Factor-Augmented Error Correction Model with Time Varying Coefficients

A Dissertation Presented

by

Xue Hao

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Applied Mathematics and Statistics

Stony Brook University

August 2015

Stony Brook University

The Graduate School

Xue Hao

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Wei Zhu
Dissertation Advisor
Professor, Department of Applied Mathematics and Statistics

Xuefeng Wang
Chairperson of Defense
Assistant Professor, Department of Preventive Medicine

Song Wu
Member
Assistant Professor, Department of Applied Mathematics and Statistics

Keli Xiao
Outside Member
Assistant Professor, College of Business

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Factor-Augmented Error Correction Model with Time Varying Coefficients

by

Xue Hao

Doctor of Philosophy

in

Applied Mathematics and Statistics

Stony Brook University

2015

Factor-based models have been extensively used in economic and financial time series analyses. The Factor-augmented Error Correction Model (FECM) is a successful generalization of the Factor-augmented Vector Autoregression Model and the Error Correction Model for large panel nonstationary time series data. By combining the factors and error correction terms together, the FECM is able to utilize both the aggregated panel information summarized through the Dynamic Factor Model as well as the long-term equilibrium information introduced by the cointegration relationship. In this thesis we extend the FECM by allowing time-varying model parameters. There are ample evidences from both theoretical and empirical studies supporting the notion that the parameters of economic and financial models often change over time. By relaxing the parameters to be time-varying, the model will be more adaptable to complicated and realistic data structures, such as those with potential structural instability after a recession or crisis. We conclude this thesis by applying the newly developed time-varying FECM to provide more suitable models for PPNR (Pre-Provision Net Revenue) studies, part of the required modeling process in CCAR (Comprehensive Capital Analysis and Review) -- commonly known as the Federal Reserve’s Stress Test on big banks and other financial institutes.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

Factor-based time series analysis has received considerable attention in quantitative studies of economic and financial issues in the past fifteen years (Stock and Watson, 2010). Due to the advancement in data gathering technologies and the expansion of data over time, the traditional time series models are increasingly challenged by higher data dimensionality. Factor models provide an effective way of synthesizing information contained in large sets -- allowing the application of advanced models on big data feasible (Bai, 2004). In recent years, factor model procedures have been adapted to multivariate time series models, such as Vector Autoregression (VAR) and Vector Error Correction Models (VECM). The combined factor-based time series models have become powerful tools in macroeconometrics empirical analysis – with further generalization, they will become increasingly indispensible. In this thesis, we will generalize and adapt this framework to U.S. bank stress test.

The US Federal Reserve Bank was compelled to install the stress test after the 2007-2009 global financial crisis and the accompanying great recession in the general market with the characteristics of severe disruption of financial markets, a stubbornly high unemployment, and a slow recovery of economy. It drew the Government’s attention to renewed calls for active macroprudential regulation aimed at preventing the build-up of risks in the financial system, while at the same time reducing the social and economic costs of financial instability (Covas, Rump, and Zakrajšek, 2014). To this end, as part of the effort, bank stress tests have become an indispensable part of the toolkit used by banks and regulators from central banks to conduct regulation and supervision (Hirtle, Schuermann, and Stiroh , 2009; Hanson, Kashyap, and Stein, 2011; Greenlaw, Kashyap, Schoenholtz, and Shin, 2012). Few papers have explicitly constructed suitable models for bank stress test analysis. In this thesis, we use the Factor-Augmented Error Correction Model (FECM) to conduct a top-down stress test exercise – with our focus on modeling the Pre-Provision Net Revenue (PPNR) study with macroeconomic variables. We extend the FECM by allowing for time-varying coefficients to better supporting the notion that the parameters of economic and financial models do change over time.

1.1 Factor-Based Time Series Models

Statistical analysis in economics and finance is facing increasing difficulties raised by growing data volume in both length and dimension. On one hand, the numbers of observations in financial time series are increasing along with the passage of time. On the other hand, statistical agencies have been collecting a greater amount of related macroeconomic, financial, and sectoral variables for much of the postwar period (Stock and Watson, 2010). Thus, researchers face data sets that have hundreds or even thousands of series (large N) with increasing observations (large T), resulting in the curse of dimensionality for classic time series models and analyses.

The Dynamic Factor Model (DFM) has gained popularity in the past fifteen years because of their ability to simultaneously and consistently model data sets with both large dimensions and lengths. Early work of dynamic factor models can be traced to Geweke (1977), who originally proposed the DFM as a time-series extension of factor models previously developed for cross-sectional data. The premise of dynamic factor models is that a few latent factors can explain a substantial amount of information in many time series. Sargent and Sims (1977) showed that a small number of factors, as few as two factors, could explain a large fraction of the variance of many important U.S. quarterly macroeconomic variables, including output, employment, and prices. This empirical finding has since been confirmed by many follow-up studies (Giannone, Reichlin, and Sala, 2004; Watson, 2004).

The frequency domain methods in the seminal work of Geweke (1977) and Sargent and Sims (1977) were useful in searching evidence of a dynamic factor structure and estimating the importance of the factor. However, those methods failed to estimate latent factors directly and hence could not be used for further analysis. Subsequent work on DFMs focused on estimating latent factors directly through time domain methods, which can be divided into three generations of evolving models (Stock and Watson, 2010). The first generation consisted of low-dimensional (small N) parametric models estimated in the time domain using the Gaussian maximum likelihood estimation (MLE) and the Kalman filter (Engle and Watson, 1981, 1983; Stock and Watson, 1989; Sargent, 1989; Quah and Sargent, 1993). This parametric method limited the number of series that can be handled due to the restriction of its nonlinear optimization procedure. The second generation nonparametric methods solve the dimensionality problem and are suitable for data sets with large N. They estimate the factors by using cross-sectional averaging methods, primarily principal components and related methods (Stock and Watson, 2002b; Bai, 2003; Bai and Ng, 2006a). The third generation of methods for estimating the factors combines the first and the second generation models by merging the statistical efficiency of the state space approach with the robustness and convenience of the principal components approach (Giannone, Reichlin, and Small, 2008; Doz, Giannone, and Reichlin, 2012). This merger overcomes the dimensionality problems faced by the first generation methods, and improves the second generation methods in handling missing data and capturing persistent and small common components.

Another issue involved in the factor model estimation is determining the number of factors. There are several methods available for estimating the number of stationary and nonstationary factors based on the principle components estimation. The screen plots introduced by Cattell (1966), which is a plot of the ordered eigenvalues against the rank of that eigenvalues, are useful visual diagnostic measures that allow one to assess the marginal contribution of principal components intuitively. Ahn and Horenstain (2013) proposed a group of theoretical measures that corresponds to finding the edge of the cliff in the screen plot. Another popular strand is to estimate the number of factors using the information criteria. Bai and Ng (2002) developed a family of factor number estimators that are motivated by the information criteria used in model selection. Their simulation results suggested that their proposed criterion outperform the traditional information criteria such as the Bayesian Information Criterion (BIC) on estimating static factor numbers in stationary data. Bai (2004) later developed these information criteria to accommodate models with nonstationary data and nonstationary factor.

With reliable estimates of the factors and the number of factors in hand, the estimated factors can be used as data in second stage analysis. Uses of the factors include but are not limited to multistep forecasts (Stock and Watson, 1999, 2002a, 2006, 2009; Boivin and Ng, 2005; Eickmeier and Ziegler, 2008), conducting instrumental variables and generalized method of moments (GMM) analysis (Favero, Marcellino, and Neglia, 2005; Beyer, Farmer, Henry, and Marcellino, 2005, 2008; Kapetanios and Marcellino, 2010; Bai and Ng, 2010), Factor-Augmented Vector Autoregression (FAVAR) (Bernanke, Boivin and Eliasz, 2005; Stock and Watson, 2005), and dynamic stochastic general equilibrium (DSGE) modeling (Boivin and Giannoni, 2006).

In this thesis, we focus on taking advantage of factors in nonstationary time series analysis. Typically, the factors extracted from the dynamic factor models are utilized in the vector error correction models, forming the Factor-Augmented Error Correction Model (FECM). The Factor-augmented Error Correction Model (FECM) was first proposed by Banerjee and Marcellio (2008) as a generalization of the Factor-augmented Vector Autoregression (FAVAR) Model. The method of FAVAR assumes that the data have been transformed to eliminate unit roots and trends. Typically this is accomplished by differencing the original time series (Bernanke, Boivin and Eliasz, 2005; Stock and Watson, 2010; Korobilis, 2012). By introducing the error correction term, which is omitted in the FAVAR model, the FECM is enabled to analyze large panel nonstationary time series data sets directly with the long-term equilibrium and cointegration information considered in the model. Furthermore, we extend the Factor-augmented Error Correction Model to allowing time-varying coefficients (TVC) in order to better capture the evolution of relations between factors and variables of interest in the model. We apply this extended TVC-FECM technique to the U.S. bank stress test analysis with historical data from the past 25 years, and expect to draw comprehensive conclusions about banks performances under different economic environments.

1.2 U.S. Bank Stress Tests

The last recession caused by the subprime mortgage crisis had enormous and persistent impact on US and global economy and financial markets. The nationwide banking emergency has drawn the attention of central banks and other regulators to seek more effective and powerful ways to enhance the capabilities of preventing the build-up of risks in the banking system.

