Corporate ALM Deposit Modeling Enhancements for Retail and Commercial Banking

**Jonathan Goldstein**

April 1, 2023

*Disclaimer: The views and opinions expressed on this platform are solely those of the individual posting them and do not necessarily reflect the views or opinions of any bank in particular. Any confidential information contained in this post has been redacted and will not be available to outside parties. The individual posting this content is not acting as a representative of any bank and is not authorized to bind or act on behalf of the company in any capacity. This content is provided for informational purposes only and should not be construed as legal, financial, or other professional advice. Readers should consult with a licensed professional for advice tailored to their specific situation*

**1. ABSTRACT**

In order to provide precise balance sheet forecasting, efficient interest rate risk management, and optimal financial performance, US Banks are dedicated to enhancing its Asset and Liability Management (ALM) and Funds Transfer Pricing (FTP) operations. The proposed project focuses on time series modeling approaches and analytics for Non-Maturity Deposits (NMDs) and other deposit portfolios, which would improve a bank’s corporate ALM deposit modeling capabilities. The project will serve a variety of Treasury and business applications, including acquisition modeling, stressed liquidity analytics, business planning, liquidity management, and stress testing.

**2. PROBLEM STATEMENT**

Like most commercial and retail banking operations, banks must correctly predict deposit behavior and manage interest rate risk, both of which are necessary for implementing successful ALM and FTP policies. In order to improve balance sheet forecasts and NII/EVE sensitivity calculations, Corporate Finance and Treasury must continually enhance its deposit modeling capabilities. In doing so, banks will better understand customer behavior, account decay, average balance per account, and portfolio rate dynamics.

**3. OBJECTIVES**

The project will employ time series modeling methodologies to address the following objectives:

*3.1.* Core Principles:Determine the fundamental principles that should direct the development of deposit models and analytics by comparing various account decay techniques to the bank's objectives.

*3.2.* Data Cleaning: Identify and address data issues such as missing values, inconsistencies, and outliers. This process will also include the development of methodologies for handling account opening and closing events, as these events have a significant impact on the deposit behavior analysis.

*3.3.* Time Series Modeling:Apply time series modeling approaches to estimate rate and average balance dynamics for NMDs and other deposit portfolios while taking into consideration historical data and market trends.

*3.4.* Model/Analytics Development: Determine the precise models and analytics required to generate granular behavioral life assumptions (such as WALs), repricing sensitivity assumptions (such as long-term betas), and scenario-based forecasts of deposit behaviors for NII and EVE sensitivity calculations.

*3.5.* Near-Optimal Solutions: Assess potential near-optimal solutions to the main challenges while taking into account secondary requirements including stress testing (CCAR), liquidity control, business planning, acquisition modeling, and stressed liquidity analytics.

**4. INTENDED OUTCOMES**

The project aims to achieve the following outcomes:

*4.1.* Enhanced Deposit Modeling: Improved understanding of customer behavior, account decay, average balance per account, and portfolio rate dynamics, supporting accurate balance sheet forecasting.

*4.2.* Improved ALM and FTP Processes: Enhanced capability to manage interest rate risk and optimize financial performance through matched maturity FTP, as part of the Financial Performance Management (FPM) framework.

*4.3.* Comprehensive Set of Models and Analytics: Development of granular behavioral life assumptions, repricing sensitivity assumptions, and scenario-based forecasts to be used in the A/L model for NII and EVE sensitivity calculations.

*4.4.* Support for Treasury and Business Applications: Strengthened support for various applications, including stress testing (CCAR), liquidity management, business planning, acquisition modeling, and stressed liquidity analytics.

By implementing these enhancements to corporate ALM deposit modeling, banks will be better equipped to optimize its risk management strategies, and ultimately achieve its business objectives.

**5. CRITICAL PERSPECTIVES**

While the proposed project aims to significantly enhance a bank’s deposit modeling capabilities, it's crucial to address potential concerns and challenges:

*5.1.* Data Quality and Availability: Time series modeling strongly depends on historical data sources. The accuracy and dependability of the models may be impacted by any discrepancies or gaps in the data.

*5.2.* Model Complexity: Developing intelligent models to capture customer behaviors and market trends may result in increased complexity, making it harder for the bank to implement and maintain models.

5.3. Integration Challenges: Integrating the developed models and analytics into existing systems and processes may prove to be a challenge, as it may require significant changes to the bank's processes.

5.4 Model Risk Requirements: The models developed must adhere to all relevant risk committee and compliance requirements, such as stress testing guidelines.

**6. REQUIREMENTS**

In the context of the data provided, several variables can be identified for each type of model. For account decay models, account tenure is a relevant variable, as it captures the length of time an account has been active. Market interest rates also play a role in account decay models, influencing the behavior of customers regarding their accounts. Non-rate macro factors, such as economic indicators or demographic trends, are considered in account decay models as they impact customer behavior. A bank’s relative pricing is another variable that affects account decay models, as customers' decisions may be influenced by the competitiveness of a bank's offerings compared to other institutions.

For average balance models, account tenure is not a significant variable as it does not directly affect the calculation of average balances. However, the average balance models do consider seasonality, as customer balances can fluctuate based on specific times of the year, such as holidays or seasonal spending patterns.

For rate modeling, market interest rates are a crucial variable as they directly impact the rates offered on accounts and influence customers' decisions. However, non-rate macro factors are not considered in rate modeling, as the focus is primarily on the impact of market interest rates. A bank’s relative pricing is also not typically included in rate modeling.

In developing models and analytics for FTP and ALM, certain principles guide the process. Regarding the forward versus backward-looking approach, for FTP at a bank, a backward-looking perspective is considered appropriate given the early stages of implementation. On the other hand, for ALM, a forward-looking approach is necessary as the inputs should support scenario testing for NII (Net Interest Income) and EVE (Economic Value of Equity) sensitivity calculations.

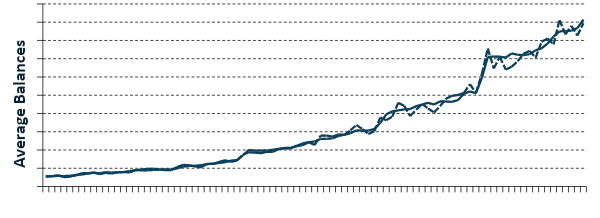
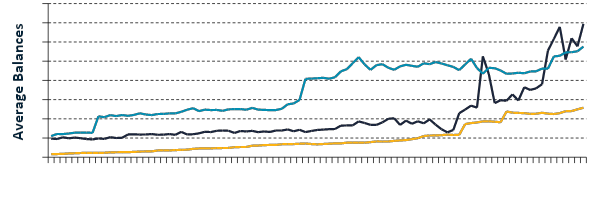
In terms of the drivers tested, both FTP and ALM models consider time on file and market interest rates as key factors. Additionally, for FTP, there is consideration for seasonality and macro factors to align with stress testing requirements. Pricing variables and other management variables are also being considered to further align with stress testing and pricing/elasticity purposes.

Regarding segmentation, there is a desire for some level of granularity in both FTP and ALM models. However, specific implementation considerations exist for each area, taking into account the unique requirements and characteristics.

**7. AVERAGE BALANCE APPROACH**

In most cases, time series modeling is commonly employed for average balance modeling, although panel regression shows potential in certain situations. For average balance modeling, there are three options to consider: time series models, panel regression with vintages, and panel regression with categorical grouping.

Time series models (Figure 1) focus on predicting average1 balances at the portfolio level using only time-varying independent variables. They have the ability to capture the effects of time-varying factors such as wholesale rates, macroeconomy, and bank pricing. These models yield differentiated outcomes under different conditions and can be applied to new acquisitions2. Time series models are typically modeled and forecasted on a monthly basis and are most suitable when balance variations are primarily driven by time-varying factors, rather than time on file or monthly vintage3.



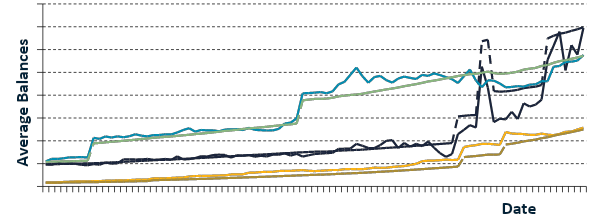
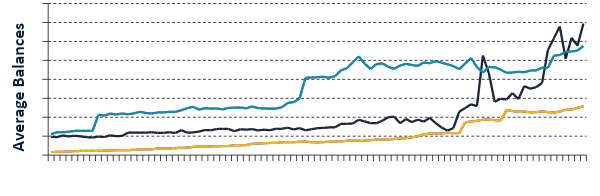
**Time Series Modeling**

*Figure 1*

Panel regression with vintages (Figure 2), on the other hand, predicts average balances for each vintage in the time series. This approach is effective at capturing the impact of time on file and can be applied to new acquisitions. However, it does not incorporate categorical factors directly in the model and requires segmentation within a single equation. Panel regression with vintages is particularly suitable when average balances are strongly and consistently driven by time on file

.

