Predicting the Next-Quarter Corporate Profit with Vector Autoregression

Chenning Xu

1   Introduction

In capitalism, profits motivate firms to employ labor and produce output for sale (Veblen 1904). Due to the fundamental uncertainty of the future, firms expect future profits and make investments accordingly, often disproportionately influenced by recent sales and profit conditions (Keynes 1936). Changes in current and expected future corporate profits significantly affect economic outlooks. Additionally, the valuation of equity in firms is often based on the discounted value of future profits. Lenders extend loans to firms based on their earnings capacity, and the soundness of these loans depends on the profits firms realize in the future (Minsky 1982). Therefore, profit is seen as one of the most critical concepts in economics and finance. The Kalecki-Levy profit equation is an accounting identity that shows how aggregate corporate profit is determined by economic variables such as investment and government deficits (Levy, Farnham, and Rajan 2008). This project aims to use the variables in the Kalecki-Levy equation to build a predictive model for the next-quarter corporate profit.

2   Theoretical Background

2.1   Kalecki-Levy Profit Equation

Levy, Farnham, and Rajan (2008) presented a stylized Kalecki-Levy equation for open economies (signs of RHS variables are adjusted here according to the National Income and Product Accounts (NIPA) accounting convention):

$$
\begin{aligned}
\text{Profit after tax} = \ & \text{Net Investment} \\
& - \text{Personal Saving} \\
& + \text{Foreign Saving} \\
& - \text{Government Saving} \\
& + \text{Net Dividend}
\end{aligned}
\tag{1}
$$

The Kalecki-Levy equation, as an accounting identity, is tautologically true. Here is the proof:

$$I = S = PS + CS + FS + GS$$

$$I = PS + (CP - ND) + FS + GS$$

$$CP = I - PS - FS - GS + ND$$

where I is net investment (gross investment - depreciation), S is net saving, PS is personal saving, CS is

corporate saving, *FS* is foreign saving, *GS* is government saving, *CP* is net corporate profit, and *ND* is net dividend (total dividend excluding dividends paid to firms). Since in the NIPA account, *Foreign Saving* is recorded as a negative figure, after flipping the sign of it, we get Equation 1. For technical details, please refer to BEA (2024).

Since the above analysis is purely an accounting analysis, it does not imply any behavioral relations or causality. However, Minsky argued that it is the RHS that determines corporate profit, not vice versa. The straightforward explanation for investment causing profit is that firms cannot control future profit but can determine how much they invest in the next quarter based on their expectations of future economic conditions. For a detailed analysis, please refer to Minsky (1986, Chapter 2).

2.2   Research Motivation

Given the Kalecki-Levy framework for profit determination, a key motivation of this study is to explore how the relationships defined in these identities might serve as practical tools for prediction, rather than merely theoretical constructs. While the original equations hold by definition, leveraging them in a predictive model allows us to investigate whether observable trends in the determinants of profit—and profit itself—can reliably signal changes in future corporate profit.

For example, does a consistent increase in net investment over the past year suggest momentum that could continue into the next quarter, thereby boosting profits? Similarly, could a high personal saving rate over several quarters indicate an upcoming decline, indirectly increasing corporate profit? Shifting focus to prediction invites us to examine not only if these relationships hold as identities but if they possess empirical stability and value in forecasting applications.

Although prior research, such as Trofimov (2022), utilized panel VAR to assess the impact of profit determinants on margins—finding a positive effect from trade surpluses and a negative one from government deficits—such studies have not applied these equations in a forward-looking, predictive context. By employing a rolling-window Vector Auto-Regression (VAR) model, this project specifically aims to predict corporate profit for the next quarter. This method leverages both historical values of profit itself and its key determinants, building on the identity to test its predictive utility, as discussed in the next section.

3   Data

The quarterly and annual series of corporate profit, personal saving, foreign saving, government saving, and net dividend are collected from NIPA (https://www.bea.gov/itable/national-gdp-and-personal-income). The monthly series of the S&P 500 price and earnings are collected from Robert Shiller (http://www.econ.yale.edu/~shiller/data.htm). Currently, only the quarterly series of NIPA is explored.