As a response to the 2007-2009 Great Recession, the Dodd–Frank Wall Street Reform and Consumer Protection Act was passed in 2010 that brought about the most significant changes to financial regulations in the United States. It requires the Federal Reserve to conduct an annual assessment of the large bank holding companies (BHCs) and selected financial institutions in the United States considering their capital adequacy to continue operations under economic and financial stress environment. The assessment consists mainly of two programs with different emphasis: the Comprehensive Capital Analysis and Review (CCAR) and the Dodd-Frank Act Stress Testing (DFAST), requiring large, complex U.S. bank holding companies (BHCs) with $50 billion or greater in total consolidated assets to participate starting 2011, and BHCs and state member banks with total consolidated assets of more than $10 billion to participate starting 2013 respectively.

As stated in the "Comprehensive Capital Analysis and Review 2015: Assessment Framework and Results" published by Federal Reserve Board in mid-March 2015, the main objective of the Comprehensive Capital Analysis and Review (CCAR) is to assess the largest U.S. BHCs' capital adequacy, capital adequacy process, and their planned capital distributions. As part of CCAR, the Federal Reserve evaluates whether BHCs have sufficient capital to continue operations throughout times of economic and financial market stress and whether they have robust, forward-looking capital planning processes that would thoroughly account for their unique risks. At the same time, the Dodd-Frank Act stress testing (DFAST), which is a complementary exercise to CCAR, is a forward-looking quantitative company-run evaluation of the impact of stressful economic and financial market conditions on BHC capital. DFAST helps the Federal Reserve and the financial companies supervised by the Federal Reserve to assess whether institutions have sufficient capital to absorb losses and to support normal operations during adverse economic conditions. While DFAST is complementary to CCAR, both efforts are built upon similar processes, data, supervisory exercises, and requirements. In the empirical application of this thesis, we apply our newly developed methods to a selected group of largest U.S BHCs as representatives, which are participants of both programs.

Typically, the CCAR procedure is divided into two stages. In the first stage, namely late-October to the beginning of January, the participating BHCs prepare and submit their capital plan and supporting documentation to the Federal Reserve. In the second stage, the Federal Reserve conducts a two-pronged approach to evaluate both the institution-specific and industry-wide risks, and publishes the results of its supervisory stress test under both the supervisory severely adverse and adverse scenarios by the end of March.

In CCAR, scenario analysis is an integral part of this supervisory assessment procedure. Scenario analysis involves the application of historical or hypothetical scenarios to assess the impact of various events on the performance of banks (Guerrieri and Welch, 2012). In the instance of 2015 CCAR, a total of 31 largest U.S. bank holding companies were asked to submit capital plans for a full nine-quarter planning horizon through the end of 2016, reflecting the hypothesized macroeconomic baseline and stress scenarios provided by the Federal Reserve in "2015 Supervisory Scenarios for Annual Stress Tests Required under the Dodd-Frank Act Stress Testing Rules and the Capital Plan Rule" published in late-October 2014. Three supervisory scenarios – baseline, adverse and severely adverse scenarios – are characterized by 28 variables including domestic and international macroeconomic features as listed in the Table 1.1. The scenarios start in the fourth quarter of 2014 (2014:Q4) and extend through the fourth quarter of 2017 (2017:Q4).

The participating banks are asked to evaluate their capital structures under these hypothetical scenarios and submit their capital plan to the Federal Reserve. The assessment of

Domestic (16 Variables)	
Economic Activity and Prices	
Real GDP Growth	Percentage changes (at an annual rate) in real and nominal Gross Domestic Product (GDP);
Nominal GDP Growth	
Unemployment Rate	Unemployment rate of the civilian non-institutional population aged 16 years and over;
Real Disposable Income Growth	Percentage changes (at an annual rate) in real and nominal disposable personal income;
Nominal Disposable Income Growth	
CPI Inflation Rate	Percentage change (at an annual rate) in the Consumer Price Index (CPI).
Asset Prices or Financial Conditions	
House Price Index	Indices of house prices;
Commercial Real Estate Price Index	Commercial property prices;
Dow Jones Total Stock Market Index	Equity prices;
Market Volatility Index	U.S. stock market volatility.
Interest Rates	
3-month Treasury Rate	The rate on the 3-month Treasury bill;
5-year Treasury Yield	The yield on the 5-year Treasury bond;
10-year Treasury yield	The yield on the 10-year Treasury bond;
BBB Corporate Yield	The yield on a 10-year BBB corporate security;
Mortgage Rate	The interest rate associated with a conforming, conventional, fixed-rate 30-year mortgage;
Prime Rate	The prime rate.
International (12 Variables)	
The three variables for each country or country block:	Percentage change (at an annual rate) in real GDP;
	Percentage change (at an annual rate) in the CPI or local equivalent;
	Level of the U.S. dollar/foreign currency exchange rate.
The four countries or country blocks included:	The euro area (the 18 European Union member states that have adopted the euro as their common currency);
	The United Kingdom;
	Developing Asia (the nominal GDP-weighted aggregate of China, India, South Korea, Hong Kong Special Administrative Region, and Taiwan);
	Japan.

Table 1.1 28 Variables in Supervisory Scenarios for 2015 CCAR (Source: 2015 Supervisory Scenarios for Annual Stress Tests Required under the Dodd-Frank Act Stress Testing Rules and the Capital Plan Rule)

the expected uses and sources of capital over the planning horizon in a capital plan mandated by the Federal Reserve is initiated by estimating of projected revenues, losses, reserves, and pro forma regulatory capital ratios, including tier 1 leverage, common equity tier 1 capital, tier1 risk-based capital, and total risk-based capital ratios, over the planning horizon under baseline conditions and stressed scenarios. The key requirement for a bank to pass the stress test is that its projected tier 1 common capital ratio (T1CR) – which is calculated as the ratio of tier 1 capital to total risk-weighted assets using the definition of in 12 CFR 225, appendix A – under the severely adverse scenario remains above a specified minimum threshold over the forecast horizon. In the CCAR 2015, this threshold is set to be 5 percent. In addition, each institution also has to maintain all other regulatory capital ratios above a minimum of 4 to 8 percent, respectively.

The Federal Reserve employs a two-pronged approach in its supervisory stress test based on the projected capital plans submitted from large BHCs. In the "bottom-up" models, proprietary granular data on institution-specific portfolios, which contain detailed information about individual loan characteristics that is provided by BHCs to the Federal Reserve confidentially, are used to estimate bank-level losses and revenues. The "top-down" models, from the other side, use macroeconomic variables and bank balance sheet data to estimate institution-specific and industry-wide loans, revenues and capital measures. The results of the top-down stress test models are useful benchmarks for the aggregated results from the bottom-up model under different macroeconomic scenarios (Covas, Rump and Zakrajšek, 2013).

In this thesis, with the help of large panel macroeconomic information aggregated in the factors and the limitation of publicly available data, we particularly focus on the top-down stress test models. In a sense, we are analyzing financial variables based on aggregate macroeconomic information. Intuitively, even slow moving variables, such as the unemployment rate, may contain useful information for modeling some of the banking measures that exhibit little high-

