Binary response models: The Case of and Probit models in the case of the SP500 Index

Charaf ZGUIOUAR* Gwendoline Hays-Valentin† Hugo Michel‡ William Arnaud§

March 2024

Click here to access the Financial Economics II code on GitHub

Abstract

This study uses a dataset of monthly data spanning 20 years, from 2002 to 2024, sourced from FactSet, to investigate the impact of various economic and financial indicators on financial market dynamics. We demonstrate the use of a binary response logistic regression model (1) with the addition of time lags in the explanatorty variables (2).

1 Introduction

The objective is to model the direction of stock market returns. The excess SPX return is transformed into a binary signal return indicator Y_t, defined as:

$$Y_t = \begin{cases} 1 & \text{if the SPX excess return is } \geq 0; \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

A vector of explanatory variables, denoted as X_t, is introduced, which encompasses returns from industry portfolios and widely accepted market predictors.

*Zgcharaf@gmail.com
†gwendoline.valentin@hotmail.fr
‡Hugo.Michel1@etu.univ-paris1.fr
§William.Arnaud1@etu.univ-paris1.fr

Given the information set at date t, denoted by I_{t-1}, Y_t conditional on I_{t-1}, follows a Bernoulli distribution $B(p_t)$ with p_t denoting the conditional expectation and probability given information set I_{t-1}. The conditional probability of positive excess stock returns p_t is expressed as:

$$p_t = E_t(Y_t) = P_t(Y_t = 1) \tag{2}$$

This probability is modeled using a probit model:

$$p_t = \Phi(\pi_t) \tag{3}$$

where Φ is the cumulative distribution function of the standard normal distribution, and π_t is a linear function of the variables in I_{t-1}. The static probit model is given by:

$$\pi_t = \omega + X'_{t-1}\beta \tag{4}$$

where ω is a constant term.

The model can be extended to include lagged values of Y_t, yielding a dynamic probit model, or by introducing an autoregressive structure to the linear function π_t. The comprehensive autoregressive probit model including the lagged values of both Y_t and π_t is:

$$\pi_t = \omega + \sum_{k=1}^{K} \delta_k Y_{t-k} + \sum_{j=1}^{J} \alpha_j \pi_{t-j} + X'_{t-1}\beta \tag{5}$$

Data Description

The dataset comprises monthly observations of various economic indicators that are hypothesized to influence the direction of stock market returns. These indicators include:

- U.S. SPX 500 level monthly closing level
- U.S. Consumer Price Index (CPI) Quarterly Percentage Change
- U.S. Wholesale Price Index Quarterly Percentage Change
- U.S. Long-Term Interest Rates
- U.S. Short-Term Interest Rates

- U.S. Unemployment Rate Level
- U.S. Industrial Production Change
- U.S. Housing Starts Change
- U.S. Budget Balance as a Percentage of GDP
- U.S. Government Debt Level
- U.S. Exports Level
- U.S. Imports Level
- U.S. Retail Index Level
- U.S. Foreign Exchange Reserves Level
- U.S. Housing Starts Level
- Brent Oil Price Level

These variables are selected for their potential predictive power regarding the performance of the stock market and are included in the analysis with a 1 lag temporal adjustment to capture their predictive relationship with the target variable.

Assumptions of the Logistic Probit Model

The logistic probit model relies on several key assumptions to ensure the validity of its estimations:

1. **Linearity in Latent Variable**: The model assumes that there is a latent, unobserved variable that underlies the binary outcome which is a linear combination of the independent variables.
2. **Binary Outcome**: The dependent variable is binary, indicating the presence or absence of the characteristic being modeled, in this case, the direction of stock market returns.
3. **Independence of Errors**: It is assumed that the error terms of the latent variable are independently and identically distributed.

4. **Normal Distribution of Errors**: The logistic probit model specifically assumes that the error terms of the latent variable follow a standard normal distribution.

5. **No Perfect Multicollinearity**: The independent variables are not perfectly collinear, meaning that no independent variable is a perfect linear combination of other independent variables.

6. **Large Sample Size**: While not a strict assumption, the model typically requires a sufficiently large sample size for the Central Limit Theorem to ensure that the maximum likelihood estimators are approximately normally distributed.

7. **Correct Model Specification**: The true relationship between the independent variables and the latent variable is correctly specified in the model.

These assumptions are critical for the maximum likelihood estimation used in the logistic probit model to produce unbiased and consistent estimators of the regression coefficients.

