Forecasting the U.S. Unemployment Rate: Another Look

Han Xiao, Rong Chen and John Guerard*

Abstract. Twenty years ago, an eminent group of economic forecasters and statisticians published a seminal work of forecasting the U.S. unemployment rate, certainly one of the most important economic measures of the U.S. economy. Montgomery, Zarnowitz, Tiao, and Tsay (MZTT, 1998) reported that linear and nonlinear time series models are useful in predicting unemployment rates relatively accurately, in both short term (1 month) and medium term (5 months) prediction. One interesting and important finding was that the weekly unemployment claims was a statistically significant input in forecasting the U.S. unemployment rate over the 1959 – 1993 time period. The weekly unemployment claims time series is a component of the U.S. Leading Economic Indicators (LEI). In this paper we replicated and extended the MZTT analysis for the 1959 to 2019 time period. We report out-of-sample one-step to twelve-step ahead monthly prediction performance of various models for the 1990-2019 period, using a no-change (random walk) model as a forecasting benchmark. Results obtained from this study include: (1) weekly unemployment claims are indeed a useful and statistically significant input in a transfer function model to forecast the unemployment rate; (2) the leading economic indicators time series is a statistically significant input in a transfer function model to forecast the unemployment rate; (3) a seasonal ARIMA (SARIMA) model outperforms the no-change benchmark for all forecasting horizons; (4) the SARIMA and transfer function models are statistically significantly better forecasting models than a null, or no-change, forecast, particularly in the Global Financial Crisis (GFC), 2008 -2019 time period. Improved upon MZTT, this paper provides a set of analysis that serves as an updated benchmark for comparison of forecasting methods and approaches on unemployment prediction.

Keywords: Time series forecasting; transfer function modeling; the unemployment rate

*Han Xiao is Associate Professor, Department of Statistics, Rutgers University, Piscataway, NJ 08854. E-mail: hxiao@stat.rutgers.edu. Rong Chen is Professor, Department of Statistics, Rutgers University, Piscataway, NJ 08854. E-mail: rongchen@stat.rutgers.edu. John Guerard is on the Scientific Advisory Board of McKinley Capital Management, LLC Anchorage, AK, and an Affliate Instructor at University of Washington, Seattle, WA. Email: jguerard@mckinleycapital.com. John Guerard is the corresponding author. Xiao's research is supported in part by National Science Foundation grant DMS-1454817, DMS-2027855, and a research grant from NEC Labs America. Chen's research is supported in part by National Science Foundation grants DMS-1737857, IIS-1741390 and DMS-2027855.

1 Introduction

The unemployment rate is one of the most important measures of an economy and has been extensively studied since early twentieth century. The unemployment rate has been linked to traditional business cycles (Burns and Mitchell, 1946; Harberler, 1937; Keynes, 1936; Klein, 1950; Matthews, 1959; Mitchell, 1913; Mitchell and Gay, 1927; Moore, 1961; Persons, 1931; Samuelson, 1948; Zarnowitz, 1992). It was generally agreed that during an expansion, employment, production, prices, money, wages, interest rates, and profits are usually rising, while the reverse occurs during a contraction.

The impact of business cycle on unemployment is widely recognized. Rapid increases in the unemployment rate were often accompanied by recessions declared by the NBER business cycle committee. Thus, the unemployment rate was often studied in terms of the (traditional) business cycle. Economists have also studied Keynes's general theory of the trade cycle and unemployment, including Brunner and Meltzer (1993); Hahn and Solow (1995); Harberler (1937); Keynes (1936) and Lucas (1983), mainly considering the trade cycle as being one of under-consumption. It is noted that an important influence on employment can be a reduction in wages which is accompanied by decreases in prices and incomes, which will diminish the need for cash. Two final distinctions in the Keynesian trade cycle are theories of prolonged associated unemployment with voluntary and involuntary unemployment, which decomposed unemployment into voluntary, or frictional, unemployment, and involuntary unemployment, which is zero "when full employment" is reached; and equilibrium can exist with unemployment only with rigid wages, in the downward direction. The question of rigid wages was addressed in Hahn and Solow (1995).

Based on Solow Growth Model (e.g. Romer (2019)) and the rational expectations-based model of Lucas (1983), unemployment rate (over frictional unemployment) is considered as a deadweight loss to society, which monetary and fiscal policy seeks to eliminate. Kydland and Prescott (1982), in discussion of Real Business Cycles, argues that stocks, driven by technology advances, caused growth and effectiveness of labor. Relying upon the real business cycle model, various economic models for unemployments were developed, including Efficiency-wage model, the Shapiro-Stiglitz Model, and the European unemployment model (e.g. Romer (2019)).

One of the most used hypotheses of unemployment over the past fifty years has been that the percentage rate of change of money wages is inversely associated with the level and the percentage rate of change in unemployment (the Phillips curve), which was empirically demonstrated in Phillips (1958) and Lipsey (1965). The expectations-augmented Phillips curve of Friedman (1968) posited that the inflation rate depends on expected inflation and the deviation of unemployment from its natural rate. Ball et al. (1988) and Ball and Mazumder (2019) provided an empirical analysis of the Phillips curve. The Phillips curve analysis of Friedman (1968) introduced the concept of the natural rate of unemployment, defined as the rate of unemployment to which the economy would converge in the long-run. An implication of the definition is that expansionary monetary policy would produce higher inflation, but might not be able to lower unemployment (Estrella and Mishin, 1999; Taylor, 1999).

Okun (1962) reported a negative short-run correlation between unemployment and output for the United States that has become a staple of macroeconomic textbooks. An et al. (2019) tested

Okun's Law for 70 countries using IMF data for the 1990 -2015 time period, and reported that a linear combination of Okun-based unemployment forecasts and WEO unemployment forecasts can deliver significant forecast accuracy gains for developing economies .

Economic theory indicates that an appropriate level of unemployment is the key of economic development. A high unemployment rate leads to human suffering and many other negative societal consequences, while an extreme low unemployment rate leads to an increase of labor cost and subsequent inflation. One of the certain mission of government monetary policy is to control unemployment rate at the ideal level while controlling inflation (reference). Hence it is vital to be able to accurately forecast the unemployment rate in the near and long term future, based on limited current information of the economy.