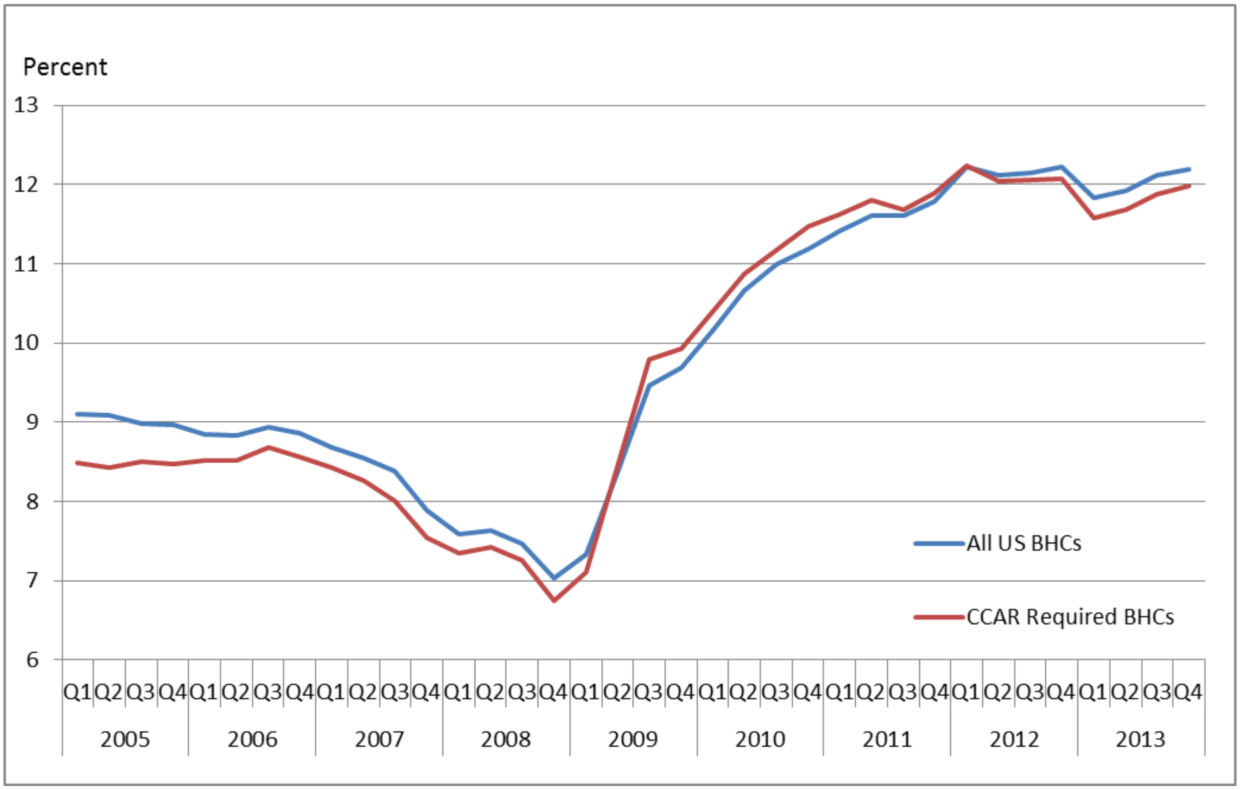

Figure 1.1 Capital Adequacy (Tier 1 Common Capital Ratio) of the U.S. Bank Holding Companies: The red line shows the aggregate tier 1 common ratio (T1CR) for the 31 BHCs that participated in the CCAR 2015; the blue line shows the aggregate T1CR of all U.S. BHCs that reported in form FR Y-9C, *Consolidated Financial Statements for Holding Companies*.T1CR is defined as the ratio of tier 1 common capital to total risk-weighted assets, both are defined in 12 CFR 225, appendix A. The 31 participated BHCs are: Ally Financial Inc.; American Express Company; Bank of America Corporation; The Bank of New York Mellon Corporation; BB&T Corporation; BBVA Compass Bancshares, Inc.; BMO Financial Corp.; Capital One Financial Corporation; Citigroup Inc.; Comerica Incorporated; Deutsche Bank Trust Corporation; Discover Financial Services; Fifth Third Bancorp; The Goldman Sachs Group, Inc.; HSBC North America Holdings Inc.; Huntington Bancshares Inc.; JPMorgan Chase & Co.; Keycorp; M&T Bank Corporation; Morgan Stanley; MUFG Americas Holdings Corporation; Northern Trust Corporation The PNC Financial Services Group, Inc.; RBS Citizens Financial Group, Inc.; Regions Financial Corporation; Santander Holdings USA, Inc.; State Street Corporation; SunTrust Banks, Inc.; U.S. Bancorp; and Wells Fargo & Co.; Zions Bancorporation.

frequency variation, for example, the tier 1 common capital ratios (Guerrieri and Welch, 2012). A few papers confirmed this idea: Bellotti and Crook (2009) found that macroeconomic variables were statistically significant in their credit account default forecasting models; Simons and Rolwes (2009) used a macroeconomic-based model to estimating probabilities of default and found convincing relations between macroeconomic variables and the default behavior of Dutch firms. As to the U.S. bank stress tests, Grover and McCracken (2014) used factor-based models and reached the conclusion that macroeconomic variables provided in the Federal Reserves stressed scenarios are useful for identifying stress of bank at the industry-wide level.

According to data recorded in form FR Y-9C, Consolidated Financial Statements for Holding Companies, the 31 selected CCAR participating large BHCs consist of 77.16% capital of all U.S. bank holding companies, in terms of the consolidated assets. Figure 1 shows that the aggregate tier 1 common capital ratio (T1CR) of CCAR participating BHCs, as a representative of the U.S. banking system, recovered to almost double of the nadir of the 2007-2009 financial crisis since the CCAR implement. Covas, Rump and Zakrajšek (2013) suggested that this significant improvement in the capital position and loss-absorbing capacity of the U.S. banking system is a key metric indicating the success of the Federal Reserve conducted stress tests. The stress tests boosted the issuance of common equity and increased retained earnings. Moreover, based on the outcomes of the stress tests, the Federal Reserve imposed restrictions on dividend payouts and share repurchases that would partly affect the financial decisions that the “stressed” institutions undertook. These promising outcomes drove the increase of the T1CR, which implied substantially enhanced resiliency of the banking sector since the end of the recession.

Chapter 2 Factor-Augmented Error Correction Model

The Factor-Augmented Error Correction Model (FECM) was first introduced by Banerjee and Marcellio (2009) as a way to model key macroeconomic variables of interest jointly with factors extracted from large panel time series variables in level. Starting from the Dynamic Factor Models for non-stationary data, with established literature discussing their applications in economics and finance realms, the FECM brings the concept of cointegration and error correction, two important strands of the econometrics literature, into the analysis. In this chapter, we first reproduce the derivation of the FECM with time constant coefficient (TCC).

2.1 Model Description and Notations

Suppose we have a small set of variables of interest denoted by Y_t. For example, in the bank stress testing, we are interested in a group of balance sheet variables that reflect the revenue, losses, reserves and capital structures of the financial institutions under different economic scenarios. The standard approaches used extensively by regulatory authorities around the world are autoregression models assuming linear relations between the variables of interest and bank characteristics or macroeconomic variables (Duane, Schuermann, Reynolds and Wyman, 2013; Covas, Rump and Zakrajšek, 2013; Grover and McCracken, 2014). To be specific, $Y_t = [Y_{1t}, \dots, Y_{Mt}]'$ is an M-dimensional time series and is supposed to be linearly affected by the N-dimensional observed variables $X_t = [X_{1t}, \dots, X_{Nt}]'$:

$$Y_t = BX_t + e_t \tag{2.1}$$

where e_t is zero-mean I(0) process that can be serially and cross-correlated, and B is the loading matrix. In our empirical application of bank stress tests analysis, $X_t = [X_{1t}, \dots, X_{Nt}]'$ is a selected group of *N* macroeconomic variables or sector indicators providing a representation of the economy, such as product output, prices index, interest rates, and so on. In general, Y_t may contain subset of X_t, in which case the corresponding elements in the idiosyncratic component e_t is constant zero.

Typically, large amount of macroeconomic and financial variables are observed in levels. For example, in our empirical analysis, 148 out of 156 macroeconomic time series are identified as nonstationary by the Augmented Dickey–Fuller (ADF) unit root test. To better incorporate the long term equilibrium information and possible cointegration of the original level data, it is a natural choice of us to consider a Vector Error Correction Model (VECM) under the stationary assumption of idiosyncratic term e_t in (2.1). However, direct derivation of VECM from (2.1) will become heavily parameterized when a large set of observation data X_t is incorporate in the model, as is the case in bank stress tests faced by the banks and regulatory central banks. As a result, it makes the model very difficult or even impossible to solve. To avoid this problem of dimensionality, we decompose the *N*-dimensional observed time series X_t into a lower dimensional vector of factors with the help of Dynamic Factor Model (DFM), and hence derive the Factor-Augmented Error Correction Model (FECM).

Consider the following generalized Dynamic Factor Model (DFM) as in Banerjee, Marcellino and Masten (2015) for a large set of variables:

$$X_{it} = \lambda_i(L)F_t + \psi(L)c_t + \epsilon_{it} \tag{2.2}$$

where $i = 1, \dots N, t = 1, \dots, T,\;\; F_t = (F_{1t}, \dots, F_{rt})'$ is an r-dimensional vector of I(1) latent factors, which follows a random walk. $c_t = \left(c_{1t}, \dots, c_{q_c t}\right)'$is a q_c-dimensional vector of I(0) latent factors. For any $t < 0, F_t = c_t = 0$.

$$\lambda_i(L) = \textstyle\sum_{j=0}^{n} \lambda_{ij} L^j \;\text{ and } \psi(L) = \sum_{l=0}^{m} \psi_{il} L^l$$

are lag polynomials of finite order p and m respectively, with L denotes the lag operator. The loading λ_{ij} and ψ_{ij} can be either deterministic or mutually independent stochastic variables. For

$$\lambda_i = \lambda_i(1) \text{ and } \;\psi_i = \psi_i(1),$$

we assume the loadings satisfy

$$E\|\lambda_i\|^4 \le M < \infty \text{ and } E\|\psi_i\|^4 \le M < \infty,$$

$$\tfrac{1}{N}\textstyle\sum_{i=0}^{N} \lambda_i {\lambda_i}' \xrightarrow{p} \Sigma_\Lambda \text{ and } \tfrac{1}{N}\sum_{i=0}^{N} \psi_i {\psi_i}' \xrightarrow{p} \Sigma_\Psi \text{ as } N \to \infty,$$

where Σ_Λ and Σ_Ψ are positive definite non-random matrices.

The ϵ_{it} is a zero-mean I(0) idiosyncratic component. The idiosyncratic component ϵ_{it} can be serially and cross-correlated, but is independent of loading λ_{ij} and ψ_{ij} for all points.

