Misclassification-Errors-Adjusted Sahm Rule for Early Identification of Economic Recession

Shuaizhang Feng
Jiandong Sun

Working Paper

2020-029

05/2020

Misclassification-errors-adjusted Sahm Rule for Early Identification of Economic Recession

Shuaizhang Feng* Jiandong Sun†

April 19, 2020

Abstract

Accurate identification of economic recessions in a timely fashion is a major macroeconomic challenge. The most successful early detector of recessions, the Sahm rule, relies on changes in unemployment rates, and is thus subject to measurement errors in the U.S. labor force statuses based on survey data. We propose a novel misclassification-error-adjusted Sahm recession indicator and provide empirically-based optimal threshold values. Using historical data, we show that the adjusted Sahm rule offers earlier identification of economic recessions. Based on the newly released U.S. unemployment rate in March 2020, our adjusted Sahm rule diagnoses the U.S. economy is already in recession, while the original Sahm rule does not.

Key Words: *Economic recession, Sahm rule, Misclassification errors, Unemployment rate*

*Corresponding author. School of Economics and Institute for Economic and Social Research, Jinan University, Guangzhou 510632, China. E-mail: shuaizhang.feng@jnu.edu.cn. Feng's research was supported by National Science Fund for Distinguished Young Scholars (Project Number: 71425005).

†Institute for Economic and Social Research, Jinan University, Guangzhou 510632, China. E-mail: jiandongsun11@gmail.com

1 Introduction

The novel coronavirus (Covid-19) pandemic has been ripping through America not just with skyrocketing numbers of confirmed cases and deaths, but also its disruptive power on the U.S. economy. While most think an imminent economic recession is unavoidable, and many suspect that the U.S. economy is already in one, it is another thing to scientifically identify it. In the U.S., according to the National Bureau of Economic Research (NBER), a recession refers to a significant decline in economic activity, lasting more than a few months, normally visible in real gross domestic product, real income, employment, industrial production, and wholesale-retail sales.[1] An economic recession, which lasts from the peak and the subsequent trough of a business cycle, is officially declared by NBER. However, the NBER's approach to determining the dates of turning points (peak and trough) is retrospective, and heavily relies on a set of economic indicators, which are not real-time and always need to be revised for several times.[2] As a result, it would usually take the NBER several months to identify a recession after it has already occurred. Considering that recession may bring enormous damage to the economy, it is important for policymakers to initiate prompt and efficient monetary or fiscal policies to reduce its negative effects. Although the NBER's procedure guarantees relatively precise dates of recessions, it is too slow for policymaking.

To be sure, researchers have been always using economic data to predict economic activities and trying to identify recessions as soon as they can. Asset price changes, in particularly the yield spread, have been mostly frequently used in this endeavor.[3] Other leading predictors include interest rate spreads (Stock and Watson 1989), gross domestic product and gross domestic income (Nalewaik 2012), the Conference Board Leading Economic Index (Lahiri and Yang 2015; Levanon et al. 2015) and so on. Methodologically, it has been found that forecasting a binary recession indicator with binary models is more stable than forecasting output growth with continuous models (Estrella et al. 2003). Recent studies employ Markovian models (Chauvet and Hamilton 2006) or dynamic binary models (Kauppi and Saikkonen 2008). Tian and Shen (2019) show that the Markovian models outperform the Probit models in detecting a recession. In addition, in light of the need of timely policy decisions, there has been increasing interest in assessing current economic conditions or predicting the very near future (Chauvet and Piger 2008; Chen et al. 2015; Giannone et al. 2008; Hamilton 2011).

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

[2] See https://www.nber.org/cycles/jan08bcdc_memo.html.

[3] See Stock and Watson (2003) for more detailed discussion. See also in Ang et al. (2006); Estrella (2005); Estrella and Hardouvelis (1991); Estrella and Mishkin (1998); Rudebusch and Williams (2009).

More recently, Sahm (2019) proposes a so-called "Sahm rule" to predict recessions, that is, if the three-months moving average of the national unemployment rate (U-3) rises 0.5 percentage points or more relative to its low during the previous 12 months, then the economy is already in a recession.[4] This rule correctly signals a recession in 2-4 months after it actually occurred since 1970, which represents a significant time saving compared to NBER. The Sahm rule also compares favorably with other prediction methodologies in terms of accuracy and timeliness, as well as simplicity, as it does not invoke sophisticated econometrics methods.

However, the official unemployment rate used to calculate the Sahm recession indicator, the difference between three-months moving average of unemployment rate and its previous 12 months low, is subject to the well-known issue of misclassification in labor force status (LFS) (Abowd and Zellner 1985; Poterba and Summers 1986). Feng and Hu (2013) show that ignoring misclassification errors in LFS leads to substantial underestimation of the unemployment rate. More importantly, they show that the degree of underestimation is larger when the level of unemployment is higher. In this sense, it is possible that at the very beginning of a recession, the rise in the official unemployment rate, which is subject to misclassification errors, is less than the increase in the underlying true unemployment rate. This may delay the signal of recessions and weaken the predictive timeliness of the Sahm rule.