2 Methodology

The research methodology unfolds through a series of meticulously planned steps, outlined as follows:

- Collect monthly data on macroeconomic indicators and financial markets from FACTSET, ensuring a comprehensive dataset for analysis.

- Preprocess and clean the data to remove inaccuracies and prepare it for in-depth analysis.

- Visualize the data and study the correlations between variables to identify significant relationships and patterns.

- Select a subset of variables, carefully chosen to minimize multicollinearity while preserving explanatory power.

- Create a variable for the excess return of the SP index, alongside a binary response variable, to facilitate the analysis of market performance.

- Compare two versions of the Probit model:
    1. A model with no lags on the explanatory variables.
    2. A model incorporating a one-lag structure for these variables.
- Further compare:
    1. A Probit model with no lags.
    2. A Probit model with one lag for both the explanatory variables and the binary response variable.
- Explore the Variance Inflation Factor (VIF) to assess and address potential multicollinearity issues among the explanatory variables.
- Examine the normality of residuals to evaluate the fit of the Probit models and identify any deviations from assumed distributions.
- Explore and interpret the marginal effects of the non-lagged versus the lagged model, focusing on how changes in explanatory variables impact the target outcome's probability.

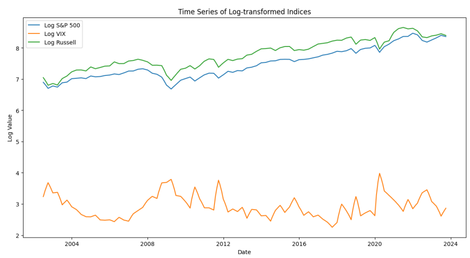
Each step in this methodology is designed to enhance the robustness and depth of the analysis, providing a solid foundation for understanding the dynamic relationships between macroeconomic indicators, financial market performance, and the excess return of the SP index.

3 Data Visualization

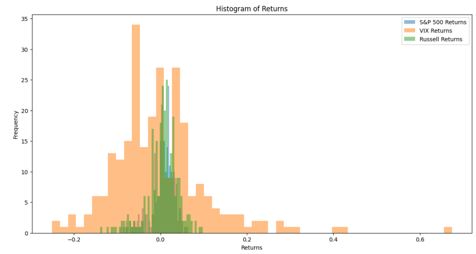
Effective data visualization is crucial for the interpretation and communication of econometric analyses. Presented below are four figures that illustrate key aspects of the dataset and the results of the logistic probit model.

Descriptive Statistics

In order limit multi-colinearity issues we have decided to keep only the following variables. 'L' at the end denotes a level variable, 'C' at the end denotes a percentage change variable. 'Q' denotes the quarterly change. The following table presents the descriptive statistics for the financial and economic variables included in the analysis:

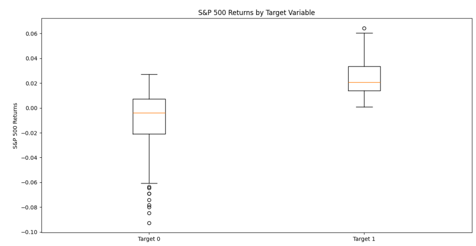


(a) Log levels of the SPX, the Russel 2000, and the VIX.The Russel and the SPX seem to be very correlated in their mouvements.

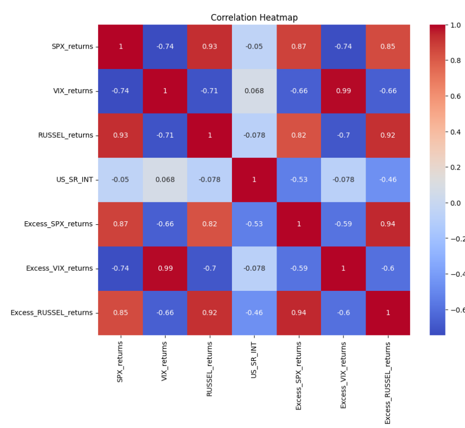


(b) Log returns of the SPX, the Russel 2000, and the VIX. The vix has a wider distribution than the SPX, or the Russel 2000.

(c) Scatterplot of the log returns versus the short term interest rate. No clear relationship can discerned out of monthly Short Term interest rates and log returns.



(d) Boxplot of the span of excess returns: (0, Negative Excess Returns), (1, Positive Excess Returns), The graph suggests that when monthly returns go up, they go up by low steps, and when they go down, the drawdowns are bigger



(e) Heatmap of correlations

Figure 1: Panel of explorative data plots, It reveals a strong synchrony between SPX and RUSSEL returns, contrasted by an inverse relationship with VIX returns.