It is convenient to write the above generalized DFM into a restricted DFM in which the factors are dynamic but the relation between the dynamic factors and the observable variables is static. Following the definition in Bai (2004) and Banerjee, Marcellino and Masten (2015), let us denote

$$\tilde{\lambda}_{ik} = \lambda_{ik} + \lambda_{ik+1} + \dots + \lambda_{in}, \quad k = 0, \dots, n.$$

$$\widetilde{\Psi}_i = [\psi_{i0}, \dots, \psi_{im}].$$

Then, we write the static form DFM as follows:

$$X_t = \Lambda F_t + \Psi G_t + \epsilon_t \tag{2.3}$$

in which

$$X_t = [X_{1t}, \dots, X_{Nt}]',$$

$$\Lambda = [\Lambda_1', \dots, \Lambda_N']', \quad where\ \Lambda_i = \tilde{\lambda}_{i0},$$

$$\Psi = [\Psi_1', \dots, \Psi_N']', \quad where\ \Psi_i = \left[\widetilde{\Psi}_i, -\tilde{\lambda}_{i1}, \dots, -\tilde{\lambda}_{in}\right],$$

$$G_t = [c_t{}', c_{t-1}{}', -1\ c_{t-m}{}', -F_t', \dots, \Delta F_{t-n+1}']',$$

$$\epsilon_t = [\epsilon_{1t}, \dots, \epsilon_{Nt}]'.$$

The static DFM (2.3) isolates the I(0) factors G_t from I(1) factors F_t, and eliminates the appearances of the lags of the factors. From the discussion in Bai (2004), without loss of generality, we assume the r-dimensional I(1) factors F_t does not have cointegration among itself. In fact, a k-dimensional cointegrated I(1) factors can always be expressed by r non-cointegrated I(1) factors and $k - r$ stationary I(0) factors. In details, suppose F_t is a vector of cointegrated I(1) factors. We can always find an invertible matrix P such that $PF_t = (\xi_t', \eta_t')'$, where ξ_t is a vector of non-cointegrated I(1) factors and η_t is a vector of I(0) factors that are linear combinations of F_t. Thus, the presence of I(0) factors in the above static DFM can accommodate the cointergrated non-stationary factors. For simplicity, denote the dimension of G_t as $q = r(p + 1) + q_c(m + 1)$.

Bringing the previous static DFM (2.3) into the equation (2.1), we obtain

$$Y_t = B\Lambda F_t + B\Psi G_t + B\epsilon_t + e_t$$

$$= \Lambda_{\mathrm{B}} F_t + \Psi_{\mathrm{B}} G_t + \epsilon_{Bt} \qquad (2.4)$$

where $\Lambda_{\mathrm{B}} = B\Lambda$ is a M-by-r loading matrix, $\Psi_{\mathrm{B}} = B\Psi$ is a M-by-q loading matrix and $\epsilon_{Bt} = B\epsilon_t + e_t$ is still a zero-mean I(0) process. The equation (2.4) is in the form of static DFM with both I(1) and I(0) factors on its right. Since F_t are non-conintegrated I(1) factors, for each given $i = 1, \dots, M$, the process Y_{it} is I(1) unless the *i*-th row of $\Lambda_{\mathrm{B}} = [\lambda_{B1}, \dots, \lambda_{BM}]$, $\lambda_{Bi} = 0$. It is clear that Y_{it} and F_t are cointegrated for each i since the remaining terms $\Psi_{\mathrm{B}} G_t$ and ϵ_{Bt} are stationary. The idiosyncratic component ϵ_{Bt} of equation (4) can be serially and cross-correlated, but independent of elements in loading matrices Λ_{B} and Ψ_{B}. Specially, for the I(0) process $\epsilon_{Bt} = (\epsilon_{B1t}, \dots, \epsilon_{BMt})'$, we assume that

$$\epsilon_{Bt} = \Gamma(L)\epsilon_{Bt-1} + \upsilon_t \qquad (2.5)$$

where υ_t are orthogonal white noise errors, and the roots of $\Gamma(L)$ lie inside the unit disc.

As shown in Banerjee, Marcellino and Masten (2014b), the serial correlation of the idiosyncratic component ϵ_{Bt} can be eliminated from the equation by pre-multiplying $I - \Gamma(L)L$ to (2.4). This leads to the error correction form of the DFM:

$$\Delta Y_t = -\big(I - \Gamma(1)\big)(Y_{t-1} - \Lambda_{\mathrm{B}} F_{t-1} - \Psi_{\mathrm{B}} G_{t-1}) + \Lambda_{\mathrm{B}} \Delta F_t + \Gamma_1(L)\Lambda_B \Delta F_{t-1}$$

$$+\Psi_{\mathrm{B}} \Delta G_t + \Gamma_1(L)\Psi_{\mathrm{B}} \Delta G_{t-1} - \Gamma_1(L)\Delta X_{t-1} + \upsilon_t, \qquad (2.6)$$

where the factorization of $\Gamma(L)$:

$$\Gamma(L) = \Gamma(1) - \Gamma_1(L)(1 - L)$$

is utilized. We denote the equation (2.6) as Factor-Augmented Error Correction Model (FECM).

Notice here we put the component of the I(0) factors G_t in the error correction term. Banerjee, Marcellino and Masten (2015) pointed out that it is best to include $G_t$in the cointegration space, since the model need to incorporate the information in the I(0) factors G_t. In fact, as we will discuss in details in the next chapter, the consistent estimation of the factors is only up to a rotation so that we will need to include the I(0) factors $G_t$in the cointegration space to allow the possibility that the estimated I(1) factors to be cointegrated.

2.2 Vector Factor-Augmented Error Correction Model

To complete the model, we go back to the beginning of this chapter. Following Banerjee, Marcellino and Masten (2015), we assume the I(1) factors F_t follow a random walk process:

$$F_t = F_{t-1} + \epsilon_t^F,$$

while the I(0) factors c_t are described in a VAR(1) expression

$$c_t = \rho_c c_{t-1} + \epsilon_t^c,$$

where ρ_c is a diagonal matrix with values on the diagonal in absolute term strictly less than one. ϵ_t^F and ϵ_t^c are independent of elements λ_{ij}, ψ_{il} in the loading matrices and idiosyncratic error ϵ_{it} for any i, j, t. ϵ_t^F and ϵ_t^c are not necessarily be $i.i.d.$ and they are allowed to be serially and cross correlated. We represent ϵ_t^F and ϵ_t^0 in a stable vector process:

$$\begin{bmatrix} \epsilon_t^F \\ \epsilon_t^c \end{bmatrix} = A(L) \begin{bmatrix} \epsilon_{t-1}^F \\ \epsilon_{t-1}^c \end{bmatrix} + \begin{bmatrix} u_t \\ w_t \end{bmatrix},$$

where u_t and w_t are zero-mean white noise innovations to dynamic nonstationary and stationary factors respectively. We then express the model as

$$\begin{bmatrix} \epsilon_t^F \\ \epsilon_t^0 \end{bmatrix} = [I - A(L)L]^{-1} \begin{bmatrix} u_t \\ w_t \end{bmatrix},$$

under the stability assumption. Note under the previous assumptions, we have $E\|\epsilon_t^F\|^4 \leq M < \infty$. $\sum_{t=0}^{T} F_t F_t'$ converges at rate T^2, while $\sum_{t=0}^{T} c_t c_t'$ converges at the standard rate T. These imply that the cross-product matrices $\sum_{t=0}^{T} F_t c_t'$ and $\sum_{t=0}^{T} c_t F_t'$ converge at rate $T^{2/3}$. As a result, the elements of the matrix composed of these four elements jointly converge to form a positive definite matrix.

From the previous derivations in this section, we now write the VAR expression for the vector of nonstationary and stationary factors:

$$\begin{bmatrix} F_t \\ c_t \end{bmatrix} = \left[\begin{bmatrix} I & 0 \\ 0 & \rho_c \end{bmatrix} + A(L)\right] \begin{bmatrix} F_{t-1} \\ c_{t-1} \end{bmatrix} - A(L) \begin{bmatrix} I & 0 \\ 0 & \rho_c \end{bmatrix} \begin{bmatrix} F_{t-2} \\ c_{t-2} \end{bmatrix} + \begin{bmatrix} u_t \\ w_t \end{bmatrix}$$

$$= C(L) \begin{bmatrix} F_{t-1} \\ c_{t-1} \end{bmatrix} + \begin{bmatrix} u_t \\ w_t \end{bmatrix},$$

where the pre-multiplier $C(L) = \begin{bmatrix} C_{11}(L) & C_{12}(L) \\ C_{21}(L) & C_{22}(L) \end{bmatrix}$ is a matrix with block sizes corresponding to the partition between F_t and c_t of r and q_c. With the definition of G_t in (2.3), we can write the previous VAR representation in static form factors F_t and G_t:

$$
\begin{bmatrix}
I_r & 0 & \dots & 0 & 0 & \cdots & \cdots & 0 \\
0 & I_q & \cdots & 0 & \vdots & \ddots & \cdots & \vdots \\
\vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \ddots & \vdots \\
0 & 0 & \dots & I_q & 0 & \cdots & \cdots & 0 \\
-I_r & 0 & \cdots & 0 & I_r & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & I_r & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & 0 & 0 & \ddots & 0 \\
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & I_r
\end{bmatrix}
\begin{bmatrix}
F_t \\ c_t \\ c_{t-1} \\ \vdots \\ c_{t-m} \\ \Delta F_t \\ \Delta F_{t-1} \\ \vdots \\ \Delta F_{t-p+1}
\end{bmatrix} =
$$

$$
\begin{bmatrix}
C_{11}(L) & C_{12}(L) & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\
C_{21}(L) & C_{22}(L) & 0 & \cdots & 0 & 0 & \cdots & \vdots & 0 \\
0 & I_q & 0 & \cdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\vdots & \cdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & I_q & 0 & \cdots & \cdots & \cdots & 0 \\
-I_r & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & 0 & I_r & \cdots & 0 & \vdots \\
\vdots & \vdots & \ddots & \vdots & 0 & 0 & \ddots & 0 & \vdots \\
0 & 0 & \cdots & \cdots & 0 & 0 & \cdots & I_r & 0
\end{bmatrix}
\begin{bmatrix}
F_{t-1} \\ c_{t-1} \\ c_{t-2} \\ \vdots \\ c_{t-m-1} \\ \Delta F_{t-1} \\ \Delta F_{t-2} \\ \vdots \\ \Delta F_{t-p}
\end{bmatrix} +
\begin{bmatrix}
I_r & 0 \\ 0 & I_q \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ 0 & 0
\end{bmatrix}
\begin{bmatrix} u_t \\ w_t \end{bmatrix}
$$

