Exploring the influencing factors of *Real Output Per Person* in the Business Sector in the United States

Guanzhi Wang, Jingxuan Wang, Qiaowei Jiang

1) Abstract

In this project, we explore the influencing factors of *Real Output Per Person* within the business sector and forecast based on historical data. Our explorations are aimed at improving general understanding of this economic indicator. The five influencing factors we have taken into consideration are *unit labor cost, labor share, employment, unit nonlabor payment, and implicit price deflator.* All of the data are quarterly and seasonally adjusted indexes with 2012 as the baseline, derived from the U.S. Bureau of Labor Statistics.

After exploring the characteristics, and providing an overview of each of the variables, we then analyze these variables' correlation with *real output per person*. Then, we build and run five unique regression models in order to find the adequate one in conduct inference and find the association between our response variable and predictors. We also use transform regression model to eliminate the autoregressive errors in the model. Finally, we adopt the ARIMA and VAR model to both better understand the characteristics of the prediction and better forecast, based on the model.

The results from the regression models suggest that none of the regression models are adequate enough for forecasting. The results of the ARIMA model shows that it is reasonable for people to use ARIMA (0,2,1) with *labor share* and *unit labor costs* to predict *real output per person.* And the results of the VAR suggest we could use VAR (3) to predict as well.

2) Introduction

Within economics, there are many time-series variables which are related to each other and deserve further exploration. In this project, we want to use time series analysis to explore the influencing factors of *real output per person* within the business sector. Business sector output is an important indicator that is closely related to gross domestic product (GDP); *real output per person* is a good measurement of labor productivity and lends itself to important government decision-making. Our explorations can serve as a reference for understanding the relationship between *real output per person* and other predictors, as well as for using historical data to conduct economic forecasting.

3) Methods and Materials

The ultimate goal of our project is to investigate which factors are significantly correlated with *real output per person* within the business sector and how these factors could influence the *real output per person* index. To achieve the goal, we need to find and understand appropriate data; this includes the characteristics of each variable and the relationship between the response variable and predictor variables.

We source our data from the U.S. Bureau of Labor Statistics and use five predictor variables: Business Sector: Unit Labor Costs; Business Sector: Labor Share; Business Sector: Employment; Business Sector: Unit Non-labor Payment; Business Sector: Implicit Price Deflator. Indeed, all predictors are related to the business sector.

All variables of our data were quarterly collected, from Jan 1, 1947 to July 1, 2019. We have 291 observations for each variable. As shown in Figure 1 of the plots of variables, there are no obvious seasonal trends that can be found. We believe this is because the predictors and response variables are in the form of an index. Also, the HoltWinters method is applied to explore all 5 predictors. Figures 2 through 6 show no special seasonal effect (or trend) in the graphs. ACF plots were also made to look at the autocorrelations inside the variables. Figure 7 suggests that strong, positive autocorrelation can be found for all variables, indicating that we need to process the data a step further. Thus, Figure 8 suggests that, after applying the first differences, most autocorrelations were mitigated. To determine the relationships between variables, we made CCF plots. From the CCFs in Figures 9 through 10, each of the corresponding two variables are strongly correlated. These correlations can be seen in the correlation matrix in Figure 11. Moreover, Figure 12's scatter-plot of all variables shows that there is multicollinearity among the variables.

Regression models

To further investigate the association between our response variable (*real output per person*) and our selected predictor variables, we will run linear regressions and conduct inference to check if the results are statistically significant. Since the scatterplot and correlation matrix Figures suggest some strong correlation between the predictor variables, a model that includes all of the predictors will likely contain face multicollinearity. Therefore, we include in our regression model the two predictors which are not strongly correlated with each other, yet still correlated with the response variable in the opposite direction. The three models are below. Because all of our variables have significant autocorrelation within at least 20 lags, we manage to find a good fit of the regression model that could capture and remove the trend in the model. Model 2 includes single time variable and model 3 is a polynomial model with time variables raised to the power of five.

Model 1: $Y= b_0 + b_1X_1 + b_2X_2$

Model 2: $Y= b_0 + b_1X_1 + b_2X_2 + b_3t$

Model 3 (Polynomial model): $Y= b_0 + b_1X_1 + b_2X_2 + b_3t + b_4t^2 + b_5t^3 + b_6t^4 + b_7t^5$

Transform regression models

We first conduct a Durbin-Watson test to check for the presence of autocorrelation at lag 1 in the residuals from our regression models. The null hypothesis states that there is no autocorrelation within our model and the alternative hypothesis suggests there is positive autocorrelation.