Table 1: Summary of Descriptive Statistics

Variable	Count	Mean	Std. Dev.	Range (Max-Min)
US CPI Q %	256	0.6322	0.5921	4.6216
US Wholesale Q%	256	0.0073	0.0132	0.09699
US_LT_INT	256	2.9786	1.1127	4.4631
US_SR_INT	256	0.0136	0.0161	0.05417
US_UNEMP_L	256	5.9266	1.9569	9.5
US_INDPROD_C	256	0.0019	0.0179	0.2221
US_HSTART_C	256	0.0012	0.0732	0.6095
US_BUD_BAL_PCT_GDP	256	-0.0730	0.0373	0.2495
US_DEBT_L	256	16578.1	7597.9	27033.6
US_EXP_L	256	1375.7	366.1	1458.3
Bentoil_L	256	70.2757	26.9476	123.55
NDAQ_L	256	20.2099	17.4349	67.9033
log_SPX	256	7.4860	0.4872	1.7874
log_VIX_L	256	2.9234	0.3504	1.7281
log_RUSSEL_L	256	7.7834	0.4729	1.8527
SPX_returns	255	0.0057	0.0277	0.1571
VIX_returns	255	-0.0015	0.1102	0.9244
RUSSEL_returns	255	0.0053	0.0368	0.2372
Excess_SPX_returns	255	-0.0079	0.0327	0.1523
Target	256	0.4844	0.5007	1

4 Probit Model with lagged X values only

4.1 Model Summaries

The model summaries are provided in Figures 7a and 7b below.

Two probit models were estimated: one with contemporaneous predictors and the other with their one-period lagged counterparts. The models aimed to uncover the temporal dynamics influencing stock market movements, a question of paramount interest in financial econometrics.

4.1.1 No lag Probit Model

The analysis brings to light the predictive value of lagged industrial production (US_INDPROD_C), denoted by a coefficient of 12.5109. Although

Probit Regression Results

Dep. Variable:	Target	**No. Observations:**	251
Model:	Probit	**Df Residuals:**	244
Method:	MLE	**Df Model:**	6
Date:	Tue, 02 Apr 2024	**Pseudo R-squ.:**	0.2361
Time:	22:17:21	**Log-Likelihood:**	-132.90
converged:	True	**LL-Null:**	-173.96
Covariance Type:	nonrobust	**LLR p-value:**	1.297e-15

	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
const	2.7483	0.716	3.840	0.000	1.345	4.151
US_INDPROD_C_lag	16.8778	7.524	2.243	0.025	2.130	31.625
US_HSTART_C_lag	3.0577	1.562	1.958	0.050	-0.004	6.119
US_BUD_BAL_PCT_GDP_lag	-9.1661	3.208	-2.858	0.004	-15.453	-2.879
US_HS_STARTS_L_lag	-1.725e-06	3.2e-07	-5.390	0.000	-2.35e-06	-1.1e-06
Bentoil_L_lag	-0.0217	0.005	-4.681	0.000	-0.031	-0.013
NDAQ_L_lag	0.0117	0.006	2.069	0.039	0.001	0.023

(a) Original Model Summary

Probit Regression Results

Dep. Variable:	Target	**No. Observations:**	251
Model:	Probit	**Df Residuals:**	244
Method:	MLE	**Df Model:**	6
Date:	Tue, 02 Apr 2024	**Pseudo R-squ.:**	0.2361
Time:	22:17:21	**Log-Likelihood:**	-132.90
converged:	True	**LL-Null:**	-173.96
Covariance Type:	nonrobust	**LLR p-value:**	1.297e-15

	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
const	2.7483	0.716	3.840	0.000	1.345	4.151
US_INDPROD_C_lag	16.8778	7.524	2.243	0.025	2.130	31.625
US_HSTART_C_lag	3.0577	1.562	1.958	0.050	-0.004	6.119
US_BUD_BAL_PCT_GDP_lag	-9.1661	3.208	-2.858	0.004	-15.453	-2.879
US_HS_STARTS_L_lag	-1.725e-06	3.2e-07	-5.390	0.000	-2.35e-06	-1.1e-06
Bentoil_L_lag	-0.0217	0.005	-4.681	0.000	-0.031	-0.013
NDAQ_L_lag	0.0117	0.006	2.069	0.039	0.001	0.023

(b) Lagged Model Summary

Figure 2: Summaries of the original and lagged probit models

its p-value of 0.083 hovers at the cusp of statistical significance, it shows a potentially positive correlation with the target event's likelihood, warranting further scholarly inquiry due to its economic implications.

