |
| $LFP\_C$ (%) | 66.1 $(64.1, 70.1)$ | 65.8 $(65.1, 66.8)$ | 65.8 $(65.5, 66.0)$ |
| $LFP\_R$ (%) | 64.8 $(63.9, 65.8)$ | 64.8 $(64.5, 65.1)$ | 64.8 $(64.7, 64.9)$ |
| $LFP\_T$ (%) | 65.8 $(64.9, 66.8)$ | 65.8 $(65.5, 66.1)$ | 65.8 $(65.7, 65.9)$ |
| $\Pr(U_t = U_{t-1} \| U_t^* = U_{t-1}^*)$ | 0.937 (0.932, 0.941) | 0.937 (0.935, 0.938) | 0.937 (0.936, 0.937) |

Note: number of repetitions is 500 for each column. For each statistic listed we report mean as well as 95% lower and confidence bounds (in parentheses) based on the generated data.

Table A3.1: Simulation results for setup #2 case 1 (results for unemployment rates with $\lambda \in \{0, 0.5, 1\}$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $Unemp\_C(\%)$ | $Unemp\_R(\%)$ | $Unemp\_T(\%)$ |
|---|---|---|---|---|---|
| 1.0 | 1.0 | 0.5 | 7.94 (6.98, 9.03) | 6.72 (6.52, 6.90) | 7.98 |
| 1.0 | 1.0 | 0.0 | 7.98 (6.99, 9.14) | 6.79 (6.59, 7.00) | 7.98 |
| 1.0 | 0.5 | 1.0 | 8.26 (7.12, 9.56) | 6.12 (5.95, 6.32) | 7.98 |
| 1.0 | 0.5 | 0.5 | 8.13 (7.07, 9.24) | 6.19 (6.01, 6.38) | 7.98 |
| 1.0 | 0.5 | 0.0 | 8.13 (6.97, 9.47) | 6.27 (6.08, 6.47) | 7.98 |
| 1.0 | 0.0 | 1.0 | 8.38 (7.18, 9.72) | 5.59 (5.41, 5.76) | 7.98 |
| 1.0 | 0.0 | 0.5 | 8.37 (7.08, 9.67) | 5.66 (5.49, 5.86) | 7.98 |
| 1.0 | 0.0 | 0.0 | 8.35 (7.27, 9.61) | 5.74 (5.56, 5.91) | 7.98 |
| 0.5 | 1.0 | 1.0 | 7.92 (6.88, 8.98) | 6.78 (6.59, 6.98) | 7.98 |
| 0.5 | 1.0 | 0.5 | 7.86 (6.86, 8.91) | 6.85 (6.66, 7.04) | 7.98 |
| 0.5 | 1.0 | 0.0 | 7.82 (6.73, 8.97) | 6.92 (6.73, 7.11) | 7.98 |
| 0.5 | 0.5 | 1.0 | 8.04 (6.91, 9.16) | 6.25 (6.08, 6.44) | 7.98 |
| 0.5 | 0.5 | 0.0 | 7.98 (7.02, 9.25) | 6.39 (6.22, 6.57) | 7.98 |
| 0.5 | 0.0 | 1.0 | 8.31 (7.09, 9.53) | 5.71 (5.53, 5.89) | 7.98 |
| 0.5 | 0.0 | 0.5 | 8.21 (6.88, 9.61) | 5.78 (5.58, 5.96) | 7.98 |
| 0.5 | 0.0 | 0.0 | 8.20 (7.01, 9.57) | 5.87 (5.68, 6.05) | 7.98 |
| 0.0 | 1.0 | 1.0 | 7.75 (6.76, 8.89) | 6.90 (6.70, 7.12) | 7.98 |
| 0.0 | 1.0 | 0.5 | 7.75 (6.67, 8.98) | 6.98 (6.79, 7.18) | 7.98 |
| 0.0 | 1.0 | 0.0 | 7.70 (6.57, 8.87) | 7.04 (6.84, 7.23) | 7.98 |
| 0.0 | 0.5 | 1.0 | 7.90 (6.86, 9.03) | 6.37 (6.19, 6.59) | 7.98 |
| 0.0 | 0.5 | 0.5 | 7.83 (6.83, 9.06) | 6.44 (6.25, 6.62) | 7.98 |
| 0.0 | 0.5 | 0.0 | 7.91 (6.83, 9.20) | 6.52 (6.33, 6.71) | 7.98 |
| 0.0 | 0.0 | 1.0 | 8.18 (6.95, 9.71) | 5.84 (5.67, 6.02) | 7.98 |
| 0.0 | 0.0 | 0.5 | 8.09 (6.88, 9.44) | 5.91 (5.75, 6.09) | 7.98 |

Note: Results showing unemployment rates when assumption 1 is relaxed such that $U_t$ depends on both $U_t^*$ and $U_{t-1}^*$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A3.2: Simulation results for setup #2 case 1 (results for unemployment rates with $\lambda \in \{-0.5, 0.5, 1.5\}$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $Unemp\_C(\%)$ | $Unemp\_R(\%)$ | $Unemp\_T(\%)$ |
|---|---|---|---|---|---|
| 1.5 | 1.5 | 0.5 | 7.96 (7.04, 9.04) | 7.13 (6.93, 7.33) | 7.98 |
| 1.5 | 1.5 | −0.5 | 7.93 (6.97, 9.03) | 7.26 (7.06, 7.46) | 7.98 |
| 1.5 | 0.5 | 1.5 | 8.33 (7.36, 9.50) | 5.93 (5.76, 6.11) | 7.98 |
| 1.5 | 0.5 | 0.5 | 8.31 (7.32, 9.39) | 6.08 (5.91, 6.26) | 7.98 |
| 1.5 | 0.5 | −0.5 | 8.22 (7.22, 9.51) | 6.21 (6.03, 6.39) | 7.98 |
| 1.5 | −0.5 | 1.5 | 9.14 (7.60, 10.41) | 4.85 (4.67, 5.02) | 7.98 |
| 1.5 | −0.5 | 0.5 | 8.86 (7.57, 10.48) | 5.00 (4.84, 5.17) | 7.98 |
| 1.5 | −0.5 | −0.5 | 8.99 (7.43, 10.23) | 5.15 (4.99, 5.31) | 7.98 |
| 0.5 | 1.5 | 1.5 | 7.79 (6.76, 8.88) | 7.24 (7.04, 7.47) | 7.98 |
| 0.5 | 1.5 | 0.5 | 7.76 (6.81, 8.91) | 7.37 (7.19, 7.57) | 7.98 |
| 0.5 | 1.5 | −0.5 | 7.71 (6.59, 8.96) | 7.51 (7.30, 7.72) | 7.98 |
| 0.5 | 0.5 | 1.5 | 8.12 (7.16, 9.18) | 6.18 (6.00, 6.34) | 7.98 |
| 0.5 | 0.5 | −0.5 | 7.96 (6.91, 9.19) | 6.46 (6.26, 6.65) | 7.98 |
| 0.5 | −0.5 | 1.5 | 8.97 (7.27, 10.28) | 5.10 (4.93, 5.27) | 7.98 |
| 0.5 | −0.5 | 0.5 | 9.19 (7.06, 10.46) | 5.25 (5.09, 5.43) | 7.98 |
| 0.5 | −0.5 | −0.5 | 8.81 (7.04, 10.18) | 5.40 (5.23, 5.57) | 7.98 |
| −0.5 | 1.5 | 1.5 | 7.51 (6.54, 8.53) | 7.48 (7.27, 7.70) | 7.98 |
| −0.5 | 1.5 | 0.5 | 7.49 (6.49, 8.66) | 7.64 (7.44, 7.85) | 7.98 |
| −0.5 | 1.5 | −0.5 | 7.46 (6.39, 8.80) | 7.76 (7.55, 7.97) | 7.98 |
| −0.5 | 0.5 | 1.5 | 7.80 (6.74, 8.92) | 6.43 (6.25, 6.61) | 7.98 |
| −0.5 | 0.5 | 0.5 | 7.77 (6.65, 9.00) | 6.57 (6.39, 6.76) | 7.98 |
| −0.5 | 0.5 | −0.5 | 7.73 (6.46, 9.10) | 6.71 (6.52, 6.91) | 7.98 |
| −0.5 | −0.5 | 1.5 | 8.59 (6.73, 9.82) | 5.36 (5.18, 5.54) | 7.98 |
| −0.5 | −0.5 | 0.5 | 8.32 (6.67, 9.78) | 5.50 (5.32, 5.67) | 7.98 |

Note: Results showing unemployment rates when assumption 1 is relaxed such that $U_t$ depends on both $U_t^*$ and $U_{t-1}^*$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A3.3: Simulation results for setup #2 case 1 (results for unemployment rates with $\lambda \in [-3, 4]$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $Unemp\_C(\%)$ | $Unemp\_R(\%)$ | $Unemp\_T(\%)$ |
|---|---|---|---|---|---|
| $-1.5$ | 4.0 | 1.0 | 7.41 (6.64, 8.86) | 9.39 (9.17, 9.62) | 7.98 |
| 4.0 | 0.5 | $-3.0$ | 8.42 (7.23, 9.83) | 6.20 (6.00, 6.41) | 7.98 |
| $-0.5$ | 0.5 | 2.0 | 7.81 (6.73, 8.93) | 6.35 (6.18, 6.54) | 7.98 |
| $-0.5$ | $-2.0$ | $-2.5$ | 47.30 (6.12, 93.44) | 4.33 (4.15, 4.48) | 7.98 |
| $-0.5$ | $-2.5$ | 0.0 | 46.16 (6.91, 92.39) | 3.39 (3.25, 3.53) | 7.98 |
| 3.0 | 3.5 | $-3.0$ | 8.38 (7.35, 10.22) | 8.76 (8.55, 8.95) | 7.98 |
| 2.5 | 2.0 | 1.0 | 8.08 (7.21, 9.03) | 7.33 (7.13, 7.55) | 7.98 |
| 3.0 | $-3.0$ | $-2.5$ | 28.49 (0.65, 81.23) | 2.41 (2.30, 2.52) | 7.98 |

Note: Results showing unemployment rates when assumption 1 is relaxed such that $U_t$ depends on both $U_t^*$ and $U_{t-1}^*$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A4.1: Simulation results for setup #2 case 1 (results for LFP rates with $\lambda \in \{0, 0.5, 1\}$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $LFP\_C(\%)$ | $LFP\_R(\%)$ | $LFP\_T(\%)$ |
|---|---|---|---|---|---|
| 1.0 | 1.0 | 0.5 | 65.78 (65.03, 66.56) | 65.15 (64.83, 65.45) | 65.78 |
| 1.0 | 1.0 | 0.0 | 65.78 (65.13, 66.64) | 65.25 (64.94, 65.54) | 65.78 |
| 1.0 | 0.5 | 1.0 | 65.93 (65.24, 66.77) | 64.84 (64.57, 65.15) | 65.78 |
| 1.0 | 0.5 | 0.5 | 65.83 (65.20, 66.65) | 64.96 (64.69, 65.26) | 65.78 |
| 1.0 | 0.5 | 0.0 | 65.81 (65.18, 66.68) | 65.07 (64.76, 65.38) | 65.78 |
| 1.0 | 0.0 | 1.0 | 65.94 (65.15, 66.81) | 64.65 (64.33, 64.93) | 65.78 |
| 1.0 | 0.0 | 0.5 | 65.91 (65.17, 66.93) | 64.77 (64.50, 65.06) | 65.78 |
| 1.0 | 0.0 | 0.0 | 65.90 (65.23, 66.80) | 64.89 (64.64, 65.15) | 65.78 |
| 0.5 | 1.0 | 1.0 | 65.80 (65.14, 66.62) | 64.86 (64.57, 65.21) | 65.78 |
| 0.5 | 1.0 | 0.5 | 65.77 (65.10, 66.52) | 65.00 (64.71, 65.28) | 65.78 |
| 0.5 | 1.0 | 0.0 | 65.72 (65.03, 66.58) | 65.11 (64.83, 65.39) | 65.78 |
| 0.5 | 0.5 | 1.0 | 65.81 (65.09, 66.60) | 64.68 (64.39, 64.97) | 65.78 |
| 0.5 | 0.5 | 0.0 | 65.79 (65.11, 66.72) | 64.94 (64.62, 65.22) | 65.78 |
| 0.5 | 0.0 | 1.0 | 65.91 (65.15, 66.81) | 64.51 (64.22, 64.78) | 65.78 |
| 0.5 | 0.0 | 0.5 | 65.89 (65.16, 66.94) | 64.64 (64.35, 64.96) | 65.78 |
| 0.5 | 0.0 | 0.0 | 65.84 (65.04, 66.80) | 64.75 (64.46, 65.05) | 65.78 |
| 0.0 | 1.0 | 1.0 | 65.73 (65.04, 66.55) | 64.73 (64.44, 65.04) | 65.78 |
| 0.0 | 1.0 | 0.5 | 65.71 (65.00, 66.54) | 64.85 (64.51, 65.15) | 65.78 |
| 0.0 | 1.0 | 0.0 | 65.69 (65.03, 66.59) | 64.97 (64.65, 65.30) | 65.78 |
| 0.0 | 0.5 | 1.0 | 65.78 (65.12, 66.52) | 64.54 (64.22, 64.87) | 65.78 |
| 0.0 | 0.5 | 0.5 | 65.74 (65.03, 66.51) | 64.66 (64.35, 64.97) | 65.78 |
| 0.0 | 0.5 | 0.0 | 65.78 (65.10, 66.77) | 64.80 (64.50, 65.09) | 65.78 |
| 0.0 | 0.0 | 1.0 | 65.91 (65.17, 67.02) | 64.37 (64.08, 64.69) | 65.78 |
| 0.0 | 0.0 | 0.5 | 65.84 (65.06, 66.85) | 64.49 (64.18, 64.81) | 65.78 |