In this paper, we revisit the approaches used in the seminal paper of Montgomery, Zarnowitz, Tsay, and Tiao (Montgomery et al., 1998), hereafter denoted as MZTT, on forecasting U.S. unemployment rate. Based on monthly observations in the period of 1959 to 1993, they demonstrated that time series models are useful in predicting 1-month to 5-month ahead unemployment rates and they compared the out-sample prediction performance of a range of models. More than 20 years have passed and it is useful and necessary to re-exam these models *post publication*, as well as to investigate new phenomenon aroused since, including the financial crisis of 2018. We pay special attention to the relationship of U.S unemplyment rate and the lagged U.S. weekly unemployment claim, as well as the US Leading Economic Indicator (LEI) series constructed by The Conference Board.

Because of its importance, there is an extensive literature on forecasting unemployment rate. MZTT modelled the U.S. unemployment rate as a function of the weekly unemployment claims time series, 1948–1993. MZTT reported that nonlinear models, such as Threshold AR models (Tong, 1990; Tong and Lim, 1980), reduced forecasting errors as much as 28%, being more effective in periods of economic contraction and rising unemployment. The majority of the reduced forecasting error occurs in the first quarter following the forecasting origin while one-third of the forecasting error occurs in the second quarter. The transfer function modeling using unemployment claims in MZTT did not reduce the forecasting error. Thomakos and Guerard Jr (2004) re-examined several sets of transfer function modeling sets, including the MZTT U.S. unemployment rate, weekly unemployment claims, and LEI relationships during the 1952 – 1998 time period. Thomakos and Guerard reported leading indicator and the initial unemployment claims provided RMSE reductions of 7.6% and 8.6% respectively over the no-change model. The RMSE reduction reductions were not statistically significant. The coefficients of the TF and VAR models were generally statistically significant and of the correct sign in the Thomakos and Guerard rolling regressions.

Guerard et al. (2020) and Vinod and Guerard (2020) studied the U.S. real GDP and unemployment rate forecasting, and reported that an adaptive averaging autoregressive model and the adaptive learning model produced forecasts has significantly smaller root mean square errors and lower mean absolute errors than the no change model forecasts reported in Mincer and Zarnowitz (1969) during the 1959 to 2018 time period. Guerard et al. (2020) used Kyriazi et al. (2019) adaptive learning model and AutoMetrics, the automatic time series modelling and forecasting system of Hendry and Krolzig (2005), Castle et al. (2013), Hendry and Doornik (2014), Castle and Hendry

(2019); Doornik and Hendry (2015) with its emphasis on structural breaks. Guerard et al. (2020) reported that the transfer functions using LEI and weekly unemployment claims of approximately 25% relative to no-change model forecasts for the 1959-2018 time period. The RMSE and MAE were statistically significant, although the transfer function model RMSE and MAE reductions were slightly less than the adaptive learning model forecasts.

Vinod and Guerard (2020) applied the transfer function methodology and Granger-causality test based on Chen and Lee (1990) to report the casual relationship in which LEI led the U.S. employment rate for the 1959 – 2018 time period, but only concurrent statistical association during the 1959–1993 time period. Several sets of regression and time series-based models have established that transfer functions with LEI and weekly unemployment claims data were highly statistically significant for modelling the MZTT unemployment rate data. The 1994 to 2018 time period was one of economic growth, despite the Global Finance Crisis of 2008–2009. This result is important because the authors did not find the MZTT result that economic contractions and rising unemployment were primarily periods of RMSE reductions.

Following MZTT, in this paper, we employ the SARIMA and transfer function modeling using the leading economic indicators and one of its components, weekly unemployment claims, to forecast the U.S. unemployment rate. We report that the SARIMA and the LEI-based transfer function models significantly outperform the no-change (or simple random walk) models. The U.S. unemployment rate was approaching a 60-year low during 2018. Moreover, the LEI was at an all-time high. There have been statistically significant relationships among the unemployment rate, weekly unemployment claims, and the LEI time series, reported in MZTT, Guerard et al. (2020), and Vinod and Guerard (2020). These time series have been addressed before, but our results are more statistically significant using more recently developed time series modelling techniques and software. The MZTT results are validated, post-publication, and additional time series modeling enhances the statistical significance of the seminal study results. Our results are very consistent with the forecasting analysis of Guerard et al. (2020) and Vinod and Guerard (2020),

The paper is organized as follows. In Section 2 we provide the detailed of the data set used in the analysis. Section 3 lists the models entertained. Several other models (including xxx) were entertained but were found to be less desirable hence omitted here. In Section 4 we report the detailed findings. Section 5 concludes.

2 Data Description

The composite indexes of leading (LEI), coincident, and lagging indicators produced by The Conference Board are summary statistics for the U.S. economy. Wesley Clair Mitchell of Columbia University constructed the indicators in 1913 to serve as a barometer of economic activity. The leading indicator series was developed to turn upward before aggregate economic activity increased, and decrease before aggregate economic activity diminished. Historically, the cyclical turning points in the leading index have occurred before those in aggregate economic activity, cyclical turning points in the coincident index have occurred at about the same time as those in aggregate economic activity, and cyclical turning points in the lagging index generally have occurred after those in aggregate economic activity.

The Conference Board's components of the composite leading index for the year 2002 reflected the work and variables shown in Zarnowitz (1992) list, which continued work of the Mitchell (1913, 1951); Mitchell and Gay (1927), Burns and Mitchell (1946), and Moore (1961).[1] Specifically, LEI is an equally weighted index in which its components are standardized to produce constant variances. The detailed information of its components and its construction can be found at The Conference Board website.[2]

The monthly unemployment rate we study is from 1959/02 to 2019/12. For the transfer function modeling, we use the LEI and the monthly initial claims for unemployment (ICSA) as predictors. These three series are plotted in Figure 1.

3 Models and their estimated parameters

3.1 SARIMA Model

The first model we consider is th seasonal ARIMA model (Box et al., 2015). It does not involve any exogenous variables. Various model selection procedures indicate that SARIMA$(2, 1, 2) \times (1, 0, 1)_{12}$ is a reasonable model. The fitted model to the whole time series is given by

$$(1 - 1.10B + 0.18B^2)(1 - 0.54B^{12})\Delta x_t = (1 - 1.09B + 0.33B^2)(1 - 0.80B^{12})a_t,$$

where a_t are assumed to be independent and identically distributed (IID), following a normal distribution, with estimated variance 0.026. The estimated coefficients and their standard errors are given in the table below.