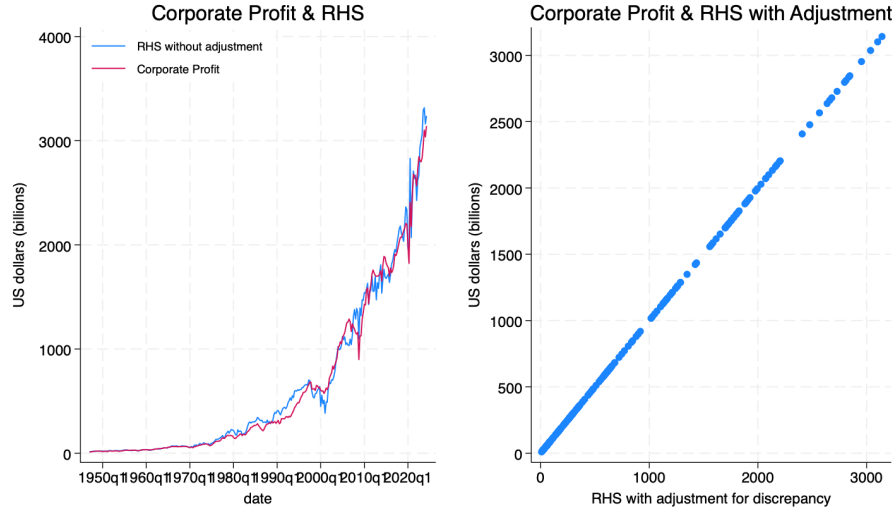## 3.1 Exploratory Data Analysis: Identity

Figure 1: Identity between Corporate Profit and RHS

This graph shows that the Kalecki-Levy equation holds true in the NIPA. The left graph shows the close relationship between corporate profit and the sum of the RHS. The right graph shows the exact identity between corporate profit and the RHS after adjusting for the statistical discrepancy term in the NIPA. Exploratory data analysis is presented below.

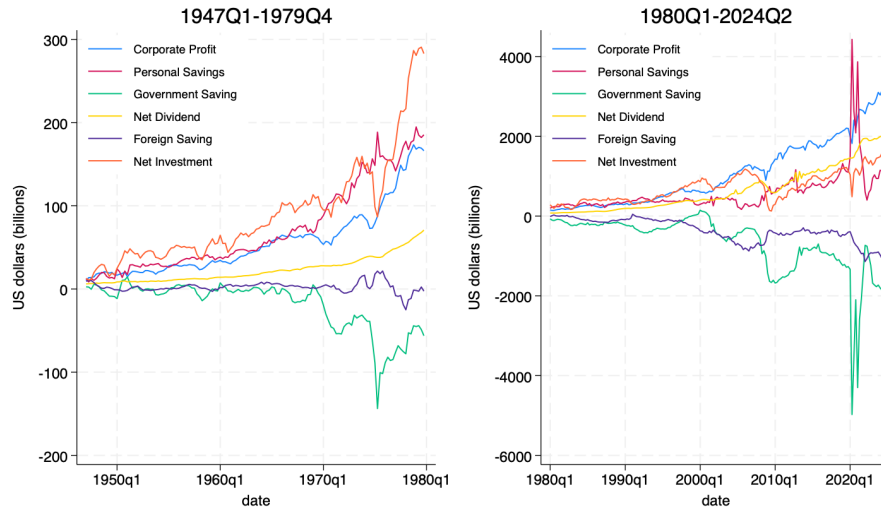## 3.2 Exploratory Data Analysis: Stationarity