Note: Results showing LFP rates when assumption 1 is relaxed such that $U_t$ depends on both $U_t^*$ and $U_{t-1}^*$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A4.2: Simulation results for setup #2 case 1 (results for LFP rates with $\lambda \in \{-0.5, 0.5, 1.5\}$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $LFP\_C(\%)$ | $LFP\_R(\%)$ | $LFP\_T(\%)$ |
|---|---|---|---|---|---|
| 1.5 | 1.5 | 0.5 | 65.77 (65.17, 66.56) | 65.46 (65.16, 65.74) | 65.78 |
| 1.5 | 1.5 | −0.5 | 65.75 (65.14, 66.57) | 65.70 (65.43, 65.96) | 65.78 |
| 1.5 | 0.5 | 1.5 | 65.92 (65.29, 66.66) | 64.85 (64.54, 65.17) | 65.78 |
| 1.5 | 0.5 | 0.5 | 65.88 (65.22, 66.71) | 65.09 (64.80, 65.39) | 65.78 |
| 1.5 | 0.5 | −0.5 | 65.84 (65.17, 66.73) | 65.34 (65.04, 65.63) | 65.78 |
| 1.5 | −0.5 | 1.5 | 66.12 (65.36, 67.04) | 64.49 (64.20, 64.78) | 65.78 |
| 1.5 | −0.5 | 0.5 | 66.04 (65.29, 67.12) | 64.74 (64.46, 65.05) | 65.78 |
| 1.5 | −0.5 | −0.5 | 66.00 (65.15, 67.17) | 64.98 (64.69, 65.26) | 65.78 |
| 0.5 | 1.5 | 1.5 | 65.73 (65.12, 66.42) | 64.91 (64.61, 65.21) | 65.78 |
| 0.5 | 1.5 | 0.5 | 65.71 (65.03, 66.50) | 65.16 (64.86, 65.47) | 65.78 |
| 0.5 | 1.5 | −0.5 | 65.68 (65.01, 66.53) | 65.42 (65.13, 65.71) | 65.78 |
| 0.5 | 0.5 | 1.5 | 65.86 (65.23, 66.67) | 64.56 (64.27, 64.84) | 65.78 |
| 0.5 | 0.5 | −0.5 | 65.77 (65.05, 66.70) | 65.06 (64.76, 65.36) | 65.78 |
| 0.5 | −0.5 | 1.5 | 66.05 (65.30, 67.16) | 64.22 (63.94, 64.53) | 65.78 |
| 0.5 | −0.5 | 0.5 | 65.98 (65.17, 67.10) | 64.45 (64.13, 64.75) | 65.78 |
| 0.5 | −0.5 | −0.5 | 65.94 (65.15, 67.13) | 64.71 (64.40, 64.98) | 65.78 |
| −0.5 | 1.5 | 1.5 | 65.67 (64.98, 66.44) | 64.64 (64.34, 64.94) | 65.78 |
| −0.5 | 1.5 | 0.5 | 65.62 (64.94, 66.38) | 64.89 (64.60, 65.22) | 65.78 |
| −0.5 | 1.5 | −0.5 | 65.61 (64.87, 66.47) | 65.15 (64.84, 65.45) | 65.78 |
| −0.5 | 0.5 | 1.5 | 65.77 (65.11, 66.67) | 64.28 (63.99, 64.59) | 65.78 |
| −0.5 | 0.5 | 0.5 | 65.73 (65.09, 66.52) | 64.52 (64.20, 64.83) | 65.78 |
| −0.5 | 0.5 | −0.5 | 65.73 (64.96, 66.72) | 64.79 (64.51, 65.10) | 65.78 |
| −0.5 | −0.5 | 1.5 | 65.93 (65.01, 66.95) | 63.92 (63.64, 64.22) | 65.78 |
| −0.5 | −0.5 | 0.5 | 65.89 (65.14, 66.92) | 64.18 (63.89, 64.47) | 65.78 |

Note: Results showing LFP rates when assumption 1 is relaxed such that $U_t$ depends on both $U_t^*$ and $U_{t-1}^*$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A4.3: Simulation results for setup #2 case 1 (results for LFP rates with $\lambda \in [-3, 4]$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $LFP\_C(\%)$ | $LFP\_R(\%)$ | $LFP\_T(\%)$ |
|---|---|---|---|---|---|
| $-1.5$ | 4.0 | 1.0 | 65.68 (65.03, 66.70) | 65.03 (64.69, 65.31) | 65.78 |
| 4.0 | 0.5 | $-3.0$ | 65.91 (65.09, 66.88) | 66.40 (66.10, 66.68) | 65.78 |
| $-0.5$ | 0.5 | 2.0 | 65.80 (65.13, 66.54) | 64.16 (63.87, 64.46) | 65.78 |
| $-0.5$ | $-2.0$ | $-2.5$ | 66.58 (64.58, 72.15) | 64.39 (64.09, 64.72) | 65.78 |
| $-0.5$ | $-2.5$ | 0.0 | 69.33 (64.71, 90.91) | 63.58 (63.28, 63.89) | 65.78 |
| 3.0 | 3.5 | $-3.0$ | 66.05 (65.26, 67.35) | 67.29 (66.99, 67.58) | 65.78 |
| 2.5 | 2.0 | 1.0 | 65.81 (65.17, 66.57) | 65.79 (65.47, 66.05) | 65.78 |
| 3.0 | $-3.0$ | $-2.5$ | 73.67 (65.39, 96.76) | 65.02 (64.73, 65.33) | 65.78 |

Note: Results showing LFP rates when assumption 1 is relaxed such that $U_t$ depends on both $U_t^*$ and $U_{t-1}^*$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A5.1: Simulation results for setup #2 case 2 (results for unemployment rates with $\lambda \in \{0, 0.5, 1\}$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $Unemp\_C(\%)$ | $Unemp\_R(\%)$ | $Unemp\_T(\%)$ |
|---|---|---|---|---|---|
| 1.0 | 1.0 | 0.5 | 8.25 (7.12, 9.51) | 6.64 (6.46, 6.82) | 7.98 |
| 1.0 | 1.0 | 0.0 | 8.43 (7.21, 9.70) | 6.61 (6.43, 6.80) | 7.98 |
| 1.0 | 0.5 | 1.0 | 8.00 (7.02, 9.04) | 6.32 (6.12, 6.50) | 7.98 |
| 1.0 | 0.5 | 0.5 | 8.22 (7.14, 9.33) | 6.30 (6.10, 6.50) | 7.98 |
| 1.0 | 0.5 | 0.0 | 8.45 (7.38, 9.77) | 6.27 (6.09, 6.45) | 7.98 |
| 1.0 | 0.0 | 1.0 | 8.01 (7.14, 9.13) | 6.03 (5.85, 6.22) | 7.98 |
| 1.0 | 0.0 | 0.5 | 8.24 (7.18, 9.42) | 6.00 (5.83, 6.17) | 7.98 |
| 1.0 | 0.0 | 0.0 | 8.44 (7.25, 9.74) | 5.97 (5.80, 6.17) | 7.98 |
| 0.5 | 1.0 | 1.0 | 7.81 (6.84, 8.95) | 6.68 (6.49, 6.89) | 7.98 |
| 0.5 | 1.0 | 0.5 | 8.05 (7.02, 9.26) | 6.66 (6.45, 6.83) | 7.98 |
| 0.5 | 1.0 | 0.0 | 8.27 (7.11, 9.58) | 6.63 (6.44, 6.84) | 7.98 |
| 0.5 | 0.5 | 1.0 | 7.80 (6.88, 8.98) | 6.35 (6.16, 6.53) | 7.98 |
| 0.5 | 0.5 | 0.0 | 8.24 (6.98, 9.44) | 6.28 (6.10, 6.47) | 7.98 |
| 0.5 | 0.0 | 1.0 | 7.79 (6.73, 8.94) | 6.06 (5.87, 6.26) | 7.98 |
| 0.5 | 0.0 | 0.5 | 8.06 (6.91, 9.31) | 6.02 (5.83, 6.21) | 7.98 |
| 0.5 | 0.0 | 0.0 | 8.22 (6.98, 9.73) | 5.98 (5.80, 6.17) | 7.98 |
| 0.0 | 1.0 | 1.0 | 7.67 (6.61, 9.00) | 6.72 (6.54, 6.90) | 7.98 |
| 0.0 | 1.0 | 0.5 | 7.88 (6.72, 9.34) | 6.67 (6.48, 6.84) | 7.98 |
| 0.0 | 1.0 | 0.0 | 8.10 (6.73, 9.59) | 6.62 (6.43, 6.82) | 7.98 |
| 0.0 | 0.5 | 1.0 | 7.61 (6.57, 8.75) | 6.37 (6.18, 6.57) | 7.98 |
| 0.0 | 0.5 | 0.5 | 7.85 (6.70, 9.12) | 6.32 (6.14, 6.52) | 7.98 |
| 0.0 | 0.5 | 0.0 | 8.10 (6.74, 9.70) | 6.28 (6.10, 6.48) | 7.98 |
| 0.0 | 0.0 | 1.0 | 7.60 (6.47, 8.86) | 6.07 (5.90, 6.26) | 7.98 |
| 0.0 | 0.0 | 0.5 | 7.87 (6.78, 9.23) | 6.04 (5.85, 6.22) | 7.98 |

Note: Results showing unemployment rates when assumption 1 is relaxed such that $U_t$ depends on both $U_t^*$ and $U_{t-1}$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A5.2: Simulation results for setup #2 case 2 (results for unemployment rates with $\lambda \in \{-0.5, 0.5, 1.5\}$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $Unemp\_C(\%)$ | $Unemp\_R(\%)$ | $Unemp\_T(\%)$ |
|---|---|---|---|---|---|
| 1.5 | 1.5 | 0.5 | 8.35 (7.29, 9.50) | 6.99 (6.79, 7.19) | 7.98 |
| 1.5 | 1.5 | −0.5 | 8.82 (7.63, 10.35) | 6.96 (6.76, 7.15) | 7.98 |
| 1.5 | 0.5 | 1.5 | 7.96 (7.18, 8.87) | 6.29 (6.10, 6.47) | 7.98 |
| 1.5 | 0.5 | 0.5 | 8.39 (7.47, 9.42) | 6.26 (6.09, 6.44) | 7.98 |
| 1.5 | 0.5 | −0.5 | 8.82 (7.65, 10.06) | 6.21 (6.02, 6.41) | 7.98 |
| 1.5 | −0.5 | 1.5 | 8.04 (7.09, 9.01) | 5.72 (5.56, 5.90) | 7.98 |
| 1.5 | −0.5 | 0.5 | 8.45 (7.30, 9.72) | 5.71 (5.52, 5.88) | 7.98 |
| 1.5 | −0.5 | −0.5 | 8.89 (7.79, 10.06) | 5.65 (5.47, 5.80) | 7.98 |
| 0.5 | 1.5 | 1.5 | 7.66 (6.57, 8.79) | 7.09 (6.90, 7.29) | 7.98 |
| 0.5 | 1.5 | 0.5 | 8.09 (6.84, 9.50) | 7.05 (6.87, 7.25) | 7.98 |
| 0.5 | 1.5 | −0.5 | 8.50 (7.16, 9.95) | 6.98 (6.78, 7.19) | 7.98 |
| 0.5 | 0.5 | 1.5 | 7.63 (6.61, 8.71) | 6.38 (6.19, 6.57) | 7.98 |
| 0.5 | 0.5 | −0.5 | 8.53 (7.28, 10.03) | 6.23 (6.04, 6.41) | 7.98 |
| 0.5 | −0.5 | 1.5 | 7.66 (6.64, 8.73) | 5.81 (5.64, 5.98) | 7.98 |
| 0.5 | −0.5 | 0.5 | 8.07 (6.91, 9.32) | 5.76 (5.57, 5.93) | 7.98 |
| 0.5 | −0.5 | −0.5 | 8.45 (7.21, 9.93) | 5.67 (5.50, 5.84) | 7.98 |
| −0.5 | 1.5 | 1.5 | 7.35 (6.28, 8.54) | 7.14 (6.95, 7.34) | 7.98 |
| −0.5 | 1.5 | 0.5 | 7.71 (6.49, 9.22) | 7.06 (6.84, 7.27) | 7.98 |
| −0.5 | 1.5 | −0.5 | 8.19 (6.49, 10.05) | 6.93 (6.72, 7.12) | 7.98 |
| −0.5 | 0.5 | 1.5 | 7.22 (6.26, 8.36) | 6.41 (6.23, 6.61) | 7.98 |
| −0.5 | 0.5 | 0.5 | 7.63 (6.44, 8.90) | 6.33 (6.15, 6.51) | 7.98 |
| −0.5 | 0.5 | −0.5 | 8.10 (6.62, 9.86) | 6.19 (6.01, 6.37) | 7.98 |
| −0.5 | −0.5 | 1.5 | 7.25 (6.08, 8.65) | 5.87 (5.68, 6.04) | 7.98 |
| −0.5 | −0.5 | 0.5 | 7.65 (6.45, 8.98) | 5.78 (5.60, 5.96) | 7.98 |

Note: Results showing unemployment rates when assumption 1 is relaxed such that $U_t$ depends on both $U_t^*$ and $U_{t-1}$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A5.3: Simulation results for setup #2 case 2 (results for unemployment rates with $\lambda \in [-3, 4]$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $Unemp\_C(\%)$ | $Unemp\_R(\%)$ | $Unemp\_T(\%)$ |
|---|---|---|---|---|---|
| −1.5 | 4.0 | 1.0 | 6.75 (5.83, 8.13) | 8.80 (8.61, 9.01) | 7.98 |
| 4.0 | 0.5 | −3.0 | 10.77 (9.13, 12.65) | 5.96 (5.78, 6.14) | 7.98 |
| −0.5 | 0.5 | 2.0 | 7.05 (6.03, 7.99) | 6.45 (6.27, 6.63) | 7.98 |
| −0.5 | −2.0 | −2.5 | 53.20 (6.92, 92.75) | 4.72 (4.54, 4.88) | 7.98 |
| −0.5 | −2.5 | 0.0 | 66.78 (6.16, 93.02) | 4.96 (4.81, 5.13) | 7.98 |
| 3.0 | 3.5 | −3.0 | 7.75 (6.77, 9.72) | 8.57 (8.37, 8.78) | 7.98 |
| 2.5 | 2.0 | 1.0 | 8.38 (7.40, 9.37) | 7.29 (7.11, 7.49) | 7.98 |
| 3.0 | −3.0 | −2.5 | 71.94 (9.71, 91.14) | 4.44 (4.31, 4.60) | 7.98 |

Note: Results showing unemployment rates when assumption 1 is relaxed such that $U_t$ depends on both $U_t^*$ and $U_{t-1}$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A6.1: Simulation results for setup #2 case 2 (results for LFP rates with $\lambda \in \{0, 0.5, 1\}$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $LFP\_C(\%)$ | $LFP\_R(\%)$ | $LFP\_T(\%)$ |
|---|---|---|---|---|---|
| 1.0 | 1.0 | 0.5 | 65.97 (65.21, 66.86) | 65.10 (64.81, 65.42) | 65.78 |
| 1.0 | 1.0 | 0.0 | 66.07 (65.26, 67.05) | 65.18 (64.89, 65.50) | 65.78 |
| 1.0 | 0.5 | 1.0 | 65.71 (65.13, 66.47) | 64.89 (64.59, 65.18) | 65.78 |
| 1.0 | 0.5 | 0.5 | 65.84 (65.21, 66.67) | 64.98 (64.68, 65.27) | 65.78 |
| 1.0 | 0.5 | 0.0 | 65.99 (65.29, 66.87) | 65.08 (64.81, 65.37) | 65.78 |
| 1.0 | 0.0 | 1.0 | 65.60 (64.98, 66.33) | 64.79 (64.49, 65.11) | 65.78 |
| 1.0 | 0.0 | 0.5 | 65.74 (65.07, 66.47) | 64.88 (64.60, 65.19) | 65.78 |
| 1.0 | 0.0 | 0.0 | 65.83 (65.14, 66.79) | 64.97 (64.67, 65.27) | 65.78 |
| 0.5 | 1.0 | 1.0 | 65.78 (65.03, 66.56) | 64.84 (64.57, 65.15) | 65.78 |
| 0.5 | 1.0 | 0.5 | 65.92 (65.12, 66.79) | 64.92 (64.62, 65.24) | 65.78 |
| 0.5 | 1.0 | 0.0 | 66.06 (65.27, 67.02) | 65.01 (64.73, 65.29) | 65.78 |
| 0.5 | 0.5 | 1.0 | 65.67 (65.06, 66.40) | 64.73 (64.42, 65.02) | 65.78 |
| 0.5 | 0.5 | 0.0 | 65.94 (65.21, 66.83) | 64.90 (64.62, 65.18) | 65.78 |
| 0.5 | 0.0 | 1.0 | 65.56 (64.94, 66.32) | 64.62 (64.33, 64.92) | 65.78 |
| 0.5 | 0.0 | 0.5 | 65.70 (65.06, 66.50) | 64.70 (64.40, 64.99) | 65.78 |
| 0.5 | 0.0 | 0.0 | 65.83 (65.12, 66.82) | 64.79 (64.51, 65.06) | 65.78 |
| 0.0 | 1.0 | 1.0 | 65.75 (64.98, 66.70) | 64.66 (64.37, 64.99) | 65.78 |
| 0.0 | 1.0 | 0.5 | 65.92 (65.17, 66.89) | 64.76 (64.45, 65.05) | 65.78 |
| 0.0 | 1.0 | 0.0 | 66.06 (65.24, 67.11) | 64.84 (64.52, 65.11) | 65.78 |
| 0.0 | 0.5 | 1.0 | 65.65 (65.00, 66.52) | 64.55 (64.28, 64.82) | 65.78 |
| 0.0 | 0.5 | 0.5 | 65.81 (65.07, 66.80) | 64.63 (64.30, 64.94) | 65.78 |
| 0.0 | 0.5 | 0.0 | 65.96 (65.12, 67.02) | 64.71 (64.39, 65.02) | 65.78 |
| 0.0 | 0.0 | 1.0 | 65.56 (64.92, 66.26) | 64.46 (64.17, 64.73) | 65.78 |
| 0.0 | 0.0 | 0.5 | 65.68 (64.99, 66.60) | 64.54 (64.24, 64.83) | 65.78 |