Durbin-Watson test

H0: r=0 (no autocorrelation)

H1: r>0 (positive autocorrelation)

To detrend and eliminate the autoregressive errors embedded in our multiple regression models, we try three transformed models using Hildreth-Lu, first difference, and Cochrane-Orcutt methods. Then, we conduct DW-tests again to check if the autoregressive errors are eliminated. Finally, we conduct a second transformation based on the existing transformation model we have to further remove the autoregressive errors. After checking the histogram and QQ-plot of the transformed model, we use a log of response variable in our final model to control for the skewness and outlier effect. After fitting the model, we check the rest of the assumptions for our linear regression, which includes that there is no autocorrelation in errors and normality of errors; we do this through using ACF plot, Durbin-Watson test (both positive and negative), and plotting a QQ-plot and histogram. The two models are below:

Model 4 (transformed model): $Y' = b_0 + b_1X_1' + b_2X_2'$

Model 5 (second transformed model): $\log(Y)'' = b_0 + b_1X_1'' + b_2X_2''$

ARIMA Method

ARIMA Model

In this part, we explore *real output per person* with the ARIMA model, which may possibly include autoregressive terms, moving average terms, and differencing operation. We aim to use ARIMA to help us to achieve better forecasting results.

Since all of our data are seasonally adjusted, the seasonal ARIMA models are not considered in this part. For the non-seasonal ARIMA models, the elements are specified in order (AR order, differencing, MA order). The three items that should be considered to determine the first guess at an ARIMA model are: the Unit Root Test, the ACF and the PACF.

Phillips-Perron Unit Root Test

H0: a time series is integrated of order 1

H1: a time series is stationary

In this case, the p-value of the Phillips-Perron Unit Root Test is 0.9831, so we cannot reject H0, and conclude that the time series is not stationary. For the ARIMA model, the parameter differencing is at least larger than 1.

ACF and PACF

The ACF and PACF plots are considered together to help determine the AR order and the MA order. But we will also consider other combinations of the order to select the best one based on Estimates of $\sigma 2$ (Error variance) and AIC (Akaike Information Criterion).

ARIMA Model with predictors

We will also consider ARIMA model with predictor based on the ARIMA model we have selected in the previous part. Based on the regression parts stated before, we will avoid collinearity by only taking *labor share* and *unit labor costs* into consideration. Our model selection process for ARIMA is as follows:

1) Select the best ARIMA Model based on Estimates of σ2 and AIC
2) Add *labor share* and *Unit Labor Costs* in the model selected in part 1)
3) Select the model from parts 1) and 2) based on estimates of σ2, AIC, residual plots and Ljung-Box Q Test.

VAR model

Since arima model could only predict only one variable, we consider using Vector Autoregressive Models (VAR) to predict all variables. The unit roots were tested in previous stage, so we do not need to do the test again. To determine the order p, we directly use "VARselect" function in R. Then, we will fit the data with "VAR" function and test the Granger causality:

H0: Real Output Per Person do not Granger-cause employment, implicit price deflator, labor share unit labor costs and unit nonlabor payments.

H1: Real Output Per Person do Granger-cause employment, implicit price deflator, labor share unit labor costs and unit nonlabor payments.

After that, we will also check the stability of our selected VAR model and make the 30 steps further forecast.

4) Results

Regression models

As shown in Table 1, the two regression models that contain time variables all indicate that the time variables are statistically significant, which suggests that the variables are time series variables and that there are significant time trends in our models. As we add more time variables to higher power in the model, the goodness of fit of the model increase from 0.95 to 0.99.

However, the DW-test shows a small p-value, meaning that despite the high value of goodness of fit, we cannot infer too much from these models because of autoregressive errors embedded in the regression. When plotting the residuals of the models, Figures 13 through 15 show that the residuals have some clustered patterns and, therefore, have positive autocorrelation.

Transform regression models

For transform model 4 (which uses three methods), Table 2 shows that the coefficients of both of the predictors are significant. However, DW-tests of all three of them still have p-values less than 0.05. So, the null hypothesis of the DW-test is rejected, which means that there is evidence of autocorrelation in the model, suggesting that none of these three methods were able to remove the autoregressive errors. Figures 16 through 18 and 19 through 21 show the plot of residuals and ACF of the transformed models with the three methods respectively. The ACF of residuals are significant at many lags. Figures 22 through 24 show the histogram of the three transformed models, suggesting that there are outliers outside of the overall pattern of the distribution, since the histograms' tail positively skews to the right. Table 3 shows the results from the second transformation and with log on the response variable instead. The Cochrane-Orcutt methods suggests a significance for both of the predictors, while the Hildreth-Lu method suggests that only the unit labor cost is significant, and the first difference method suggests that only labor share is significant. As shown from figure 25, 29, 33, the residuals for all three second transform models in the plot indicate that the residuals are rather independent. The DW-test for positive autocorrelation also has p-value of 1, which shows that there is no positive autocorrelation.