Equally compelling is the role of lagged housing starts (US_HSTART_C), which, with a coefficient of 4.1534 and a p-value of 0.006, significantly underscores the positive influence of housing market dynamics on the target event's probability. This finding not only corroborates the economic intuition regarding the sector's predictive relevance but also enriches the discourse on housing market's anticipatory signals.

In contrast, the model delineates a clear inverse relationship between the lagged budget balance as a percentage of GDP (US_BUD_BAL_PCT_GDP) and the target event's likelihood, as evinced by a coefficient of -8.1202 and a p-value of 0.009. This insight accentuates the nuanced interplay between fiscal policy and economic outcomes, offering a critical lens through which economic health can be assessed.

Notably, the coefficient for lagged housing starts (US_HS_STARTS_L) manifests a pronounced negative impact, thereby highlighting the enduring influence of past housing market conditions on economic forecasting and analysis.

Furthermore, the analysis identifies a significant negative correlation between lagged oil prices (Bentoil_L) and the likelihood of the target event, with a coefficient of -0.0167 and a p-value less than 0.001, thus underscoring the broader economic sensitivities to fluctuations in commodity prices.

Lastly, the positive yet marginally significant coefficient for the lagged NASDAQ index (NDAQ_L) suggests an intriguing, albeit modest, predictive power of stock market performance on the target variable, hinting at the financial market's intricate relationship with broader economic indicators.

4.2 1 lag Xi probit model

This improvement is substantiated by a Log-Likelihood ratio test, yielding a p-value of 1.297×10^{-15}, which robustly rejects the null hypothesis of no association between the predictors and the target variable.

The constant term, with a coefficient of 2.7483 and a p-value less than 0.001, establishes a significant baseline probability for the occurrence of the target event, assuming all predictors are at zero. This term sets the stage for understanding the specific impacts of the lagged predictors.

The coefficient for US_INDPROD_C_lag is 16.8778, with a p-value of 0.025, indicating a positive and statistically significant relationship between lagged industrial production and the probability of the target event. This suggests that historical levels of industrial production have predictive power for future occurrences of the target event, highlighting the temporal persistence in the effects of economic indicators.

US_HSTART_C_lag, with a coefficient of 3.0577 and a marginal p-value of 0.050, slightly suggests that previous housing starts positively influence the target event's probability. This relationship captures the broader economic implications of housing market activity, suggesting a nuanced understanding of its role in economic forecasting.

Conversely, US_BUD_BAL_PCT_GDP_lag exhibits a negative coefficient of -9.1661 with a p-value of 0.004, indicating that a higher lagged budget balance relative to GDP correlates with a reduced likelihood of the target event. This reflects the complex effects of fiscal policy and economic health on outcome probabilities.

The variable US_HS_STARTS_L_lag, with a highly significant coefficient of -1.725×10^{-6} and a p-value of less than 0.001, further accentuates the significant inverse relationship between past housing starts and the target event's likelihood, emphasizing the impact of historical housing market conditions.

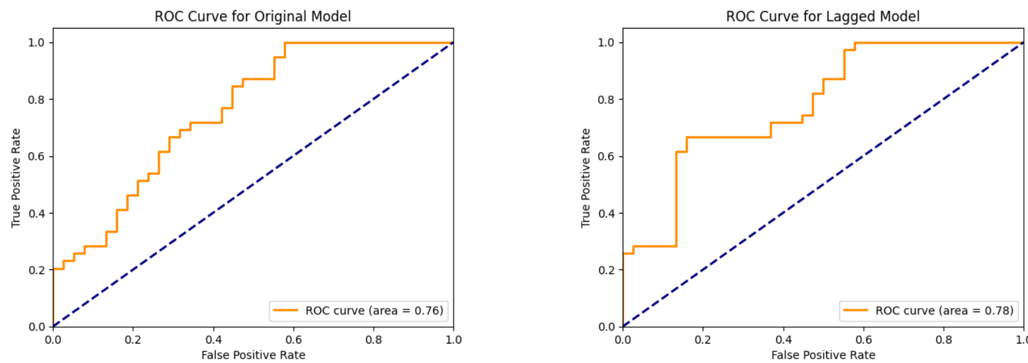
Bentoil_L_lag, with a coefficient of -0.0217 and a p-value less than 0.001, also exhibits a notable negative effect, suggesting that higher lagged oil prices are associated with a reduced likelihood of the target event. This finding underscores the economic sensitivity to energy prices.