Note: Results showing LFP rates when assumption 1 is relaxed such that $U_t$ depends on both $U_t^*$ and $U_{t-1}$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A6.2: Simulation results for setup #2 case 2 (results for LFP rates with $\lambda \in \{-0.5, 0.5, 1.5\}$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $LFP\_C(\%)$ | $LFP\_R(\%)$ | $LFP\_T(\%)$ |
|---|---|---|---|---|---|
| 1.5 | 1.5 | 0.5 | 66.02 (65.27, 66.89) | 65.40 (65.09, 65.70) | 65.78 |
| 1.5 | 1.5 | −0.5 | 66.32 (65.44, 67.47) | 65.57 (65.26, 65.89) | 65.78 |
| 1.5 | 0.5 | 1.5 | 65.60 (65.03, 66.20) | 64.96 (64.68, 65.23) | 65.78 |
| 1.5 | 0.5 | 0.5 | 65.84 (65.20, 66.61) | 65.15 (64.83, 65.43) | 65.78 |
| 1.5 | 0.5 | −0.5 | 66.12 (65.39, 66.93) | 65.34 (65.06, 65.62) | 65.78 |
| 1.5 | −0.5 | 1.5 | 65.43 (64.87, 66.02) | 64.77 (64.49, 65.10) | 65.78 |
| 1.5 | −0.5 | 0.5 | 65.65 (65.00, 66.39) | 64.97 (64.67, 65.27) | 65.78 |
| 1.5 | −0.5 | −0.5 | 65.87 (65.16, 66.79) | 65.13 (64.83, 65.44) | 65.78 |
| 0.5 | 1.5 | 1.5 | 65.73 (65.02, 66.55) | 64.87 (64.55, 65.18) | 65.78 |
| 0.5 | 1.5 | 0.5 | 66.02 (65.09, 67.04) | 65.06 (64.77, 65.34) | 65.78 |
| 0.5 | 1.5 | −0.5 | 66.32 (65.41, 67.32) | 65.24 (64.94, 65.54) | 65.78 |
| 0.5 | 0.5 | 1.5 | 65.56 (64.93, 66.31) | 64.63 (64.34, 64.93) | 65.78 |
| 0.5 | 0.5 | −0.5 | 66.12 (65.28, 67.18) | 64.98 (64.68, 65.28) | 65.78 |
| 0.5 | −0.5 | 1.5 | 65.38 (64.77, 65.99) | 64.45 (64.08, 64.72) | 65.78 |
| 0.5 | −0.5 | 0.5 | 65.58 (64.97, 66.28) | 64.63 (64.33, 64.90) | 65.78 |
| 0.5 | −0.5 | −0.5 | 65.83 (65.03, 66.75) | 64.79 (64.50, 65.07) | 65.78 |
| −0.5 | 1.5 | 1.5 | 65.71 (64.97, 66.55) | 64.54 (64.23, 64.82) | 65.78 |
| −0.5 | 1.5 | 0.5 | 65.99 (65.17, 66.98) | 64.72 (64.41, 65.01) | 65.78 |
| −0.5 | 1.5 | −0.5 | 66.33 (65.27, 67.72) | 64.84 (64.55, 65.12) | 65.78 |
| −0.5 | 0.5 | 1.5 | 65.49 (64.85, 66.26) | 64.29 (63.98, 64.58) | 65.78 |
| −0.5 | 0.5 | 0.5 | 65.75 (65.02, 66.76) | 64.44 (64.17, 64.74) | 65.78 |
| −0.5 | 0.5 | −0.5 | 66.12 (65.19, 67.41) | 64.60 (64.31, 64.92) | 65.78 |
| −0.5 | −0.5 | 1.5 | 65.31 (64.74, 65.97) | 64.11 (63.82, 64.39) | 65.78 |
| −0.5 | −0.5 | 0.5 | 65.57 (64.85, 66.44) | 64.27 (63.97, 64.56) | 65.78 |

Note: Results showing LFP rates when assumption 1 is relaxed such that $U_t$ depends on both $U_t^*$ and $U_{t-1}$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A6.3: Simulation results for setup #2 case 2 (results for LFP rates with $\lambda \in [-3, 4]$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $LFP\_C(\%)$ | $LFP\_R(\%)$ | $LFP\_T(\%)$ |
|---|---|---|---|---|---|
| $-1.5$ | 4.0 | 1.0 | 65.52 (64.84, 66.46) | 64.86 (64.54, 65.15) | 65.78 |
| 4.0 | 0.5 | $-3.0$ | 67.22 (66.05, 68.69) | 66.26 (65.98, 66.57) | 65.78 |
| $-0.5$ | 0.5 | 2.0 | 65.39 (64.78, 66.10) | 64.21 (63.93, 64.48) | 65.78 |
| $-0.5$ | $-2.0$ | $-2.5$ | 66.30 (64.74, 69.48) | 64.50 (64.21, 64.77) | 65.78 |
| $-0.5$ | $-2.5$ | 0.0 | 65.33 (64.53, 66.49) | 64.09 (63.80, 64.40) | 65.78 |
| 3.0 | 3.5 | $-3.0$ | 65.58 (64.81, 66.89) | 67.17 (66.85, 67.47) | 65.78 |
| 2.5 | 2.0 | 1.0 | 66.01 (65.33, 66.76) | 65.77 (65.47, 66.08) | 65.78 |
| 3.0 | $-3.0$ | $-2.5$ | 65.97 (65.05, 67.25) | 65.68 (65.39, 65.98) | 65.78 |

Note: Results showing LFP rates when assumption 1 is relaxed such that $U_t$ depends on both $U_t^*$ and $U_{t-1}$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A7: Simulation results for setup #2 case 3 (results for unemployment rates).

| $p$ | $Unemp\_C(\%)$ | $Unemp\_R(\%)$ | $Unemp\_T(\%)$ |
|---|---|---|---|
| 0.0 | 7.96 (6.90, 9.13) | 6.33 (6.14, 6.50) | 7.98 |
| 0.1 | 8.04 (6.97, 9.29) | 6.32 (6.12, 6.52) | 7.98 |
| 0.2 | 8.02 (6.93, 9.27) | 6.33 (6.16, 6.52) | 7.98 |
| 0.3 | 7.99 (6.88, 9.33) | 6.33 (6.15, 6.50) | 7.98 |
| 0.4 | 8.04 (6.76, 9.32) | 6.32 (6.11, 6.51) | 7.98 |
| 0.5 | 8.02 (7.03, 9.24) | 6.31 (6.13, 6.51) | 7.98 |
| 0.6 | 8.01 (6.92, 9.21) | 6.32 (6.16, 6.49) | 7.98 |
| 0.7 | 8.03 (6.97, 9.22) | 6.32 (6.13, 6.50) | 7.98 |
| 0.8 | 7.99 (6.93, 9.23) | 6.32 (6.12, 6.51) | 7.98 |
| 0.9 | 8.02 (6.80, 9.26) | 6.32 (6.14, 6.50) | 7.98 |
| 1.0 | 8.02 (6.87, 9.33) | 6.31 (6.15, 6.51) | 7.98 |

Note: Results showing unemployment rates when assumption 1 is relaxed such that when $U^*_{t-1} = U^*_t$ we replace the value of $U_t$ with $U_{t-1}$ with probability p. Sample size is 100,000. The last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A8: Simulation results for setup #2 case 3 (results for LFP rates).

| $p$ | $Unemp\_C(\%)$ | $Unemp\_R(\%)$ | $Unemp\_T(\%)$ |
|---|---|---|---|
| 0.0 | 65.75 (65.06, 66.58) | 64.80 (64.50, 65.10) | 65.78 |
| 0.1 | 65.82 (65.14, 66.75) | 64.81 (64.53, 65.11) | 65.78 |
| 0.2 | 65.81 (65.12, 66.61) | 64.81 (64.50, 65.11) | 65.78 |
| 0.3 | 65.78 (65.07, 66.75) | 64.80 (64.50, 65.11) | 65.78 |
| 0.4 | 65.83 (65.08, 66.79) | 64.82 (64.53, 65.15) | 65.78 |
| 0.5 | 65.80 (65.13, 66.67) | 64.81 (64.53, 65.13) | 65.78 |
| 0.6 | 65.82 (65.18, 66.61) | 64.82 (64.53, 65.09) | 65.78 |
| 0.7 | 65.83 (65.16, 66.67) | 64.81 (64.49, 65.09) | 65.78 |
| 0.8 | 65.77 (65.08, 66.72) | 64.81 (64.53, 65.08) | 65.78 |
| 0.9 | 65.79 (65.06, 66.59) | 64.80 (64.47, 65.13) | 65.78 |
| 1.0 | 65.80 (65.11, 66.67) | 64.81 (64.53, 65.07) | 65.78 |

Note: Results showing LFP rates when assumption 1 is relaxed such that when $U^*_{t-1} = U^*_t$ we replace the value of $U_t$ with $U_{t-1}$ with probability p. Sample size is 100,000. The last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A9.1: Simulation results for setup #3 (results for unemployment rates with $\lambda \in \{0, 0.5, 1\}$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $Unemp\_C(\%)$ | $Unemp\_R(\%)$ | $Unemp\_T(\%)$ |
|---|---|---|---|---|---|
| 1.0 | 1.0 | 0.5 | 8.02 (6.92, 9.26) | 6.35 (6.18, 6.53) | 8.04 |
| 1.0 | 1.0 | 0.0 | 8.09 (6.83, 9.43) | 6.37 (6.18, 6.55) | 8.06 |
| 1.0 | 0.5 | 1.0 | 8.15 (7.03, 9.30) | 6.29 (6.12, 6.47) | 7.93 |
| 1.0 | 0.5 | 0.5 | 8.27 (7.21, 9.50) | 6.31 (6.13, 6.49) | 7.96 |
| 1.0 | 0.5 | 0.0 | 8.30 (7.16, 9.53) | 6.33 (6.14, 6.51) | 8.00 |
| 1.0 | 0.0 | 1.0 | 8.33 (7.22, 9.46) | 6.25 (6.06, 6.43) | 7.87 |
| 1.0 | 0.0 | 0.5 | 8.36 (7.20, 9.53) | 6.26 (6.08, 6.44) | 7.89 |
| 1.0 | 0.0 | 0.0 | 8.39 (7.28, 9.71) | 6.28 (6.10, 6.47) | 7.93 |
| 0.5 | 1.0 | 1.0 | 7.82 (6.75, 9.17) | 6.34 (6.15, 6.51) | 8.01 |
| 0.5 | 1.0 | 0.5 | 7.92 (6.83, 9.18) | 6.36 (6.18, 6.55) | 8.05 |
| 0.5 | 1.0 | 0.0 | 7.94 (6.83, 9.33) | 6.39 (6.20, 6.57) | 8.08 |
| 0.5 | 0.5 | 1.0 | 7.99 (6.88, 9.22) | 6.30 (6.10, 6.48) | 7.95 |
| 0.5 | 0.5 | 0.0 | 8.05 (6.90, 9.31) | 6.33 (6.15, 6.51) | 8.00 |
| 0.5 | 0.0 | 1.0 | 8.13 (7.05, 9.31) | 6.26 (6.08, 6.45) | 7.87 |
| 0.5 | 0.0 | 0.5 | 8.18 (7.03, 9.32) | 6.28 (6.09, 6.47) | 7.91 |
| 0.5 | 0.0 | 0.0 | 8.22 (7.21, 9.39) | 6.29 (6.09, 6.48) | 7.94 |
| 0.0 | 1.0 | 1.0 | 7.67 (6.60, 8.94) | 6.36 (6.16, 6.55) | 8.02 |
| 0.0 | 1.0 | 0.5 | 7.69 (6.60, 8.91) | 6.37 (6.19, 6.56) | 8.06 |
| 0.0 | 1.0 | 0.0 | 7.76 (6.55, 9.09) | 6.40 (6.22, 6.59) | 8.10 |
| 0.0 | 0.5 | 1.0 | 7.77 (6.70, 8.89) | 6.31 (6.12, 6.50) | 7.96 |
| 0.0 | 0.5 | 0.5 | 7.80 (6.71, 9.09) | 6.34 (6.15, 6.55) | 7.99 |
| 0.0 | 0.5 | 0.0 | 7.94 (6.74, 9.14) | 6.36 (6.17, 6.55) | 8.04 |
| 0.0 | 0.0 | 1.0 | 7.94 (6.81, 9.14) | 6.27 (6.07, 6.46) | 7.89 |
| 0.0 | 0.0 | 0.5 | 7.96 (6.82, 9.18) | 6.29 (6.11, 6.47) | 7.92 |