In this paper, we examine the robustness of the Sahm recession indicator to misclassification errors in LFS. Previous studies have widely discussed the issue of misclassification errors in LFS, which arises from the intrinsic difficulties in classifying LFS of some specific groups of people, like marginally-attached worker and involuntary part-time workers, as well as other practical challenges in classifying LFS with survey data. Using a latent variable approach, Feng and Hu (2013) correct for misclassification errors and estimate the corrected true unemployment rate. We apply their method and use the corrected unemployment statistics to re-calculate the Sahm recession indicator. We find that for the historical recessions since the late 1970s, our corrected recession indicator always rises more promptly than the original Sahm recession indicator after the onset of a recession. We also propose optimal threshold values for the identification of recessions and show that our misclassification-error-adjusted approach improves predictive timeliness of the Sahm rule.

[4] See also in https://fred.stlouisfed.org/series/SAHMREALTIME.

2 Methods

We use the method proposed by Hu (2008) and used in Feng and Hu (2013) to correct for misclassification errors in unemployment rate.[5] According to the 4-8-4 rotational group structure of the Current Population Survey (CPS), suppose we observe an i.i.d. sample of self-reported labor status U for three periods $\{U_{t+1}, U_t, U_{t-9}\}_i$ for individual i. For example, if U_t stands for one's LFS in January 2020, then U_{t+1} and U_{t-9} denote his or her LFS in February 2020 and in April 2019, respectively. Although each person appears eight times in CPS, we choose these three-months data $(t+1, t, t-9)$ for three reasons: (i) we want the three months to be close enough to minimize sample attrition; (ii) we want the three months to cover the eight-months break in the 4-8-4 rotation structure of CPS to ensure that there are enough variations in the LFS; (iii) the assumption regarding the dynamics of latent true LFS (Assumption 2 below) is more likely to be satisfied if we use the data reported a while ago, e.g., nine months earlier.

We assume that the latent true LFS U_t^* has the same support as the self-reported LFS U_t as follows:

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

Let $f(\cdot)$ stand for probability mass functions of its arguments. Let $\Omega_{\neq t}$ denote all the variables in all the periods except period t, i.e., $\Omega_{\neq t} = \{(U_\tau, U_\tau^*), \tau \neq t\}$. We assume that the misclassification errors distribution satisfies a local independence assumption as follows:

Assumption 1 $f(U_t|U_t^*, \Omega_{\neq t}) = f(U_t|U_t^*)$.

This assumption implies that misclassification errors may be correlated with the true LFS, and correlated with all other variables only through the true LFS. In addition, we simplify the dynamics of the latent true LFS as follows:

Assumption 2 $f\left(U_{t+1}^*|U_t^*, U_{t-9}^*\right) = f\left(U_{t+1}^*|U_t^*\right)$.

This assumption implies that the true LFS in period $t-9$ has no prediction power on the true LFS in the period $t+1$ beyond the true LFS in the current period t. Under Assumption 1 and 2, the

[5]See Feng and Hu (2013) for more technical details.

relationship between observed probabilities and unobserved ones is as follows:

$$f\left(U_{t+1}, U_t, U_{t-9}\right) = \sum_{U_t^*} f\left(U_{t+1}|U_t^*\right) f\left(U_t|U_t^*\right) f\left(U_t^*, U_{t-9}\right). \tag{1}$$

By integrating U_{t+1} out, we obtain

$$f\left(U_t, U_{t-9}\right) = \sum_{U_t^*} f\left(U_t|U_t^*\right) f\left(U_t^*, U_{t-9}\right). \tag{2}$$

We then use the identification results in Hu (2008) to show that all the unobservables on the RHS of Equation (1) may be identified. Define $M_{U_t|U_t^*} \equiv \left[f_{U_t|U_t^*}\left(i|j\right)\right]_{i,j}$, $M_{U_t^*,U_{t-9}} \equiv \left[f_{U_t^*,U_{t-9}}\left(j,k\right)\right]_{j,k}$, $M_{1,U_t,U_{t-9}} \equiv \left[f_{U_{t+1},U_t,U_{t-9}}\left(1,i,k\right)\right]_{i,k}$, and $D_{1|U_t^*} \equiv diag\left[f_{U_{t+1}|U_t^*}\left(1|j\right)\right]_j$. We can show that Equation (1) and (2) are equivalent to

$$M_{1,U_t,U_{t-9}} = M_{U_t|U_t^*} D_{1|U_t^*} M_{U_t^*,U_{t-9}}, \tag{3}$$

and

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

To solve the unknown misclassification probabilities, we need following technical assumption:

Assumption 3 *Matrix $M_{U_t,U_{t-9}}$ is invertible.*

This assumption is testable, as it is imposed on observed probabilities. Under Assumption 3, we can derive following equation by eliminating $M_{U_t^*,U_{t-9}}$ in Equation (3) and (4):

$$M_{1,U_t,U_{t-9}} M_{U_t,U_{t-9}}^{-1} = M_{U_t|U_t^*} D_{1|U_t^*} M_{U_t|U_t^*}^{-1}. \tag{5}$$

This implies that the observed matrix on the LHS of Equation (5) has an eigenvalue-eigenvector decomposition on the RHS. In order to identify a unique decomposition, we need the following two additional assumptions:

Assumption 4 $f_{U_{t+1}|U_t^*}\left(1|k\right)$ *are different for a different k.*

Assumption 5 $f_{U_t|U_t^*}\left(k|k\right) > f_{U_t|U_t^*}\left(j|k\right)$ *for $j \neq k$.*

Assumption 4 can be also tested directly, as $f_{U_{t+1}|U_t^*}(1|k)$ for $k \in \{1, 2, 3\}$ are eigenvalues of observed matrix $M_{1,U_t,U_{t-9}} M_{U_t,U_{t-9}}^{-1}$. Assumption 5 implies that people are more likely to report their true LFS than any other possible status. These two assumptions guarantee that the eigenvalues are distinctive and that the eigenvectors can be ordered by the value of true labor force status.

3 Data

We use the public-use monthly Current Population Surveys datasets from January 1978 to February 2020 to estimate misclassification probabilities. Because of the 4-8-4 rotational group structure, the monthly CPS files can be matched to form longitudinal panels, which enables us to obtain the joint distribution of self-reported LFS in three periods. The matching method in this paper is the same as Feng and Hu (2013). We first follow the algorithm proposed by Madrian and Lefgren (2000) to match CPS monthly files and adjust for sample attrition for the matched files. We then pool different periods of matched data together to increase sample sizes. Specifically, the misclassification probabilities for period t, $f_{U_t|U_t^*}$, is estimated based on pooled matched samples from period $t-60$ to $t-1$, as in Feng et al. (2018).

We use seasonally-adjusted unemployment rate and labor force participation rate to calculate the reported LFS distribution f_{U_t}, which are officially released by the U.S. Bureau of Labor Statistics. Since we focus on predictive timeliness, we will use real-time labor force statistics available in a given month,[6] which can be retrieved from the Archival Federal Reserve Economic Data, Federal Reserve Bank of St. Louis.[7]

Finally, given that we may have identified the misclassification probabilities $f_{U_t|U_t^*}$ in $M_{U_t|U_t^*}$ and obtained the observed distribution f_{U_t}, we may then identify the distribution of the latent true LFS $f_{U_t^*}$ from following equation,

$$f_{U_t} = \sum_{U_t^*} f_{U_t|U_t^*} f_{U_t^*}, \tag{6}$$

and therefore, the underlying true unemployment rate.

[6]The seasonal factors for labor force statistics will be re-estimated at the end of each calendar year, so such annual adjustment may update estimates in recent years. See more details in https://www.bls.gov/web/empsit/cps-seas-adjustment-methodology.pdf.

[7]See https://alfred.stlouisfed.org/release?rid=50. Note that the earliest vintage for labor force participation rate is 1997, so we use this vintage for labor force participation rate series during 1979-1996.

4 Results

Figure 1 shows the seasonally-adjusted monthly reported and corrected unemployment rates, as well as the NBER-defined periods of economic recession. The reported values are directly from the BLS, and the corrected ones are calculated using the latent variable approach outlined in the previous sections. The shaded areas indicate economic recessions as per the definition of NBER, which is the period between peak month (included) and subsequent trough month. It is clearly shown that during recessions, which generally post higher levels of unemployment, the differences between the corrected and reported unemployment rates are also much bigger than otherwise. This suggests that rises in the Sahm recession indicator might have been suppressed and not truly reflecting changes in unemployment rates.

We then compare the real-time Sahm recession indicators based on both reported unemployment rates and our corrected ones. The results are shown in Figure 2. We find that during the recessions, the indicator based on the corrected unemployment rate are higher than that based on the reported unemployment rate, while in expansions, the two indicators almost coincide. In addition, when recession is coming, the indicator based on our corrected unemployment rate always rises ahead of the original Sahm's indicator.