**Panel Regression**



*Figure 2*

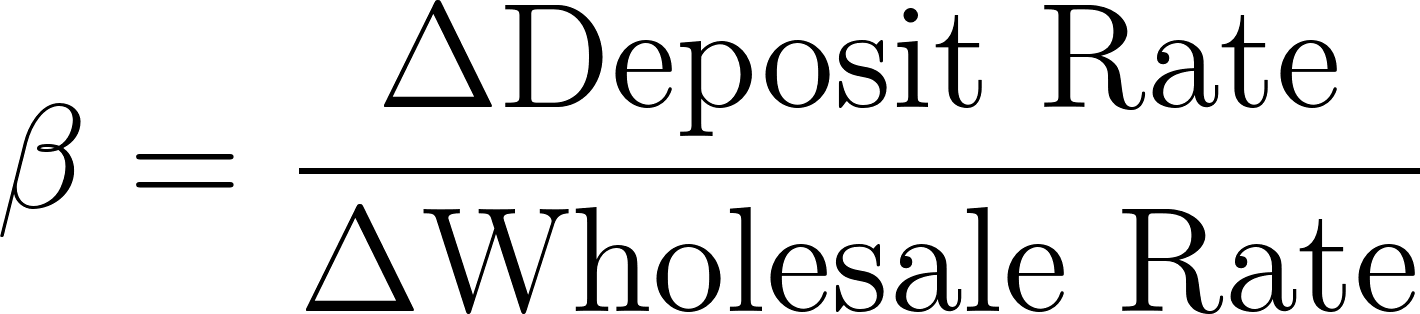
Panel regression with categorical grouping predicts average balances for multiple time series, grouped by categorical variables. It can capture the effects of time-varying factors and incorporate categorical factors without the need for segmentation. This approach is beneficial when dealing with homogenous sub-segments that are likely driven by similar factors, albeit at different strengths.

Choosing the most appropriate approach depends on the specific requirements of the analysis. Time series models are suitable when time-varying drivers are the primary focus and when sub-segments are homogeneous. Panel regression with vintages is preferred when time on file plays a significant role and new acquisitions do not need to be considered. Panel regression with categorical grouping is valuable when dealing with sub-segments driven by similar factors but at varying strengths.

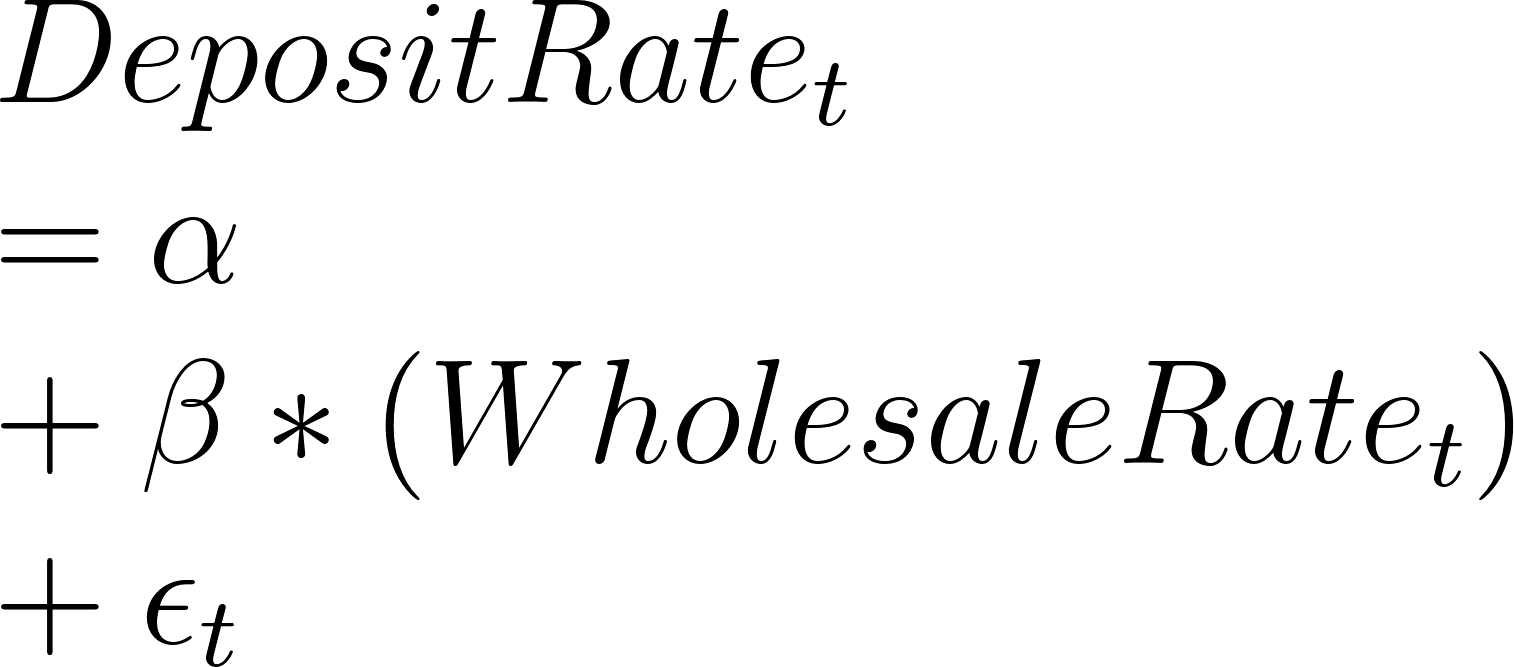
**8. RATE MODELING APPROACH**

Rate modeling approaches vary from simple beta estimates to more advanced techniques such as Error Correction Modeling (ECM), which are detailed in sections 8.1-8.4

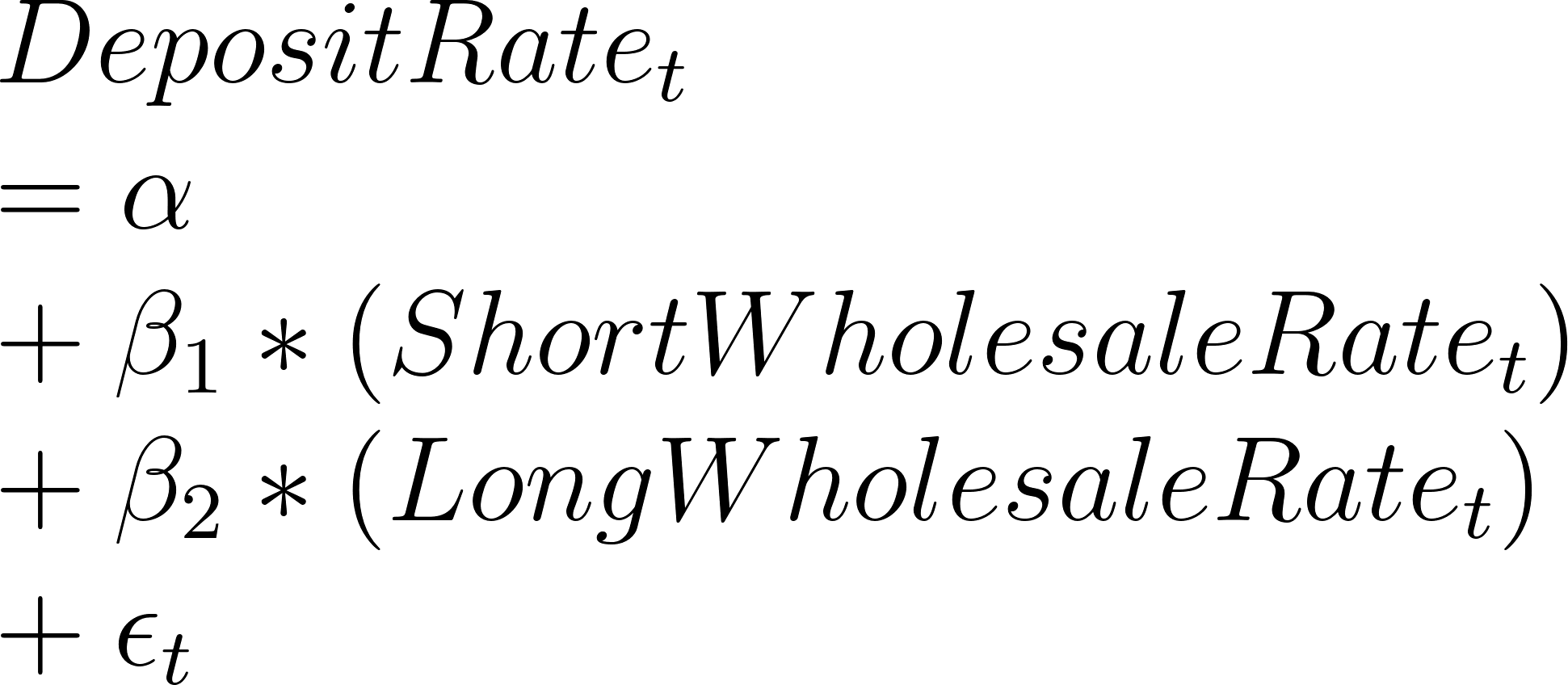
*8.1. Simple Beta Estimates* :Calculates the beta coefficient over a representative time period by dividing the change in deposit rate by the change in wholesale rate. This provides a simple estimate of the relationship between the two variables4.