By pre-multiplying the inverse of the initial matrix, the reduced form VAR for static factors is written as

$$
\begin{bmatrix} F_t \\ G_t \end{bmatrix} = M(L) \begin{bmatrix} F_{t-1} \\ G_{t-1} \end{bmatrix} + R \begin{bmatrix} u_t \\ w_t \end{bmatrix} \tag{2.7}
$$

where $M(L) = \begin{bmatrix} M_{11}(L) & M_{12}(L) \\ M_{21}(L) & M_{22}(L) \end{bmatrix}$ is a matrix with block sizes corresponding to the partition between F_t and G_t and R is a q – by – $(r + q_c)$ matrix accounts for dynamic singularity of G_t. Subtracting a vector of the lag of factors F_t and G_t on both sides of (2.7), we get

$$
\begin{bmatrix} \Delta F_t \\ \Delta G_t \end{bmatrix} = (M(L) - I) \begin{bmatrix} F_{t-1} \\ G_{t-1} \end{bmatrix} + R \begin{bmatrix} u_t \\ w_t \end{bmatrix} \tag{2.8}
$$

We combine this result with FECM function in (2.6) to obtain a vector FECM. Grouping the coefficients together, we write the vector FECM concisely as

$$\begin{bmatrix}\Delta Y_t\\ \Delta F_t\\ \Delta G_t\end{bmatrix} = \alpha\beta' \begin{bmatrix}Y_{t-1}\\ F_{t-1}\\ G_{t-1}\end{bmatrix} + \Phi_1 \begin{bmatrix}\Delta Y_{t-1}\\ \Delta F_{t-1}\\ \Delta G_{t-1}\end{bmatrix} + \cdots + \Phi_p \begin{bmatrix}\Delta Y_{t-p}\\ \Delta F_{t-p}\\ \Delta G_{t-p}\end{bmatrix} + \begin{bmatrix}v_t\\ Ru_t\\ Rw_t\end{bmatrix} \tag{2.9}$$

where the last term $\nu_t = [v_t', (Ru_t)', (Rw_t)']'$ is a zero-mean white noise process. Once the unobserved factors of F_t and G_t are identified, equation (2.9) is in the form of standard Vector Error Correction Model (VECM). We can then employ the typical estimation methods, such as the one provided by Johansen (1995), to estimate the remaining parameters in the model. In the following part of this thesis, we use the term FECM to refer to the vector form equation (2.9).

2.3 Comments and Discussion

Banerjee, Marcellino and Masten (2015) pointed out that from model (2.6) and our assumption of non-cointegrated structure on I(1) factors F_t, the cointegration is only between each individual interested variable and the factors. However, treating the cointegrated coefficient β as unrestricted, the expression (2.9) omits the potentially cross-equations correlations in the original model. For a similar reason, the loading matrix α and short term coefficents $\Phi_1, \dots, \Phi_p$ are also left unrestricted. The lag order in (2.9) cannot be directly recovered from the structure of $\Gamma(L)$ in (2.5) and $M(L)$ in (2.7). In our empirical applications, it is chosen based on the VAR model that is equivalent to the VECM as shown in Tsay (2010, page 432 – 433), via appropriate information criteria. Banerjee, Marcellino and Masten (2015) concluded that the extent of the potential mis-specification of (2.9) depends mainly on the structure of the matrix $\Gamma(L)$ in (2.5). When the true $\Gamma(L)$ is diagonal, which eliminates the cross-correlation among the idiosyncratic errors of (2.4), the approximated expression (2.9) is very close to the original model.

The derived FECM can accommodate the case in which ϵ_{Bit} in (2.4) and (2.5) is I(1) for some i (Banerjee, Marcellino and Masten , 2015). By fitting the framework of Bai and Ng (2004), the space spanned by F_t and G_t can be consistently estimated jointly instead of separately, using data in differences. However, we give the preference to stationary idiosyncratic component ϵ_{Bit} assumption in our model definition. This is not a strong restriction and the empirical illustrations support our choice (Banerjee, Marcellino and Masten , 2015). In our empirical experiment, 156 U.S. quarterly macroeconomic time series, covering the period from the first quarter of 1990 (1990:Q1) to the fourth quarter of 2014 (2014:Q4), are selected in the dataset. By applying the ADF unit root test to the estimated idiosyncratic components after extracting 5 factors from the dataset, 154 out of 156 residual series reject the unit root null hypothesis at significance level of 5%. The only two exceptions also report considerably small p-values of 0.0585 and 0.1221, respectively. Hence, we are confident in restricting the idiosyncratic components to be stationary processes.

From a theoretical point of view, the advantages of the FECM are two-folded. First, it is obvious that the FECM includes more information brought by factors from variables that are missing in the ECM. When the FECM is the true data generating process (DGP), the influence of this missing information is substantial, resulting in a dominant better performance in FECM. Simulation results from Banerjee, Marcellino and Masten (2015) suggested that even though the ECM outperforms the FECM in the cases when the underlying DGP is ECM, the relative loss form the use of FECM is rather small. Second, comparing to the Factor-Augmented Vector Autoregression (FAVAR) model in Bernanke, Boivin and Eliasz (2005), that extends the VAR with factors involved, FECM has an additional error correction term. The error correction term, representing the cointegration space of nonstationary time series, allows long-run components of variables to obey equilibrium constraints (Engle and Granger, 1987). The appearance of cointegration and error correction term in large macroeconomic dataset is statistically significant in empirical studies (Banerjee, Marcellino and Masten , 2014b). While the FAVAR is nested in the FECM specification by imposing α =0 in (2.9), the omitted error correction term cannot be recovered in the FAVAR when the true DGP follows FECM settings.

Note that from (2.4), it has

$$Y_{t-1} - \Lambda_{\mathrm{B}} F_{t-1} - \Psi_{\mathrm{B}} G_{t-1} = \epsilon_{Bt-1}$$

that it would look like the error correction term omitted in the FAVAR comparing to (2.6) could be replaced by additional lags of I(0) idiosyncratic term. Substituting this equation into (2.6), however, we obtain an expression with data in its first difference:

$$\Delta Y_t = \Lambda_{\mathrm{B}} \Delta F_t + \Psi_{\mathrm{B}} \Delta G_t + \Delta \epsilon_{Bt}, \qquad (2.10)$$

containing a non-invertible moving average (MA) component. While conventional structural analysis in a FAVAR framework relies on inverting a system like (2.10) (Stock and Watson, 2005; Lüetkepohl, 2014), unless we allow the number of factors to be infinite or include the explicit non-invertible MA structure of the error term, the standard FAVAR model generates biased results when the data is I(1).

The analytical and Monte Carlo simulation results in Banerjee, Marcellino and Masten (2014a) also suggested that the FECM is virtually always better than the FAVAR. This is consistent with the findings in our empirical application. Table 2.1 lists the Mean Square Error (MSE) of FECM relative to the MSE of FAVAR model of the in-sample forecasting results for the 9 series we considered in the empirical experiment using FAVAR and FECM respectively. In both models, the input factors are identical and extracted from the 156 macroeconomic time series with 100 quarterly observations from 1990:Q1 to 2014:Q4. The orders of lags are selected based on Bayesian Information Criterion (BIC) and the ranks of cointegration in FECM are determined by Johansen's trace test.

Variable	MSE of FECM relative to MSE of FAVAR	FAVAR	FECM	
		Lags	Lags	Cointegration Rank
Net Interest Income	0.757106	2	1	2
Total Noninterest Income	0.574723	2	1	3
Compensation Expense	0.596774	2	1	3
Fixed Assets Expense	0.778127	2	1	3
Total Noninterest Expense	0.637431	2	1	2
Net Interest Income of Citigroup	1.016468	2	1	2
Total Noninterest Income of Citigroup	0.719711	2	1	3
Net Interest Income of Wells Fargo	0.466292	2	1	4
Total Noninterest Income of Wells Fargo	1.07384	2	1	2

Table 2.1 MSE of FECM relative to MSE of FAVAR in forecasting PPNR components

Table 2.1 shows that from the 9 forecasting results, the FECM dominantly outperform the FAVAR in terms of MSE, with only two exceptions in which the methods are comparable. These results are sufficient evidences of the fact that the error correction terms are indispensable in our model fitting. Hence the FECM is expected to work substantially better than the FAVAR model.