We next turn to the issue of optimal threshold values for the identification of a recession. To do so, we consider the predictive performance in recession and expansion periods, respectively. For each possible threshold value x, we define the loss function as follows:

$$Loss_i(x) = \frac{1}{T_i} \sum_t L\left(R_t, I\left(s_t \geq x\right)\right) \tag{7}$$

where i is an business cycle status, 1 for recession periods, 0 for expansion periods, T_i is the total number of months for each status, R_t is equal to 1 if current month t is in recession as defined by the NBER, otherwise 0, and $I\left(\cdot\right)$ is an indicator function, equal to 1 when indicator s_t hits the threshold x. In the recession periods,

$$L\left(1, I\left(s_t \geq x\right)\right) = \begin{cases} 1, \; I\left(s_t \geq x\right) \neq 1 \\ 0, \; I\left(s_t \geq x\right) = 1 \end{cases}$$

while in the expansion periods,

$$L\left(, I\left(s_t \geq x\right)\right)=\begin{cases}1, & I\left(s_t \geq x\right) \neq 0 \\ 0, & I\left(s_t \geq x\right)=0\end{cases}$$

Therefore, the loss value represents the proportion of wrong "guesses" for each business cycle periods. There is a trade-off in determining an optimal threshold, that is, the loss value would increase with the threshold in recessions but decrease in expansions. As the threshold increases, we are more likely to reject false alarms during the expansion periods, but are less likely to promptly identify the starting date of a recession when it comes.

Table 1 shows the loss value of the original and corrected recession indicators under each threshold, both in recession and expansion periods. In recessions, the loss values of our corrected indicator are less than those of the original Sahm indicator for any fixed threshold. This shows that the corrected indicator is better in identifying recessions. To determine the optimal threshold, we follow Claudia Sahm's original insight and choose a value that is large enough to avoid false alarms, i.e., we choose the lowest value that would give zero loss value during the expansion periods. The idea is that we can tolerate some (unavoidable) delayed identification of true recessions but we will exclude any false claims in order not to confuse the two. By this standard, according to Table 1, the optimal threshold value would be 0.47 for the original Sahm recession indicator and 0.53 for our corrected indicator.

The original threshold used by Sahm is 0.5, not 0.47. Nevertheless, given the available historical data, 0.5 and 0.47 does not make any difference as the loss values are the same during both the recession and expansion periods. According to the original Sahm rule, with the threshold of 0.5, 43.6% of the recession months are incorrectly classified as not in recession, which would generally result in delayed identification. In comparison, the probability of incorrect classification is only 12.8% during the recession periods using our new misclassification-error-adjusted Sahm rule with the threshold of 0.53. Both methods would not give any false alarms when the economy is actually in expansion.

Figure 3 compares the timeliness of the original Sahm recession indicator with our corrected one over the previous five U.S. recessions. In general, it would take the NBER from half a year to one year to precisely identify the peak or trough of a recession. The original Sahm recession indicator identifies all the recessions within four months after they have already begun, which significantly improves upon NBER. Our corrected indicator outperforms the original one, and further substantially improves the timeliness of prediction in all cases. For example, for the last recession that began in December 2007, the NBER only announced it one year later in December 2008. The original Sahm indicator identifies the recessions in April 2008 with only four months lag, while our corrected indicator identifies it in December 2007, as soon as the employment data for the month were released.

Finally, we apply our adjusted Sahm rule to the newly released official U.S. unemployment data for March 2020. Figure 1 shows that the official unemployment rate was 4.4, up from 3.5 in February, while our correct unemployment rate went from 5.3 to 6.9. The original Sahm recession indicator is 0.3, which would not qualify as recession using either Sahm's original cutoff point of 0.5, or our optimal cutoff point of 0.47 (Figure 2). However, the adjusted Sahm recession indicator, based on the misclassification-error-adjusted unemployment rates, is 0.54, which is just above the optimal threshold of 0.53. Therefore, while the original Sahm rule would not identify the U.S. economy as already in a recession, our adjusted Sahm rule would. In addition, Table 2 shows the probabilities of a recession using the original real-time Sahm recession indicator and our misclassification-errors-adjusted one.[8] From a historical view, given that the original Sahm recession indicator in March 2020 is 0.3, the probability of a recession now is only 50%, even within 12 months, while our adjusted recession indicator is 0.54, indicating the U.S. economy is already in recession.

5 Conclusion

This paper examines the robustness of Sahm recession indicator to misclassification errors in LFS. We employ the latent variable approach used in Feng and Hu (2013) to correct for bias in unemployment rate due to misclassification errors, and re-calculate the Sahm recession indicator based on corrected unemployment rate. We find that bias in unemployment rate due to misclassification errors does affect the predictive timeliness of Sahm recession indicator. We then propose a more proper threshold for our adjusted indicator. Using historical records, our misclassification-errors adjustment substantially improves the predictive timeliness of Sahm's recession indicator. Based on the newly released U.S. unemployment rate in March 2020, our adjusted Sahm rule detects that the U.S. economy is already in recession, while the original one does not.

[8] Table 2 replicates the results in https://www.brookings.edu/blog/up-front/2019/06/06/how-will-we-know-when-a-recession-is-coming.

References

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

Ang, A., M. Piazzesi, and M. Wei (2006). What does the yield curve tell us about GDP growth? *Journal of Econometrics 131*(1-2), 359–403.