Note: Results showing unemployment rates when assumption 2 is relaxed such that $U^*_{t+1}$ depends on both $U^*_t$ and $U^*_{t-9}$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A9.2: Simulation results for setup #3 (results for unemployment rates with $\lambda \in \{-0.5, 0.5, 1.5\}$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $Unemp\_C(\%)$ | $Unemp\_R(\%)$ | $Unemp\_T(\%)$ |
|---|---|---|---|---|---|
| 1.5 | 1.5 | 0.5 | 8.19 (7.16, 9.48) | 6.38 (6.20, 6.56) | 8.09 |
| 1.5 | 1.5 | −0.5 | 8.30 (7.25, 9.61) | 6.41 (6.22, 6.63) | 8.15 |
| 1.5 | 0.5 | 1.5 | 8.36 (7.38, 9.49) | 6.26 (6.06, 6.45) | 7.89 |
| 1.5 | 0.5 | 0.5 | 8.43 (7.33, 9.62) | 6.30 (6.13, 6.47) | 7.95 |
| 1.5 | 0.5 | −0.5 | 8.53 (7.26, 9.82) | 6.33 (6.15, 6.51) | 8.01 |
| 1.5 | −0.5 | 1.5 | 8.74 (7.66, 9.87) | 6.18 (6.01, 6.36) | 7.76 |
| 1.5 | −0.5 | 0.5 | 8.77 (7.54, 10.13) | 6.21 (6.02, 6.40) | 7.81 |
| 1.5 | −0.5 | −0.5 | 8.82 (7.54, 10.07) | 6.25 (6.06, 6.43) | 7.88 |
| 0.5 | 1.5 | 1.5 | 7.69 (6.69, 8.81) | 6.36 (6.17, 6.56) | 8.05 |
| 0.5 | 1.5 | 0.5 | 7.75 (6.68, 8.92) | 6.40 (6.21, 6.59) | 8.11 |
| 0.5 | 1.5 | −0.5 | 7.94 (6.70, 9.27) | 6.44 (6.24, 6.63) | 8.19 |
| 0.5 | 0.5 | 1.5 | 7.93 (6.88, 9.14) | 6.28 (6.11, 6.48) | 7.91 |
| 0.5 | 0.5 | −0.5 | 8.14 (7.05, 9.36) | 6.36 (6.17, 6.54) | 8.04 |
| 0.5 | −0.5 | 1.5 | 8.31 (7.32, 9.46) | 6.20 (6.03, 6.40) | 7.78 |
| 0.5 | −0.5 | 0.5 | 8.30 (7.14, 9.55) | 6.23 (6.04, 6.41) | 7.83 |
| 0.5 | −0.5 | −0.5 | 8.40 (7.09, 9.71) | 6.27 (6.08, 6.46) | 7.90 |
| −0.5 | 1.5 | 1.5 | 7.31 (6.26, 8.44) | 6.39 (6.21, 6.57) | 8.08 |
| −0.5 | 1.5 | 0.5 | 7.35 (6.17, 8.64) | 6.42 (6.24, 6.60) | 8.14 |
| −0.5 | 1.5 | −0.5 | 7.49 (6.36, 8.83) | 6.47 (6.29, 6.66) | 8.21 |
| −0.5 | 0.5 | 1.5 | 7.58 (6.51, 8.69) | 6.31 (6.13, 6.50) | 7.94 |
| −0.5 | 0.5 | 0.5 | 7.62 (6.57, 8.81) | 6.35 (6.17, 6.55) | 8.02 |
| −0.5 | 0.5 | −0.5 | 7.65 (6.56, 8.82) | 6.37 (6.18, 6.56) | 8.07 |
| −0.5 | −0.5 | 1.5 | 7.85 (6.76, 9.06) | 6.23 (6.02, 6.42) | 7.81 |
| −0.5 | −0.5 | 0.5 | 7.93 (6.76, 9.17) | 6.27 (6.11, 6.44) | 7.88 |

Note: Results showing unemployment rates when assumption 2 is relaxed such that $U^*_{t+1}$ depends on both $U^*_t$ and $U^*_{t-9}$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A9.3: Simulation results for setup #3 (results for unemployment rates with $\lambda \in [-3, 4]$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $Unemp\_C(\%)$ | $Unemp\_R(\%)$ | $Unemp\_T(\%)$ |
|---|---|---|---|---|---|
| $-1.5$ | 4.0 | 1.0 | 6.45 (5.36, 7.84) | 6.51 (6.31, 6.68) | 8.26 |
| 4.0 | 0.5 | $-3.0$ | 9.12 (7.88, 10.36) | 6.29 (6.09, 6.48) | 7.97 |
| $-0.5$ | 0.5 | 2.0 | 7.53 (6.47, 8.65) | 6.31 (6.13, 6.49) | 7.93 |
| $-0.5$ | $-2.0$ | $-2.5$ | 8.79 (7.46, 10.17) | 6.25 (6.05, 6.45) | 7.86 |
| $-0.5$ | $-2.5$ | 0.0 | 8.88 (7.55, 10.32) | 6.12 (5.94, 6.31) | 7.64 |
| 3.0 | 3.5 | $-3.0$ | 7.71 (6.82, 8.90) | 6.44 (6.27, 6.61) | 8.21 |
| 2.5 | 2.0 | 1.0 | 8.28 (7.23, 9.44) | 6.29 (6.11, 6.49) | 7.96 |
| 3.0 | $-3.0$ | $-2.5$ | 11.03 (9.51, 12.62) | 5.99 (5.80, 6.15) | 7.46 |

Note: Results showing unemployment rates when assumption 2 is relaxed such that $U^*_{t+1}$ depends on both $U^*_t$ and $U^*_{t-9}$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A10.1: Simulation results for setup #3 (results for LFP rates with $\lambda \in \{0, 0.5, 1\}$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $LFP\_C(\%)$ | $LFP\_R(\%)$ | $LFP\_T(\%)$ |
|---|---|---|---|---|---|
| 1.0 | 1.0 | 0.5 | 65.92 (65.18, 66.76) | 64.98 (64.67, 65.26) | 65.96 |
| 1.0 | 1.0 | 0.0 | 66.04 (65.29, 66.85) | 65.05 (64.74, 65.36) | 66.04 |
| 1.0 | 0.5 | 1.0 | 65.88 (65.16, 66.70) | 64.87 (64.55, 65.16) | 65.84 |
| 1.0 | 0.5 | 0.5 | 65.96 (65.30, 66.96) | 64.95 (64.66, 65.26) | 65.93 |
| 1.0 | 0.5 | 0.0 | 66.03 (65.31, 66.99) | 65.02 (64.71, 65.31) | 66.00 |
| 1.0 | 0.0 | 1.0 | 65.87 (65.19, 66.68) | 64.87 (64.57, 65.15) | 65.82 |
| 1.0 | 0.0 | 0.5 | 65.91 (65.23, 66.76) | 64.94 (64.67, 65.27) | 65.90 |
| 1.0 | 0.0 | 0.0 | 66.00 (65.19, 66.88) | 65.01 (64.70, 65.28) | 65.98 |
| 0.5 | 1.0 | 1.0 | 65.75 (65.06, 66.57) | 64.76 (64.47, 65.06) | 65.73 |
| 0.5 | 1.0 | 0.5 | 65.85 (65.18, 66.74) | 64.83 (64.53, 65.15) | 65.81 |
| 0.5 | 1.0 | 0.0 | 65.88 (65.19, 66.76) | 64.90 (64.61, 65.22) | 65.89 |
| 0.5 | 0.5 | 1.0 | 65.77 (65.10, 66.67) | 64.74 (64.44, 65.03) | 65.70 |
| 0.5 | 0.5 | 0.0 | 65.87 (65.14, 66.75) | 64.89 (64.60, 65.20) | 65.86 |
| 0.5 | 0.0 | 1.0 | 65.76 (65.13, 66.63) | 64.73 (64.43, 65.04) | 65.68 |
| 0.5 | 0.0 | 0.5 | 65.81 (65.12, 66.65) | 64.79 (64.52, 65.08) | 65.75 |
| 0.5 | 0.0 | 0.0 | 65.87 (65.20, 66.65) | 64.88 (64.56, 65.16) | 65.85 |
| 0.0 | 1.0 | 1.0 | 65.63 (64.94, 66.43) | 64.61 (64.31, 64.89) | 65.58 |
| 0.0 | 1.0 | 0.5 | 65.69 (64.93, 66.56) | 64.69 (64.40, 64.98) | 65.67 |
| 0.0 | 1.0 | 0.0 | 65.74 (64.92, 66.65) | 64.76 (64.44, 65.05) | 65.74 |
| 0.0 | 0.5 | 1.0 | 65.60 (64.92, 66.45) | 64.60 (64.28, 64.90) | 65.55 |
| 0.0 | 0.5 | 0.5 | 65.68 (65.01, 66.46) | 64.67 (64.41, 64.94) | 65.64 |
| 0.0 | 0.5 | 0.0 | 65.76 (64.98, 66.70) | 64.74 (64.48, 65.07) | 65.72 |
| 0.0 | 0.0 | 1.0 | 65.61 (64.92, 66.46) | 64.57 (64.26, 64.87) | 65.52 |
| 0.0 | 0.0 | 0.5 | 65.66 (64.97, 66.44) | 64.67 (64.40, 64.99) | 65.62 |

Note: Results showing LFP rates when assumption 2 is relaxed such that $U^*_{t+1}$ depends on both $U^*_t$ and $U^*_{t-9}$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A10.2: Simulation results for setup #3 (results for LFP rates with $\lambda \in \{-0.5, 0.5, 1.5\}$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $LFP\_C(\%)$ | $LFP\_R(\%)$ | $LFP\_T(\%)$ |
|---|---|---|---|---|---|
| 1.5 | 1.5 | 0.5 | 66.15 (65.43, 67.03) | 65.14 (64.83, 65.44) | 66.14 |
| 1.5 | 1.5 | −0.5 | 66.30 (65.55, 67.20) | 65.28 (64.97, 65.60) | 66.30 |
| 1.5 | 0.5 | 1.5 | 65.97 (65.33, 66.76) | 64.94 (64.63, 65.24) | 65.91 |
| 1.5 | 0.5 | 0.5 | 66.11 (65.44, 66.94) | 65.10 (64.82, 65.39) | 66.08 |
| 1.5 | 0.5 | −0.5 | 66.21 (65.46, 67.13) | 65.24 (64.94, 65.55) | 66.23 |
| 1.5 | −0.5 | 1.5 | 65.95 (65.33, 66.69) | 64.91 (64.61, 65.20) | 65.86 |
| 1.5 | −0.5 | 0.5 | 66.06 (65.38, 66.91) | 65.06 (64.76, 65.34) | 66.02 |
| 1.5 | −0.5 | −0.5 | 66.15 (65.43, 67.04) | 65.20 (64.91, 65.49) | 66.17 |
| 0.5 | 1.5 | 1.5 | 65.70 (65.08, 66.55) | 64.68 (64.39, 65.00) | 65.66 |
| 0.5 | 1.5 | 0.5 | 65.83 (65.11, 66.69) | 64.85 (64.54, 65.14) | 65.84 |
| 0.5 | 1.5 | −0.5 | 65.99 (65.20, 66.91) | 64.99 (64.68, 65.29) | 66.00 |
| 0.5 | 0.5 | 1.5 | 65.69 (65.03, 66.64) | 64.66 (64.38, 64.95) | 65.62 |
| 0.5 | 0.5 | −0.5 | 65.95 (65.19, 66.93) | 64.96 (64.67, 65.25) | 65.95 |
| 0.5 | −0.5 | 1.5 | 65.67 (65.01, 66.47) | 64.63 (64.31, 64.93) | 65.56 |
| 0.5 | −0.5 | 0.5 | 65.76 (65.12, 66.61) | 64.78 (64.47, 65.07) | 65.72 |
| 0.5 | −0.5 | −0.5 | 65.90 (65.21, 66.76) | 64.93 (64.62, 65.19) | 65.90 |
| −0.5 | 1.5 | 1.5 | 65.43 (64.75, 66.23) | 64.41 (64.13, 64.70) | 65.38 |
| −0.5 | 1.5 | 0.5 | 65.54 (64.81, 66.42) | 64.56 (64.27, 64.84) | 65.54 |
| −0.5 | 1.5 | −0.5 | 65.69 (64.96, 66.63) | 64.72 (64.40, 65.01) | 65.71 |
| −0.5 | 0.5 | 1.5 | 65.45 (64.82, 66.26) | 64.38 (64.07, 64.69) | 65.33 |
| −0.5 | 0.5 | 0.5 | 65.52 (64.85, 66.34) | 64.53 (64.25, 64.83) | 65.49 |
| −0.5 | 0.5 | −0.5 | 65.66 (64.95, 66.44) | 64.70 (64.41, 65.00) | 65.68 |
| −0.5 | −0.5 | 1.5 | 65.39 (64.75, 66.13) | 64.35 (64.07, 64.63) | 65.28 |
| −0.5 | −0.5 | 0.5 | 65.51 (64.83, 66.26) | 64.49 (64.21, 64.80) | 65.43 |

Note: Results showing LFP rates when assumption 2 is relaxed such that $U^*_{t+1}$ depends on both $U^*_t$ and $U^*_{t-9}$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

Table A10.3: Simulation results for setup #3 (results for LFP rates with $\lambda \in [-3, 4]$).

| $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | $LFP\_C(\%)$ | $LFP\_R(\%)$ | $LFP\_T(\%)$ |
|---|---|---|---|---|---|
| $-1.5$ | 4.0 | 1.0 | 65.13 (64.39, 66.09) | 64.21 (63.90, 64.51) | 65.19 |
| 4.0 | 0.5 | $-3.0$ | 66.72 (65.90, 67.60) | 65.92 (65.62, 66.18) | 66.94 |
| $-0.5$ | 0.5 | 2.0 | 65.36 (64.72, 66.19) | 64.29 (64.00, 64.58) | 65.23 |
| $-0.5$ | $-2.0$ | $-2.5$ | 65.86 (65.12, 66.83) | 64.89 (64.62, 65.20) | 65.85 |
| $-0.5$ | $-2.5$ | 0.0 | 65.53 (64.87, 66.31) | 64.50 (64.20, 64.81) | 65.42 |
| 3.0 | 3.5 | $-3.0$ | 66.46 (65.89, 67.30) | 65.96 (65.66, 66.27) | 67.01 |
| 2.5 | 2.0 | 1.0 | 66.25 (65.50, 67.16) | 65.27 (64.97, 65.59) | 66.26 |
| 3.0 | $-3.0$ | $-2.5$ | 66.55 (65.76, 67.55) | 65.72 (65.43, 65.99) | 66.66 |

Note: Results showing LFP rates when assumption 2 is relaxed such that $U^*_{t+1}$ depends on both $U^*_t$ and $U^*_{t-9}$. Sample size is 100,000. For each combination of $\lambda_1$, $\lambda_2$ and $\lambda_3$, the last three columns report mean values from 500 repetitions and 95% lower and upper bounds (in parentheses).

# 4 Evaluation of Assumptions 3 & 4 using CPS data

Assumption 3 requires the observed matrix $M_{U_t,U_{t-9}|x}$ to be invertible. We therefore test the assumption using CPS data directly. Based on the pooled matched CPS sample, we calculate the determinants directly and using bootstrap to derive standard errors. Table A11 present the results. For each demographic group, we can always reject the null that the determinant is zero at the 1% significance level. Therefore, it seems that assumption 3 holds with CPS data.

Under Assumptions 1, 2, and 3, Assumption 4 requires the eigenvalues of $M_{1,U_t,U_{t-9}|x}M^{-1}_{U_t,U_{t-9}|x}$, which is also an observed matrix to be distinctive. Therefore we may similar test this assumption using CPS data. We first derive the three eigenvalues of $M_{1,U_t,U_{t-9}|x}M^{-1}_{U_t,U_{t-9}|x}$ and then rank them in ascending order, such that Eig1≤Eig2≤Eig3. We then calculate the two differences Eig2-Eig1 and Eig3-Eig2 and derive standard errors via bootstrapping. The results are shown in Table A12. Once again, for all demographic groups, the t-values are large and we reject the null that at least two eigenvalues are the same at the 1% level.