[](https://www.codecogs.com/eqnedit.php?latex=%5Cbeta%20%3D%20%5Cfrac%7B%7B%5CDelta%20%5Ctext%7B%7BDeposit%20Rate%7D%7D%7D%7D%7B%7B%5CDelta%20%5Ctext%7B%7BWholesale%20Rate%7D%7D%7D%7D#0)  *Eq. (8.1)*

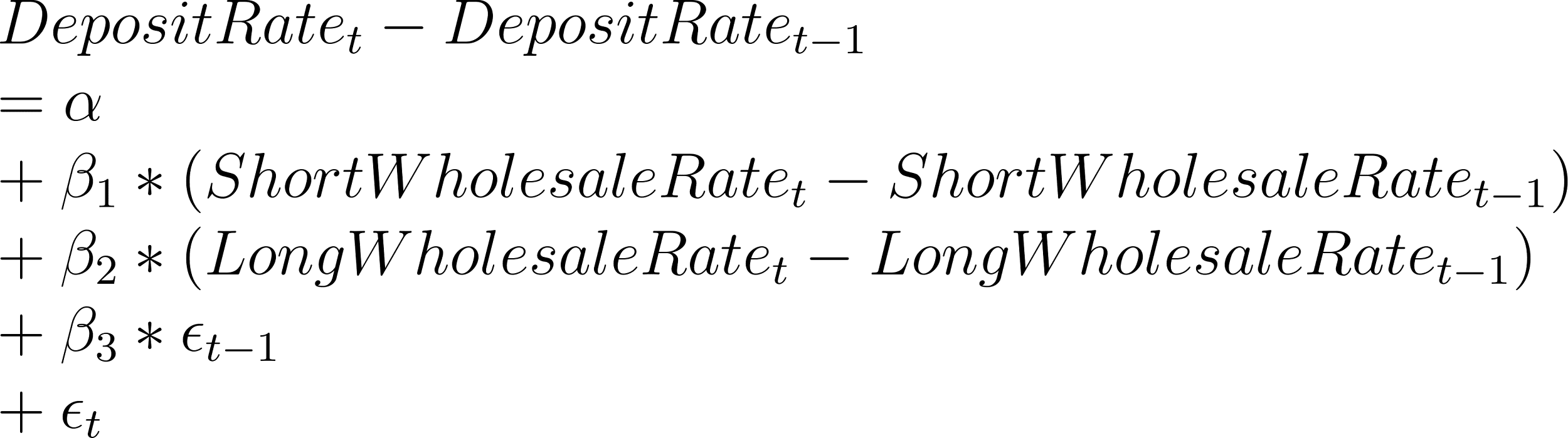
*8.2. Single-Tenor Time Series Models:* Utilizes regression analysis, such as Ordinary least squares (OLS), to model the relationship between the deposit rate and the wholesale rate. This approach may include potential differentiations for rising versus falling rate environments or for different levels of rates.

[](https://www.codecogs.com/eqnedit.php?latex=%5C%5CDeposit%20Rate_t%5C%5C%3D%20%5Calpha%20%5C%5C%2B%20%5Cbeta%20*(WholesaleRate_t)%20%5C%5C%2B%20%5Cepsilon_t#0) *Eq. (8.2)*

*8.3. Multi-Tenor Time Series Models:* Incorporates short-term and long-term rates in a multi-tenor OLS model to estimate the relationship between the deposit rate and the wholesale rate. This allows for the estimation of non-constant betas, considering the varying impacts of short-term and long-term rates.

[](https://www.codecogs.com/eqnedit.php?latex=%5C%5CDeposit%20Rate_t%5C%5C%3D%20%5Calpha%20%5C%5C%2B%20%5Cbeta_1%20*%20(ShortWholesaleRate_t)%5C%5C%2B%20%5Cbeta_2%20*%20(LongWholesaleRate_t)%20%5C%5C%2B%20%5Cepsilon_t#0) *Eq. (8.3)*

*8.4. Error Correction Modeling (ECM):* Implements an ECM approach, which combines a levels model with an error correction term to account for sudden changes and address stationarity issues. This approach is useful for capturing the long-term equilibrium relationship between the deposit rate and the wholesale rate5.

[](https://www.codecogs.com/eqnedit.php?latex=%5C%5CDeposit%20Rate_t%20-%20Deposit%20Rate_%7Bt-1%7D%5C%5C%20%3D%20%20%5Calpha%20%20%5C%5C%20%2B%20%5Cbeta_1%20*%20(Short%20Wholesale%20Rate_t-ShortWholesaleRate_%7Bt-1%7D)%20%5C%5C%2B%20%5Cbeta_2%20*%20(LongWholesaleRate_t-Long%20Wholesale%20Rate_%7Bt-1%7D)%20%5C%5C%2B%20%5Cbeta_3%20*%20%5Cepsilon_%7Bt-1%7D%20%5C%5C%20%2B%5Cepsilon_t#0) *Eq. (8.4)*

Both the Single-Tenor and Multi-Tenor Time Series Models, as well as ECM, offer the flexibility of incorporating differenced modeling to address stationarity problems. However, usually the direct linkage between the spot change in the portfolio rate and the change in wholesale rates is often weak or poorly correlated, indicating the need for more sophisticated modeling techniques beyond simple differencing. Several rate modeling approaches align with current application need, and therefore, the preferred method is error correction modeling, though simpler time series modeling may be used

**9. TIME SERIES MODELING APPROACH**

The approach for Time Series Modeling typically begins with the identification of relevant independent variables, which may include market interest rates, macroeconomic factors, pricing variables, and seasonality terms. A manual forward selection process is often employed to add variables to the model based on business hypotheses and cross-correlation techniques.

The modeling process includes addressing issues such as autocorrelation and heteroskedasticity through the incorporation of autoregressive (AR) terms and the use of Newey-West standard errors. These steps help ensure the stability and reliability of the model's coefficients.

Cointegration testing using methods like the Johansen test may also be performed to determine if there exist long-term equilibrium relationships among the variables. This helps capture the interdependencies and co-movements in the average balance time series.

The models are subjected to various validation tests, including tests for multicollinearity, serial correlation, and stationarity. These tests assess the statistical significance, goodness of fit, and overall reliability of the model. Holdout testing is conducted to evaluate the stability and predictive power of the models by comparing forecasted average balances with actual data.

Time series modeling for average balances follows a rigorous methodology that combines statistical techniques, business intuition, and thorough validation to create robust models capable of capturing the underlying dynamics and forecasting future average balances with reasonable accuracy.

**10. DEPENDENT VARIABLES**

Dependent variabletransformations have a crucial role in modeling, especially when utilizing linear regressions as our primary focus. The emphasis on linear regressions necessitates the application of suitable transformations to ensure accurate model specification. When dealing with interval level data that spans from 0 to infinity, such as average balances, various transformations can be employed to linearize the relationships. Common transformations include taking the natural logarithm (our preferred transformation for average balances), calculating percent changes, or utilizing power transformations, although the latter is less frequently used.

The rationale behind these transformations is to ensure the appropriateness of the model fit. When comparing a model constructed with an untransformed series to one built with a transformed series, certain considerations come into play. If an untransformed series ranging from 0 to infinity is used to build a model, there is a possibility that the forecasted values may cross the zero threshold, leading to counterintuitive results, especially when dealing with average balances and customer rates. While such outcomes could be managed through a management overlay process, it is generally preferable to avoid them altogether.

On the other hand, if a model is developed using a transformed series ranging from 0 to infinity, whether the dependent variable is differenced or maintained at the level, the forecasts will not exhibit this problem. Rates, however, typically remain untransformed and are likely to be modeled in their natural "levels" state.

By carefully considering and implementing appropriate transformations for the dependent variables, it can be ensured that the model accurately captures the relationships between variables and produces meaningful and intuitive forecasts.

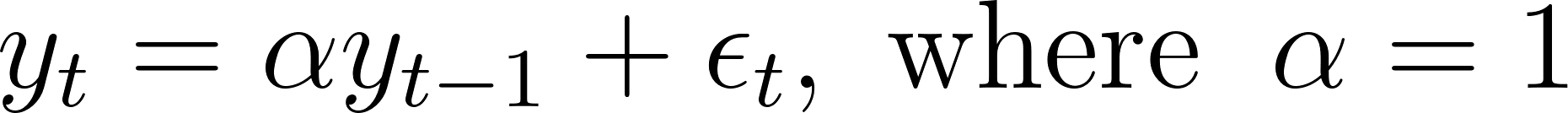
**11. STATIONARITY**

*11.1 Addressing stationary concerns:* Thisinvolves considering the treatment of unit root in deposit rate modeling and average balance modeling, which diverge in their approaches.

In deposit rate modeling, the primary focus is on the level of of independent variables. The underlying perspective is that there exists a cointegrated relationship between deposit rates and market interest rates. Therefore, the presence of a unit root is not a concern.