Chauvet, M. and J. D. Hamilton (2006). Dating business cycle turning points. In C. Milas, P. Rothman, and D. van Dijk (Eds.), *Nonlinear Time Series Analysis of Business Cycles*, pp. 1–54. The Netherlands: Elsevier.

Chauvet, M. and J. Piger (2008). A comparison of the real-time performance of business cycle dating methods. *Journal of Business & Economic Statistics 26*(1), 42–49.

Chen, T., E. P. K. So, L. Wu, and I. K. M. Yan (2015). The 2007–2008 US recession: What did the real-time Google Trends data tell the United States? *Contemporary Economic Policy 33*(2), 395–403.

Estrella, A. (2005). Why does the yield curve predict output and inflation? *The Economic Journal 115*(505), 722–744.

Estrella, A. and G. A. Hardouvelis (1991). The term structure as a predictor of real economic activity. *The Journal of Finance 46*(2), 555–576.

Estrella, A. and F. S. Mishkin (1998). Predicting US recessions: Financial variables as leading indicators. *Review of Economics and Statistics 80*(1), 45–61.

Estrella, A., A. P. Rodrigues, and S. Schich (2003). How stable is the predictive power of the yield curve? Evidence from Germany and the United States. *Review of Economics and Statistics 85*(3), 629–644.

Feng, S. and Y. Hu (2013). Misclassification errors and the underestimation of the US unemployment rate. *The American Economic Review 103*(2), 1054–1070.

Feng, S., Y. Hu, and J. Sun (2018). On the robustness of alternative unemployment measures. *Economics Letters 166*, 1–5.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics 55*(4), 665–676.

Hamilton, J. D. (2011). Calling recessions in real time. *International Journal of Forecasting 27*(4), 1006–1026.

Hu, Y. (2008). Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution. *Journal of Econometrics 144*(1), 27–61.

Kauppi, H. and P. Saikkonen (2008). Predicting US recessions with dynamic binary response models. *The Review of Economics and Statistics 90*(4), 777–791.

Lahiri, K. and L. Yang (2015). A further analysis of the conference board's new Leading Economic Index. *International Journal of Forecasting 31*(2), 446–453.

Levanon, G., J.-C. Manini, A. Ozyildirim, B. Schaitkin, and J. Tanchua (2015). Using financial indicators to predict turning points in the business cycle: The case of the leading economic index for the United States. *International Journal of Forecasting 31*(2), 426–445.

Madrian, B. C. and L. J. Lefgren (2000). An approach to longitudinally matching Current Population Survey (CPS) respondents. *Journal of Economic and Social Measurement 26*(1), 31–62.

Nalewaik, J. J. (2012). Estimating probabilities of recession in real time using GDP and GDI. *Journal of Money, Credit and Banking 44*(1), 235–253.

Poterba, J. M. and L. H. Summers (1986). Reporting errors and labor market dynamics. *Econometrica: Journal of the Econometric Society*, 1319–1338.

Rudebusch, G. D. and J. C. Williams (2009). Forecasting recessions: the puzzle of the enduring power of the yield curve. *Journal of Business & Economic Statistics 27*(4), 492–503.