Figure 2: Time-Series of the Variables in Kalecki-Levy Equation

This graph shows that all six variables in the Kalecki-Levy equation trend with time and are therefore not stationary. For time series analysis, detrending the variables is necessary.
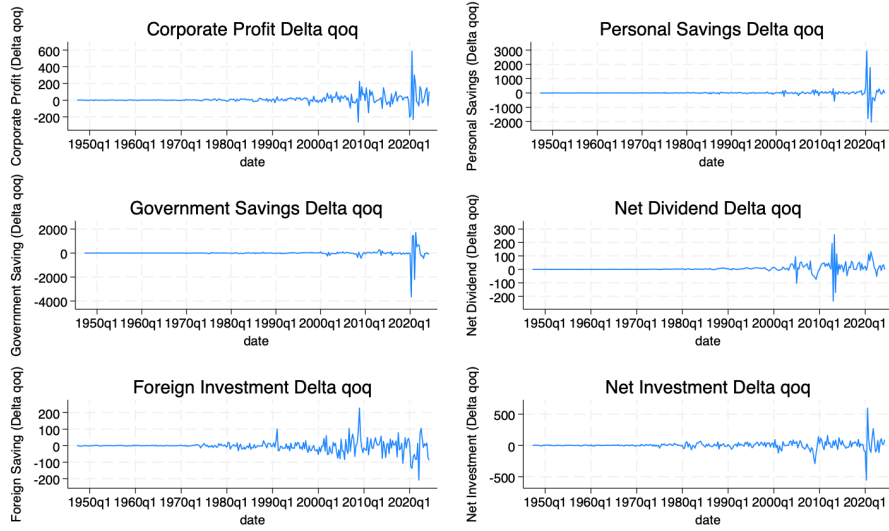

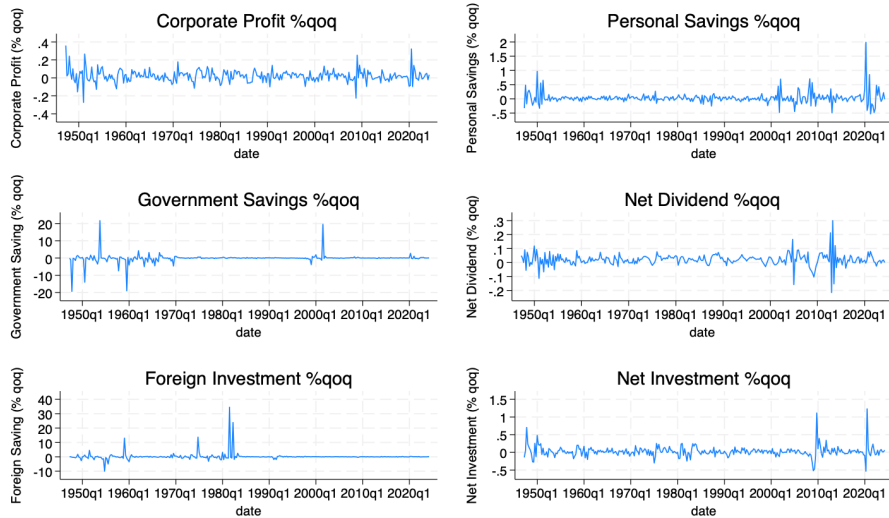Figure 3: Detrending the variables - Level change QoQ

Figure 4: Detrending the variables - % change QoQ

The results of Phillips-Perron unit root tests reject the null hypothesis with p-values under 0.001 for all six variables in both approaches, indicating that they do not follow a random walk. No obvious trend could be observed in both graphs. However, two approaches to detrending come with their advantages and disadvantages.

Figure 3 shows the quarter-on-quarter difference of the six variables. For some variables, such as corporate profit, net dividend, foreign saving, and net investment, there appears to be increasing standard deviation.

This approach preserves the identity of the six variables, as taking differences does not alter the relative magnitude. Additionally, the magnitude of these variables is preserved.

Figure 4 shows the quarter-on-quarter percentage change of the six variables. there appears to be no increasing standard deviation. However,this approach produces highly volatile outcomes, with some changes over 1000%. And when these variables change signs, the percentage change will always be negative, not necessarilly indicating a decline of magnitude. Besides, the information of Kalecki-Levy identity is lost. If the relative magnitude of the RHS variables changes over time, using percentage change may not capture structural shifts.