Chapter 3 Factor Identification and Estimation

As mentioned in the previous chapter, the FECM in (2.9) can be treated as a standard VECM condition on the estimated factor spaces. Hence, the cointegration spaces and the remaining loading coefficients can be handled with approaches for standard VECM. In this chapter, we turn our attention to factor estimation.

3.1 Dynamic Factor Model

In section 2.1, the factors used in the subsequent FECM derivation are firstly described in a dynamic factor model of equation (2.2). Factor models provide an effective framework for synthesizing information contained in a substantial amount of time series. Factor models in high dimension are popular and have been successfully utilized in economic monitoring and forecasting, business cycle analysis, consumer theory analysis, asset pricing and so on (Bai, 2002; Baltagi, Kao and Wang, 2015).

A large amount of theory and econometrics analysis about high dimensional factor models has been intensively studied in the past fifteen years. Stock and Watson (2002) used the principle component method in estimating and forecasting the static factor model, and provided theorems showing that the factor space can be consistently estimated when both dimensions N and T go to infinity. Bai and Ng (2002) developed a group of panel criteria to obtain the consistent estimation of the number of factors, which improved the performance of traditional information criteria. Bai (2003) made the breakthrough by developing the inferential theory, including the theoretical discussion of the rate of convergence and the limiting distribution of the estimated factors, factor loadings, and common components, for large dimensional factor models with principal component estimators.

In the empirical applications we are interested in, a potentially large number of observed macroeconomic variables in levels are modeled as being simultaneously driven by a small number of unobserved underlying stationary and nonstationary factors. The latter is also known as cross-section common stochastic trends. While the previous works are mainly focused on the stationary observation and factors, Bai (2004) extended the theory and estimation methods to models with nonstationary dynamic factors and stationary idiosyncratic errors. Bai and Ng (2004) accommodate the situations when nonstationarity presents in both common and idiosyncratic components. The principal components estimator was employed in their studies.

3.2 Principal Components Estimation

Among the three generations of the factor estimation approaches, the second generation of nonparametric estimations using cross-sectional averaging methods is extensively used in large dimension factor models. The key advantages of the principal components estimator is that the space spanned by the factors is consistently estimated and the estimated factors are precisely enough to be treated as data in subsequent regressions, if the number of time series N is sufficiently large (Stock and Watson, 2010). Moreover, the principal components estimator is extremely easy to compute and is asymptotically equivalent to the maximum likelihood estimator when normality is assumed (Bai, 2004). The asymptotic properties and calculation efficiency make the principal components estimator the ideal estimation method for large data cases.

Specify the factor model as

$$X_t = \Lambda F_t + \epsilon_t \tag{3.1}$$

where $X_t = [X_{1t}, \dots, X_{Nt}]'$, $F_t = (F_{1t}, \dots, F_{rt})'$, $\epsilon_t = [\epsilon_{1t}, \dots, \epsilon_{Nt}]'$, $\Lambda = [\lambda_1, \dots, \lambda_N]'$. The vector representation is given by

$$X = F\Lambda' + \epsilon$$

where $X = [X_1, \dots, X_T]'$ is a $T \times N$ matrix observation, $F = [F_1, \dots, F_T]'$ and $\epsilon = [\epsilon_1, \dots, \epsilon_T]'$.

Let k denote the number of factors that are estimated, which is not necessarily equal to the true number of factors r at this stage, the principal component estimators are obtained by solving the optimization problem

$$\min_{\Lambda^k, F^k} S(k) = \min_{\Lambda^k, F^k} (NT)^{-1} \ \Sigma_{i=1}^{N} \Sigma_{t=1}^{T} \left(x_{it} - \lambda_i^{k\prime} F_t^k \right)^2 \tag{3.2}$$

where the superscripts on Λ^k and F^k signify the number of estimated factors. The results return by this minimization problem F^k and Λ^k are $T \times k$ matrix of estimated factors and its corresponding $N \times k$ matrix of estimated loadings. Notice that in the vector form of the factor model, F and are not separately identifiable. For an arbitrary $r \times r$ invertible matrix $A, F\Lambda' = FA^{-1}A\Lambda' = F^*\Lambda^{*\prime}$, where $F^* = FA^{-1}$ and $\Lambda^{*\prime} = A\Lambda'$. Hence, we impose one of the following normalization to the optimization problem

1) $\Lambda^{k\prime}\Lambda^k/N = I_k$ and $F^{k\prime}F^k$ is diagonal,

2) $F^{k\prime} F^k/T = I_k$ and $\Lambda^{k\prime}\Lambda^k$ is diagonal.

Since the roles of F and Λ in the common components are equivalent, there are two ways to obtain the estimates. From the perspective of computational efficiency, when $T > N$, we concentrate out F_t^k and solve (3.2) subject to the first normalization. The problem is identical to maximizing $tr(\Lambda^{k'}(X'X)\Lambda^k)$. The estimated loading matrix $\overline{\Lambda^k}$ is $\sqrt{N}$ times the eigenvectors corresponding to the k largest eigenvalues of $N \times N$ matrix $X'X$. The estimated factor equals to $\overline{F^k} = X\overline{\Lambda^k}/N$ under the normalization. When $T < N,$ we concentrate out Λ^k and solve (3.2) subject to the second normalization. The estimated factor $\widetilde{F^k}$ is $\sqrt{T}$ times the eigenvectors corresponding to the k largest eigenvalues of $T \times T$ matrix XX' and the corresponding estimated loading matrix is $\widetilde{\Lambda^k} = X'\widetilde{F^k}/T$.

3.3 Determining the Number of Factors

We suppose the DFM in (2.2) follows Assumption A-D, I and J as in Bai (2004), which are nested in our previous settings in the derivation with additional restrictions holding in general cases. By Theorem 5 in Bai (2004), the number of non-cointegrated I(1) factors *r* can be consistently estimated via their information criteria based on principle components estimators of data in levels. In addition, the differenced data satisfy all assumptions of Bai and Ng (2002), making it possible for us to consistently estimate the total number of non-stationary and stationary factors *(r+q)*. Note that with the stationary assumption, the idiosyncratic term is over-differenced as shown in (2.10) when using data in differences. However, over-differencing does not violate any conditions for the consistent estimator (Bai and Ng, 2002; Bai, 2004). Once consistently estimated, the number of factors can be treated as known in the estimation procedure of the factor space. For example, the estimated factor numbers can be used as the pre-specified k in the principal components estimators.

The family of information criteria developed by Bai and Ng (2002) for stationary data and Bai (2004) for nonstationary data in levels, trade off the benefit of including an additional factor against the cost of increased sampling variability arising from estimating another parameter (Stock and Watson, 2010). The consistent estimator is obtained by minimizing a penalized sum of squares, with k less than some selected threshold $kmax$, in the form of:

$$PC(k) = S\left(k, \widehat{F^k}, \widehat{\Lambda^k}\right) + kg(N,T),$$

where $S\left(k, \widehat{F^k}, \widehat{\Lambda^k}\right)$ is the least squares objective function in(3.2) evaluated at the principle components estimators $\left(\widehat{F^k}, \widehat{\Lambda^k}\right)$. The sum of squared residuals does not depends on which estimate of (F^k, Λ^k) is used. That is, the two estimators $\left(\widetilde{F^k}, \widetilde{\Lambda^k}\right)$ and $\left(\overline{F^k}, \overline{\Lambda^k}\right)$ are not differentiated. The term $g(N,T)$ is a penalty function satisfying certain converging conditions for stationary and nonstationary data respectively. The value of penalty term in the panel criterion increases linearly in the number of factors.

To be specific, from Bai and Ng (2002), the penalty function for stationary data satisfies the converging conditions that when $N, T \to \infty$,

(1) $g(N,T) \to 0$ and,

(2) $C_{NT}^2 \cdot g(N,T) \to \infty$, where $C_{NT}^2 = \min\{\sqrt{N}, \sqrt{T}\}$.

Based on these conditions, Bai and Ng (2002) proposed three criteria:

$$PC_{p1}(k) = S(k, \widehat{F^k}) + k\hat{\sigma}^2 \left(\frac{N+T}{NT}\right) \ln\left(\frac{NT}{N+T}\right);$$

$$PC_{p2}(k) = S(k, \widehat{F^k}) + k\hat{\sigma}^2 \left(\frac{N+T}{NT}\right) \ln C_{NT}^2 \,;$$

$$PC_{p3}(k) = S(k, \widehat{F^k}) + k\hat{\sigma}^2 \left(\frac{lnC_{NT}^2}{C_{NT}^2}\right);$$

where $S(k, \widehat{F^k})$ is the sum squares of the residuals as defined in (3.2) and $\hat{\sigma}^2 = S(kmax, \widehat{F^{kmax}})$, $C_{NT}^2 = \min\{\sqrt{N}, \sqrt{T}\}$. Apparently, these three criteria may be affected by the choice of the threshold *kmax*, which is undesirable in practice. Bai and Ng (2002) further proposed an improved version of criteria that would get rid of the threshold *kmax:*

$$IC_{p1}(k) = \ln(S(k, \widehat{F^k})) + k\left(\frac{N+T}{NT}\right) \ln\left(\frac{NT}{N+T}\right);$$

$$IC_{p2}(k) = \ln(S(k, \widehat{F^k})) + k\left(\frac{N+T}{NT}\right) \ln C_{NT}^2 \,;$$

$$IC_{p3}(k) = \ln\left(S(k, \widehat{F^k})\right) + k\left(\frac{lnC_{NT}^2}{C_{NT}^2}\right).$$

Although, the three criteria in each group are asymptotically equivalent, they could have different properties in finite samples. Bai and Ng (2002) recommended the first two criteria of each group, namely PC_{p1}, PC_{p2} and IC_{p1}, IC_{p2} according to their simulation results.