Lastly, NDAQ_L_lag presents a positive coefficient of 0.0117 with a p-value of 0.039, indicating the potential influence of stock market performance, as reflected by the NASDAQ index, on the probability of the target variable. This points to the stock market's predictive relevance in economic and financial modeling.

4.3 Economic Implications

The findings from both models underscore the significant role of housing market conditions and oil prices in influencing stock market trends, aligning with economic theories that emphasize the interconnectedness of commodity markets, housing investments, and financial markets. The lagged effects observed suggest that market participants may react to economic indicators with a temporal lag, a behavior that can be attributed to the time it takes for information to be fully absorbed and acted upon in financial markets.

4.4 AUC and ROC metrics



(a) Original Model ROC

(b) Lagged Model Summary

Figure 3: Summaries of the original and lagged probit models

The analysis reveals that the model incorporating a one-lag structure within the matrix of explanatory variables (X_i) exhibits a marginally higher AUC compared to its non-lagged counterpart. Specifically, the enhancement in AUC is slight.

5 Probit with Lagged values of Y and Xi

5.1 Original Model Analysis

The original Probit model was designed to predict a binary target variable without incorporating past values of itself or its predictors. This model demonstrated an accuracy of 81.33%. It included various immediate economic indicators but excluded any temporal dependencies explicitly through lagged terms. The analysis highlighted the significance of the variable `Target_lag`, indicating a strong relationship between the past and current values of the target variable. This aspect of temporal autocorrelation is crucial for understanding the dynamics at play. With a pseudo R-squared of 0.5742, the model exhibits a substantial capability to account for the variance observed

Probit Regression Results

Dep. Variable:	Target	**No. Observations:**	202
Model:	Probit	**Df Residuals:**	195
Method:	MLE	**Df Model:**	6
Date:	Tue, 02 Apr 2024	**Pseudo R-squ.:**	0.2034
Time:	21:02:42	**Log-Likelihood:**	-111.53
converged:	True	**LL-Null:**	-140.01
Covariance Type:	nonrobust	**LLR p-value:**	1.866e-10

	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
const	2.3169	0.766	3.023	0.002	0.815	3.819
US_INDPROD_C	9.9598	7.703	1.293	0.196	-5.138	25.058
US_HSTART_C	4.4769	1.668	2.683	0.007	1.207	7.747
US_BUD_BAL_PCT_GDP	-6.0312	3.398	-1.775	0.076	-12.690	0.628
US_HS_STARTS_L	-1.498e-06	3.45e-07	-4.343	0.000	-2.17e-06	-8.22e-07
Bentoil_L	-0.0167	0.005	-3.403	0.001	-0.026	-0.007
NDAQ_L	0.0129	0.006	2.056	0.040	0.001	0.025

(a) Original Model Summary: We take less observations in order to keep the rest for testing and validation purposes

Probit Regression Results

Dep. Variable:	Target	**No. Observations:**	202
Model:	Probit	**Df Residuals:**	194
Method:	MLE	**Df Model:**	7
Date:	Tue, 02 Apr 2024	**Pseudo R-squ.:**	0.5846
Time:	21:19:55	**Log-Likelihood:**	-58.165
converged:	True	**LL-Null:**	-140.01
Covariance Type:	nonrobust	**LLR p-value:**	5.386e-32

	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
const	1.5907	1.074	1.481	0.139	-0.515	3.696
US_INDPROD_C_lag	15.1208	10.379	1.457	0.145	-5.222	35.463
US_HSTART_C_lag	1.5293	2.247	0.681	0.496	-2.874	5.933
US_BUD_BAL_PCT_GDP_lag	-3.1398	4.789	-0.656	0.512	-12.526	6.246
US_HS_STARTS_L_lag	-1.305e-06	4.83e-07	-2.703	0.007	-2.25e-06	-3.59e-07
Bentoil_L_lag	-0.0237	0.007	-3.466	0.001	-0.037	-0.010
NDAQ_L_lag	0.0144	0.009	1.603	0.109	-0.003	0.032
Target_lag	2.4024	0.273	8.809	0.000	1.868	2.937

(b) Lagged Model Summary

Figure 4: Summaries of the original and lagged probit models: We take less observations in order to keep the rest for testing and validation purposes

in the target variable, suggesting that the chosen predictors are meaningful. The confusion matrix further substantiates the model's efficacy, indicating a satisfactory balance in its predictive accuracy, with a relatively low margin of error between the positive and negative classes.