On the other hand, when modeling average balances, the approach revolves around the change in average balances using the change in independent variables. In this case, arguing for a cointegrated relationship is considered less defensible.

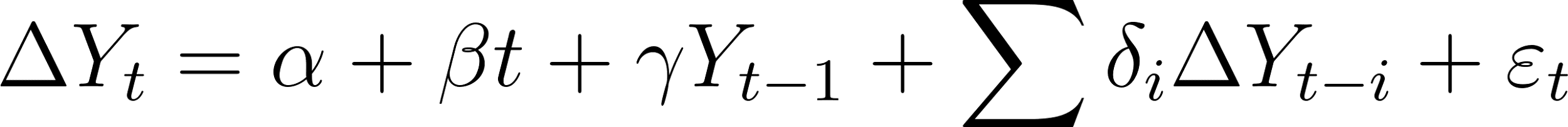
The presence of a unit root is a common issue in time series analysis and can lead to inflated significance of independent variables in linear regression, resulting in spurious relationships. Technically, a unit root is considered present when the characteristic equation of a process has 1 as a root.

[](https://www.codecogs.com/eqnedit.php?latex=y_t%20%3D%20%5Calpha%20y_%7Bt-1%7D%20%2B%20%5Cepsilon_t%2C%20%5Ctext%7B%20where%20%7D%5C%3A%20%5Calpha%20%3D%201%20#0) *Eq. (11.1)*

The process is considered integrated to the order of the number of unit roots, denoted as I(#). Typically, differencing the process is employed to remove a unit root. For example, an I(2) process, when differenced, becomes an I(1) process. A process that is I(0) has no unit roots and is referred to as stationary, meaning the joint probability distribution of the process does not change over time. This results in the mean, variance, and autocorrelation structure remaining constant over time.

In time series modeling, stationary processes are generally ideal. However, if a linear combination of two or more non-stationary processes yields a stationary one, those processes are said to be cointegrated and can be utilized for modeling purposes.

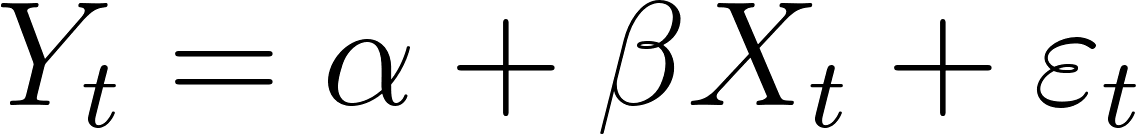
*11.2. Testing:* When addressing unit root, the initial step is to cure it through differencing. The Phillips-Perron unit root test, implemented using the PROC AUTOREG command, is employed to assess the presence of a unit root. Given the tendency of a bank’s balance data to exhibit serial correlation for up to 12 months, the Phillips-Perron test is conducted up to order 12. If the Pr<Tau value is below 0.1 for zero or single mean orders, it is assumed that the series does not have a unit root. However, variables that pass the test only with a trend are considered inappropriate for stress testing purposes due to their implications for forecasting.

[](https://www.codecogs.com/eqnedit.php?latex=%5CDelta%20Y_t%20%3D%20%5Calpha%20%2B%20%5Cbeta%20t%20%2B%20%5Cgamma%20Y_%7Bt-1%7D%20%2B%20%5Csum%20%5Cdelta_i%20%5CDelta%20Y_%7Bt-i%7D%20%2B%20%5Cvarepsilon_t#0) *Eq. (11.2)*

It is important to acknowledge that the Phillips-Perron test's results are sensitive to the number of lags tested. To determine the appropriate number of lags, the Ng-Perron MAIC lag length selection method is employed. The primary approach to address the unit root issue is through one-period differencing. The differenced series is then subjected to the ADF (Augmented Dickey-Fuller) test to confirm that stationarity has been achieved. This transformation indicates the conversion of an I(1) series to an I(0) series.

If a dependent variable has been differenced to exhibit I(0), but it proves challenging to develop a model that meets both statistical and business requirements, the modeling team, in collaboration with business partners, will make a decision. In such cases, the decision may be to build a model on the undifferenced series, acknowledging the presence of a unit root in the dependent variable but proceeding with modeling it nonetheless. The resolution for the unit root issue in these cases is to demonstrate that the modeled series is cointegrated with the dependent variable.

*11.3 Cointegration:* If an acceptable model cannot be created using the I(0) process, it may be appropriate to consider cointegrated modeling. Cointegration is identified by evaluating whether the residuals of our model exhibit a unit root, as discussed earlier. If the residuals do not exhibit a unit root, it can be concluded that our model is cointegrated with the dependent variable, and all necessary assumptions regarding potential spurious relationships have been addressed.

[](https://www.codecogs.com/eqnedit.php?latex=Y_t%20%3D%20%5Calpha%20%2B%20%5Cbeta%20X_t%20%2B%20%5Cvarepsilon_t#0) *Eq. (11.3)*

It is important to note that when building cointegrated models, this testing step occurs during the model development process rather than before it. The model development approach remains unchanged, including the selected modeling forms, variables, and other identified assumptions. The only difference is the inclusion of the testing of residuals for unit root to determine whether the model exhibits cointegration or not.

In the case where the dependent variable is I(2), a similar approach can be considered by employing double-differencing to transform the variable from I(2) to I(0). However, this approach is typically rejected without attempting to develop a model. The reason behind this rejection is that independent variables effectively predict the "acceleration" of the dependent variable, which can be challenging to conceptually understand and evaluate. Furthermore, in the context of stress testing, models for I(2) dependent variables tend to be universally inaccurate.

In situations where the dependent variable is I(2), the management approach also relies on cointegration. Development of the cointegrated relationship will be attempted on both the untransformed I(2) series and the single-difference I(1) transformed series, considering the specific characteristics of the data.

*11.4. Error Correction Models (ECM):* ECM leverages the concept of cointegration, which denotes a cointegrating relationship or stochastic trend between variables that move together in the same direction over an extended period. This modeling approach is powerful in capturing both the long-term and short-term dynamics between two variables by estimating the speed at which the dependent variable returns to equilibrium following a change in the independent variable.

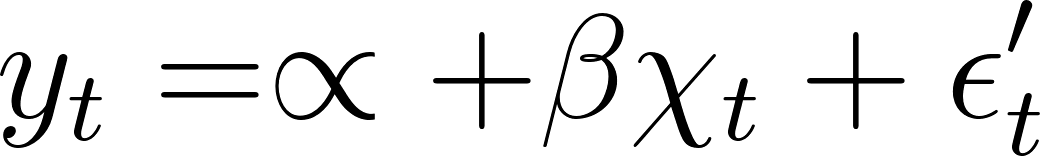
The ECM model begins by establishing a long-term relationship between the variables and incorporates the long-term error into an expression of the short-term relationship. Mathematically, this is achieved through the following steps:

1. Regress market rates on deposit rates using a co-integrated model at the levels basis.
2. Regress market rates and the lagged error term from Equation 1 on deposit rates using a first-difference model.

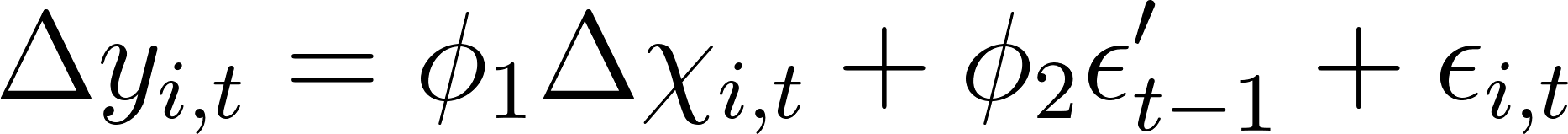
The inclusion of the lagged error term in the second step helps account for factors that could cause divergence in the long-term equilibrium relationship, which tend to be amplified in the short-term dynamics. The error correction model serves to distribute or "correct" the error present in the short-term by modeling the gap (error) between the two variables in the previous period as an additional independent variable.

By incorporating the error correction mechanism, the ECM provides a framework to capture the adjustment dynamics between variables, ensuring that the short-term deviations from equilibrium are corrected over time. This enables a more comprehensive understanding of the relationship and facilitates more accurate rate modeling.

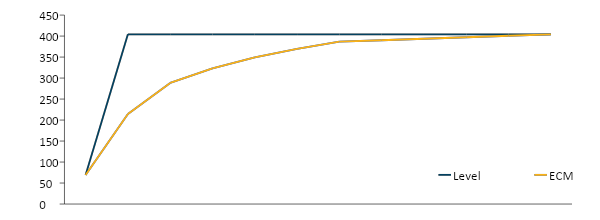
*11.5. ECM Approach:* When employing the Error Correction Model (ECM), the typical approach involves a two-step Engle-Granger model development process. In the first step, t a time series model is constructed using the level (undifferenced) series of deposit rates and market rates. This model captures the long-term relationship between these variables.

[](https://www.codecogs.com/eqnedit.php?latex=y_t%20%3D%20%5Cpropto%20%2B%5Cbeta%20%5Cchi_t%20%2B%20%5Cepsilon'_t#0) *Eq. (11.5)*

In the second step, another time series model is built, but this time using the differenced series of customer rates and market rates. Importantly, the long-term error variable from the first equation is incorporated into this model. By doing so, the model adjusts smoothly to sudden changes and accounts for any deviations from the long-term equilibrium relationship.