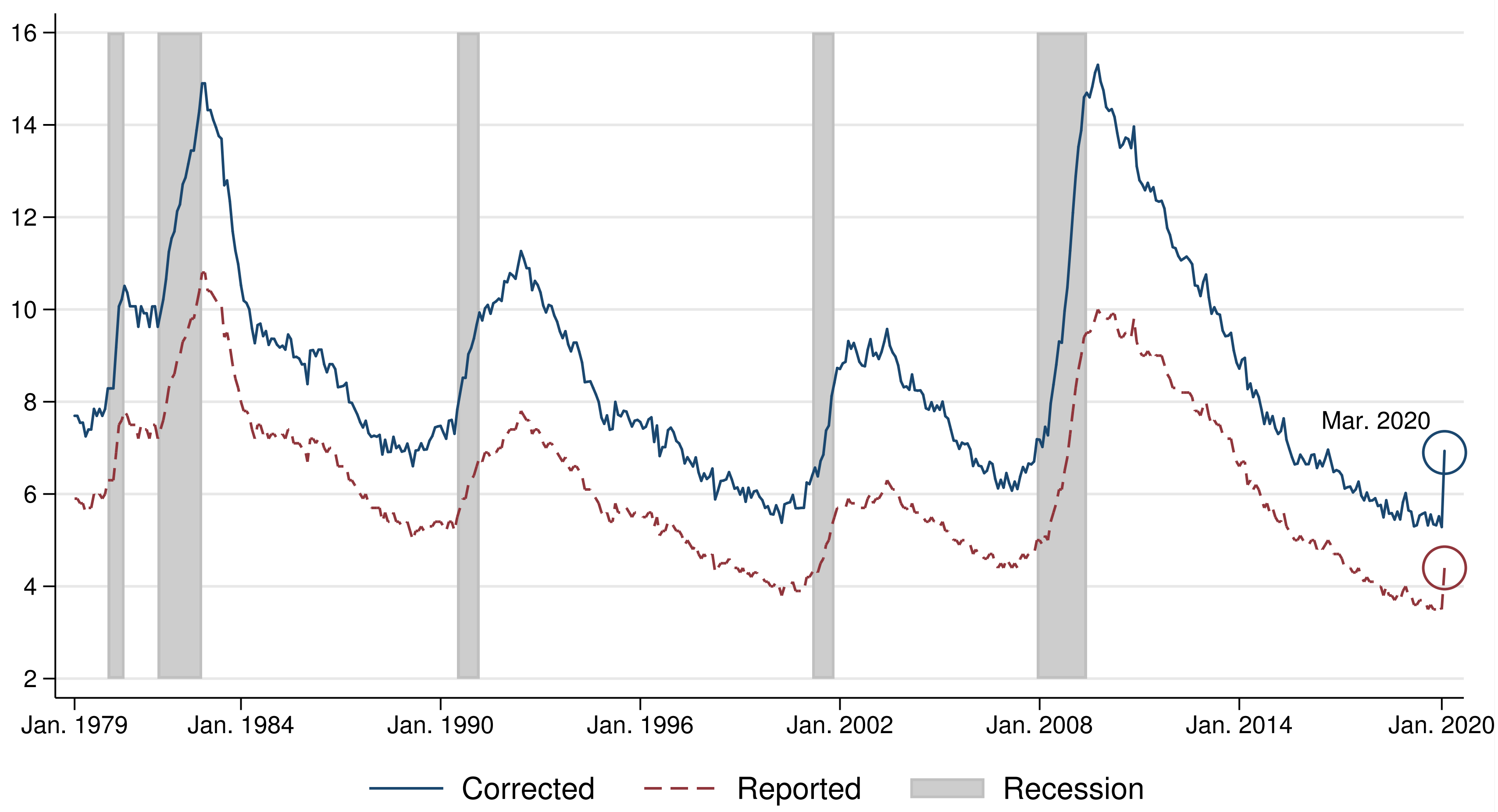
Sahm, C. (2019). Direct Stimulus Payments to Individuals. In H. Boushey, R. Nunn, and J. Shambaugh (Eds.), *Recession Ready: Fiscal Policies to Stabilize the American Economy*, pp. 67–92. Washington, DC: The Hamilton Project and the Washington Center on Equitable Growth.

Stock, J. H. and M. W. Watson (1989). New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual 4*, 351–394.

Stock, J. H. and M. W. Watson (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature 41*(3), 788–829.

Tian, R. and G. Shen (2019). Predictive power of Markovian models: Evidence from US recession forecasting. *Journal of Forecasting 38*(6), 525–551.

Figure 1: Reported and corrected unemployemnt rates



Note: The reported and corrected unemployment rates are based on current vintage. All the series are seasonally adjusted.