The problem of increasing standard deviation and difficutlty of capturing structural change could be mitigated by using a smaller training sample in the later rolling-window VAR tests. However, the problem of high volatiy and sign changing in the percentage change approach remains an issue.

4 Methodology and Analysis

4.1 VAR and Granger Causality Tests

Vector autoregressions(Békés and Kézdi 2021, 505) are performed on the entire dataset using both the level difference and percentage difference variables, followed granger causality tests("Granger Causality" 2024). The results of using percentage change variables, with eight lags of endogenous variables and time as exogenous variable are presented here. The model in math would be expressed as:

$$p_{cp,t} = \alpha_{cp} + \sum_{i=1}^{4} \left(\beta_{cp,cp,i} p_{cp,t-i} + \beta_{cp,ps,i} p_{ps,t-i} + \beta_{cp,gs,i} p_{gs,t-i} \right. \tag{2}$$
$$\left. + \beta_{cp,ni,i} p_{ni,t-i} + \beta_{cp,fi,i} p_{fi,t-i} + \beta_{cp,nd,i} p_{nd,t-i} + \gamma_{cp} \cdot \text{date}_t + \epsilon_{cp,t} \right)$$

Where: α_{cp} is the intercept for the equation; $\beta_{cp,cp,i}, \beta_{cp,ps,i}, \beta_{cp,gs,i}, \beta_{cp,ni,i}, \beta_{cp,fi,i}, \beta_{cp,nd,i}$ are the coefficients for the lagged values of the endogenous variables at lag i; γ_{cp} is the coefficient for the exogenous variable (date); $\epsilon_{cp,t}$ is the error term (residual) for $d_{cp,t}$. Due to the length limit and our narrow interest on predicting corporate profit, the equations for other variables are not shown here.

Table 1: VAR Results

Predictor	Coefficient	Predictor	Coefficient
L.c__ps	0.084** (0.033)	**L4.c__ps**	-0.043** (0.021)
L3.c__ps	0.051** (0.023)	**L2.c__gs**	0.007** (0.003)
L4.c__gs	0.002* (0.001)	**L5.c__gs**	0.002* (0.001)
L7.c__ni	0.044** (0.020)	**L2.c__fi**	-0.002*** (0.001)
L3.c__fi	0.002** (0.001)	**L5.c__fi**	0.002** (0.001)
L6.c__fi	0.001* (0.001)	**L7.c__fi**	0.001* (0.001)
L.c__nd	-0.206* (0.112)	**L2.c__nd**	-0.153* (0.087)
L4.c__nd	0.185* (0.109)	**date**	-0.000 (0.000)
** __cons **	0.028*** (0.009)	R2	0.2673

Robust std. err. in parenthses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Only the significant coefficients are presented here. The notation L# indicates the number of quarters in the past for the lagged percentage changes. For example, L3.c_ps represents the percentage change in personal saving from three quarters ago.

The results are somewhat complex to interpret. Lagged changes in corporate profit show no correlation with changes in next-quarter profit. Lagged changes in personal saving appear positively correlated with profit changes in the short run but exhibit a negative correlation over a longer horizon. Lagged changes in government saving (typically negative, indicating a deficit) and net investment display positive correlations with profit changes after a delay of at least one year. Foreign saving (usually negative, reflecting a current account deficit) is negatively correlated with profit change in the short run but shows a consistent positive correlation over a longer period. Finally, lagged changes in net dividends seem negatively correlated with profit change in the short run, switching to a positive correlation over a longer horizon.

Given that this model includes six variables with eight lags each (a total of 48 lagged variables), caution is needed when interpreting p-values, despite the use of robust standard errors. The model accounts for approximately 26% of the variation in next-quarter profit change; however, without out-of-sample testing, the risk of overfitting remains.

The Granger causality test (not reported) suggests that, except for net investment, we can reject the null hypothesis that the independent variables do not Granger-cause percentage changes in corporate profit. Interestingly, none of the variables Granger-cause government saving, indicating that it may behave autonomously within our sample.