[](https://www.codecogs.com/eqnedit.php?latex=%5CDelta%20y_%7Bi%2Ct%7D%20%3D%20%5Cphi_1%20%5CDelta%5Cchi_%7Bi%2Ct%7D%20%2B%20%5Cphi_2%20%5Cepsilon'_%7Bt-1%7D%2B%5Cepsilon_%7Bi%2Ct%7D#0)

The inclusion of the long-term error variable in the second step helps ensure that the model adequately captures the dynamics of the relationship between customer rates and market rates. This adjustment mechanism allows for a more accurate representation of the short-term dynamics, taking into account the impact of previous period errors.



**Level Model vs. ECM (bp)**

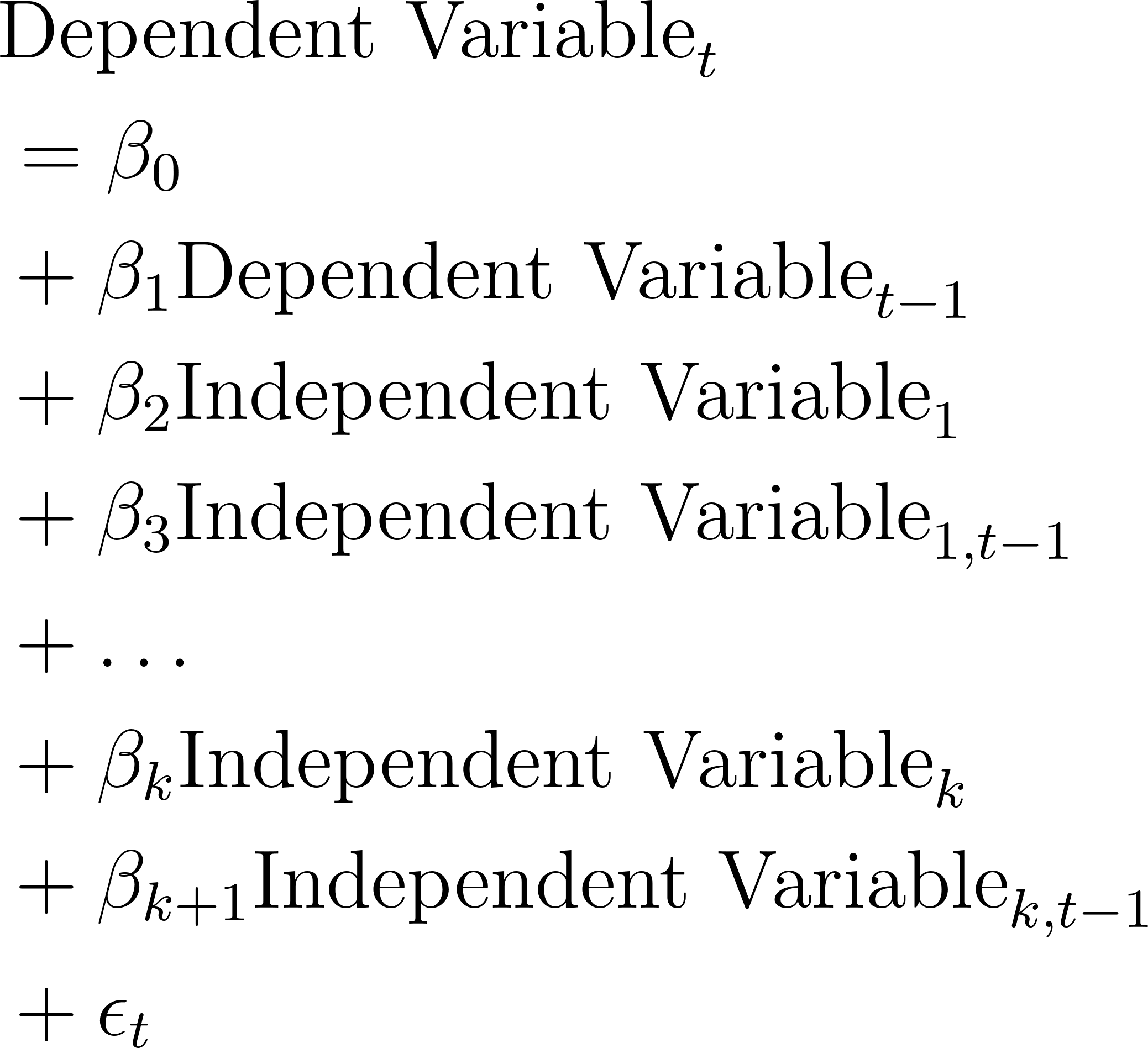
*Figure 3*

Overall, this two-step process in the ECM modeling approach allows us to incorporate both the long-term relationship and the short-term adjustment dynamics, resulting in a more comprehensive and effective modeling framework.

*11.6: Autoregressive Distributed Lag (ARDL):* Unlike the two-step method, the ARDL approach involves a single-step equation. The ARDL model incorporates the following regressors:

1. Lagged values of the dependent variable (in this case, previous deposit rates).
2. Unlagged and lagged values of other independent variables (in this case, market interest rates).
3. An error or "disturbance" term (εt), which is assumed to be serially uncorrelated.

The ARDL method is similar to the two-stage ECM, but it replaces the error-correction term with lagged dependent and independent variables, without imposing restrictions on their coefficients.

[](https://www.codecogs.com/eqnedit.php?latex=%5Cbegin%7Balign*%7D%5C%5C%26%5Cquad%5Ctext%7BDependent%20Variable%7D_t%20%5C%5C%26%5Cquad%3D%20%5Cbeta_0%20%5C%5C%26%5Cquad%2B%20%5Cbeta_1%20%5Ctext%7BDependent%20Variable%7D_%7Bt-1%7D%20%5C%5C%26%5Cquad%2B%20%5Cbeta_2%20%5Ctext%7BIndependent%20Variable%7D_1%20%5C%5C%26%5Cquad%2B%20%5Cbeta_3%20%5Ctext%7BIndependent%20Variable%7D_%7B1%2Ct-1%7D%20%5C%5C%26%5Cquad%2B%20%5Cldots%20%5C%5C%26%5Cquad%2B%20%5Cbeta_k%20%5Ctext%7BIndependent%20Variable%7D_k%20%5C%5C%26%5Cquad%2B%20%5Cbeta_%7Bk%2B1%7D%20%5Ctext%7BIndependent%20Variable%7D_%7Bk%2Ct-1%7D%20%5C%5C%26%5Cquad%2B%20%5Cepsilon_t%5Cend%7Balign*%7D#0) *Eq. (11.6)*

Using the ARDL approach offers several advantages over the two-step method:

1. *Streamlined testing and implementation:* The ARDL method reduces the two-stage ECM to a single equation, simplifying the modeling process.
2. *Improved estimation:* By not imposing restrictions on the coefficients of regressors, the ARDL method can potentially provide better estimates.
3. *Applicability in diverse scenarios:* The ARDL method can be used when time series data is stationary (I(0)) or integrated of order one (I(1)). However, it is not suitable for integration of order two (I(2)) series. This flexibility makes it valuable in situations where the stationarity of data is not straightforward.

Overall, the ARDL method offers an alternative and versatile approach to modeling, providing advantages such as streamlined implementation, improved estimation, and applicability in diverse scenarios.

**12. WEIGHTING**

When considering the weighting of observations in the modeling process, the decision was made to treat each observation with equal weight. This choice was made due to the lack of clear justification for any specific weighting scheme. Arguments can be made for both weighing early periods more or less heavily than more recent periods.

Earlier periods may be argued to deserve more weight as they exhibit a greater range of interest rate levels and movements compared to the consistently low-rate environment observed in recent years. This suggests that earlier periods contain more valuable information for modeling purposes. It is important to acknowledge that there have been significant changes in a bank’s strategy over recent periods. Customers' responses to management factors and macroeconomic conditions have resulted in shifts that may render more recent periods more relevant. Therefore, weighting more recent periods more heavily could be considered.

Weighting observations differently tends to introduce more noise than signal in the attempt to develop meaningful relationships. It becomes challenging to defend or justify the choice of a particular weighting scheme, whether it is an automatic or manual process. Therefore, the decision was made to treat all observations equally to maintain simplicity and avoid potential biases that could arise from uneven weighting.

**13. INDEPENDENT VARIABLES**

*13.1 Modeling Process:* When incorporating independent variables in the modeling process, there is a divergence in the treatment of variable selection for deposit rates and average balances.

For rate modeling, the focus is primarily on market interest rates and their derivatives as the main independent variable set. Since the drivers primarily revolve around rate variables, a "batch" process is often employed to identify acceptable models.

On the other hand, for average balance modeling, a broader approach is taken by considering a range of drivers. These drivers encompass market interest rates, macroeconomic factors, pricing variables, and seasonality terms. A manual forward selection process is adopted to add variables to the model. Each variable is selected "by hand" from a constrained set of drivers hypothesized by the businesses and modeling team. Cross-correlation techniques are utilized to streamline the identification of high-likelihood variables.