Table A11: Testing assumption 3

| Demographic group | Determinant | S.E. | t-value |
|---|---|---|---|
| (1) Male/White/Age≤40 | 6.5e-004 | (1.4e-005) | 45.8 |
| (2) Male/White/Age > 40 | 8.5e-004 | (2.1e-005) | 41.1 |
| (3) Male/Nonwhite/Age≤40 | 2.3e-003 | (6.7e-005) | 33.8 |
| (4) Male/Nonwhite/Age > 40 | 1.6e-003 | (6.6e-005) | 24.0 |
| (5) Female/White/Age≤40 | 6.2e-004 | (1.5e-005) | 40.7 |
| (6) Female/White/Age > 40 | 5.2e-004 | (1.3e-005) | 38.9 |
| (7) Female/Nonwhite/Age≤40 | 1.6e-003 | (5.5e-005) | 28.7 |
| (8) Female/Nonwhite/Age > 40 | 9.8e-004 | (4.6e-005) | 21.1 |

Note: S.E. are standard errors based on 500 bootstrap repetitions.

Table A12: Testing assumption 4

| Demographic group | Eig2-Eig1 | Eig3-Eig2 |
|---|---|---|
| (1) Male/White/Age≤40 | 0.14 | 0.78 |
| | (0.01) | (0.01) |
| | [16.5] | [93.0] |
| (2) Male/White/Age > 40 | 0.15 | 0.83 |
| | (0.01) | (0.01) |
| | [18.2] | [100.6] |
| (3) Male/Nonwhite/Age≤40 | 0.11 | 0.82 |
| | (0.01) | (0.01) |
| | [11.1] | [82.9] |
| (4) Male/Nonwhite/Age > 40 | 0.12 | 0.85 |
| | (0.01) | (0.01) |
| | [8.3] | [60.9] |
| (5) Female/White/Age≤40 | 0.16 | 0.77 |
| | (0.01) | (0.01) |
| | [15.5] | [76.8] |
| (6) Female/White/Age > 40 | 0.15 | 0.82 |
| | (0.01) | (0.01) |
| | [14.3] | [76.9] |
| (7) Female/Nonwhite/Age≤40 | 0.07 | 0.86 |
| | (0.01) | (0.01) |
| | [6.3] | [81.2] |
| (8) Female/Nonwhite/Age > 40 | 0.14 | 0.82 |
| | (0.02) | (0.02) |
| | [7.6] | [44.9] |

Note: Numbers reported in parentheses are standard errors based on 500 bootstrap repetitions. Numbers reported in square brackets are associated t-values.

# 5 Additional results on misclassification probabilities

This section provides some additional results and robustness checks on the estimated misclassification probabilities.

## 5.1 Testing differences in misclassification probabilities between demographic groups

First, we formally test for differences in the misclassification probabilities between demographic groups. Table A13 reports the results with all statistically significant differences listed. The first panel compares males vs. females, controlling for the race and age categories. When employed, males are more likely to misreport as unemployed but less likely to misreport as not-in-labor-force. The differences are always statistically significant at the 5% significance level except for the comparison between young nonwhite males and young nonwhite females. When unemployed, the differences are mostly insignificant, with the only exception being that old white males are less likely to (mis)report as being not-in-labor-force compared to old white females. In addition, when not-in-labor-force, males are more likely to be misclassified as employed.

Panel 2 of Table A13 compares whites with nonwhites. When employed, whites are less likely to be misclassified, either to unemployed or to not-in-labor-force. However, unemployed young whites are more likely to misreport as employed. We also found that young white females are much less likely to misreport as not-in-labor-force compared to young nonwhite females, with the difference in probabilities being 18.6% and statistically significant.

The last panel in Table A13 compares young people (aged 40 and less) with older people (aged over 40). In general, young people are more likely to misreport when they are employed or not-in-labor-force, as the first and last two columns show. Compared with older white females, young white females are less likely to misreport as being not-in-labor-force when they are actually unemployed.

Some previous studies have made strong assumptions regarding between-group misclassification errors. For example, in order to achieve identification, Sinclair and Gastwirth (1998) assume that males and females have the same misclassification error probabilities (see also Sinclair and Gastwirth 1996). Our results suggest that the equality assumptions of misclassification probabilities across different subgroups, which are essential for identification in the H-W models, are unlikely to hold in reality.

## 5.2 Comparing with existing estimates

Our results are broadly consistent with those in the existing literature. Table A14 compares our weighed average estimates of misclassification probabilities with some of those obtained in the previous literature. Note that all the estimates share the same general pattern: the biggest misclassification probabilities happen when unemployed individuals misreport their labor force status as either not-in-labor-force ($\Pr(U_t = 3|U_t^* = 2)$) or employed ($\Pr(U_t = 1|U_t^* = 2)$), while the other misclassification probabilities are all below 3%. Our point estimates of $\Pr(U_t = 3|U_t^* = 2)$ and $\Pr(U_t = 1|U_t^* = 2)$ are somewhat higher than many of the existing estimates. However, several previous estimates of $\Pr(U_t = 3|U_t^* = 2)$ and $\Pr(U_t = 1|U_t^* = 2)$ in Table A14 have large standard errors so that our point estimates are well within their 95% confidence intervals. Due to our methodological advantage and the large sample size, we are able to produce much more precise estimates.

Table A13: Comparing misclassification probabilities (%) across demographic groups

| | Pr(2\|1) | Pr(3\|1) | Pr(1\|2) | Pr(3\|2) | Pr(1\|3) | Pr(2\|3) |
|---|---|---|---|---|---|---|
| Male vs. Female | | | | | | |
| (1)-(5) | 0.3* | -0.8* | 1.5 | 6.4 | 1.6* | 0.0 |
| | (0.07) | (0.12) | (2.03) | (4.86) | (0.50) | (0.41) |
| (2)-(6) | 0.1* | -0.5* | -1.5 | -9.4* | 0.3* | 0.1 |
| | (0.04) | (0.08) | (1.88) | (4.07) | (0.09) | (0.08) |
| (3)-(7) | 0.1 | -0.3 | 1.5 | -11.3 | 2.8* | 4.3* |
| | (0.13) | (0.21) | (1.98) | (9.42) | (0.84) | (1.26) |
| (4)-(8) | 0.2* | -0.3* | 1.6 | -3.1 | -0.0 | -0.7* |
| | (0.10) | (0.15) | (2.66) | (8.07) | (0.18) | (0.19) |
| White vs. Nonwhite | | | | | | |
| (1)-(3) | -0.2 | -0.9* | 6.7* | -0.9 | 0.9 | -4.3* |
| | (0.11) | (0.15) | (1.78) | (4.85) | (0.56) | (1.32) |
| (2)-(4) | -0.2* | -0.6* | 1.0 | -3.1 | 0.1 | 0.1 |
| | (0.08) | (0.11) | (2.15) | (5.91) | (0.17) | (0.14) |
| (5)-(7) | -0.5* | -0.5* | 6.7* | -18.6* | 2.2* | 0.0 |
| | (0.10) | (0.19) | (2.21) | (9.42) | (0.79) | (0.08) |
| (6)-(8) | -0.1 | -0.3* | 4.0 | 3.2 | -0.2* | -0.7* |
| | (0.07) | (0.13) | (2.45) | (6.83) | (0.11) | (0.15) |
| Young vs. Old | | | | | | |
| (1)-(2) | 0.5* | 0.4* | 3.6* | -1.7 | 4.6* | -0.1 |
| | (0.06) | (0.08) | (1.71) | (3.46) | (0.43) | (0.41) |
| (3)-(4) | 0.4* | 0.7* | -2.1 | -3.9 | 3.8* | 4.3* |
| | (0.13) | (0.17) | (2.20) | (6.82) | (0.40) | (1.27) |
| (5)-(6) | 0.3* | 0.7* | 0.6 | -17.4* | 3.4* | 0.0 |
| | (0.05) | (0.12) | (2.18) | (5.31) | (0.27) | (0.08) |
| (7)-(8) | 0.6* | 0.8* | -2.0 | 4.4 | 1.0 | -0.7* |
| | (0.11) | (0.19) | (2.48) | (10.35) | (0.75) | (0.15) |

Note: $\Pr(i|j)$ stands for $\Pr(U_t = i|U_t^* = j)$. The numbers in parentheses in the first column refer to demographic groups defined as follows: (1) Male/White/Age≤40; (2) Male/White/Age > 40; (3) Male/Nonwhite/Age≤40; (4) Male/Nonwhite/Age > 40 ; (5) Female/White/Age≤40; (6) Female/White/Age > 40; (7) Female/Nonwhite/Age≤40 ; (8) Female/Nonwhite/Age > 40. Numbers in parentheses are standard errors, and '*' signifies statistical significance at the 5% level.

Table A14: Comparing misclassification probabilities (%) with those in previous studies

| | Pr(2\|1) | Pr(3\|1) | Pr(1\|2) | Pr(3\|2) | Pr(1\|3) | Pr(2\|3) |
|---|---|---|---|---|---|---|
| PS | 0.54 | 1.72 | 3.78 | 11.46 | 1.16 | 0.64 |
| | (0.07) | (0.18) | (0.70) | (1.09) | (0.13) | (0.09) |
| BB1 | 0.40 | 0.00 | 4.60 | 27.90 | 2.60 | 0.00 |
| | (0.10) | (n.a.) | (15.20) | (5.30) | (1.50) | (n.a.) |
| BB2 | 0.40 | 0.80 | 8.60 | 17.00 | 1.10 | 0.90 |
| | (0.10) | (0.10) | (1.00) | (1.20) | (0.10) | (0.10) |
| SG1 | 0.00 | 0.80 | 6.35 | 16.80 | 1.87 | 0.96 |
| | (0.47) | (0.38) | (10.61) | (5.38) | (0.65) | (0.40) |
| SG2 | 0.00 | 0.96 | 11.13 | 10.00 | 2.02 | 1.09 |
| | (0.98) | (0.25) | (12.58) | (2.46) | (0.34) | (0.24) |
| SG3 | 0.00 | 0.96 | 9.74 | 10.84 | 2.27 | 1.03 |
| | (0.69) | (0.31) | (7.17) | (2.21) | (0.44) | (0.29) |
| This paper | 0.6 | 1.5 | 17.3 | 20.2 | 2.9 | 0.2 |
| | (0.02) | (0.03) | (0.59) | (1.39) | (0.10) | (0.09) |

Note: $\Pr(i|j)$ stands for $\Pr(U_t = i|U_t^* = j)$. 'PS' refers to estimates by Poterba and Summers (1986) (from their Table III); 'BB1' refers to the estimates of Biemer and Bushery (2000) using H-W model for year 1996 (from their Table 5); 'BB2' refers to the estimates of Biemer and Bushery (2000) using MLCA model for year 1996 (from their Table 5); 'SG1' refers to estimates in Sinclair and Gastwirth (1998) for years with low levels of unemployment (1988-1990) (from their Table 5); 'SG2' refers to estimates in Sinclair and Gastwirth (1998) for years with moderate levels of unemployment (1981, 1984-1986) (from their Table 5); 'SG3' refers to estimates in Sinclair and Gastwirth (1998) for years with high levels of unemployment (1982-1983) (from their Table 5); 'This paper' refers to our weighted estimates, which are copied from the last two rows of Table 1.

### 5.3 Robustness check: pooling different periods of data

We then do several robustness checks for the estimated misclassification probabilities. First, in this version of the paper we have updated the sample period from Jan 1996-Dec 2009 to Jan 1996- Aug 2011, representing an increase of 20 sample months. Nevertheless, we still keep the estimated misclassification matrix in the previous version using data up to Dec 2009. The implicit assumption is that misclassification behaviors are relatively stable over time. Therefore, it is not necessary to update misclassification probabilities as new monthly data come out. And readers interested can just update the corrected unemployment series using the misclassification probabilities reported in this paper, without having to be involved in the more complicated procedure of estimating the misclassification probabilities again. Therefore, it is important to test whether updating the joint distributions to August 2011 would make any difference.

Table A15 report the results. For each demographic group, the first row lists misclassification probabilities when we pool data up to Aug 2011. The second row is the differences between the first row and our baseline case (which are reported in Table 1 of the paper when we pool data up to Dec 2009). The third row lists standard errors of the difference. Overall the differences are small and only in a few cases we see statistically significant differences.

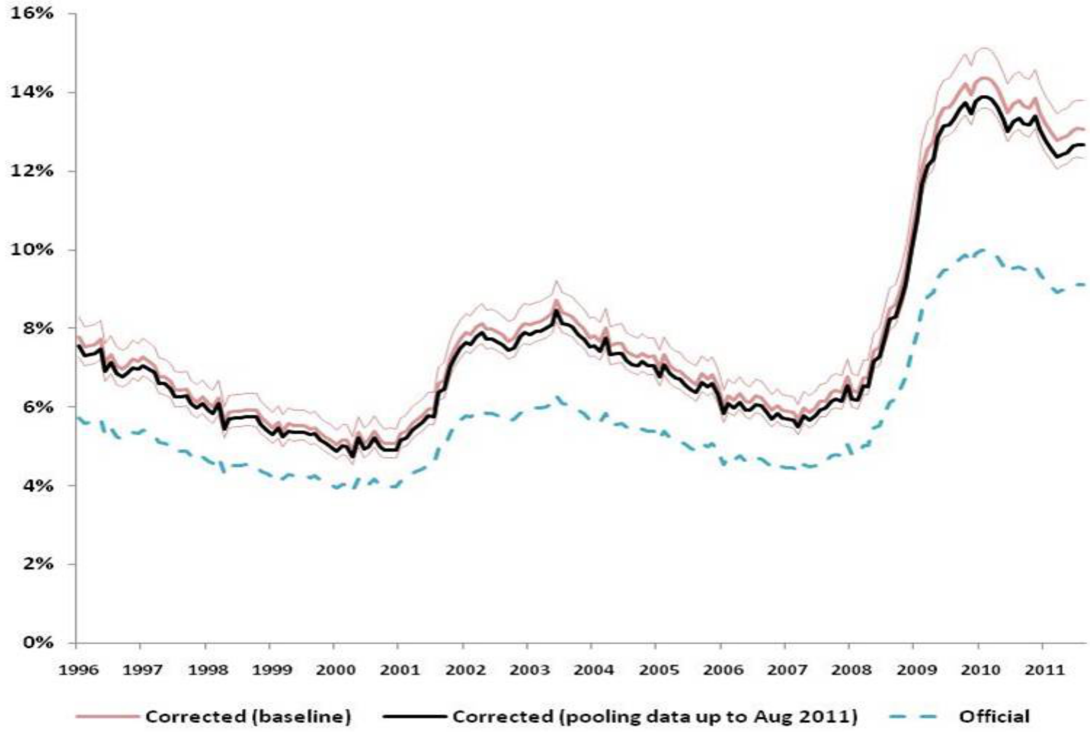
It is perhaps even more important to examine whether levels and trends of the corrected unemployment rates are sensitive to choices made when estimating the misclassification matrices. Figure A1 depicts corrected unemployment rates based on the misclassification matrices reported in Table A15 as well as the baseline case. Note that the two corrected unemployment series are very close to each other and not statistically significantly different based on the confidence intervals for the whole period.