Parameter	ϕ_1	ϕ_2	θ_1	θ_2	Φ_1	Θ_1
Estimate	1.10	-0.18	-1.09	0.33	0.54	-0.80
SE	0.27	0.26	0.26	0.21	0.07	0.05

All the coefficients except for ϕ_2 and θ_2 are highly significant. If we remove one of the two lag-2 coefficients ϕ_2 and θ_2, the other becomes highly significant, and the AIC gets bigger. Therefore, we choose to use the current model. Diagnostics show that the model is adequate and the residuals are white.

[1]Geoffrey Moore passed in 2000. Pami Dua edited a collection of papers to honor Moore, entitled Business Cycles and Economic Growth (Oxford University Press, New York). The reader is referred to papers in the volume by Victor Zarnowitz, John Guerard, Lawrence Klein, and S. Ozmucur, and D. Ivanova and Kajal Lahiri. The Conference Board index of leading indicators was composed of the following variables in 2001; Average weekly hours (mfg.); Average weekly initial claims for unemployment insurance; Manufacturers' new orders for consumer goods and materials; Vendor performance; Manufacturers' new orders of non-defense capital goods; Building permits of new private housing units; Index of stock prices; Money supply; Interest rate spread; and the Index of consumer expectations.

[2]The reader is referred to Zarnowitz (1992), an IJF Associate Editor for many years, for his seminal development of underlying economic assumption and theory of the LEI and business cycles. The Zarnowitz et al. (2001a), Zarnowitz et al. (2001b) and Zarnowitz (2004) publications drove development of the LEI at The Conference Board.

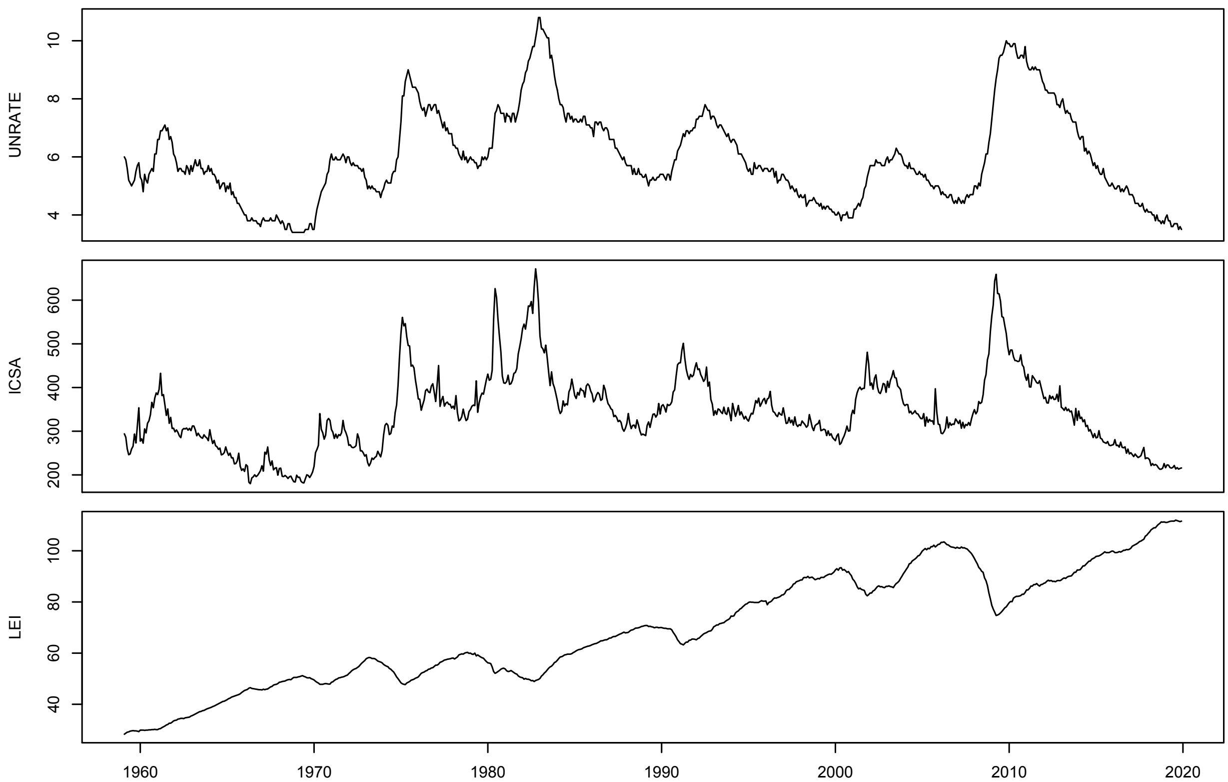

Figure 1: Monthly unemployment rate (UNRATE), initial claims (ICSA) and LEI.

3.2 TAR Model

The second model we employ is the thresholded autoregressive (TAR) model of Tsay (1989), fitted to differenced unemployment rate Δx_t. Again, after a careful analysis and model building excises, we choose Δx_{t-3} as the threshold variable, and use two regimes depending on whether $\Delta x_{t-3} \leq c$ or not. The estimated threshold value c is 0.1, indicating that the dynamic of the unempolyment rate are different, depending whether the change of the rate three months ago is greater than 10% or not. In each regime, an AR(6) model is fitted. The estimated model is given by

$$\Delta x_t = \begin{cases} -0.07\Delta x_{t-1} + 0.20\Delta x_{t-2} + 0.12\Delta x_{t-3} & \\ \quad +0.20\Delta x_{t-4} + 0.13\Delta x_{t-5} + 0.04\Delta x_{t-6} + a_t^{(1)} & \text{when } \Delta x_{t-3} \leq 0.1 \\ 0.17\Delta x_{t-1} + 0.16\Delta x_{t-2} + 0.19\Delta x_{t-3} & \\ \quad +0.05\Delta x_{t-4} - 0.09\Delta x_{t-5} + 0.04\Delta x_{t-6} + a_t^{(2)} & \text{when } \Delta x_{t-3} > 0.1 \end{cases}$$

where $a_t^{(1)}$ and $a_t^{(2)}$ are assumed to be IID Normal, with estimated variances 0.023 and 0.031, respectively. During the entire sample period, the proportion of times the process is in Regime 1 is about 0.78. The model is "self-excited" as it does not depend on exogenous variables, but changes it dynamic based on its own past.