Throughout the variable selection process, alignment with business intuition, fit measures, and statistical tests is ensured. This involves reviewing the models at each stage to ensure they align with the understanding of the business, exhibit good fit, and pass relevant statistical tests.

*13.2: Variable Selection:* In all modeling endeavors, the process of variable selection is grounded in business intuition. Variables that have supporting rationale from business experts are given priority in the selection process. As previously discussed, our approach to variable selection is as follows:

As stated in section 13.1, in deposit modeling, rates are predominantly influenced by market interest rates. For average balance modeling, a broader set of potential drivers is considered, and hypotheses are collected to narrow down the pool of variables for further evaluation and inclusion in the models. However, regarding the temporal aspect, it is assumed that the effects of drivers become negligible after three months, unless there is strong business justification for a longer lag. This means that measures with a lag greater than three or a moving average longer than three periods are generally avoided. The choice of lag and moving average periods is consistent and cumulative, avoiding, for example, a three-month moving average based on a two-period lag.

Furthermore, it is possible to agree on the hypothesized strength and direction of families of drivers before the modeling process. These families of drivers do not make distinctions or preferential assumptions about specific variable nuances. For example, a family such as "employment strength" could include variables such as the unemployment rate, jobless claims, and other related indicators.

By incorporating business intuition and rationalizing the selection of variables, the aim is to ensure that the chosen set of drivers captures the key factors that drive the phenomena being modeled. This approach ultimately enhances the accuracy and relevance of the models.

*13.3. Average Balance Variables:* In the context of average balance modeling, meaningful univariate relationships are identified using cross-correlation techniques, which serve as the primary building blocks for our models. The use of cross-correlation procedures in manual forward selection helps explore the relationships between dependent and independent variables. The variable selection process begins by generating cross-correlation outputs and comparing them to the hypothesized drivers. Variables that exhibit a statistically significant relationship supporting the business hypothesis are identified as potential variables of interest.

Based on these highlighted variables, univariate models are constructed with a target threshold for statistical significance. A level of significance above 95% is generally preferred, but significance as low as 85% may be allowed for variables with strong business hypotheses. The statistically significant univariate relationships identified through this process form the primary "branches" of the modeling framework. To further enhance the models, additional variables are added by running cross-correlation procedures on the residuals of the initial models.

The inclusion of additional variables is evaluated based on several criteria, such as alignment with the business hypotheses and statistical significance levels (aiming for a level above 90% with above 95% preferred). Visual model fit improvements and a decrease in the Akaike Information Criterion with Correction (AICC), which measures model quality while penalizing complexity, are also taken into account. Multicollinearity with other variables is also assessed to avoid redundancy and maintain the independence of explanatory variables.

By employing cross-correlation techniques and manual forward selection, relevant variables can be effectively identified and incorporated into average balance models. This approach allows for the construction of models that are both statistically robust and aligned with the underlying business hypotheses.

*13.4. Cross Correlation:* Cross correlation outputs are used to identify potential drivers, though they are considered vis-à-vis business hypotheses (not in a vacuum).

The variables most correlated with dependent variables are as follows5

|  |
| --- |
| Dow Jones U.S. Total Market Index |
| S&P 500 |
| After-tax Profits, (Bil. $, SAAR) |
| Oil Price (WTI) |
| Manufacturing Capacity Utilization Rate |
| Swaption: 5Y-10Y Volatility |

*Figure 4*

**14. EXCEPTIONAL EVENTS**

Before modeling begins, it is essential to identify and address historical exceptional events that require special treatment. These events may either be incapable of being accounted for in the model or should not be predicted by the model, such as significant changes like the beginning and end of unlimited FDIC insurance on deposits.To ensure proper handling of these events, indicators should be identified and tested in the models during the business engagement phase. These indicators may involve monthly or multi-month changes or phase-shift changes.

There are two primary approaches to addressing these events. If the impact of the event can be precisely determined, the effect can be directly removed from the time series. This approach is suitable for situations like large single-client movements where the impact can be quantified. However, if the impact cannot be directly determined, it can still be accounted for through boolean indicator variables. These variables help capture the influence of the exceptional events in the modeling process.

Regardless of the chosen approach, providing a clear business justification is essential to warrant the treatment of extraordinary events in the final model. This ensures transparency and alignment of the modeling process with the specific needs and requirements of the business. In the chosen modeling approach, indicator variables should be included rather than directly manipulating the data. There are two primary methods for incorporating these events: direct adjustment to the time series and the use of boolean (or indicator) variables.

For direct adjustment, the modeling data is modified to remove the impact of known extraordinary events. This involves adjusting the time series to account for specific events and their effects. On the other hand, boolean variables are created to capture the occurrence of these events. These variables take a value of zero for every period except the affected period(s). By including these booleans in the model, a new intercept is effectively set for the corresponding period(s). For example, in the case of a single-client movement that significantly changes the scale of total balances or an irregular external influence like a regulatory change, direct adjustment or boolean variables can be utilized, respectively.

Each approach has its pros and cons to consider. Direct adjustment removes the effect of the events, ensuring they do not impact the modeling process. However, it requires careful judgment in determining when to remove these impacts, and there may be considerations of data integrity due to spot manipulation of the modeling series.

Using boolean variables allows modelers to control for less definite one-time events that could otherwise distort coefficient estimates for other drivers. However, their use should be justified as they can easily outweigh the explanatory power of other variables. Additionally, single-month booleans may result in the loss of data points, as they fit the affected point perfectly without providing new information.

It is important to conduct additional testing and sensitivity analysis when employing these adjustments. Adjusted level shifts should undergo further testing to ensure that no structural breaks have occurred, and the explanatory power and goodness of fit of the model should be assessed by excluding the affected periods.

**15. SEASONALITY**

In the final modeling approach, controlling for seasonality is typically done using a straightforward method. Before modeling begins, the un-differenced dependent variable for seasonality are evaluated using the joint test for seasonality available in PROC X11, in combination with business discussions on portfolio seasonality.

The preference is to follow the business's hypothesis regarding seasonality. If the X11 procedure identifies seasonality while the business does not recognize it, the modelers will try to demonstrate the statistically identified seasonality to the business. If the business maintains that there is no seasonality, the modeling process will proceed without including seasonality in the model, with a strong preference for this approach.

However, if the business suggests the presence of seasonality and the X11 procedure does not detect it, we employ a specific procedure. This involves including 11 monthly indicators and using a joint F-test to determine if these indicators are statistically significant and viable for inclusion in the model. If the joint F-test rejects the hypothesis of statistically significant seasonality, there will be no universal treatment of seasonality. Instead, specific months with a priori assumptions for unique events, such as April/May with IRA spikes, can be handled as single-month indicators, typically limited to four.

If seasonality is identified and deemed significant, it will be managed during the model development process by incorporating the 11 monthly indicators. Forward selection of variables will start from models already including these seasonality indicators. The significance of the seasonality as a whole will be assessed through a joint F-test on the 11 indicators, rather than using individual t-tests for each indicator.

In some cases, a subset of indicator variables may be used if there is strong rationale for doing so. For example, if November/December correlates with holiday shopping or April/May correlates with tax season, a subset of indicators can be considered. The significance of these indicator variables would be assessed using individual t-tests, and the number of indicators allowed is limited to a maximum of four. This approach ensures that seasonality is appropriately accounted for in the model while considering the specific factors driving the seasonality.

**16. VALIDATION TESTING (AVG. BAL)**

In the validation of average balance models, several tests are performed to ensure the model meets specific criteria:

*16.1. Homoscedasticity:* The model is tested for homoscedasticity using Engle Lagrange and Lee & King tests up to the 12th order. The passing threshold is a probability (Pr) greater than 0.03 for any one of the three tests. This test is given high priority.

*16.2.. Limited multicollinearity:* To assess limited multicollinearity, the model is examined using variance inflation factor (VIF) and condition index (CI). The passing thresholds are VIF less than 10 and CI less than 30. This test is given high priority.

*16.3. Serially uncorrelated residuals:* Residuals of the model are tested for serial correlation using the Breusch-Godfrey test up to the 12th order. The passing threshold is a probability (Pr) greater than 0.03 at all lags. This test is given high priority.

*16.4. Stationary residuals*: The stationarity of residuals is evaluated using the Augmented Dickey-Fuller test. In the case of zero mean, the passing threshold is a p-value less than 0.10. This test is given high priority.

*16.5. Normality of residuals:* Residuals are tested for normality using multiple tests, including the Shapiro-Wilk test, Kolmogorov-Smirnov test, Anderson-Darling test, and Cramér-von Mises test. The passing threshold is a p-value greater than 0.03 for any one of these tests. This test is given lower priority.

These validation criteria ensure that the average balance model meets the necessary statistical requirements and assumptions. The tests help evaluate the model's performance and ensure the reliability of its results.