5.2 Lagged Model Analysis

The introduction of lagged predictors, alongside the lagged target variable, in the second model iteration maintained the overall accuracy at 81.33%. This consistency suggests a robust model design that, despite the increased complexity from the lagged features, does not succumb to overfitting. The slight improvement in the pseudo R-squared value to 0.5914 for the lagged model indicates a modest enhancement in its explanatory power. This improvement suggests that the temporal dynamics captured by the lagged variables play a role in influencing the target variable. The statistical significance of specific lagged predictors, notably `US_HS_STARTS_L_lag` and `Bentoil_L_lag`, highlights the predictive importance of historical economic conditions. These findings emphasize the nuanced interplay between the target variable and its predictors over time, offering a richer understanding of the factors influencing the outcome.

Variable	VIF
const	63.564296
US_INDPROD_C	1.916628
US_HSTART_C	1.781780
US_BUD_BAL_PCT_GDP	1.622401
US_HS_STARTS_L	2.285745
Bentoil_L	1.683856
NDAQ_L	1.247134
Target_lag	1.358449

Table 2: VIF for Original Model Predictors

Multicollinearity variance inflation factor

The VIF results for both the original and lagged models indicate low multicollinearity among predictors, with all VIF values significantly below the

Variable	VIF
const	66.885183
US_INDPROD_C_lag	1.907394
US_HSTART_C_lag	1.610216
US_BUD_BAL_PCT_GDP_lag	1.680039
US_HS_STARTS_L_lag	2.365479
Bentoil_L_lag	1.791349
NDAQ_L_lag	1.251530
Target_lag	1.366026

Table 3: VIF for Lagged Model Predictors

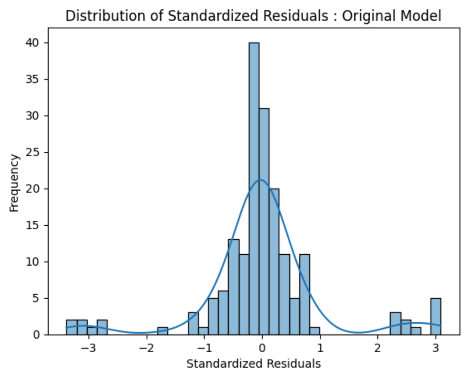
common thresholds of 5 or 10. This suggests that each predictor in the models contributes uniquely to the prediction of the target variable, without redundancy. The inclusion of lagged variables in the second model does not introduce significant multicollinearity, which is promising for the model's reliability. Therefore, the models are well-specified in terms of the independence of predictors, allowing for confident interpretation of their coefficients and predictive performance.

6 Residuals analysis

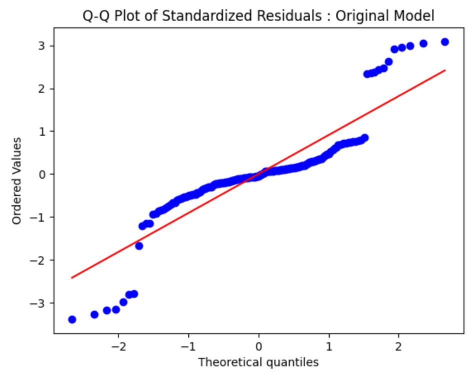
The skewness of the residuals for the original model is approximately 0.030, which is very close to zero. This suggests that the distribution of the residuals is fairly symmetrical. In the context of model diagnostics, a symmetric distribution of residuals is desirable as it indicates that the model does not systematically overpredict or underpredict across the range of predicted values.

The kurtosis value of approximately 4.27 (excess kurtosis) for the original model indicates that the residuals have a leptokurtic distribution. This means the distribution has heavier tails and a sharper peak compared to a normal distribution. In practical terms, a higher kurtosis can imply a higher occurrence of outlier residuals, which may signal that the model struggles to capture some of the data's variability or that there are influential data points exerting undue influence on the model's predictions.