Figure 2: Real-time Sahm recession indicator: reported v.s. corrected

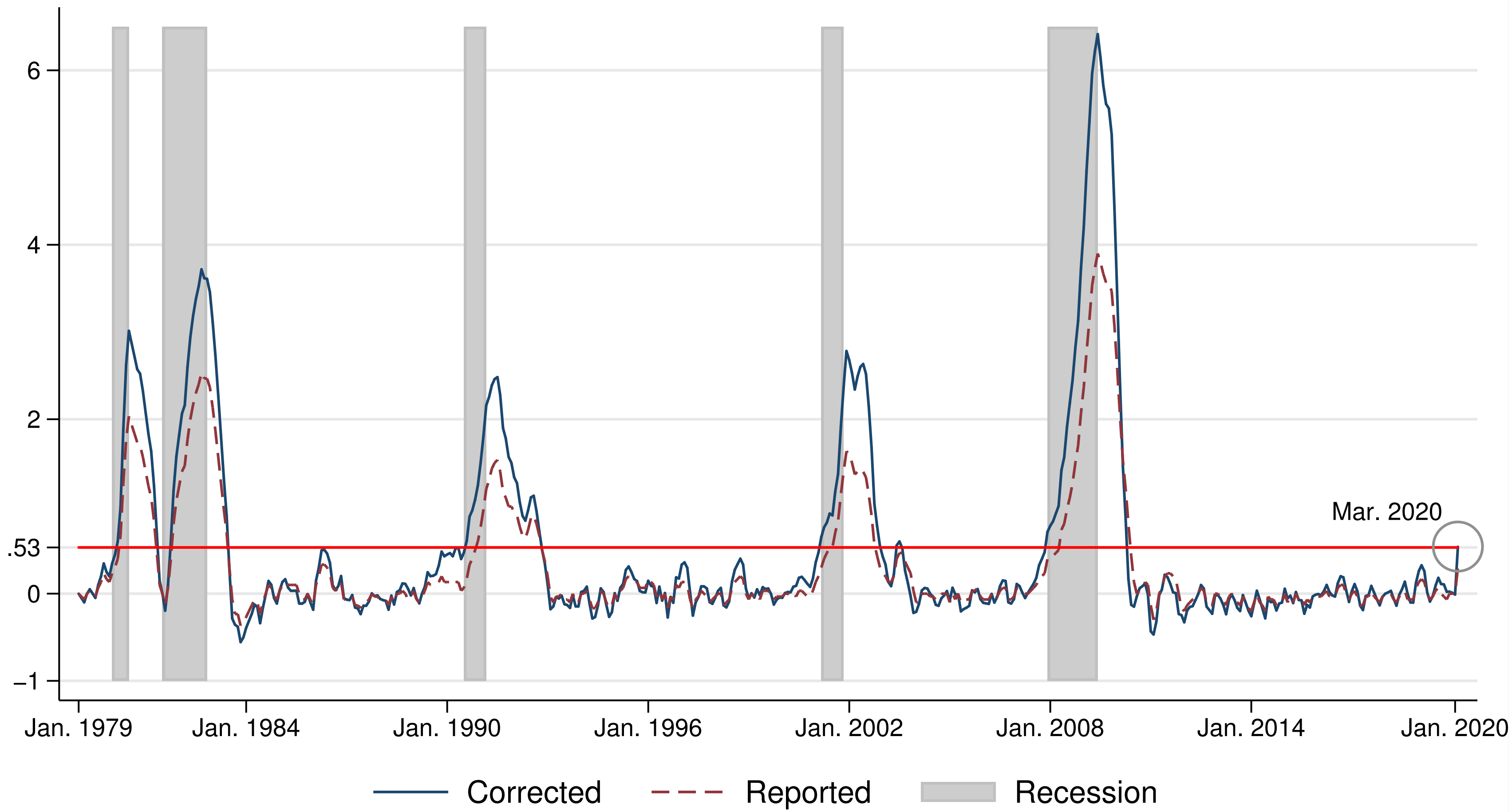
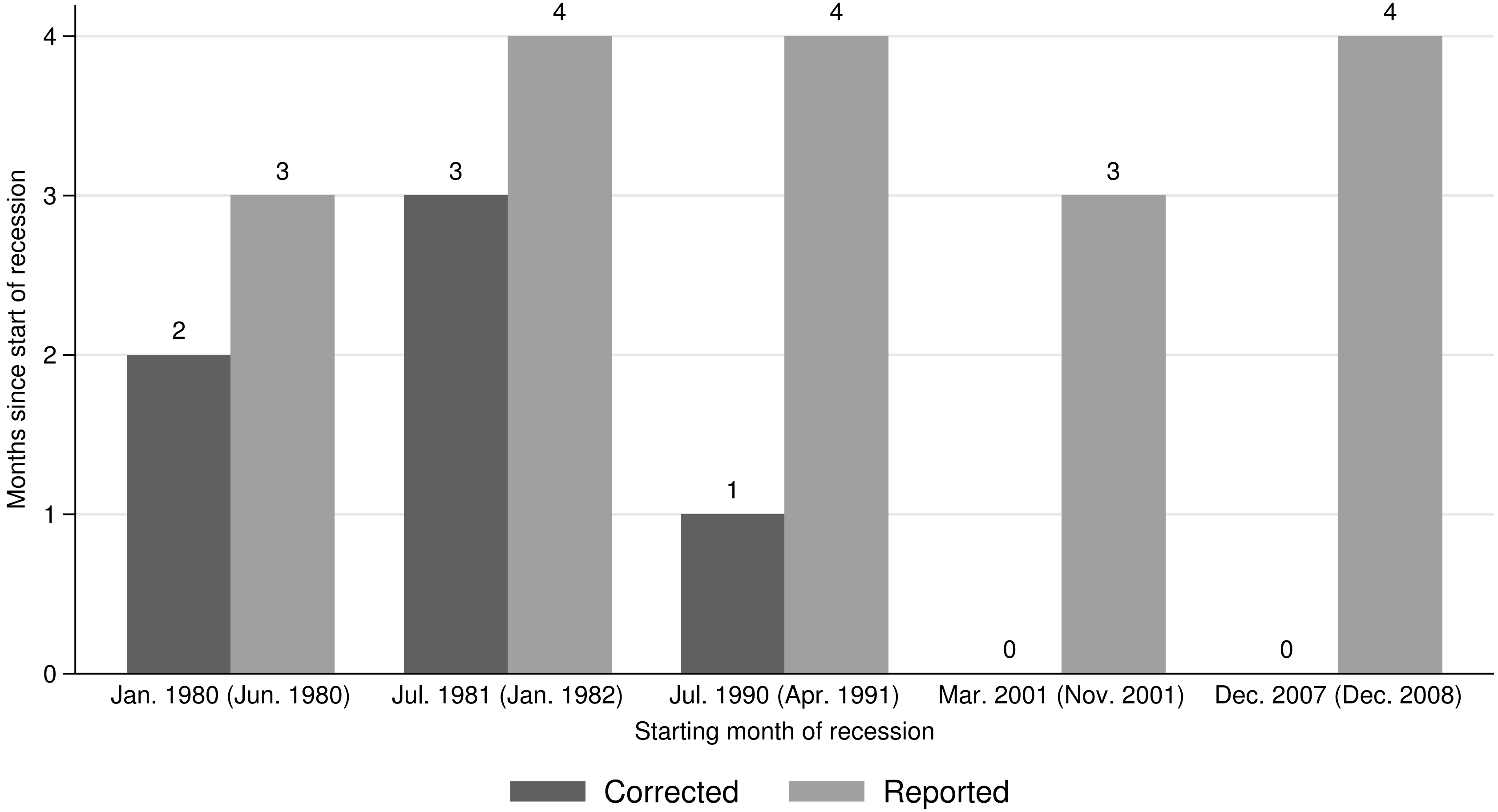