4.2 Rolling-Window VAR

To mimic real-life prediction process, rolling-window var tests are performed, using smaller sets of training data to predict the next-quarter corproate profit change. This approach mitigates the problem of increasing standard deviation and the risk of structureal changes. Results of tests using both level and percentage chaneg data are reported here. The math formula for the VAR using level difference and more tests of other specifications can be foud in the robustness check section.

Table 2: Rolling-Window VAR

	R2
R2 for level 4-80	.4872432
R2 for control level 4-80	.1167403
R2 for pct 4-80	.1888433
R2 for pct control 4-80	.0173584
R2 for level 8-100	.6149069
R2 for control level 8-100	.1208536
R2 for pct 8-100	.2844621
R2 for pct control 8-100	.019117

Table 4 presents results from two rolling-window VAR models and their corresponding control models, which include only lagged changes in corporate profit and time. The "Level 8-100" model uses level differences, 8 lags, and a 100-observation training sample. The R^2 values for the "Level 8-100" and "Level 4-80" models are notably high, at 0.615 and 0.487, respectively, indicating that these models explained over 61% and 48%

of the variation in next-quarter profit changes. Both models significantly outperformed their control models.

Models based on percentage changes underperformed compared to those using level changes, potentially because the structural information from the right-hand side (RHS) variables enhances predictive power, or because percentage changes are too volatile for effective training. Nonetheless, the percentage change models still significantly outperformed their control models. Interestingly, the level control models performed much better than the percentage change control models, suggesting that level differences may indeed provide an advantage in model training.
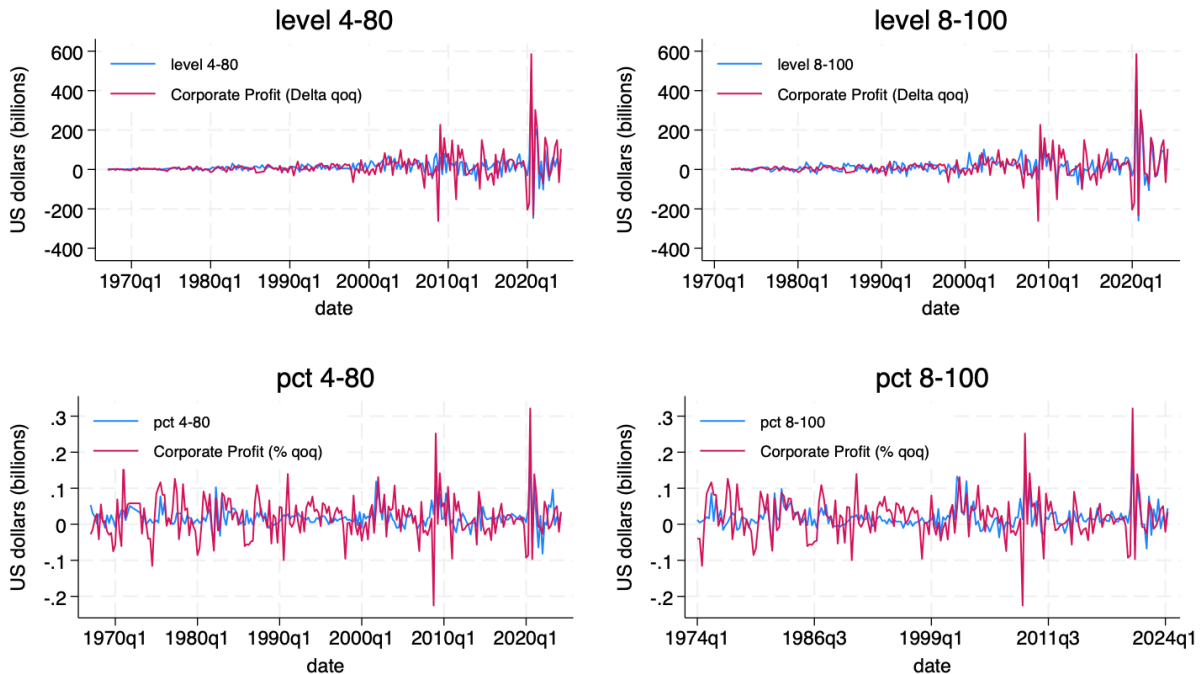