Table A15: Check whether Misclassification probabilities (%) change if pool data from different time periods

| Demographic group | Pr(2\|1) | Pr(3\|1) | Pr(1\|2) | Pr(3\|2) | Pr(1\|3) | Pr(2\|3) |
|---|---|---|---|---|---|---|
| (1) Male/White/Age≤40 | 1.0 | 1.3 | 17.6 | 17.0 | 5.5 | 0.0 |
| | -0.1 | 0.0 | 2.5* | 0.2 | 0.4 | 0.0 |
| | (0.0) | (0.0) | (1.2) | (3.7) | (0.4) | (0.4) |
| (2) Male/White/Age > 40 | 0.5 | 0.9 | 14.3 | 17.2 | 1.4 | 0.1 |
| | -0.1 | -0.0 | 2.2* | 1.6 | 0.0 | 0.0 |
| | (0.0) | (0.0) | (1.1) | (2.7) | (0.1) | (0.1) |
| (3) Male/Nonwhite/Age≤40 | 1.1 | 2.2 | 11.8 | 15.1 | 5.0 | 4.3 |
| | -0.0 | 0.0 | 1.5 | 3.0 | 0.1 | 0.0 |
| | (0.1) | (0.1) | (1.4) | (6.1) | (0.4) | (1.9) |
| (4) Male/Nonwhite/Age > 40 | 0.7 | 1.4 | 13.9 | 18.0 | 1.3 | 0.0 |
| | 0.0 | 0.0 | 1.6 | 4.0 | -0.0 | 0.0 |
| | (0.1) | (0.1) | (2.1) | (6.7) | (0.2) | (0.2) |
| (5) Female/White/Age≤40 | 0.6 | 2.1 | 15.6 | 13.9 | 4.4 | 0.0 |
| | -0.0 | 0.0 | 2.9 | -3.2 | 0.1 | 0.0 |
| | (0.1) | (0.1) | (1.7) | (5.4) | (0.2) | (0.1) |
| (6) Female/White/Age > 40 | 0.4 | 1.4 | 14.3 | 29.0 | 1.1 | 0.0 |
| | -0.0 | -0.0 | 3.7* | -0.8 | -0.0 | 0.0 |
| | (0.0) | (0.0) | (1.5) | (3.9) | (0.0) | (0.0) |
| (7) Female/Nonwhite/Age≤40 | 1.1 | 2.5 | 9.9 | 36.4 | 1.5 | 0.0 |
| | -0.0 | 0.1 | 1.9 | -7.0 | 0.7 | 0.0 |
| | (0.1) | (0.1) | (1.7) | (10.3) | (1.0) | (0.1) |
| (8) Female/Nonwhite/Age > 40 | 0.5 | 1.8 | 10.9 | 27.1 | 1.2 | 0.4 |
| | -0.0 | -0.0 | 2.9 | -2.0 | 0.0 | 0.3 |
| | (0.1) | (0.1) | (2.3) | (8.1) | (0.1) | (0.3) |

Note: for each subgroup, the first row lists misclassification probabilities when we pool data up to Aug 2011. the second row is difference between the first row and numbers reported in table 1 (where we pool data up to Dec 2009). the standard errors of the differences are reported in the third row. * signifies the difference is statistically significant at the 5% level.

Figure A1: Corrected unemployment rates using estimated misclassification probabilities when pool data up to Aug 2011




Note: Figure showing seasonally-adjusted corrected unemployment rate series when we pool data up to August 2011 in estimating the misclassification matrix, in addition to the baseline corrected unemployment rate series and official unemployment rate series showing in Figure 1. The thin lines signify 95% upper and lower confidence bounds for the baseline corrected series.

### 5.4 Robustness check: misclassification probabilities dependent on labor market conditions

Next, we divide our sample into three sub-periods based on the US business cycles and allow misclassification probabilities to be different for each sub-period. The first sub-period goes from the beginning of our study period (January 1996) to October 2001, which is roughly the end of the 2001 recession. The second sub-period goes from November 2001 to November 2007, corresponding to the expansion period between two recessions (the 2001 recession and the most recent 2007-09 recession). The third sub-period goes from December 2007 to the end of our study period (Aug 2011), which includes the 2007-2009 recession and its aftermath. Compared with the first two sub-periods, sub-period 3 is characterized by much higher levels of unemployment and presumably reflecting considerably weaker labor market conditions. Therefore we are able to test directly whether misclassification probabilities are affected by labor market conditions.

Table A16 reports the misclassification probabilities for each sub-period. There does seem to be differences in misclassification behaviors among different sub-periods characterized by different labor market conditions. For example, column (3) shows the probability of reporting employed while the true status is unemployed. Note that for all the demographic sub-groups, the probabilities of misreporting in sub-period 3 is considerably lower than in sub-periods 1 and 2. This shows that when labor market are weak and the pool of unemployed people includes a larger share of job losers and others whose status is unambiguous, then misreporting of unemployment tend to be less prevalent. In Table A17 we test the statistical significance of the differences between misclassification probabilities in different sub-periods. It has been that shown that there do exist some significant differences.

Nevertheless, we are able to show that such differences in the misclassification probabilities do not lead to estimated corrected unemployment rates to significantly differ from our baseline series. Figure A2 show corrected unemployment rate series constructed by using different misclassification matrices for each sub-period as reported in Table A16, in addition to our baseline corrected unemployment rate series and the official unemployment rate series. The two corrected series are very close to each other, and the one constructed using three different misclassification matrices are within the 95% confidence bounds of the baseline series. In addition, the two corrected series are quite different from the official series in terms of levels and cyclical patterns. Thus we can conclude that the cyclical pattern of differences between the official and corrected unemployment rates shown in Figure 1 is not an artifact of not allowing the prevalence of reporting errors to vary with labor market conditions.

In summary, the misclassification probabilities may be statistically significantly different, but they is economically insignificant because the corrected unemployment rates are not statistically different from our baseline estimates.

Table A16: misclassification probabilities (%) for the three different subperiods

| Demographic group | Pr(2\|1) | Pr(3\|1) | Pr(1\|2) | Pr(3\|2) | Pr(1\|3) | Pr(2\|3) |
|---|---|---|---|---|---|---|
| (1) Male/White/Age≤40 | 0.8 | 1.3 | 29.2 | 7.5 | 6.7 | 2.9 |
| | 0.9 | 1.4 | 19.5 | 15.5 | 6.1 | 0.0 |
| | 1.2 | 1.1 | 11.6 | 27.6 | 3.8 | 0.0 |
| (2) Male/White/Age > 40 | 0.4 | 0.9 | 17.4 | 21.6 | 1.5 | 0.1 |
| | 0.4 | 0.9 | 16.0 | 20.6 | 1.3 | 0.2 |
| | 0.6 | 0.9 | 12.3 | 11.0 | 1.2 | 0.1 |
| (3) Male/Nonwhite/Age≤40 | 1.4 | 2.1 | 16.9 | 8.9 | 4.9 | 3.0 |
| | 0.9 | 2.3 | 12.8 | 19.0 | 5.3 | 5.9 |
| | 1.2 | 2.1 | 8.1 | 8.8 | 4.2 | 3.1 |
| (4) Male/Nonwhite/Age > 40 | 0.7 | 1.5 | 18.6 | 21.5 | 1.6 | 0.3 |
| | 0.6 | 1.5 | 16.3 | 20.3 | 1.2 | 0.0 |
| | 0.6 | 1.2 | 11.1 | 11.4 | 1.0 | 0.0 |
| (5) Female/White/Age≤40 | 0.7 | 2.0 | 20.6 | 16.6 | 4.1 | 0.0 |
| | 0.5 | 2.2 | 18.3 | 9.3 | 4.7 | 0.3 |
| | 0.7 | 1.9 | 10.2 | 19.1 | 3.8 | 0.0 |
| (6) Female/White/Age > 40 | 0.3 | 1.4 | 22.8 | 28.1 | 1.1 | 0.0 |
| | 0.3 | 1.4 | 17.2 | 30.9 | 1.0 | 0.0 |
| | 0.4 | 1.4 | 10.2 | 25.6 | 1.1 | 0.0 |
| (7) Female/Nonwhite/Age≤40 | 1.2 | 2.4 | 11.9 | 43.8 | 0.0 | 0.0 |
| | 1.0 | 2.8 | 11.7 | 19.4 | 3.4 | 0.0 |
| | 0.9 | 2.1 | 6.2 | 51.9 | 0.0 | 0.0 |
| (8) Female/Nonwhite/Age > 40 | 0.2 | 1.7 | 18.2 | 14.3 | 1.3 | 0.9 |
| | 0.5 | 1.9 | 12.4 | 26.6 | 1.2 | 0.2 |
| | 0.6 | 1.7 | 6.8 | 40.8 | 1.2 | 0.0 |

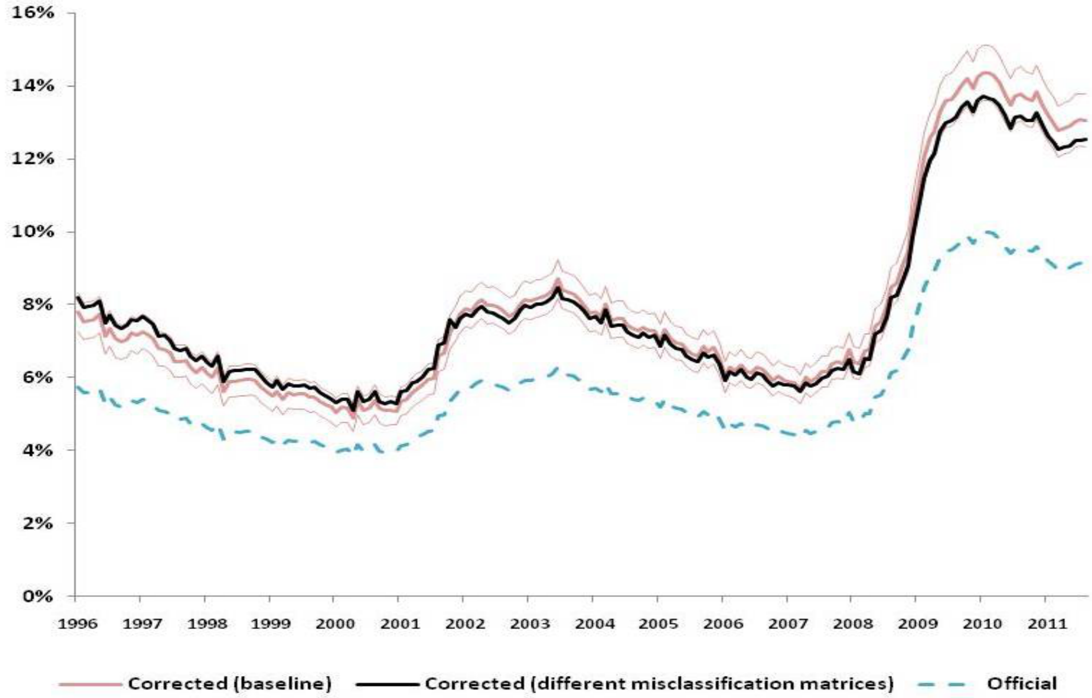
Note: $\Pr(i|j)$ stands for $\Pr(U_t = i|U_t^* = j)$. Each panel represents a demographic group as defined. Within each panel, the three rows represent misclassification probabilities for sub-period 1, 2, and 3, respectively.

Table A17: Testing whether misclassification probabilties are the same for the three different subperiods

| Demographic group | Pr(2\|1) | Pr(3\|1) | Pr(1\|2) | Pr(3\|2) | Pr(1\|3) | Pr(2\|3) |
|---|---|---|---|---|---|---|
| (1) Male/White/Age≤40 | -0.1 | -0.1 | 9.7* | -7.9 | 0.6 | 2.9* |
| | -0.3* | 0.3* | 7.9* | -12.1 | 2.3* | 0.0 |
| | 0.4* | -0.1 | -17.6* | 20.1* | -3.0* | -2.9* |
| (2) Male/White/Age > 40 | -0.1 | -0.0 | 1.4 | 0.9 | 0.1 | -0.1 |
| | -0.2* | -0.0 | 3.7* | 9.6 | 0.1 | 0.1* |
| | 0.3* | 0.0 | -5.1 | -10.6 | -0.2 | -0.0 |
| (3) Male/Nonwhite/Age≤40 | 0.5* | -0.2* | 4.1 | -10.1* | -0.4 | -2.8* |
| | -0.2* | 0.2 | 4.7* | 10.2 | 1.1 | 2.8* |
| | -0.3* | -0.0 | -8.8* | -0.1 | -0.7 | 0.1 |
| (4) Male/Nonwhite/Age > 40 | 0.0 | -0.1 | 2.3 | 1.2 | 0.5 | 0.3 |
| | 0.0 | 0.3* | 5.2* | 8.9 | 0.2 | 0.0 |
| | -0.0 | -0.3* | -7.5* | -10.1 | -0.7 | -0.3 |
| (5) Female/White/Age≤40 | 0.2* | -0.3* | 2.3 | 7.3 | -0.6 | -0.3 |
| | -0.2* | 0.4* | 8.0* | -9.8* | 1.0 | 0.3* |
| | -0.0 | -0.1* | -10.4* | 2.5 | -0.4 | 0.0 |
| (6) Female/White/Age > 40 | 0.0 | -0.1 | 5.6* | -2.8 | 0.1 | 0.0 |
| | -0.1 | 0.1* | 7.1* | 5.3 | -0.1 | 0.0 |
| | 0.1 | 0.0 | -12.7* | -2.5 | 0.1 | -0.0 |
| (7) Female/Nonwhite/Age≤40 | 0.2 | -0.4* | 0.3 | 24.4* | -3.4* | 0.0 |
| | 0.1 | 0.8* | 5.5* | -32.5* | 3.4* | 0.0 |
| | -0.4* | -0.3* | -5.8* | 8.0* | 0.0 | 0.0 |
| (8) Female/Nonwhite/Age > 40 | -0.4* | -0.1* | 5.8* | -12.3 | 0.1 | 0.7 |
| | -0.0 | 0.1* | 5.6* | -14.1* | 0.0 | 0.2* |
| | 0.4* | 0.0 | -11.5* | 26.5* | -0.1 | -0.9 |

Note: $\Pr(i|j)$ stands for $\Pr(U_t = i|U_t^* = j)$. within each panel, the first line is the difference between subperiod 1 and subperiod 2, the second row is the difference between subperiod 2 and subperiod 3, and the third row is the difference between subperiod 3 and subperiod 1. signifies statistical difference at the 5% level.

Figure A2: Corrected unemployment rates when using three different misclassification matrices for each subperiod




Note: Figure showing seasonally-adjusted corrected unemployment rate series when misclassification probabilities are allowed to vary in different sub-periods, in addition to the baseline corrected unemployment rate series and official unemployment rate series showing in Figure 1. The thin lines signify 95% upper and lower confidence bounds for the baseline corrected series.

## 5.5 Robustness check: using different matching weights

Lastly, we check the role played by matching weights when we derive joint LFS distributions from matching CPS monthly data. It is well known that attrition is a serious issue in matched CPS samples and the matched sample may not be representative of the US population along important dimensions (see for example: Paracchi and Welch, 1995). Therefore we calculate and use matching weights when matching three CPS monthly data sets. We first run a Logit model to estimate the probability of attrition, then use the predicted probability to construct matching weights. Under the assumption that the probability of attrition is determined by the variables we included, our method is consistent.