Figure 3: Months that Sahm recession indicator triggers relative to the start of recessions



Note: In parentheses is the NBER's announcing month of a recession.

Table 1: Loss value (in percentage) for each threshold

Threshold	Recession		Non-recession	
	Reported	Corrected	Reported	Corrected
0	2.6	2.6	50.6	47.8
0.05	2.6	2.6	25.4	34.4
0.1	7.7	5.1	16.5	23.2
0.15	7.7	5.1	8.4	16.2
0.2	10.3	5.1	4.2	11.7
0.25	12.8	5.1	1.7	9.5
0.3	15.4	5.1	1.1	7.5
0.35	20.5	7.7	0.6	5.3
0.4	23.1	7.7	0.3	3.9
0.45	38.5	10.3	0.3	2.8
0.46	38.5	10.3	0.3	2.5
0.47	**43.6**	10.3	**0.0**	2.2
0.48	43.6	12.8	0.0	1.7
0.49	43.6	12.8	0.0	1.4
0.5	43.6	12.8	0.0	1.1
0.5	43.6	12.8	0.0	1.1
0.51	48.7	12.8	0.0	0.8
0.52	48.7	12.8	0.0	0.3
0.53	48.7	**12.8**	0.0	**0.0**
0.54	53.8	12.8	0.0	0.0
0.55	53.8	12.8	0.0	0.0
0.6	56.4	17.9	0.0	0.0
0.65	56.4	23.1	0.0	0.0
0.7	61.5	23.1	0.0	0.0
0.75	64.1	25.6	0.0	0.0
0.8	66.7	30.8	0.0	0.0

Note: Numbers calculated based on data from January 1979 to February 2020. Months between the NBER's announcing dates of the peak and the subsequent trough are excluded (97 out of 494), as in those months the economy is still considered in recessions.

Table 2: Probabilities (in percentage) of recession by Sahm recession indicator, January 1979 to February 2020

Range of indicator	Recession now	Recession in 1 month	Recession in 3 months	Recession in 6 months	Recession 12 months	Number of months
Panel A: the original Sahm recession indicator						
< 0	0.6	0.6	0.6	2.2	5.1	178
$[0, 0.1)$	1.6	2.4	3.2	7.3	20.2	124
$[0.1, 0.2)$	2.2	6.7	20.0	28.9	33.3	45
$[0.2, 0.3)$	15.4	23.1	30.8	30.8	38.5	13
$[0.3, 0.4)$	50.0	50.0	50.0	50.0	50.0	6
$[0.4, 0.5)$	88.9	88.9	88.9	88.9	88.9	9
≥ 0.5	100.0	100.0	100.0	100.0	100.0	22
Panel B: the corrected recession indicator						
< 0	0.5	0.5	0.5	1.1	3.7	188
$[0, 0.1)$	1.1	1.1	1.1	4.5	15.7	89
$[0.1, 0.2)$	0.0	2.4	4.9	17.1	24.4	41
$[0.2, 0.3)$	0.0	0.0	6.7	6.7	26.7	15
$[0.3, 0.4)$	7.1	14.3	42.9	42.9	50.0	14
$[0.4, 0.53)$	12.5	25.0	37.5	56.3	68.8	16
≥ 0.53	100.0	100.0	100.0	100.0	100.0	34

Note: Each entry means the probability of a recession (now or in the near future) under each range of recession indicator. Months between the NBER's announcing dates of the peak and the subsequent trough are excluded (97 months out of 494), as in those months the economy is still considered in recessions. We also exclude months before 1979, as there were no formal announcements of business cycle turning points prior 1979.