Bai (2004) extended the discussion to new criteria for nonstationary data. When consider the data in levels, the penalty functions should satisfy the converging condition that as $N, T \to \infty$,

(1) $g(N,T)\left(\frac{\ln\ln(T)}{T}\right) \to 0$ and,

(2) $g(N,T) \to \infty$.

Then the corresponding information criteria is given as

$$IPC_1(k) = S(k, \widehat{F^k}) + k\hat{\sigma}^2\alpha_T\left(\frac{N+T}{NT}\right)\ln\left(\frac{NT}{N+T}\right);$$

$$IPC_2(k) = S(k, \widehat{F^k}) + k\hat{\sigma}^2\alpha_T\left(\tfrac{N+T}{NT}\right)\ln C_{NT}^2;$$

$$IPC_3(k) = S(k, \widehat{F^k}) + k\hat{\sigma}^2\alpha_T\left(\frac{N+T-k}{NT}\right)\ln(NT);$$

where $S(k, \widehat{F^k})$ is the sum squares of the residuals as defined in (3.2) and

$\hat{\sigma}^2 = S(kmax, \widehat{F^{kmax}})$, $C_{NT}^2 = \min\{\sqrt{N}, \sqrt{T}\}$, $\alpha_T = \frac{T}{4lnln(T)}$.

In general, the information criteria estimators of Bai and Ng (2002) tend to overestimate the correct factor numbers unless the cross correlations are sufficiently weak. An alternative way to determine the number of factors in stationary data was proposed by Ahn and Horenstein (2013) by using the maximizers of the ratio of two adjacent eigenvalues:

$$ER(k) = \frac{S(k-1, \widehat{F^{k-1}}, \widehat{\Lambda^{k-1}}) - S(k, \widehat{F^k}, \widehat{\Lambda^k})}{S(k, \widehat{F^k}, \widehat{\Lambda^k}) - S(k+1, \widehat{F^{k+1}}, \widehat{\Lambda^{k+1}})},$$

where $S(k, \widehat{F^k}, \widehat{\Lambda^k})$ is the least squares objective function in(3.2) evaluated at the principle components estimators $(\widehat{F^k}, \widehat{\Lambda^k})$. The ratio can be viewed as a formalization of Cattell's scree test. Intuitively this corresponds to finding the edge of the cliff in the screen plot (Stock and Watson, 2010). Under the same assumptions of Bai and Ng (2002), Ahn and Horenstein (2013) proved that their estimator also produced consistent estimators of the number of factors in DFM. Ahn and Horenstein's estimator only uses the eigenvalues of sample covariance matrices of response variables. This new approach does not require pre-specified penalty functions or estimated threshold values, which are somewhat arbitrary choices in Bai and Ng (2002)'s information criteria. Their Monte Carlo experiment results suggested that the new estimators outperform the Bai and Ng (2002) estimators even in samples with small N and T.

When the number of factors is known, Result A1 in Bai and Ng (2008) shows that the factor space can be consistently estimated by the principle components estimator without knowing if the data are stationary or not. The principal components factor estimates are same to the true factor multiplied by a full rank matrix H, whose rank is equal to the known factor number. In addition, the corresponding estimated loading matrix is equal to the true loading matrix times the inverse of the transposed same matrix H. That is, while the factor and loading matrix cannot be separately identified, they are estimated up to a transformation, namely FH and $\Lambda(H')^{-1}$. Hence the common components $F\Lambda'$ are directly identifiable. For many empirical analysis purposes, knowing FH is as good as knowing F (Bai, 2004).

In the FECM application, the factors are specified in the DFM of (2.3) under the assumption of I(0) idiosyncratic term. Therefore, the space spanned by the factors can be consistently estimated using principle components approach. The number of factors r and q is either known or consistently estimated by the previously mention procedures. To be specific, the number of nonstationary factors r is estimated by the ICP information criteria proposed by Bai (2004) on level data. The total number of both nonstationary and stationary factors $r+q$ is estimated by using either Bai and Ng (2002) or Ahn and Horenstein (2013) approach on differenced data. The number of stationary factors q is then easily obtained by subtracting r from the total number $r+q$. Subsequently using the principal components estimator, the nonstationary factor estimate $\widetilde{F_t}$ is the r eigenvectors of XX' corresponding to the first r largest eigenvalues subject to normalization $\widetilde{F_t}'\widetilde{F_t}/T^2 = I$, while the stationary factor estimate $\widetilde{G_t}$ is the q eigenvectors corresponding to the next q largest eigenvalues normalized with $\widetilde{G_t}'\widetilde{G_t}/T = I$ (Bai, 2004).

3.4 Breaks and Structural Instability

The early research about the large dynamic factor models is nearly all about the model with stable parameters and factor structures. In particular, the factor loadings, number of factors and the factors space are time-invariant and do not allow to change. However, in practice, when long period of time series are used in the large factor models, there is broad evidence of inevitable instability.

Stock and Watson (2002) considered the case of temporal instability in factor loadings. They showed that given the number of factors, when the shifts of the factor loadings from time to time are small and idiosyncratic, the standard principal components estimation of factors is still consistent and robust. Bates, Plagborg-Møller, Stock and Watson (2013) further provided a sufficient condition for consistent estimation of the factor space when the factor loadings have certain types of instability. Their conclusion allows for larger instabilities in factor loadings than preceding theoretical calculations to hold the consistency of principal components estimator. The basic assumptions for the factor, initial factor loadings and idiosyncratic errors in their study follow the ones in Bai and Ng (2002), while the factor loading innovations are assumed varying with the magnitude converges to zero asymptotically. Their inference theory can be in parallel derived for nonstationary factor circumstances if we substitute the assumptions in Bai (2004) for the current one, with slightly modification on proofs. The simulation and empirical application showed that instability in the factor loadings has a limited impact on estimation of the factor space. However, estimation of the number of factors based on the Bai and Ng (2002) information criteria is more substantially affected.

In recent years, there is an emerging interest of big breaks in factor model analysis, partly because of the unexpected dramatic changes in the economy environment caused by the Subprime mortgage crisis and Great Recession in 2007 – 2009. Unlike other post-war U.S. recessions, the recent Great Recession is characterized by a sever disruption of financial markets, a slow recovery, and a lasting episode of zero nominal interest rates and unconventional monetary policies (Cheng, Liao and Schorfheide, 2014).

Stock and Watson (2009) considered the case of a single large break in factor structure with the empirical analysis of U.S. macroeconomic data consider the Great Moderation starting from 1984. They evaluated two sets of factors and loadings with sub-sample from before and after the break. They showed that the full-sample principal components estimator of the factor asymptotically spans the space of the two combined factors, while the number of full-sample factors can exceed the number of subsample factor in both sets. They found that forecasts based on full-sample factors can outperform those based on subsamples. This is in contrast to Banerjee, Marcellino and Masten (2007)'s Monte Carlo results. The latter, indicated that when the instability of factor loadings is big, the factor-based forecasting become worse significantly, even though the factor space itself is invariant. Breitung and Eickmeier (2011) also mentioned that the presence of structural breaks in the factor loadings might lead to overestimation of the number of factors. This is because a factor model with unstable factor loading can be represented by an equivalent model with extra pseudo factors but stable factor loadings. This underlies the failure of consistent estimation of factor numbers on full sample for any existing method, including Bai and Ng (2002), Onatski (2008, 2009), Ahn and Horestein (2013) etc. (Baltagi, Kao and Wang, 2015). One practical solution is to split the full-sample set into sub-sample sets by breaks, and estimate the factor numbers on each subset, if the change point is known. Breitung and Eickmeier (2009), Chen, Dolado and Gonzalo (2014), Han and Inoue (2014) and Corradi and Swanson (2014) proposed a series of tests to detect the existence of large breaks in factor loadings. . Baltagi, Kao and Wang (2015) developed a least squares estimator of the change point without requiring prior knowledge of the factor numbers and observability. Meanwhile, Cheng, Liao and Schorfheide (2014) proposed a shrinkage procedure that consistently estimates the number of pre- and post-break factors without requiring knowledge of the change point.

Chapter 4 FECM with Time Varying Coefficients

The analysis in the realms of economics and finance is time sensitive. There are both theoretical and empirical evidences suggesting the relation among variables may vary over time, especially during the periods of recession triggered by financial crisis, market shock or the bursting of an economic bubble. In the empirical application of this thesis, we use a U.S. macroeconomic variables data set covering the period spanning from 1990:Q1 to 2014:Q4, that includes the extended economic boom in the United States during 1990s, the burst of the dot com bubble around 2000 and the Great Recession following the subprime mortgage crisis in 2007-2009. A potential variation of variable relation is expected which we are highly interested in modeling and analyzing.