The estimated coefficients and their standard errors are reported below.

Lag	1	2	3	4	5	6
$\Delta x_{t-3} \leq 0.1$	-0.07	0.20	0.12	0.20	0.13	0.04
SE	0.04	0.04	0.05	0.04	0.04	0.04
$\Delta x_{t-3} > 0.1$	0.17	0.16	0.19	0.05	-0.09	0.04
SE	0.07	0.08	0.06	0.08	0.08	0.07

Most coefficient in the model are significant. We choose to use the past 6 months for the prediction and keep the insignificant terms in the model for simplicity. Residual analysis of the model is satisfactory. The total residual sum of squares is 19.50, comparing with 18.65 of the SARMA model, though the TAR model uses more parameters.

3.3 Transfer Function Modeling

We also consider the transfer function models (Box et al., 2015) using the initial claims for unemployment (ICSA/1000, denoted by u_t) and the leading economic index (LEI/10, denoted by v_t). Two models are entertained, one with ICSA only and the other with both ICSA and LEI.

The selected transfer function model for Δx_t is a SARIMA$(2,0,2) \times (1,0,1)_{12}$ with lags 1 to 6 of ICSA, which is estimated as

$$\begin{aligned} &\Delta x_t = 2.02u_{t-1} + 0.12u_{t-2} - 1.48u_{t-3} - 0.44u_{t-4} + 0.48u_{t-5} - 0.70u_{t-6} + w_t, \\ &(1 - 1.18B + 0.23B^2)(1 - 0.57B^{12})w_t = (1 - 1.36B + 0.50B^2)(1 - 0.80B^{12})a_t, \end{aligned}$$

where a_t are assumed to be IID Normal, with estimated variance 0.021. Although this model only slightly reduces the residual variance comparing to the SARIMA model in Section 3.1, at the expense of six more parameters, it outperforms SARIMA model in terms out-sample prediction, to be reported later in Section 4. We summarize the estimated coefficients together with their standard errors in the following table.

Coefficients of SARIMA	ϕ_1	ϕ_2	θ_1	θ_2	Φ_1	Θ_1
Estimate	1.18	-0.23	-1.36	0.50	0.57	-0.80
SE	0.13	0.13	0.12	0.11	0.08	0.05
Coefficients of u_{t-k}	2.02	0.12	-1.48	-0.44	0.48	-0.70
SE	0.30	0.47	0.44	0.44	0.44	0.29

Most coefficients are significant at 5% level. Since the significance levels of the estimated coefficients change when different periods of data are used (for out-sample prediction excise), we keep all the terms in the model.

The estimated transfer function model of a SARIMA$(2, 0, 2) \times (1, 0, 1)_{12}$ with lags 1 to 6 of both the ICSA and the LEI is

$$\begin{aligned}
\Delta x_t = &\ 1.67u_{t-1} + 0.23u_{t-2} - 1.64u_{t-3} - 0.19u_{t-4} + 0.32u_{t-5} - 0.29u_{t-6}, \\
&- 0.21v_{t-1} + 0.10v_{t-2} - 0.28v_{t-3} + 0.34v_{t-4} - 0.26v_{t-5} + 0.31v_{t-6} + w_t, \\
(1 - 1.03B + 0.11B^2)&(1 - 0.56B^{12})w_t = (1 - 1.25B + 0.38B^2)(1 - 0.81B^{12})a_t,
\end{aligned}$$

with estimated residual variance to be 0.020. The estimated coefficients and their standard errors are reported below.

Coefficients of SARIMA	ϕ_1	ϕ_2	θ_1	θ_2	Φ_1	Θ_1
Estimate	1.03	-0.11	-1.25	0.38	0.56	-0.81
SE	0.14	0.15	0.14	0.14	0.07	0.053
Coefficients of u_{t-k}	1.67	0.23	-1.64	-0.19	0.32	-0.29
SE	0.34	0.51	0.50	0.49	0.49	0.33
Coefficients of v_{t-k}	-0.21	0.10	-0.28	0.34	-0.26	0.31
SE	0.17	0.30	0.30	0.29	0.29	0.17

Although most coefficients of u_{t-k} and v_{t-k} are not significant, we still keep all the terms in the model to explore their potential for the predictions. The reader may ask if simpler models are preferred and why we report a transfer function model forecast using both ICSA and LEI variables. We do so because both ICSA and LEI bivariate models offer forecasting improvement relative to an ARIMA model of the unemployment rate, see Guerard et al. (2020).

4 Out-sample Rolling Forecast Performance

We compare the rolling forecast performances of the models considered in Section 2. The predictions are made for the change of the unemployment rate, i.e. for the differenced series Δx_t. A null model which always predicts the unemployment rate change as 0 is also included as a benchmark. Specifically, the following methods/models are compared.

I. Null. Always predict Δx_t as 0.

II. SARIMA$(2,0,2) \times (1,0,1)_{12}$ for Δx_t, as given in Section 3.1

III. TAR. The two regime TAR model in Section 3.2

IV. Transfer-1. SARIMA$(2,0,2) \times (1,0,1)_{12}$ and 6 lagged ICSA, in Section 3.3.

V. Transfer-2. SARIMA$(2,0,2) \times (1,0,1)_{12}$, 6 lagged ICSA and 6 lagged LEI, in Section 3.3.

VI. Average. The average of the predictions given by II–IV.

For models I and II, it is straightforward to predict with an arbitrary horizon. For the TAR model in III, one-step to three step prediction can be done using the estimated model in the regime that Δx_{t+h-3} indicates, where h is the prediction horizon. For longer prediction horizon, we use simulation procedure, based on the estimated innovation distribution. For the transfer function models IV and V, using the model specified in Section 3.3 would require the knowledge of future ICSA and LEI values, or require a separate accurate prediction method for ICSA and LEI. To mitigate the difficulty, we make the following modifications. If $1 \leq h \leq 6$, we fit a model with only lag-h to lag-6 ICSA and LEI in the model, and use this model to make a h-step-ahead prediction. If $7 \leq h \leq 12$, we fit a model with only the lag-h indicator in the model, and then make a h-step-ahead forecast using this model.