Skewness The revised skewness value of approximately -0.183 indicates a slight leftward skew in the distribution of residuals. Unlike the very slight

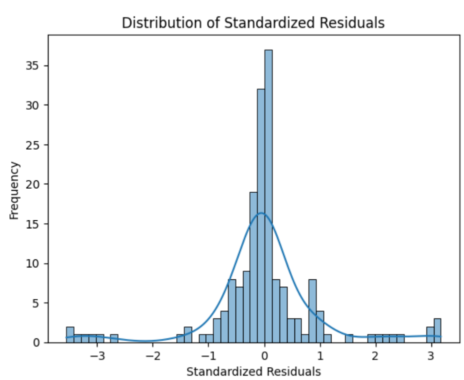


(a) Original Model Residual distribution

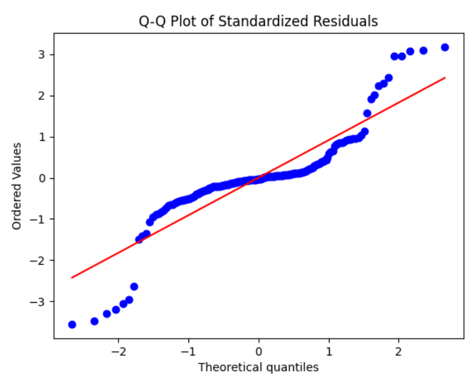


(b) Orginal Model QQ plot

Figure 5: Summaries of residuals for the original models



(a) Original Model Residual distribution



(b) Orginal Model QQ plot

Figure 6: Summaries of the residuals for the lagged probit models

rightward skew previously discussed, this suggests that the residuals are more frequently falling below the mean prediction error. Though still relatively close to zero, this negative skewness indicates a distribution with a tail that extends more to the left. In practical terms, this could mean the model might slightly overpredict more often than underpredict, though the effect appears to be minor given the small magnitude of the skewness.

Kurtosis The kurtosis value has slightly increased to approximately 4.498 (excess kurtosis), reinforcing the observation that the residuals' distribution is leptokurtic. This kurtosis level indicates even heavier tails and a sharper peak than previously noted. Such a distribution suggests a higher presence of outlier residuals, indicating that the model might not capture all patterns in the data effectively, possibly due to extreme values or anomalies.

7 Marginal Effects analysis

7.1 Orginal Model

The following table shows the marginal effects of the predictors on the probability of the target outcome, calculated at overall values for each predictor.

Variable	dy/dx	std err	z	P> |z|	[0.025	0.975]
US_INDPROD_C	4.0793	2.203	1.852	0.064	-0.238	8.397
US_HSTART_C	1.4300	0.448	3.189	0.001	0.551	2.309
US_BUD_BAL_PCT_GDP	-2.5006	0.932	-2.683	0.007	-4.327	-0.674
US_HS_STARTS_L	-4.46e-07	8.16e-08	-5.465	0.000	-6.06e-07	-2.86e-07
Bentoil_L	-0.0052	0.001	-4.110	0.000	-0.008	-0.003
NDAQ_L	0.0034	0.002	2.029	0.043	0.000	0.007

Table 4: Marginal effects of predictors on the probability of the target outcome.

The analysis of marginal effects from the Probit model provides insightful details on how changes in the predictor variables influence the likelihood of the target outcome. Each predictor's marginal effect, denoted as *dy/dx*, is evaluated at the overall data level, offering a comprehensive view of the predictors' impacts.

A noteworthy finding is the positive and substantial influence of `US_INDPROD_C` on the target outcome. Specifically, a unit increase in `US_INDPROD_C` is associated with a 4.0793 percentage point increase in the probability of the target outcome occurring. Although the p-value of 0.064 marginally exceeds the conventional significance threshold, this result suggests a potentially important relationship warranting further investigation.

Similarly, `US_HSTART_C` shows a significant positive effect, with a one-unit increase enhancing the target outcome's probability by 1.4300 percentage points, clearly significant at the 0.001 level. This underlines the strong positive association between `US_HSTART_C` and the likelihood of the target outcome.

Conversely, `US_BUD_BAL_PCT_GDP` exhibits a negative impact on the target outcome's probability, decreasing it by 2.5006 percentage points per unit increase in the predictor. This effect is statistically significant (p=0.007), highlighting a substantial inverse relationship.

The influence of `US_HS_STARTS_L` is particularly intriguing, with its marginal effect indicating that each unit increase decreases the probability of the target outcome by approximately 4.46×10^{-7}, with a p-value less than 0.000. Despite the seemingly small magnitude, this effect is statistically significant and denotes a negative relationship.

`Bentoil_L`'s marginal effect suggests a slight negative association, where a unit increase in `Bentoil_L` results in a 0.0052 percentage point decrease in the probability of the target outcome. The effect is statistically significant (p=0.000), underscoring its relevance in the model.