Nevertheless, there might be attrition based on unobservables which we are not able to incorporate when calculating the matching weights. To examine the robustness of our method, we have tried not using the matching weights, i.e., not accounting for attrition, when matching CPS monthly data sets. Table A18 presents the results, where for each demographic group, the first row list misclassification probabilities when we do not use matching weights (i.e., assuming there's no attrition in matching). The second row show differences between the first row and the baseline probabilities as shown in Table 1. Standard errors are shown in the third row. Overall, the differences in misclassification probabilities are very small and there are only a few statistically significant differences.

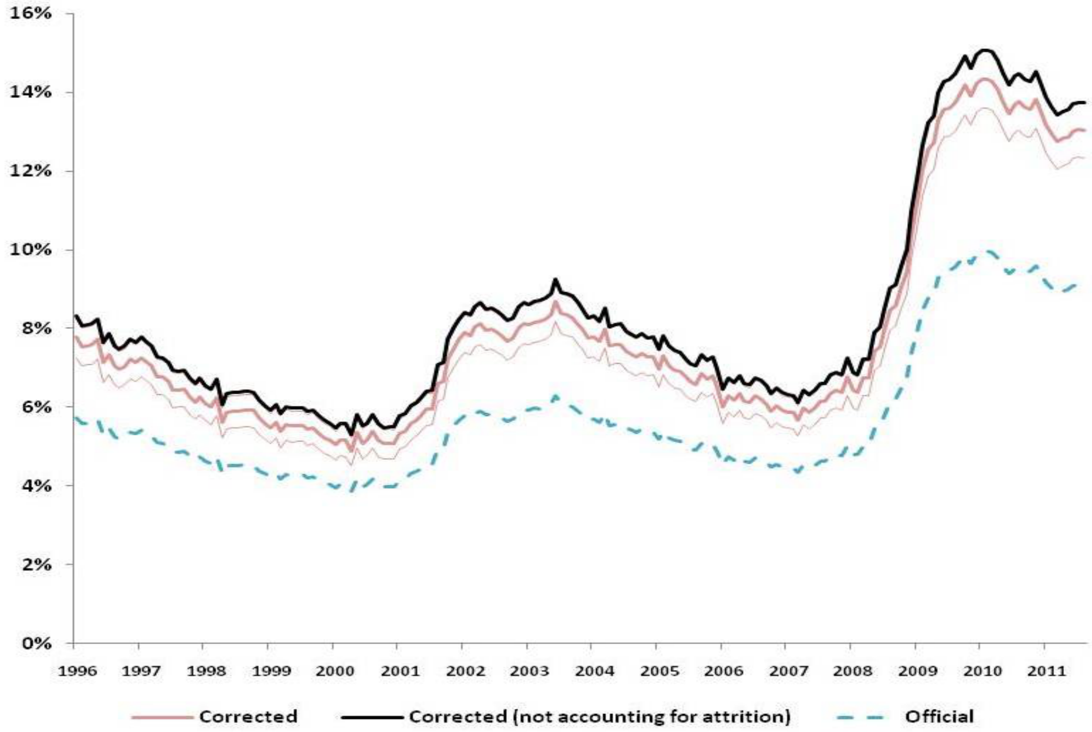
Figure A3 compares the corrected unemployment rate series using different weights. Note that the corrected unemployment rates only change very modestly when we do not account for attrition and are with the 95% confidence bounds of the baseline series. Both corrected unemployment rates are far from the official unemployment rate series. Therefore, our results are robust to the procedure to correct the weights for attrition.

Table A18: Check whether it matters or not to use matching weight to account for attrition

| Demographic group | Pr(2\|1) | Pr(3\|1) | Pr(1\|2) | Pr(3\|2) | Pr(1\|3) | Pr(2\|3) |
|---|---|---|---|---|---|---|
| (1) Male/White/Age≤40 | 0.7 | 1.1 | 22.5 | 17.6 | 6.2 | 0.0 |
| | 0.2* | 0.2* | -2.5 | -0.4 | -0.2 | 0.0 |
| | (0.1) | (0.1) | (1.6) | (4.7) | (0.5) | (0.6) |
| (2) Male/White/Age > 40 | 0.4 | 0.9 | 17.5 | 19.3 | 1.4 | 0.1 |
| | 0.0 | 0.0 | -1.0 | -0.5 | -0.0 | 0.0 |
| | (0.0) | (0.0) | (1.4) | (3.6) | (0.1) | (0.1) |
| (3) Male/Nonwhite/Age≤40 | 0.9 | 1.9 | 15.5 | 20.2 | 5.4 | 3.2 |
| | 0.2 | 0.3 | -2.1 | -2.1 | -0.4 | 1.1 |
| | (0.1) | (0.1) | (1.9) | (7.5) | (0.6) | (2.1) |
| (4) Male/Nonwhite/Age > 40 | 0.6 | 1.4 | 16.8 | 22.1 | 1.3 | 0.0 |
| | 0.1 | 0.1 | -1.3 | -0.2 | -0.1 | 0.0 |
| | (0.1) | (0.1) | (2.8) | (8.7) | (0.2) | (0.2) |
| (5) Female/White/Age≤40 | 0.5 | 1.9 | 20.9 | 12.4 | 4.4 | 0.0 |
| | 0.1 | 0.2* | -2.3 | -1.6 | -0.0 | 0.0 |
| | (0.1) | (0.1) | (2.3) | (6.8) | (0.3) | (0.1) |
| (6) Female/White/Age > 40 | 0.3 | 1.4 | 18.7 | 28.7 | 1.1 | 0.0 |
| | 0.0 | 0.0 | -0.8 | -0.5 | -0.0 | 0.0 |
| | (0.0) | (0.0) | (1.8) | (4.7) | (0.1) | (0.0) |
| (7) Female/Nonwhite/Age≤40 | 0.9 | 2.3 | 13.1 | 31.1 | 2.7 | 0.0 |
| | 0.2 | 0.2 | -1.2 | -1.7 | -0.5 | 0.0 |
| | (0.1) | (0.1) | (2.5) | (12.7) | (1.2) | (0.1) |
| (8) Female/Nonwhite/Age > 40 | 0.4 | 1.7 | 15.1 | 24.4 | 1.3 | 0.6 |
| | 0.1 | 0.0 | -1.2 | 0.7 | -0.1 | 0.0 |
| | (0.1) | (0.1) | (3.0) | (8.8) | (0.1) | (0.2) |

Note: $\Pr(i|j)$ stands for $\Pr(U_t = i|U_t^* = j)$. for each subgroup, the first row lists misclassification probabilities when we do not account for attrition when matching CPS monthly data sets. the second row is difference between the first row and numbers reported in table 1 (where we do account for attrition). the standard errors of the differences are reported in the third row. * signifies the difference is statistically significant at the 5% level.

Figure A3: Corrected unemployment rates using estimated misclassification probabilities not accounting for attrition




Note: Figure showing corrected unemployment rate series when we do not account for attrition in estimating the misclassification matrix, in addition to the baseline corrected unemployment rate series and official unemployment rate series showing in Figure 1. The thin lines signify 95% upper and lower confidence bounds for the baseline corrected series. The corrected series not accounting for attrition is indistinguishable on the graph from the 95% upper bounds of the baseline corrected series, although numbers are not identical.

# 6 Additional results on unemployment rates

First, we report monthly corrected unemployment rates from January 1996 to August 2011 in Table A19. Note that researchers can update the series when new data come in using our estimated misclassification probabilities. The numbers reported in Table A19 are not seasonally adjusted. Standard errors are also reported using bootstrapping.

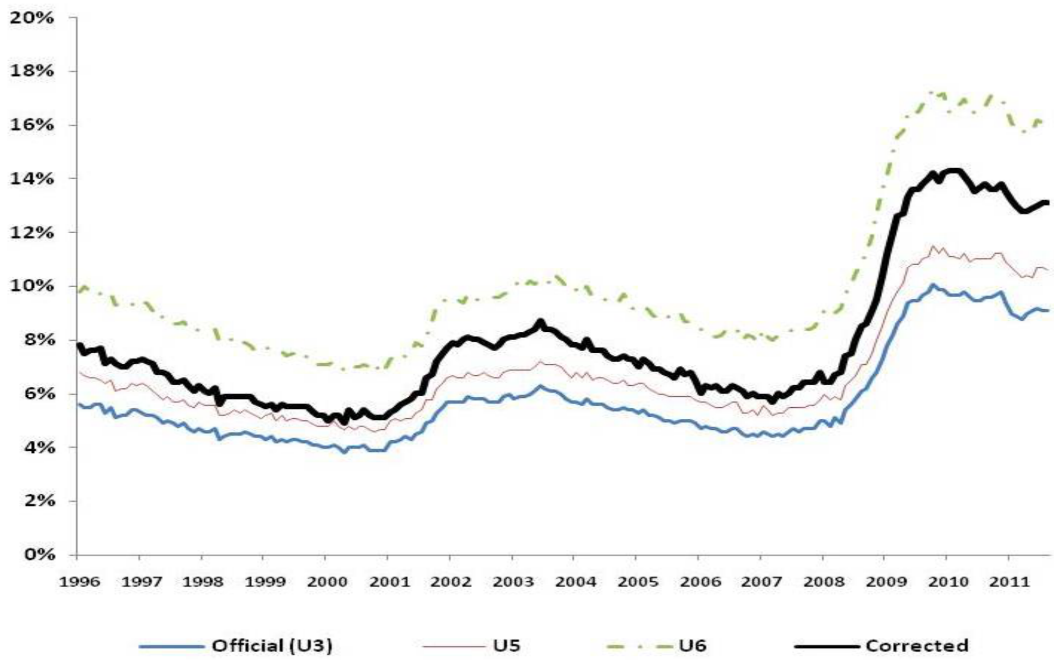
Next we compare our corrected unemployment rates with several alternative unemployment measures announced by BLS. in Figure A4 Note that our corrected series are somewhat in between two broad measures of unemployment rates that BLS report: U5 and U6. U5 basically includes all marginally attached workers (such as discouraged workers) as unemployed, while U6 includes both marginally attached workers and part-time workers for economic reasons as unemployed. Thus our corrected series at least partly correct for the LFS of the two groups of people which are difficult to classify conceptually. This can also be seen by the relatively large estimates of $\Pr(U_t = 3|U_t^* = 2)$ and $\Pr(U_t = 1|U_t^* = 2)$ using our procedure.

Table A19: Monthly corrected unemployment rates (%)

| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1996 | 8.7 | 8.2 | 7.9 | 7.3 | 7.2 | 7.4 | 7.7 | 6.9 | 6.7 | 6.6 | 6.8 | 6.7 |
| | (0.28) | (0.26) | (0.26) | (0.25) | (0.24) | (0.25) | (0.25) | (0.24) | (0.24) | (0.23) | (0.23) | (0.23) |
| 1997 | 8.2 | 7.8 | 7.4 | 6.5 | 6.3 | 6.9 | 6.8 | 6.3 | 6.2 | 5.8 | 5.7 | 5.8 |
| | (0.27) | (0.26) | (0.25) | (0.22) | (0.22) | (0.24) | (0.23) | (0.23) | (0.22) | (0.21) | (0.21) | (0.21) |
| 1998 | 6.9 | 6.7 | 6.6 | 5.3 | 5.4 | 6.2 | 6.2 | 5.8 | 5.7 | 5.4 | 5.3 | 5.1 |
| | (0.24) | (0.23) | (0.23) | (0.20) | (0.20) | (0.22) | (0.22) | (0.22) | (0.21) | (0.20) | (0.20) | (0.19) |
| 1999 | 6.3 | 6.3 | 5.8 | 5.3 | 5.0 | 5.8 | 5.8 | 5.4 | 5.2 | 4.9 | 4.9 | 4.7 |
| | (0.22) | (0.22) | (0.21) | (0.20) | (0.19) | (0.21) | (0.21) | (0.20) | (0.20) | (0.19) | (0.19) | (0.18) |
| 2000 | 5.9 | 5.8 | 5.5 | 4.6 | 4.9 | 5.4 | 5.5 | 5.3 | 4.8 | 4.6 | 4.7 | 4.6 |
| | (0.21) | (0.21) | (0.20) | (0.18) | (0.19) | (0.20) | (0.20) | (0.20) | (0.19) | (0.18) | (0.19) | (0.18) |
| 2001 | 6.2 | 6.0 | 6.0 | 5.4 | 5.3 | 6.2 | 6.3 | 6.5 | 6.3 | 6.7 | 7.2 | 7.3 |
| | (0.22) | (0.21) | (0.21) | (0.20) | (0.20) | (0.22) | (0.22) | (0.22) | (0.22) | (0.23) | (0.24) | (0.23) |
| 2002 | 8.7 | 8.4 | 8.4 | 7.9 | 7.5 | 8.3 | 8.2 | 7.8 | 7.3 | 7.2 | 7.7 | 7.7 |
| | (0.27) | (0.26) | (0.26) | (0.26) | (0.25) | (0.26) | (0.26) | (0.25) | (0.24) | (0.24) | (0.25) | (0.25) |
| 2003 | 8.9 | 8.8 | 8.5 | 8.0 | 7.9 | 9.0 | 8.7 | 8.3 | 7.9 | 7.6 | 7.7 | 7.4 |
| | (0.27) | (0.27) | (0.26) | (0.25) | (0.25) | (0.27) | (0.27) | (0.26) | (0.25) | (0.25) | (0.25) | (0.24) |
| 2004 | 8.6 | 8.3 | 8.3 | 7.3 | 7.2 | 7.9 | 7.8 | 7.2 | 6.9 | 6.8 | 7.0 | 6.9 |
| | (0.27) | (0.27) | (0.27) | (0.24) | (0.25) | (0.26) | (0.25) | (0.24) | (0.23) | (0.23) | (0.23) | (0.23) |
| 2005 | 7.8 | 7.9 | 7.3 | 6.6 | 6.5 | 7.1 | 7.0 | 6.5 | 6.5 | 6.1 | 6.5 | 6.1 |
| | (0.25) | (0.25) | (0.24) | (0.23) | (0.23) | (0.23) | (0.24) | (0.23) | (0.22) | (0.22) | (0.22) | (0.21) |
| 2006 | 6.9 | 6.9 | 6.4 | 6.0 | 5.8 | 6.4 | 6.7 | 6.2 | 5.7 | 5.3 | 5.6 | 5.6 |
| | (0.23) | (0.23) | (0.22) | (0.21) | (0.21) | (0.22) | (0.23) | (0.21) | (0.21) | (0.20) | (0.20) | (0.20) |
| 2007 | 6.7 | 6.5 | 6.0 | 5.6 | 5.5 | 6.3 | 6.5 | 6.1 | 6.0 | 5.8 | 5.9 | 6.4 |
| | (0.23) | (0.22) | (0.21) | (0.20) | (0.20) | (0.21) | (0.23) | (0.22) | (0.21) | (0.21) | (0.21) | (0.22) |
| 2008 | 7.3 | 7.1 | 7.1 | 6.3 | 7.1 | 7.8 | 8.3 | 8.4 | 8.2 | 8.4 | 9.0 | 10.0 |
| | (0.24) | (0.23) | (0.23) | (0.21) | (0.23) | (0.25) | (0.26) | (0.27) | (0.26) | (0.26) | (0.28) | (0.29) |
| 2009 | 12.1 | 12.7 | 12.9 | 12.3 | 12.9 | 13.9 | 14.0 | 13.7 | 13.6 | 13.6 | 13.4 | 13.9 |
| | (0.34) | (0.35) | (0.36) | (0.34) | (0.36) | (0.38) | (0.37) | (0.38) | (0.38) | (0.37) | (0.37) | (0.38) |
| 2010 | 15.2 | 15.1 | 14.7 | 13.6 | 13.4 | 13.8 | 14.0 | 13.7 | 13.2 | 13.0 | 13.4 | 13.1 |
| | (0.40) | (0.40) | (0.39) | (0.38) | (0.37) | (0.38) | (0.39) | (0.38) | (0.38) | (0.36) | (0.37) | (0.37) |
| 2011 | 14.1 | 13.7 | 13.2 | 12.4 | 12.5 | 13.4 | 13.4 | 13.0 | | | | |
| | (0.38) | (0.38) | (0.37) | (0.35) | (0.35) | (0.37) | (0.38) | (0.37) | | | | |

Note: Not seasonally adjusted. Numbers reported in parentheses are bootstrapped standard errors.