The rolling forecast are performed from January 1990 to December 2019. Specifically, for every month during the period, we use the available data before that month to estimate the models, then perform 1-step ahead to 12-step ahead prediction, and compare with the observed values. We report in Table 1 the mean squared forecast errors for every 5 years, as well as the mean squared errors for the 30 years together. We also use a one-sided paired t test to compare the absolute forecast errors generated by each pair of two models. In Table 2, a small p-value, denoted by "`1`" or "`2`", indicates that the method indexed by the row performs better than the one indexed by the column based on the one-sided test at 10% or 5% level respectively. For example, in the first cell of 4-th row (Transfer-1), the three "`2`" indicates that the Transfer-1 significantly outperforms the Null model in terms of 1, 2 and 5 steps ahead predictions, at 5% level. The rest five "`-`" suggests that the Transfer-1 is not better than the Null model in performing forecasts of the corresponding horizons. Note that Table 2 could have been reported as a "tournament" table giving the competition results (win/loss/tie) of different models. However, to simplify the symbols, we choose to indicate only when one model wins over the other, which contains the same information as the "tournament" table. In the literature, the Diebold-Mariano test (Diebold and Mariano, 2002) has been commonly used to compare the predictive accuracy. We choose to use the paired t-test here for the purpose of making the one-sided comparison.

From Table 1, it is seen that among models/methods I–V, the ICSA transfer functions produce the smallest forecasting errors, with forecasting MSE reductions of 24.6 percent comparing to the no-change models, during the 1990–2019 time periods. The 1-step error reduction is consistent with the Guerard et al. (2020) RMSE results. The 1-step RMSE error reductions of Transfer-1 are 2.4 and 6.7 percent, with the SARIMA and TAR models, respectively. We also compare the mean absolute prediction error (MAE), and report the results in Table 3, which is in the Appendix. Similar conclusions can be drawn based on MAE.

We report in Table 2 that the weekly unemployment claims transfer functions (Transfer-1) produces the smallest forecasting errors with the one-month to six-month forecasting horizons than all other models, including the no-change models, consistent with the Guerard et al. (2020) results. The SARIMA model reduces forecasting errors relative to the no-change models. We substantiate the TAR model reductions reported in the seminal MZTT study, although the TAR model does not dominate the transfer function models on forecasting effectiveness.

We also observe that the method VI (by taking average of the predictions given by II–V) leads to the smallest forecast errors, for 1 month to 5 months horizons. The statistically significant forecast enhancements of the composite models in the six-step-ahead forecast gives researchers a reasonable period to implement the forecasting models. The dominance of the composite forecasting model performance is consistent with the Madridakis forecasting competitions of the past 38 years. The forecasting literature has an extensive literature reporting that the average forecast is the better forecast out of sample, see Bates and Granger (1969); Clemen (1989); Makridakis and Hibon (1979) and the summary results from the M-forecasting competitions (Makridakis et al., 1982, 1993; Makridakis and Hibon, 2000; Makridakis et al., 2020). When the horizons are larger than or equal to 6 months, no methods can make significantly better predictions than the no-change model, indicating the non-predictability at large horizons.

	P1	P2	P3	P4	P5	P6	All	P1	P2	P3	P4	P5	P6	All
				One Step Ahead							Two Step Ahead			
I	2.00	1.95	1.72	4.02	2.88	1.55	2.35							
II	1.60	2.08	1.39	2.29	2.29	1.51	1.86	1.57	2.06	1.38	2.31	2.29	1.46	1.85
III	1.90	1.96	1.56	2.44	2.90	1.47	2.04	1.87	2.06	1.61	2.59	2.93	1.54	2.10
IV	1.67	1.86	1.31	2.10	2.43	1.29	1.77	1.63	2.06	1.26	2.00	2.02	1.43	1.73
V	2.16	2.05	1.52	1.81	2.41	1.26	1.87	1.84	2.23	1.34	1.53	2.28	1.40	1.77
VI	1.61	1.87	1.28	1.86	2.39	1.31	1.72	1.57	2.00	1.27	1.86	2.29	1.41	1.73
				Three Step Ahead							Four Step Ahead			
II	1.52	2.06	1.40	2.59	2.20	1.55	1.89	1.56	2.20	1.41	2.82	2.06	1.50	1.93
III	1.74	2.07	1.61	2.94	2.80	1.49	2.11	1.81	2.13	1.68	3.19	2.75	1.47	2.17
IV	1.51	2.09	1.42	2.69	2.19	1.52	1.90	1.60	2.22	1.43	2.84	2.06	1.53	1.95
V	1.83	2.25	1.45	2.00	2.60	1.49	1.94	1.85	2.88	1.47	1.91	2.26	1.60	2.00
VI	1.48	2.03	1.33	2.35	2.37	1.47	1.84	1.56	2.24	1.42	2.56	2.21	1.49	1.91
				Five Step Ahead							Six Step Ahead			
II	1.61	2.13	1.38	2.98	2.05	1.51	1.94	1.63	2.09	1.40	3.33	2.15	1.55	2.02
III	1.86	2.01	1.55	3.29	2.73	1.45	2.15	1.88	1.96	1.66	3.42	2.82	1.46	2.20
IV	1.63	2.10	1.40	2.97	2.11	1.52	1.95	1.65	2.06	1.42	3.31	2.21	1.56	2.03
V	1.67	2.26	1.32	2.38	2.23	1.58	1.91	1.70	2.15	1.43	3.24	2.29	1.65	2.08
VI	1.64	2.07	1.38	2.85	2.22	1.49	1.94	1.69	2.02	1.46	3.30	2.32	1.52	2.05
				Nine Step Ahead							Twelve Step Ahead			
II	1.71	2.12	1.48	3.78	2.12	1.67	2.15	1.71	2.11	1.45	4.01	2.25	1.76	2.22
III	1.96	1.96	1.72	3.72	2.86	1.51	2.29	2.02	1.99	1.71	3.94	3.03	1.60	2.38
IV	1.74	2.10	1.51	3.81	2.35	1.68	2.20	1.75	2.12	1.49	4.09	2.29	1.78	2.25
V	1.81	2.25	1.53	3.69	2.48	1.88	2.27	1.86	2.60	1.47	3.81	2.52	2.21	2.41
VI	1.77	2.06	1.54	3.73	2.40	1.65	2.19	1.79	2.13	1.50	3.93	2.47	1.78	2.27