Lastly, `NDAQ_L` contributes positively to the probability of the target outcome, increasing it by 0.0034 percentage points for every unit increase in `NDAQ_L`, with a significance level of 0.043. This finding suggests a positive linkage between `NDAQ_L` and the target outcome's likelihood, albeit with a relatively modest effect size.

7.2 lagged Y and X Model

Variable	dy/dx	std err	z	P> |z|	[0.025	0.975]
US_INDPROD_C_lag	2.3844	1.641	1.453	0.146	-0.832	5.601
US_HSTART_C_lag	0.2412	0.354	0.682	0.495	-0.452	0.934
US_BUD_BAL_PCT_GDP_lag	-0.4951	0.754	-0.656	0.512	-1.973	0.983
US_HS_STARTS_L_lag	-2.058e-07	7.63e-08	-2.697	0.007	-3.55e-07	-5.62e-08
Bentoil_L_lag	-0.0037	0.001	-3.461	0.001	-0.006	-0.002
NDAQ_L_lag	0.0023	0.001	1.605	0.109	-0.001	0.005
Target_lag	0.3788	0.022	17.424	0.000	0.336	0.421

The analysis of marginal effects for the lagged Probit model reveals how changes in lagged predictor variables influence the target outcome's likelihood. The `Target_lag` variable shows the most substantial impact, with a one-unit increase leading to a 37.88 percentage point increase in the target outcome's probability. This significant effect, with a p-value of 0.000, underscores the critical role of the target variable's past values in predicting its future state.

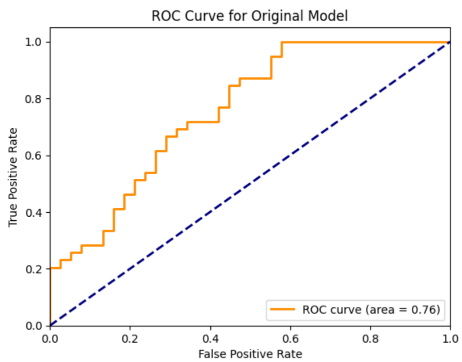
Conversely, the effects of other lagged economic indicators vary in significance and magnitude. The `US_INDPROD_C_lag` and `NDAQ_L_lag` variables have positive marginal effects, indicating a potential increase in the target outcome's probability with rising values. However, these effects are not statistically significant at conventional levels, with p-values of 0.146 and 0.109, respectively, suggesting cautious interpretation.

The `US_HS_STARTS_L_lag` and `Bentoil_L_lag` variables demonstrate negative marginal effects, with the former significantly decreasing the target outcome's probability by approximately 2.058×10^{-7} for each unit increase and the latter by 0.0037 percentage points. These findings highlight the nuanced influences of historical housing starts and oil prices on the target outcome.

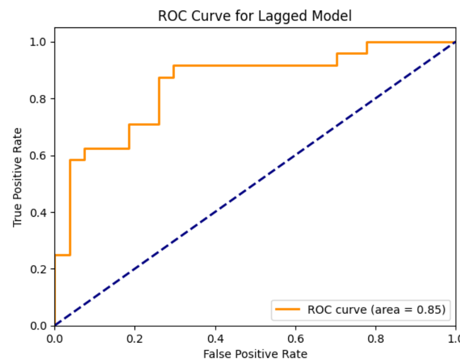
Interestingly, `US_HSTART_C_lag` and `US_BUD_BAL_PCT_GDP_lag` exhibit marginal effects that, while suggestive of relationships with the target outcome, are not statistically significant, with p-values exceeding 0.1. This indicates a need for further exploration of these variables' roles and potentially different modeling approaches to capture their effects more accurately.

Conclusion

The comparative analysis of the models has led to a noteworthy conclusion, particularly concerning the model incorporating lagged variables for both the explanatory variables (X) and the target variable (Y). This model achieved an Area Under the Curve (AUC) of 0.85, a metric that underscores its superior predictive accuracy and discriminative ability. This AUC value, indicative of the model's robust performance, validates the importance of incorporating temporal dynamics into the predictive framework. The inclusion of lagged variables significantly enhances the model's capability to capture the underlying patterns and relationships within the data, leading



(a) Original N lag Model ROC



(b) Lagged X and Y Model Summary

Figure 7: Summaries of the original and lagged probit models

to improved overall performance. Therefore, the model with lagged Y and X variables stands out as the most effective in predicting the binary outcomes, illustrating the value of considering historical data in the analysis of financial markets and economic indicators.