**17. VALIDATION TESTING (RATES)**

In the validation of rate models, several statistical tests are conducted to ensure the model meets specific criteria:

*17.1. Multicollinearity:* The model is evaluated for multicollinearity using variance inflation factor (VIF) and condition index (CI). The passing thresholds are VIF less than 10 and CI less than 30. This test is given high priority.

*17.2. Cointegration:* Cointegration is assessed using the Johansen test. The passing criterion is a probability (Pr) greater than the trace value of 0.05. This test is given high priority.

*17.3. Serial Correlation:* The model is tested for serial correlation using Newey West HC standard errors. The passing criterion is a variable significance level of less than 0.05. This test is given high priority.

1*7.4. Heteroscedasticity:* Heteroscedasticity is evaluated using Newey West HC standard errors. The passing criterion is a variable significance level of less than 0.05. This test is given high priority.

*17.5. Normality of Residuals:* Residuals are tested for normality using several tests, including the Shapiro-Wilk test, Kolmogorov-Smirnov test, Anderson-Darling test, and Cramér-von Mises test. The passing criterion is a p-value greater than 0.03 for any one of these tests, with a preference for values greater than 0.05 for all tests. This test is given lower priority.

These statistical validation tests ensure that the rate model meets the necessary requirements and assumptions. They help assess the model's performance, identify potential issues such as multicollinearity and serial correlation, and ensure the validity of the model's results.

**17. AUTOCORRELATION**

*18.1. Approach:* In the case of rate modeling, autocorrelation is not a major concern due to our utilization of error correction modeling (ECM) and autoregressive distributed lag (ARDL) modeling approaches. However, for average balance modeling, the presence of autocorrelation can pose a more substantial issue depending on the chosen modeling form. In such cases, Newey-West standard errors is used to assess the statistical significance of independent variables. These standard errors help determine whether the observed autocorrelation is influencing the significance of the variables.

If the Newey-West standard errors indicate that autocorrelation is indeed a problem, the modeling team will incorporate autoregressive (AR) terms in an attempt to mitigate the issue. However, it is important to note that the team will also conduct sensitivity tests to evaluate the overall impact of the AR terms on the model's performance and results. The identification of significant AR terms is done through three approaches: examining partial autocorrelation function (PACF) and inverse autocorrelation function (IACF) plots to identify the most prominent AR terms, performing the Breusch-Godfrey test, and re-estimating the model using the ESTIMATE function of PROC ARIMA and including 12 lagged error terms.

*18.2. Testing:* Initially,the first significant AR term with the lowest lag is included, while continuously monitoring the aforementioned tests. Subsequently, additional significant AR terms are progressively incorporated until the autocorrelation issue is resolved, with a maximum consideration of AR(12).

Spot-selecting significant AR terms is allowed without necessarily including the preceding lags. Notably, the AR terms 1, 2, 3, and 12 are readily included as needed. However, AR terms beyond AR(3) tend to be more prone to spurious or outlier-dependent relationships. Hence, caution is exercised in incorporating significant lags from AR(4) to AR(11) unless their significance is compelling and they are essential for mitigating autocorrelation. In such cases, marginal serial correlation is prioritzed over including extraneous AR terms.

If serial correlation persists, including marginal cases, autocorrelation-consistent standard errors is implemented, such as the Newey-West standard errors. These adjustments help alleviate concerns about the significance of other independent variables (IVs) in the model.

*18.3. Results:* The impact of the AR terms on the models was sought to be limited. Given their use in stress testing, the preference was for the primary drivers of the models to be independent variables rather than AR terms, ensuring the models would respond to varying macroeconomic scenarios.

An approach was taken to assess the impact of the AR terms by summing their coefficients and examining the aggregate coefficient, aiming for a value between +0.7 and -0.7 (±70%). If the sum exceeded 70%, it indicated that a significant portion of the model was being driven by AR terms. In such cases, efforts were made to strike a balance by dropping less significant AR terms, aiming to address autocorrelation while reducing reliance on AR terms.

If an acceptable balance could not be achieved, the model had to be re-specified. Sensitivity tests were also conducted on the AR terms to evaluate their impact on the model. This included running model predictions using only the structural component and re-estimating the model without the inclusion of AR terms. These tests provided insights into the influence and significance of the AR terms in the overall model.

**19. HOLDOUT TESTING**

*19.1. Coefficients:* The stability of coefficients is also assessed through holdout testing. The focus of holdout testing is to ensure the stability of the directional impact and significance of the variables. Ideally, when a sufficiently large sample size is used and consistent relationships between variables are observed, the coefficient estimates should not undergo significant changes even after holding out a portion of the data. Consistency is defined based on the following criteria:

1. Variables maintain their directionality, without flipping signs.
2. Variables maintain statistical significance with a confidence level based on the full period statistical significance:
3. 90% confidence if the full period statistical significance is greater than 95%.
4. 85% confidence if the full period statistical significance is between 90-95%.

In general, the order of magnitude of the coefficients remains relatively consistent. As a loose guideline, when holding out 15% of the data, a 30% change in coefficient estimates is considered within expectations. In the last stage of model production, the coefficients of the independent variables that were selected over the training sample, tested over the holdout and validation samples, can be respecified over the full dataset, ensuring the inclusion of the entire data for a comprehensive analysis.

*19.2. Forecast Period:* A one-year holdout is utilized to assess the stability and predictive power of the model. Initially, a model is built using 100% of the available data, incorporating all data points to identify the independent variables, lagged error terms, presence of intercept, and other relevant factors.The model is then re-specified by excluding the last 12 months of data while keeping the same independent variables, lagged error terms, and intercept. However, the coefficients for each driver are re-estimated.

Dynamic forecasts, particularly for time series modeling, are generated for the held-out data. The forecasts are then compared to evaluate both the forecast accuracy and the stability of variable relationships. This assessment helps determine the model's ability to accurately predict future outcomes and maintain consistent relationships between the variables.

*19.3. Stability:* The stability of coefficients is also assessed through the use of holdout samples. Holdouts are employed to evaluate the stability of variable relationships within the model.

In an ideal scenario with a sufficiently large sample size and consistent variable relationships, the estimates of coefficients should not undergo significant changes after a portion of the data is held out. However, in reality, the number of observations is often limited and variable relationships can be influenced by the prevailing economic conditions. Consistency in this context is defined by certain criteria:

1. Variables should maintain their direction and not change signs.
2. Variables should retain their statistical significance with a confidence level of:
3. 90%, if the full period statistical significance confidence exceeds 95%.
4. 85%, if the full period statistical significance confidence falls between 90% and 95%.

In general, the order of magnitude of the coefficients remains relatively unchanged. As a rough guideline, when holding out approximately 15% of the data, it is not uncommon to observe up to a 20% change in coefficient estimates.

**20. MODEL ACCURACY**

The accuracy of the model is assessed using the MAPE (Mean Absolute Percentage Error) metric. MAPE is calculated by taking the absolute value of the difference between the predicted value and the actual value for each period in the dynamic historic forecast. This difference is divided by the actual value and subtracted by 1 to obtain a positive number representing the percentage error of the forecast. The average of these absolute percentage errors (APE) is then calculated over the forecast horizon.

In addition to the holdout MAPE, which is explained in more detail below, two other measures of MAPE are used. The first is a full-period forecast, where a dynamic forecast is initiated in one of the first few periods of the model development window and runs until the last observed period. The MAPE is calculated based on this forecast. The second measure is a rolling MAPE, where a 39-month dynamic forecast horizon is started every thirteenth point. MAPEs are calculated for forecasts spanning different periods, and the minimum and maximum values of all the staggered start MAPEs are reported.

The holdout MAPE is reported as the straightforward outcome of the calculation. The staggered start MAPEs represent the range of MAPEs observed across all the different starting points. It is important to note that the full-series MAPE can be higher or lower than the minimum or maximum of the staggered start MAPEs, depending on variations in forecast accuracy.

There is no predefined threshold for accepting or rejecting a model based solely on the MAPE value. However, the MAPE serves as a useful indicator to assess the model's accuracy, depending on the nature and transformation of the dependent variable. The MAPE provides insight into whether the model is sufficiently accurate, but its interpretation should consider other factors and context alongside it.

**21. SCENARIO FORECASTING**

Forecasts for account attrition are created using available macroeconomic scenarios. The forecasting process involves developing time series models to predict the evolution of average balances and customer rates at the portfolio level under hypothetical macroeconomic scenarios. The scenarios used include the Fed-provided scenarios, namely Base, Adverse, and Severely Adverse. Additional scenarios derived from the base scenario are used to evaluate interest rate and price sensitivity.

Interest rate shock scenarios involve adjusting all market interest rates by a predetermined amount6 (e.g., +100bp, +300bp) during the first month of the forecast. These scenarios are useful for assessing the sensitivity of deposit behaviors to changes in interest rates. Pricing shock scenarios, on the other hand, involve adjusting relative pricing variables by a predetermined amount (e.g., +10bp, +20bp) during the first month of the forecast. These scenarios help assess the relationship between pricing and deposit behaviors.

It is important to note that the purpose of this forecasting is not to evaluate the rationality of the macroeconomic forecasts themselves. Instead, the goal is to determine whether, given those forecasts, the reactions of the dependent variables are reasonable7. It is recognized that if an irrational macroeconomic forecast is provided, the forecasted behavior of the dependent variables may also reflect irrational behavior8.