Table 1: Mean squared rolling forecase error (MSE) ($\times 10^2$) by 5-year periods. The first row gives the MSE of the Null model. The rest of the table should be understood as a 4×2 block matrix, corresponding to 1, 2, 3, 4, 5, 6, 9 and 12 step ahead forecasts respectively. In particular, the top two blocks are for 1- and 2-step predictions, and the rest are self-evident. For each of the 8 blocks, the rows correspond to models, and the columns P1 to P6 the 5-year periods: 90-94, 95-99, 00-04, 05-09, 10-14, 15-19 respectively. The column "All" gives the MSE over all 30 years.

	Null	SARIMA	TAR	Transfer-1	Transfer-2	Average
Null		--------	---1--22	-------1	------22	-------2
SARIMA	-1-12---		222222--	---1--22	---2-222	------11
TAR	--------	--------		--------	------12	--------
Transfer-1	22--2---	-1------	222222--		---2-222	------
Transfer-2	----1---	--------	-2--1---	--------		--------
Average	221-2---	221-----	222222--	--1-----	1-12-222	

Table 2: Comparing the absolute forecast errors by paired t tests. “2” and “1” stand for a p-value less than or equal to 0.05 and 0.10 respectively, and indicates that the row method is better than the column method in terms of absolute forecast errors. “-” stands for a p-value larger than 0.1. For each pairwise competition, there are 8 “-”, “1” and/or “2”, corresponding to 1, 2, 3, 4, 5, 6, 9, 12 step ahead forecasts respectively.

	1-step	2-step	3-step	4-step	5-step	6-step	9-step	12-step
Null	0.975	0.975	0.975	0.975	0.975	0.975	0.972	0.972
	0.707	0.708	0.708	0.709	0.709	0.710	0.711	0.713
SARIMA	0.981	0.981	0.975	0.978	0.975	0.967	0.967	0.967
	0.658	0.658	0.668	0.677	0.685	0.690	0.699	0.703
TAR	0.972	0.972	0.969	0.969	0.972	0.969	0.967	0.967
	0.678	0.678	0.694	0.710	0.724	0.727	0.738	0.742
Transfer-1	0.978	0.981	0.975	0.978	0.975	0.967	0.967	0.967
	0.600	0.624	0.655	0.670	0.677	0.682	0.692	0.697
Transfer-2	0.978	0.983	0.978	0.981	0.975	0.967	0.964	0.969
	0.593	0.613	0.634	0.648	0.665	0.681	0.691	0.693

Table 3: Empirical coverage probabilities of the 95% prediction intervals. The prediction intervals are constructed for the rolling forecasts from January 1990 to December 2019. The second row for each method reports the average width of the prediction intervals.

We also report in Table 3 the empirical coverage probabilities of the 95% prediction intervals corresponding to each of the five methods, as well as the average length of the prediction intervals given by each method. More details of the tables are provided in the captions. It can be seen that the empirical coverage of the prediction interval is slightly higher than the designed coverage of 95%, showing the estimated prediction variance is larger than the truth. On the other hand, the prediction interval length is less than 0.7% (or $\pm 0.35\%$) for the transfer functions models, showing that the out-of-sample predictions are quite accurate. The interval length would be smaller if we set the empirical coverage to 95%.

5 Conclusion and Suggestions for Future Research

In this study, we study the performance of a Seasonal ARIMA model, a Threshold AR model, a transfer function model using ISCA along, and a transfer function model using both ICSA and LEI, for the purpose of forecasting the U.S. unemployment rate during the 1959-2019 time period. We report additional statistically significant association regarding the relationship among transfer functions and average forecast models in the 1- to 5-step-ahead forecasts during the 1990 – 2019 time period. We replicate the seminal study of MZTT and extend the analysis to cover a much longer and post publication period. It serves as an updated benchmark analysis from that offered by MZTT. Moreover, the statistical models produce lower relative errors in the Global Financial Crisis and its aftermath time period, 2008-2019. In summary, we report on the statistically significant impact of the LEI and weekly unemployment claims time series on the unemployment rate series. The MZTT variable relationships are confirmed, in-sample, and post- publication. The unemployment rate was at a 50-year low at the time we estimated our LEI and unemployment rate analysis. It will be interesting, at the end of the pandemic, to examine the relationship between the LEI and its components and the unemployment rate.

References

An, Z., Ball, L., Jalles, J., and Lounggani, P. (2019). Do imf forecasts respect okun's law? evidence for advanced and developing economies. *International Journal of Forecasting*, 35:1131–1142.

Ashley, R. (2003). Statistically significant forecasting improvements: how much out-of-sample data is likely necessary? *International Journal of Forecasting*, 19(2):229–239.

Ball, L., Mankiw, N., and Romer, D. (1988). The new keynesian economics and the output-inflation trade-off. *Brookings Papers on Economic Activity*, pages 1–65.

Ball, L. and Mazumder, S. (2019). A phillips curve with anchored expectations and short-term unemployment. *Journal of Money, Credit, and Banking*.

Bates, J. M. and Granger, C. W. J. (1969). The combination of forecasts. *The Journal of the Operational Research Society*, 20:451–468.

Box, G. E., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Brunner, K. and Meltzer, A. (1993). *Money and the economy: Issues in monetary analysis*. Cambridge University Press, Cambridge.

Burns, A. F. and Mitchell, W. C. (1946). Measuring business cycles.

Castle, J. L., Clements, M. P., and Hendry, D. F. (2015). Robust approaches to forecasting. *International Journal of Forecasting*, 31(1):99–112.

Castle, J. L., Doornik, J. A., and Hendry, D. F. (2013). Model selection in equations with many 'small'effects. *Oxford Bulletin of Economics and Statistics*, 75(1):6–22.