Figure A4: Comparing our corrected unemployment rates with alternative measures announced by BLS




Note: Figure showing corrected unemployment rate series, official unemployment rate series as well as two alternative measures of unemployment that BLS uses. U5 classify marginally attached people as unemployed while U6 classify both marginally attached and part-time workers for economic reasons as unemployed. All series are seasonally adjusted. Sources: U5, U6 and Official unemployment rate series (U3) are from http://www.bls.gov/webapps/legacy/cpsatab15.htm.

# 7 Results on labor force participation rates

This section reports the results on the labor force participation rates. Table A20 presents results for each demographic group for the three sub-periods: January 1996 to October 2001, November 2001 to November 2007, and December 2007 to August 2011. For each demographic group and each sub-period, the corrected labor force participation rates are always higher than the reported ones, but the differences are small and not statistically significant. For example, for young white males, in the first sub-period (January 1996 to October 2001), the corrected labor force participation rate is 87.8%, which is higher than the reported rate by 1.3 percentage points. In the second sub-period, the corrected labor force participation rate is 84.9%, again higher than the reported rate of 83.6% by 1.3 percentage points. In the latest recession period, the difference between corrected and reported labor force participation rates is 1.9 percentage points. By contrast, the standard errors are close to 4%.

The last two rows of Table A20 reports LFP for the whole US population. The corrected participation rate is always slightly higher than the reported one, but the average difference is less than 2%, and not statistically significant. For the three sub-periods, the corrected labor force participation rate is 68.1%, 67.3% and 66.8%, respectively. The reported rates are only slightly lower, at 67.1%, 66.2% and 65.2%, respectively.

Table A21 reports all monthly seasonally unadjusted labor force participation rates as well as bootstrapped standard errors. Figure A5 graphically depicts both the corrected and the official seasonally adjusted series, with both somewhat flat during the period under study.

Table A20: Labor force participation rates (%) averaged over three sub-periods for different demographic groups

| Demographic group | Sub-period 1 (1996/01-2001/10) reported | Sub-period 1 (1996/01-2001/10) corrected | Sub-period 2 (2001/11-2007/11) reported | Sub-period 2 (2001/11-2007/11) corrected | Sub-period 3 (2007/12-2011/8) reported | Sub-period 3 (2007/12-2011/8) corrected |
|---|---|---|---|---|---|---|
| (1) Male/White/Age$\leq$40 | 86.5 | 87.8 | 83.6 | 84.9 | 80.8 | 82.7 |
| | (3.9) | (3.9) | (3.7) | (3.8) | (3.6) | (3.7) |
| (2) Male/White/Age $>$ 40 | 65.5 | 66.0 | 66.6 | 67.3 | 66.4 | 67.6 |
| | (2.9) | (3.0) | (3.0) | (3.0) | (3.0) | (3.0) |
| (3) Male/Nonwhite/Age$\leq$40 | 75.6 | 76.5 | 74.1 | 74.8 | 71.8 | 73.1 |
| | (3.4) | (3.5) | (3.3) | (3.4) | (3.2) | (3.4) |
| (4) Male/Nonwhite/Age $>$ 40 | 63.6 | 65.0 | 64.9 | 66.6 | 64.2 | 66.7 |
| | (2.8) | (2.9) | (2.9) | (3.0) | (2.9) | (3.0) |
| (5) Female/White/Age$\leq$40 | 72.4 | 73.2 | 69.6 | 70.2 | 68.3 | 69.0 |
| | (3.2) | (3.3) | (3.1) | (3.1) | (3.1) | (3.1) |
| (6) Female/White/Age $>$ 40 | 49.0 | 49.8 | 51.6 | 52.6 | 52.4 | 54.0 |
| | (2.2) | (2.2) | (2.3) | (2.4) | (2.3) | (2.4) |
| (7) Female/Nonwhite/Age$\leq$40 | 68.9 | 73.0 | 66.7 | 70.7 | 64.9 | 69.7 |
| | (3.1) | (3.5) | (3.0) | (3.4) | (2.9) | (3.5) |
| (8) Female/Nonwhite/Age $>$ 40 | 52.6 | 53.3 | 54.5 | 55.6 | 54.4 | 55.8 |
| | (2.4) | (2.4) | (2.4) | (2.5) | (2.4) | (2.5) |
| Total | 67.1 | 68.1 | 66.2 | 67.3 | 65.2 | 66.8 |
| | (1.3) | (1.3) | (1.3) | (1.3) | (1.2) | (1.3) |

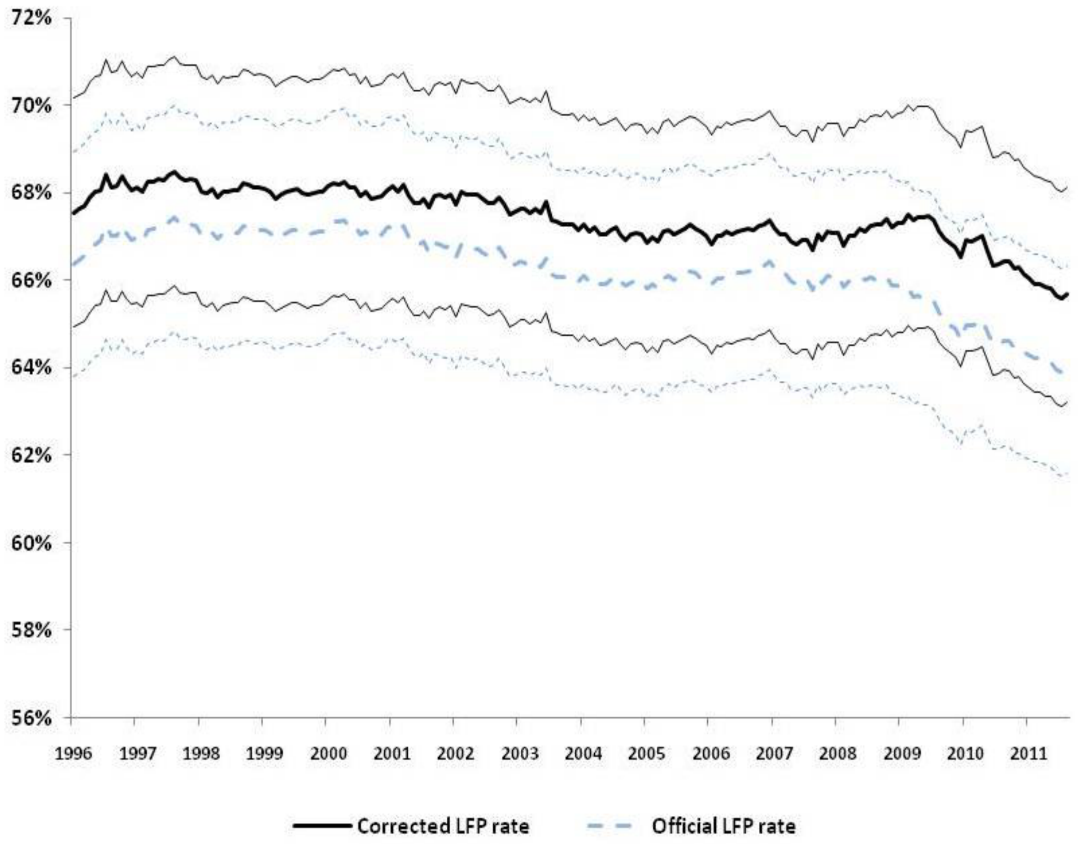
Note: Numbers reported in parentheses are bootstrapped standard errors based on 500 repetitions.

Table A21: Monthly corrected labor force participation rates (%)

| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1996 | 67.1 | 67.3 | 67.6 | 67.5 | 67.9 | 68.8 | 69.4 | 68.4 | 67.9 | 68.3 | 68.1 | 67.8 |
| | (1.33) | (1.33) | (1.33) | (1.33) | (1.34) | (1.36) | (1.37) | (1.34) | (1.33) | (1.34) | (1.34) | (1.33) |
| 1997 | 67.7 | 67.7 | 68.2 | 67.8 | 68.2 | 69.0 | 69.4 | 68.8 | 68.1 | 68.2 | 68.2 | 68.1 |
| | (1.33) | (1.33) | (1.34) | (1.33) | (1.34) | (1.35) | (1.36) | (1.34) | (1.33) | (1.33) | (1.33) | (1.33) |
| 1998 | 67.6 | 67.7 | 68.0 | 67.5 | 67.9 | 68.7 | 69.0 | 68.4 | 68.0 | 68.1 | 68.0 | 67.9 |
| | (1.32) | (1.32) | (1.32) | (1.31) | (1.32) | (1.34) | (1.34) | (1.33) | (1.32) | (1.32) | (1.32) | (1.32) |
| 1999 | 67.7 | 67.8 | 67.8 | 67.6 | 67.8 | 68.8 | 69.0 | 68.3 | 67.7 | 67.9 | 67.9 | 67.8 |
| | (1.31) | (1.32) | (1.32) | (1.31) | (1.32) | (1.34) | (1.34) | (1.32) | (1.31) | (1.31) | (1.31) | (1.31) |
| 2000 | 67.8 | 68.1 | 68.1 | 67.9 | 67.9 | 68.8 | 68.8 | 68.3 | 67.6 | 67.8 | 67.8 | 67.9 |
| | (1.31) | (1.32) | (1.32) | (1.31) | (1.31) | (1.33) | (1.33) | (1.32) | (1.31) | (1.31) | (1.31) | (1.31) |
| 2001 | 67.9 | 67.9 | 68.1 | 67.6 | 67.5 | 68.4 | 68.6 | 67.9 | 67.7 | 67.9 | 67.8 | 67.8 |
| | (1.31) | (1.31) | (1.31) | (1.30) | (1.30) | (1.32) | (1.32) | (1.31) | (1.30) | (1.31) | (1.31) | (1.30) |
| 2002 | 67.5 | 67.9 | 67.9 | 67.6 | 67.7 | 68.5 | 68.5 | 68.0 | 67.7 | 67.7 | 67.4 | 67.3 |
| | (1.30) | (1.31) | (1.31) | (1.30) | (1.30) | (1.32) | (1.32) | (1.31) | (1.30) | (1.30) | (1.30) | (1.29) |
| 2003 | 67.4 | 67.5 | 67.4 | 67.3 | 67.3 | 68.4 | 68.2 | 67.6 | 67.0 | 67.2 | 67.3 | 66.9 |
| | (1.29) | (1.29) | (1.29) | (1.29) | (1.29) | (1.31) | (1.30) | (1.29) | (1.28) | (1.28) | (1.28) | (1.28) |
| 2004 | 67.0 | 66.9 | 67.0 | 66.7 | 66.8 | 67.7 | 68.0 | 67.3 | 66.7 | 67.0 | 67.1 | 66.9 |
| | (1.28) | (1.28) | (1.28) | (1.27) | (1.28) | (1.29) | (1.30) | (1.28) | (1.27) | (1.28) | (1.28) | (1.27) |
| 2005 | 66.5 | 66.7 | 66.6 | 66.7 | 66.9 | 67.7 | 68.0 | 67.5 | 67.1 | 67.2 | 67.1 | 66.8 |
| | (1.27) | (1.27) | (1.27) | (1.27) | (1.28) | (1.29) | (1.29) | (1.28) | (1.28) | (1.28) | (1.28) | (1.27) |
| 2006 | 66.5 | 66.7 | 66.8 | 66.7 | 66.8 | 67.7 | 68.0 | 67.5 | 67.0 | 67.2 | 67.3 | 67.2 |
| | (1.27) | (1.27) | (1.27) | (1.27) | (1.27) | (1.29) | (1.29) | (1.28) | (1.27) | (1.28) | (1.28) | (1.28) |
| 2007 | 66.8 | 66.7 | 66.8 | 66.5 | 66.6 | 67.6 | 67.8 | 67.0 | 66.9 | 66.9 | 67.0 | 66.9 |
| | (1.27) | (1.27) | (1.27) | (1.26) | (1.27) | (1.28) | (1.29) | (1.27) | (1.27) | (1.27) | (1.27) | (1.27) |
| 2008 | 66.7 | 66.5 | 66.7 | 66.6 | 67.0 | 67.8 | 68.1 | 67.7 | 67.1 | 67.3 | 67.1 | 67.1 |
| | (1.27) | (1.26) | (1.27) | (1.26) | (1.27) | (1.29) | (1.29) | (1.28) | (1.27) | (1.28) | (1.28) | (1.27) |
| 2009 | 67.0 | 67.2 | 67.1 | 67.1 | 67.3 | 68.1 | 68.2 | 67.5 | 66.8 | 66.7 | 66.6 | 66.3 |
| | (1.28) | (1.28) | (1.28) | (1.28) | (1.28) | (1.30) | (1.30) | (1.29) | (1.28) | (1.28) | (1.27) | (1.26) |
| 2010 | 66.6 | 66.6 | 66.7 | 66.7 | 66.6 | 67.0 | 67.2 | 66.8 | 66.3 | 66.1 | 66.1 | 65.8 |
| | (1.27) | (1.27) | (1.27) | (1.27) | (1.27) | (1.28) | (1.28) | (1.27) | (1.27) | (1.26) | (1.26) | (1.25) |
| 2011 | 65.7 | 65.6 | 65.7 | 65.5 | 65.7 | 66.3 | 66.4 | 66.1 | | | | |
| | (1.25) | (1.25) | (1.25) | (1.25) | (1.25) | (1.26) | (1.26) | (1.26) | | | | |

Note: Not seasonally adjusted. Numbers reported in parentheses are bootstrapped standard errors.

Figure A5: Corrected and Official (Reported) Labor Force Participation Rates




Note: Figure displays corrected and official (reported) Labor Force Participation (LFP) rates for the whole population (seasonally adjusted) from January 1996 to August 2011. The thin lines signify 95% upper and lower confidence bounds.

# References

P.P. Biemer and J.M. Bushery. On the validity of markov latent class analysis for estimating classification error in labor force data. *Survey Methodology*, 26(2):139–152, 2000.

J.M. Poterba and L.H. Summers. Reporting errors and labor market dynamics. *Econometrica*, 54(6):1319–1338, 1986.

M.D. Sinclair and J.L. Gastwirth. On procedures for evaluating the effectiveness of reinterview survey methods: application to labor force data. *Journal of the American Statistical Association*, 91(435):961–969, 1996.

M.D. Sinclair and J.L. Gastwirth. Estimates of the errors in classification in the labour force survey and their effect on the reported unemployment rate. *Survey methodology*, 24(2):157–169, 1998.