Figure 5: Prediction Series

Figure 5 displays the prediction series of the four models in Table 4 alongside the actual level and percentage change in corporate profit. The two level models appear to capture the increasing standard deviation effectively, while predictions from the percentage models are consistently too stable compared to the actual series. Since predictions of level changes can be easily transformed into percentage change by deviding the prediction by the previous level of corproate profit, the rolling-window VAR using level differences thus seems preferable to using percentages. Further analysis of the level models will be conducted in the robustness check section.

5   Robustness

The above section has demonstrated the considerable predictive power of VAR using variables in the Kalecki-Levy equation, especially when using level difference. This section will present more discussion on the level difference VAR model.

The math formula for VAR using level difference is similar to Equation 2 but it could not have all the five

RHS variables due to perfect multi-colinearity. The models reported in Table 4 dropped net dividend for its relative stability. The math formular for the two models can be expressed as:

$$
\begin{aligned}
d_{cp,t} = \alpha_{cp} + \sum_{i=1}^{4} \big(& \beta_{cp,cp,i} d_{cp,t-i} + \beta_{cp,ps,i} d_{ps,t-i} + \beta_{cp,gs,i} d_{gs,t-i} \\
& + \beta_{cp,ni,i} d_{ni,t-i} + \beta_{cp,fi,i} d_{fi,t-i} + \gamma_{cp} \cdot \text{date}_t + \epsilon_{cp,t} \big)
\end{aligned}
\tag{3}
$$

Surprisingly, the granger causality test could not reject the null hypothesis for any of the four variables at 5% significance level. It is possibly due to the heteroskedasticity from increasing standard deviation and the following increased robust standard error. Here another two model using level difference will be presented as a robustness check. These two models will drop net investment instead of net dividend, since in Table 4 net investment failed to show granger causality for next-quarter net corporate profit.

Table 3: Rolling-Window VAR

	R2
R2 for alternative level 4-80	.5018473
R2 for alternative level 8-100	.5924271

The resutls are very similar to the two models in Table 4. Still, level 8-100 does better than 4-80, and much better than the two models using pct changes. To show the latter, here the predicted level changes are transformed into pct changes.
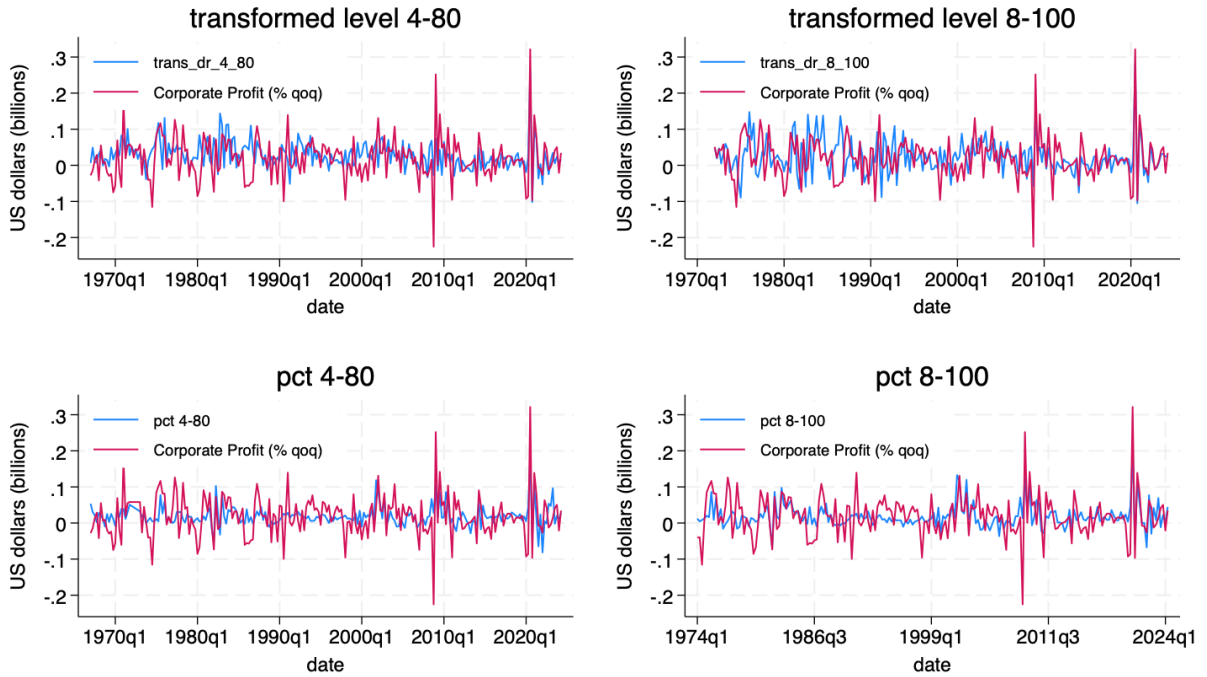