Castle, J. L. and Hendry, D. (2019). *Modelling Our Changing World*. Palgrave. London.

Castle, J. L. and Shephard, N. (2009). *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*. OUP Oxford.

Chen, C. and Lee, C.-W. J. (1990). A varma test on the gibson paradox. *The Review of Economics and Statistics*, pages 96–107.

Chen, C. and Liu, L.-M. (1993). Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, 88(421):284–297.

Clemen, R. (1989). Combining forecasts: a review and annotated bibliography with discussion. *International Journal of Forecasting*, 5:559–608.

Dhrymes, P. J. and Guerard, J. B. (1978). *Introductory econometrics*, volume 218. Springer. With contributions from Guerard, J. B., Jr.

Diebold, F. X. and Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics*, 20(1):134–144.

Doornik, J. A. and Hendry, D. F. (2015). Statistical model selection with "big data". *Cogent Economics & Finance*, 3(1):1045216.

Estrella, A. and Mishin, F. (1999). Rethinking the role of nairu in monetary policy: Implications of model formulation and uncertainty. In Taylor, J., editor, *Monetary Policy Rule*. University of Chicago Press: Chicago.

Friedman, M. (1968). The role of monetary policy. *American Economic Review*, pages 1–17.

Granger, C. W. (2001). *Essays in econometrics: collected papers of Clive WJ Granger*, volume 32. Cambridge University Press.

Granger, C. W. J. and Newbold, P. (2014). *Forecasting economic time series*. Academic Press.

Guerard, J., Thomakos, D., and Kyriazi, F. (2020). Automatic time series modeling and forecasting: A replication case study of forecasting real gdp, the unemployment rate and the impact of leading economic indicators. *Cogent Economics & Finance*, 8(1):1759483.

Guerard, J. B. (2001). A note on the forecasting effectiveness of the us leading economic indicators. *Indian Economic Review*, pages 251–268.

Guerard Jr, J. B. (2004). The forecasting effectiveness of the us leading economic indicators: Further evidence and initial g7 results. In *Business cycles and economic growth: An analysis using leading indicators*, pages 174–187. Oxford University Press.

Hahn, F. and Solow, R. (1995). *A critical essay on modern macroeconomic theory*. MIT Press. Cambridge, MA.

Harberler, G. (1937). *Prosperity and depression*. League of Nations, Geneva. Reprinted 1963. Atheneum, New York.

Hendry, D. F. and Doornik, J. A. (2014). *Empirical model discovery and theory evaluation: automatic selection methods in econometrics*. MIT Press.

Hendry, D. F. and Krolzig, H.-M. (2005). The properties of automatic gets modelling. *The Economic Journal*, 115(502):C32–C61.

Keynes, J. (1936). *The general theory of employment, interest, and money*. Macmillan, London.

Klein, L. (1950). *Economic fluctuations in the United States, 1921–41.* Wiley & Sons, New York.

Kydland, F. and Prescott, E. (1982). Time to build and aggregate fluctuations. *Econometrica*, 50:1345–1369.

Kyriazi, F., Thomakos, D. D., and Guerard, J. B. (2019). Adaptive learning forecasting, with applications in forecasting agricultural prices. *International Journal of Forecasting*, 35(4):1356–1369.

Levanon, G., Manini, J.-C., Ozyildirim, A., Schaitkin, B., and Tanchua, J. (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.

Lipsey, R. (1965). Structural and deficient demand unemployment reconsidered. In Ross, A., editor, *Employment Policy and the Labor Market*, pages 210–255. Berkeley: University of California Press.

Lucas, R. (1983). *Studies in business-cycle theory.* MIT Press. Cambridge, MA.

Makridakis, S., Andersen, A., Carbone, R., Fildes, R., and Hibon, M. (1982). The accuracy of extrapolation (time series) methods: results of a forecasting competition. *Journal of Forecasting*, 1:111–153.

Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., and Mills, T. (1993). The m2-competition: a real-time judgmentally based forecasting study. *International Journal of Forecasting*, 9:5–22.

Makridakis, S. and Hibon, M. (1979). Accuracy of forecasting: an empirical investigation. *Journal of the Royal Statistical Society, Series A*, 142:97–145.

Makridakis, S. and Hibon, M. (2000). The m3-competition: results, conclusions and implications. *International Journal of Forecasting*, 16:451–476.

Makridakis, S., Spiliotis, E., and Assimakopoulos, V. (2020). The m4 competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*, 36:54–74.

Matthews, C. (1959). *The Business Cycle*. The University of Chicago Press, Chicago.

Mincer, J. A. and Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations: Analysis of forecasting behavior and performance*, pages 3–46. NBER.

Mitchell, W. C. (1913). *Business cycles*, volume 3. University of California Press.

Mitchell, W. C. (1951). What happens during business cycles, a progress report.

Mitchell, W. C. and Gay, E. F. (1927). *Business cycles: The problem and its setting*, volume 1. National Bureau of Economic Research New York.

Montgomery, A. L., Zarnowitz, V., Tsay, R. S., and Tiao, G. C. (1998). Forecasting the us unemployment rate. *Journal of the American Statistical Association*, 93(442):478–493.

Moore, G. H. (1961). *Business cycle indicators*, volume 1. Princeton University Press Princeton, NJ.

Okun, A. (1962). Potential gnp: Its measurement and significance. In *Proceedings of the Business and Economic Statistics Section of the American Statistical Association. Alexandria, VA: American Statistical Association*, pages 89–104.

Persons, W. (1931). *Forecasting Business cycles*. Wiley & Sons, Inc, New York.

Phillips, A. (1958). The relationship between unemployment and the rate of change of money wage rates in the united kingdom, 1861-1957. *Economica*, 25:283–299.

Romer, D. (2019). *Advanced macroeconomics*. McGraw-Hill, New York, 5 edition.

Samuelson, P. (1948). *Economics*. McGraw-Hill, New York.

Taylor, J. (1999). *Monetary Policy Rules*. University of Chicago Press: Chicago.

Thomakos, D. D. and Guerard Jr, J. B. (2004). Naive, arima, nonparametric, transfer function and var models: A comparison of forecasting performance. *International Journal of Forecasting*, 20(1):53–67.