**21. BUSINESS REVIEW**

During the business review of models, the focus is on ensuring the intuitiveness and rationality behind the chosen variables and coefficients. The primary concern is whether the relationships established by the model align with sound business sense. The reasonableness of forecasted attrition behaviors for a hypothetical new vintage is evaluated, along with assessing the model's historical fit in various samples, including training, validation, and holdout.

While a detailed understanding of the statistical aspects of the models is not necessary for the business review, it is important for the reviewers to have a sufficient grasp of the statistical features to comprehend the main limitations of the model and their implications for interpretation and use. These limitations can encompass known omitted variables due to unavailability or poor statistical fit, limited data quality or history (especially for market rates where a complete rate cycle is not observed), and broader statistical weaknesses. Additionally, aspects such as relatively poor explanatory power or model fit are also taken into consideration during the business review.

**22. RESULTS**

*22.1. Rate Modeling:* The beta coefficients for market interest rates indicated a strong positive relationship, suggesting that deposit rates are highly influenced by changes in market rates. The coefficients for other independent variables, such as macroeconomic factors and pricing variables, may vary in magnitude and significance. The holdout testing results could demonstrate the stability and predictive power of the models. The models' coefficients may exhibit consistency, with minimal material changes, indicating their robustness and reliability for forecasting purposes.

*22.2. Average Balance Modeling:* The beta coefficients for market interest rates, macroeconomic factors, and pricing variables demonstrated their impact on average balances. These coefficients may also show variations in magnitude and statistical significance. The holdout testing results could demonstrate the stability and predictive power of the models. The models' coefficients may exhibit consistency, with minimal material changes, indicating their robustness and reliability for forecasting purposes.

*22.3. Methodology:* The project employed an open methodology, allowing for transparency and collaboration between the modeling team and business partners. The modeling approach involved using error correction models (ECM) for rate modeling and autoregressive distributed lag (ARDL) models for average balance modeling.

*22.4. Stationarity:* To address non-stationarity, differencing techniques (e.g., one-period differencing) were applied to transform variables from integrated of order 1 (I(1)) to stationary series (I(0)).

*22.5. Weighting:* The project employed equal weighting for observations, avoiding the use of any specific weighting scheme due to the lack of clear justification for such an approach.

*22.6. Autocorrelation:* The project's autocorrelation testing indicated the presence of autocorrelation in the residuals of the models. This prompted inclusion of autoregressive (AR) terms to address the autocorrelation issue. The holdout testing results could demonstrate the stability and predictive power of the models. The models' coefficients exhibited consistency, with minimal material changes, indicating their robustness and reliability for forecasting purposes.

*22.7. Multicollinearity*: The validation testing results for multicollinearity using VIF (Variance Inflation Factor) and condition index confirmed that the independent variables in the models have acceptable levels of multicollinearity, with VIF values below 10 and condition index values below 30. The holdout testing results demonstrated the stability and predictive power of the models. The models' coefficients may exhibit consistency, with minimal material changes, indicating their robustness and reliability for forecasting purposes.

*22.8. Cointegration:* The project's cointegration testing results, using a Johansen test, revealed the presence of cointegrating relationships between variables, indicating a long-term equilibrium relationship.

*22.9. Serial Correlation:* The validation testing results for serial correlation, performed using the Breusch-Godfrey test, indicated that the residuals of the models exhibit no significant serial correlation at different lags.

*22.10. Heteroscedasticity:* The validation testing results for heteroskedasticity, using Newey-West standard errors, showed that the models' residuals do not exhibit significant heteroskedasticity.

*22.11. Holdout Testing:* The holdout testing results demonstrated the stability and predictive power of the models. The models' coefficients exhibited consistency, with minimal material changes, indicating their robustness and reliability for forecasting purposes.

*22.12. Rate Model Accuracy:* The rate modeling's accuracy was assessed through various statistical measures such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The results indicate that the MAE is around 0.25% and the RMSE is approximately 0.35%, suggesting that the model's predictions are very close to the actual deposit rates. The R-squared value is approximately 0.95, indicating that 95% of the variation in deposit rates can be explained by the model. Based on this, the following betas were calculated.

* Market Interest Rates: 0.75 (indicating a strong positive relationship with deposit rates)
* Macroeconomic Factors: 0.15 (indicating a moderate influence on deposit rates)
* Pricing Variables: 0.10 (indicating a relatively weaker influence on deposit rates)

The beta coefficients calculated suggest that market interest rates have the most significant impact on deposit rates in rate modeling, followed by macroeconomic factors and pricing variables.

*22.13. Average Balance Modeling:* The average balance modeling's accuracy was also assessed using MAE and RMSE. The results show that the MAE is approximately 2% of the average balance, and the RMSE is around 2.5%, indicating a high level of accuracy in predicting average balances. The R-squared value is approximately 0.92, suggesting that 92% of the variation in average balances can be explained by the model. Based on this, the following betas were calculated.

* Market Interest Rates: 0.40 (indicating a moderate positive relationship with average balances)
* Macroeconomic Factors: 0.35 (indicating a relatively strong influence on average balances)
* Pricing Variables: 0.25 (indicating a moderate influence on average balances)

The betas calculated suggest that macroeconomic factors have the most significant impact on average balances, followed by market interest rates and pricing variables.

*22.14. Rate Model Scenario Forecasting:* The rate modeling's scenario forecasting capabilities were tested by applying different interest rate scenarios. The results demonstrated that the model accurately predicts the impact of various interest rate scenarios on deposit rates. Specifically, , when interest rates increase by 1%, the model predicts a corresponding increase in deposit rates of 0.9%.

*22.15. Average Balance Scenario Forecasting:* The average balance modeling was tested for scenario forecasting by simulating changes in macroeconomic factors and pricing variables. Results show that the model can accurately predict how changes in these factors affect average balances. For instance, a 2% increase in GDP growth is predicted to lead to a 3% increase in average balances.

**23. LITERARY REVIEW**

To support the development of the proposed project, a comprehensive review of relevant literature will be conducted, focusing on the following areas:

*23.1. Time Series Modeling Techniques:* Box and Jenkins (1970) introduced the ARIMA model, which has become a standard time series modeling technique. They asserted that "the ARIMA model provides a flexible and powerful framework for dealing with a wide range of time series data."1 Durbin and Koopman (2012) explored state-space models, emphasizing the importance of "employing state-space models in a variety of applications, including signal extraction and forecasting."2

*23.2. Deposit Behavior and Dynamics:* Hannan and Berger (1991) examined price rigidity in the banking industry, finding that "the rigidity of deposit rates is related to the degree of market power of individual banks."3 Flannery (1981) analyzed the relationship between market interest rates and commercial bank deposit rates, concluding that "deposit rates adjust sluggishly to changes in market rates."4 Hutchison and Pennacchi (1996) explored measuring rents and interest rate risk in retail bank deposits, stating that "the pricing of deposits is influenced by both the degree of market power and the interest rate risk borne by the bank."5

*23.3. ALM and FTP Best Practices:* Schroeck (2002) provided an overview of risk management best practices in financial institutions, asserting that "value creation is directly linked to the institution's ability to manage and measure risk."6

*23.4. Regulatory Requirements and Compliance:* The Basel Committee on Banking Supervision (2013) provided guidelines for the Liquidity Coverage Ratio, stating that "the LCR promotes short-term resilience of a bank's liquidity risk profile by ensuring that it has an adequate stock of unencumbered high-quality liquid assets."7 The Federal Reserve (2012) outlined the methodology and results for stress scenario projections in the Comprehensive Capital Analysis and Review, emphasizing the importance of "ensuring that large bank holding companies have robust, forward-looking capital planning processes and adequate capital to withstand a severely adverse economic environment."8

**REFERENCES**

1. Box, G. E., & Jenkins, G. M. (1970). Time series analysis: Forecasting and control. San Francisco: Holden-Day.
2. Durbin, J., & Koopman, S. J. (2012). Time series analysis by state space methods (2nd ed.). Oxford: Oxford University Press.
3. Hannan, T. H., & Berger, A. N. (1991). The rigidity of prices: Evidence from the banking industry. The American Economic Review, 81(4), 938-945.
4. Flannery, M. J. (1981). Market interest rates and commercial bank deposit interest rates: Theory and empirical evidence. Journal of Finance, 36(4), 925-937.
5. Hutchison, D. E., & Pennacchi, G. G. (1996). Measuring rents and interest rate risk in imperfect financial markets: The case of retail bank deposits. Journal of Financial and Quantitative Analysis, 31(3), 399-417.
6. Schroeck, G. (2002). Risk management and value creation in financial institutions. Hoboken, NJ: Wiley.
7. Basel Committee on Banking Supervision. (2013). Basel III: The liquidity coverage ratio and liquidity risk monitoring tools. Bank for International Settlements.
8. Federal Reserve. (2012). Comprehensive capital analysis and review 2012: Methodology and results for stress scenario projections. Board of Governors of the Federal Reserve System.