However, the ACF of residuals from Figures 26, 30, and 34 all indicate that there is evidence of strong negative autocorrelation at lag 1 for all three methods. Indeed, the DW-test for negative autocorrelation has p-value that is less than 0.05, suggesting that there is negative autocorrelation at lag 1 in the second transform models. Figures 27, 28, 31, 32, 35, and 36 show the histogram and QQ-plot for all second transform models with three methods. Only with the Cochrane-Orcutt method, the residuals have relatively normal distribution that meet the assumption of normality for conducting linear regression.

ARIMA

Figure 37~40 show that, when d=1, the order (2,1,0) may be best and, when d=2, the order (0,2,1) may be best. Based on Table 4, we can see the order (0,2,2) has the smallest estimate of σ2 and the order (0,2,3) has the smallest AIC. However, the difference between these indicators of order (0,2,1) and the smallest ones are approaching 0, and the ARIMA (0,2,1) is much more simple than the other two models. So, we decide to choose the ARIMA (0,2,1).

Then, based on Table 5, adding *labor share* and *Unit Labor Costs* into the ARIMA Model can promote the performance of the model significantly, as the estimations of σ2 and AIC are much smaller than the original model. Figure 41 indicates that after adding *predictors* - though the model has some significant autocorrelation in the residuals, which means the prediction intervals may not provide accurate coverage - the residuals are not significantly different from white noise. We can also conclude from Figure 42 that after adding predictors, most of the p-values for the Ljung-Box Q test all are well above 0.05, indicating non-significance. So our final model is ARIMA (0,2,1) adding *labor share* and *Unit Labor Costs.*

VAR

In table 6, we can see the best order is about 2 or 3. We pick 3 since the best AIC order is 3. We fit the VAR model with the function “var=VAR(y, p=3)”. Table 7 shows a summary of the model. Table 8 shows the results of the Granger causality test. From the test, there is no such

Granger-cause between the response variable and predictors. So, the next step is to test the stability of the model. Figure 43 demonstrates that every variable is stable in our VAR(3) model. Thus, we can say that our VAR(3) model is well suited to predict the variables.

5) Discussion and Conclusions

Regression model selection

The regression models, which include the time variables, do not adequately capture the trend in our model, and cannot be reliably used to forecast future values. The transform models with all three methods still contain autoregressive errors, which contradicts with the assumption of residuals independence for linear regression; therefore, these are also inadequate. The second transform indicates an unexpected negative significant autocorrelation, which makes the second transform model not adequate as well.

ARIMA

We use the selected ARIMA model to do forecasting in two scenarios:

1) Drop the last 7% of the data from the model, then forecast and verify

Since all of our data has 291 observations, we use the first 270 observations to fit the model and then doing forecasting given the last 21 observations of *labor.share* and *Unit Labor Costs*. Figure 44 shows the results of the forecasting. The red points are the real values while the black points are predictions using our ARIMA model. We can see that the difference between the prediction and the real value are mostly approaching 0, showing our model performs quite well, and there doesn't exist overfitting problems. Figure 45 shows the residuals obey the normal distribution. So, all the graphs support the assumption that there is no pattern in the residuals.

2) Forecast 3 years into the future using all of the data

Since we do not have future data of *labor.share* and *Unit Labor Costs*, we use the mean value of the last five years to estimate these two predictors in the next 3 years. Figure 46~47 shows the result of our forecasting; the red points are the predictions, and the blue dashed lines are the interval within standard errors. From Figure 47, we can deduce that the trend is roughly consistent with previous years. It is reasonable for people to use ARIMA(0,2,1) with *labor share* and *Unit Labor Costs* to predict *real output per person.*

VAR

Using the adequate VAR(3) model to predict the variables, we can predict 30 steps ahead of data with 0.95 CI. Figure 48 is the fanchart, which is the graph of our predictions. And the table 9 show the lower bound, upper bound and 95% CI of the predictions. VAR is really a useful model to do the multivariable time series forecasting.

Final Conclusion

To conclude, in this project, we explored 6 indicators in Business sector, analyzed every individual variable and the relationships between response variable and predictors. To achieve our goal, we finally built regression models, ARIMA models and VAR model to interpret how selected parameters influence the *Real Output Per Person* in Business Sector.