Tong, H. (1990). *Non-Linear Time Series: A Dynamic System Approach*. Oxford Press, Oxford.

Tong, H. and Lim, K. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society, Series B*, 42:245–292.

Tsay, R. S. (1988). Outliers, level shifts, and variance changes in time series. *Journal of forecasting*, 7(1):1–20.

Tsay, R. S. (1989). Testing and modeling threshold autoregressive processes. *Journal of the American statistical association*, 84(405):231–240.

Tsay, R. S. (2005). *Analysis of financial time series*, volume 543. John wiley & sons.

Tsay, R. S. and Chen, R. (2018). *Nonlinear time series analysis*, volume 891. John Wiley & Sons.

Viner, R. (1949). Koopmans on the choice of variables to be studied and of methods of measurement (with a reply from koopmans). *Review of Economics and Statistics*, 31:77–94.

Vinod, H. and Guerard, J.B., J. (2020). Tcausality studies of real gdp, unemployment, and leading indicators. In Guerard, J.B., J. and Ziemba, W. T., editors, *Handbook of Applied Investment Research*. World Scientific Publishing, Co., Singapore.

Vinod, H. D. (2008). *Hands-On Intermediate Econometrics Using R: Templates for Extending Dozens of Practical Examples (With CD-ROM)*. World Scientific Publishing Company.

Vinod, H. D. (2019). New exogeneity tests and causal paths. In *Handbook of Statistics*, volume 41, pages 33–64. Elsevier.

West, K. D. and McCracken, M. W. (1998). Regression-based tests of predictive ability. *International Economic Review*, 39(4).

Zarnowitz, V. (1992). *Business Cycles: Theory, History, Indicators, and Forecasting*. University of Chicago Press, Chicago, IL.

Zarnowitz, V. (2004). The autonomy of recent us growth and business cycles. *Business Cycles and Economic Growth: An Analysis Using Leading Indicators*, pages 44–82.

Zarnowitz, V. et al. (2001a). The old and the new in the us economic expansion. Technical report, The Conference Board, Economics Program.

Zarnowitz, V., Ozyildirim, A., and Ozyildrim, A. (2001b). On the measurement of business cycles and growth cycles. *Indian Economic Review*, pages 37–54.

6 Appendix

We provide the comparison of different models/methods in terms of the mean absolute rolling forecast error (MAE) in the Appendix.

	P1	P2	P3	P4	P5	P6	All	P1	P2	P3	P4	P5	P6	All
	One Step Ahead							Two Step Ahead						
I	1.10	1.02	0.98	1.48	1.22	0.98	1.13							
II	1.08	1.11	0.94	1.22	1.18	1.02	1.09	1.07	1.10	0.94	1.22	1.18	1.00	1.09
III	1.15	1.06	0.99	1.26	1.31	1.01	1.13	1.14	1.10	1.00	1.29	1.32	1.04	1.15
IV	1.07	1.05	0.91	1.20	1.20	0.95	1.06	1.12	1.11	0.87	1.19	1.11	0.98	1.06
V	1.20	1.14	0.96	1.09	1.20	0.95	1.09	1.15	1.24	0.92	1.04	1.18	0.99	1.09
VI	1.06	1.06	0.90	1.12	1.20	0.97	1.05	1.08	1.11	0.87	1.15	1.17	0.98	1.06
	Three Step Ahead							Four Step Ahead						
II	1.04	1.09	0.95	1.29	1.19	1.03	1.10	1.01	1.14	0.97	1.32	1.09	1.01	1.09
III	1.06	1.10	1.02	1.36	1.30	1.02	1.14	1.07	1.14	1.05	1.41	1.27	1.02	1.16
IV	1.03	1.10	0.96	1.29	1.19	1.02	1.10	1.03	1.14	0.98	1.33	1.09	1.04	1.10
V	1.12	1.17	0.98	1.16	1.26	1.01	1.12	1.12	1.35	1.01	1.17	1.16	1.08	1.15
VI	1.01	1.09	0.92	1.26	1.21	1.01	1.08	1.00	1.15	0.98	1.29	1.13	1.03	1.10
	Five Step Ahead							Six Step Ahead						
II	1.03	1.09	0.95	1.33	1.08	1.03	1.08	1.06	1.10	0.95	1.41	1.11	1.03	1.11
III	1.08	1.06	1.00	1.37	1.25	0.98	1.12	1.11	1.05	1.04	1.41	1.29	0.99	1.15
IV	1.03	1.08	0.95	1.33	1.10	1.03	1.09	1.07	1.09	0.96	1.41	1.13	1.03	1.11
V	1.07	1.11	0.94	1.25	1.11	1.05	1.09	1.10	1.11	0.97	1.40	1.14	1.06	1.13
VI	1.04	1.07	0.95	1.32	1.11	1.01	1.08	1.08	1.07	0.97	1.40	1.15	1.01	1.12
	Nine Step Ahead							Twelve Step Ahead						
II	1.08	1.10	0.97	1.48	1.08	1.09	1.13	1.09	1.10	0.96	1.52	1.13	1.11	1.15
III	1.12	1.05	1.02	1.45	1.26	0.99	1.15	1.14	1.04	1.01	1.48	1.29	1.02	1.16
IV	1.10	1.10	0.98	1.48	1.12	1.10	1.14	1.10	1.11	0.96	1.53	1.13	1.12	1.16
V	1.14	1.15	1.00	1.47	1.15	1.17	1.18	1.16	1.23	0.98	1.51	1.18	1.24	1.22
VI	1.10	1.08	0.99	1.47	1.14	1.08	1.14	1.11	1.11	0.97	1.51	1.16	1.11	1.16

Table 4: Mean absolute rolling forecase error (MAE) ($\times 10$) by 5-year periods. The first row gives the MAE of the Null model. The rest of the table should be understood as a 4×2 block matrix, corresponding to 1, 2, 3, 4, 5, 6, 9 and 12 step ahead forecasts respectively. In particular, the top two blocks are for 1- and 2-step predictions, and the rest are self-evident. For each of the 8 blocks, the rows correspond to models, and the columns P1 to P6 the 5-year periods: 90-94, 95-99, 00-04, 05-09, 10-14, 15-19 respectively. The column "All" gives the MAE over all 30 years.