Figure 6: Comparing transformed level prediction to pct prediction

This comparison shows clearly that the level prediction is capturing more variation of changes in profit than percentage predictions. This is confirmed by the statistics below.

Table 4: Comparison between level and percentage prediction

	mean	sd	min	max
transformed level 4-80	.0277941	.0376068	-.1017144	.2615596
transformed level 8-100	.0227857	.0478203	-.1050153	.2817062
pct 4-80	.0187013	.0272969	-.0813488	.1781538
pct 8-100	.018449	.032276	-.0681762	.1937686

6 Conclusion

This project has conducted VAR and granger causality tests to investigate whether the variables in the Kalecki-Levy equation have lead-lag relations, which turned out to be complicated and have no straightforward intepretation. This project designed rolling-window VAR models to mimic real-life prediction process, and gained considerable success. It has shown that using level difference is much more effective than using percentage difference in rolling-window VAR prediction.

The high R^2 of the rolling-window VAR models indicates considerable predictive power for such a model. If simply analyzing past data can yield positive results, combining this approach with external forecasts and macroeconomic insights could produce an even more effective predictive model. The findings from this study demonstrate that approaching corporate profit from a *macro perspective* holds great potential compared to the bottom-up approach, which merely aggregates the expected profits of individual firms.

This project has some limitations. For example, by using a rolling-window model, there is limited information on *how* the lagged variables predict next-quarter corporate profit, aside from their overall effectiveness. Additionally, this project can extend to building models to predict next-year corproate profit, or to predict the next-next quarter corporate profit, which the strength of VAR allows.

References

BEA. 2024. "NIPA Handbook: Concepts and Methods of the u.s. National Income and Product Accounts." 2024. https://www.bea.gov/resources/methodologies/nipa-handbook.

Békés, Gábor, and Gábor Kézdi. 2021. "Data Analysis for Business, Economics, and Policy. Higher Education from Cambridge University Press." 2021. https://doi.org/10.1017/9781108591102.

"Granger Causality." 2024. In *Wikipedia.* https://en.wikipedia.org/w/index.php?title=Granger_causality&oldid=1248890167.

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest, and Money.*

Levy, David A, Martin P Farnham, and Samira Rajan. 2008. "Where Profits Come From."

Minsky, Hyman P. 1982. "The Financial-Instability Hypothesis: Capitalist Processes and the Behavior of the Economy."

Trofimov, Ivan D. 2022. "Determinants of the Profit Rates in the OECD Economies: A Panel Data Analysis of the Kalecki's Profit Equation." *Structural Change and Economic Dynamics.*

Veblen, T. 1904. *The Theory of Business Enterprise.* The Theory of Business Enterprise. Oxford, England: Scribners.