In the past decade, time varying coefficient (TVC) approach has been extensively applied to VAR and FAVAR model for analysis of macroeconomic issues, such as the mechanism of monetary policy (Primiceri, 2005; Koop, Leon-Gonzalez and Strachan, 2008; Korobilis, 2012), accurate index of financial conditions (Koop and Korobilis, 2014) and dynamic relation of variables (Nakajima, Kasuya, and Watanabe, 2011).There are relative few literature considering time varying coefficient model which permit cointegration. Li, Song and Witt (2005) used the TVC- ECM in forecasting of Tourism Demand with all parameters changing with time except that the cointergration coefficients are fixed. Bierens and Martins (2010) studied a time varying cointegration via Chebyshev Time Polynomials while all other parameters in the VECM were time-invariant. Koop, Leon-Gonzalez and Strachan (2011) provided a comprehensive framework of Bayesian inference on time varying cointegration in VECM, with all parameters to be time varying in state space form.

In this chapter, we extend the FECM in Chapter 2 by allowing time varying coefficient and stochastic volatility.

4.1 Model Description

Recall the FECM (2.9) derived in Chapter 2. Let the coefficients that determining the cointegration relations, controlling short term dynamics and the error covariance matrix be time varying. TVC-FECM is written as:

$$\begin{bmatrix}\Delta Y_t\\ \Delta F_t\\ \Delta G_t\end{bmatrix} = \alpha_t\beta_t' \begin{bmatrix}Y_{t-1}\\ F_{t-1}\\ G_{t-1}\end{bmatrix} + \Phi_{1t}\begin{bmatrix}\Delta Y_{t-1}\\ \Delta F_{t-1}\\ \Delta G_{t-1}\end{bmatrix} + \cdots + \Phi_{pt}\begin{bmatrix}\Delta Y_{t-p}\\ \Delta F_{t-p}\\ \Delta G_{t-p}\end{bmatrix} + \begin{bmatrix}v_t\\ Ru_t\\ Rw_t\end{bmatrix} \tag{4.1}$$

where the error terms are supposed to be independent $N(0,\Omega_t)$, α_t and β_t are full rank *n*–by–*k* matrices for every time point, where $n = M + r + q$ is the sum of dimensions of interested variables and factors, and *k* is their cointegration rank. We group the involved variables of interest and factors in a vector $y_t = \begin{bmatrix}Y_t\\ F_t\\ G_t\end{bmatrix}$ and idiosyncratic term in $\delta_t = \begin{bmatrix}v_t\\ Ru_t\\ Rw_t\end{bmatrix}$. Note that in Chapter 2, we conclude that the FECM with no restriction imposing on coefficients is a reasonable approximation. Hence we can treat the target variables and extracted factor in no difference when stating the TVC-FECM and considering its estimation. We rewritten (4.1) in the vector form as:

$$\Delta y_t = \alpha_t\beta_t' y_{t-1} + \Phi_{1t}\Delta y_{t-1} + \cdots + \Phi_{pt}\Delta y_{t-p} + \delta_t \tag{4.2}$$

Under the expression of (4.2), the TVC-FECM is in the same specification of TVP-VECM in Koop, Leon-Gonzalez and Strachan (2011), with the deterministic terms ignored.

As to the time varying structure of the coefficients, the state space form allowing the coefficients to evolve according to an AR(1) or random walk is popular in TVC time series model studies (Cogley and Sargent, 2005; Primiceri, 2005; Korobilis, 2009). There are established literature studying the inference and properties of the space state form modeling. Those techniques are readily to be used in the TVC-FECM. We separate the coefficients in (4.2) into three main blocks and present their time varying structures respectively.

4.2 Time Varying Cointegration

The most challenging part in the TVC-FECM is to properly modeling the changing pattern of the cointegration. Without further restriction, it is only the cointegrating space, instead of the particular cointegrating vectors, that is identified. This is consistent with the situation in standard time-invariant VECM, in which we can always find an orthogonal matrix Ω satisfying $\Omega\Omega' = I$ such that

$$\alpha\beta' = \alpha\Omega\Omega'\beta' = \alpha\Omega(\beta\Omega)' = \alpha_*\beta_*'$$

(Tsay, 2010). Moreover, the locations and dispersions of the cointegration vectors β_t do not always translate directly to comparable locations and dispersions on the cointegration space (Koop, Leon-Gonzalez and Strachan, 2011). Therefore, the state space form of the cointegration coefficients is defined based on the cointegration space instead of the cointegration vectors β_t. In fact, Koop, Leon-Gonzalez and Strachan (2010) argued that it is not sensible to express changes on cointegrating vectors in state space form directly. They illustrated this point by an analytical example which showed that small changes on vectors of β_t may cause that the space spanned alter dramatically. This is clearly undesirable in the model. Koop, Leon-Gonzalez and Strachan (2010) proposed some simple but reasonable principles about the design of cointegrating space structure: first, the cointegrating space has a distribution centered on the cointegrating space of the previous time point; second, the change in location of the spaces from time to time should be small, accommodating the gradual evolution in long term equilibrium; third, there should be ways to express prior beliefs, if any, about the marginal distribution of the cointegrating space at each time t. Following these principles, the state equation for the time varying cointegrating spaces is given by

$$b_t^* = \rho b_{t-1}^* + \eta_t,$$

$$\eta_t \sim N(0, I_{nk}) \text{ for } t = 2, \dots, T, \tag{4.3}$$

$$b_0^* \sim N\left(0, \frac{1}{1-\rho^2} I_{nk}\right).$$

where n is the dimension of y_t, k is the number of cointegration, $i.e.$ the rank of the cointegration matrix, $|\rho| < 1$, $b_t^* = vec(\beta_t^*)$ — the vector stack of the unrestricted matrix of cointegration vectors β_t^*. Let

$$\beta_t = \beta_t^* (\kappa_t)^{-1},$$

where

$$\kappa_t = (\beta_t^{*\prime} \beta_t^*)^{1/2},$$

β_t is semi-orthogonal, which satisfying the commonly used normalization $\beta_t'\beta_t = I$ in the VECM literatures.

From (4.3), we can derive that for each time point t, the marginal prior of b_t^* is multivariate normal distributed: $b_t^* \sim N(0, cI_{nk})$ with c is some constant. This implies a uniform distribution for β_t on the Stiefel manifold and a uniform distribution for the space it spans on the Grassmann manifold (Strachan and Inder, 2004). The previous literature stressed that this uniform distribution is a sensible and proper non-informative prior for the cointegrating space. Koop, Leon-Gonzalez and Strachan (2011) further investigated the distribution of cointegration spaces at time t condition on that of time $t-1$. Equations in (4.3) implies that b_t^* given b_{t-1}^* is multivariate normal distributed and hence, the conditional density of β_t^* given β_{t-1}^* is matrix normal with mean $\rho\beta_{t-1}^*$ and covariance matrix I_{nk}. From the results in Chikuse (2003), the conditional distribution of cointegration space is therefore an orthogonal projective Gaussian distribution with parameter $F_t = \beta_{t-1}\rho^2\kappa_{t-1}^2\beta_{t-1}'$. Koop, Leon-Gonzalez and Strachan (2011) showed that, following this distribution, the cointegration space of time t is centered on the cointegration space of time $t-1$, in both modal sense and expectation sense, which leads the satisfactory of their first desirable principal.

To ensure the second principal of “gradual” evolution in cointegration space, Koop, Leon-Gonzalez and Strachan (2011) emphasized the importance of the restriction $|\rho| < 1$, which is quite different from the random walk evolution commonly specified in the VAR model for the time varying coefficients. They found that when $\rho = 1$, the b_t^* could wander far from the origin, implying the violation of their second principle, means that the variation in cointegration spaces would shrink until it imposes the identical spaces from time to time at the limit. On the other hand, the value of ρ could not be too far away from one. Simulation results showed that even

$\rho = 0.99$ would allow for implausibly huge changes in the cointegration space. Therefore, Koop, Leon-Gonzalez and Strachan (2011) recommended choosing value of ρ in [0.999, 1).

One more thing to be noticed in the state equation (4.3) is that the variances for the error term η_t are set to be constant over time. This prior distribution is typically used in TVC-VAR models and imposes a same expected change in the coefficients in every time period. It also helps ensure the gradual evaluation desired in the second principal and the constant change which often occurs in practice.

We now draw the measurement model for (4.3) to complete the state space form expression. Additionally denote

$$\alpha_t^* = \alpha_t(\kappa_t)^{-1},$$

$$\tilde{x}_t = (\alpha_t^* \otimes y_{t-1}').$$

Then,

$$\alpha_t^* \beta_t^{*\prime} y_{t-1} = (\alpha_t^* \otimes y_{t-1}') b_t^* = \tilde{x}_t b_t^*.$$

We can write (4.2) in the following representation to obtain the measurement equation:

$$\widetilde{y_t} = \Delta y_t - [\Phi_{1t} \Delta y_{t-1} + \cdots + \Phi_{pt} \Delta y_{t-p}] = \alpha_t^* \beta_t^{*\prime} y_{t-1} + \xi_t \;=\; (\alpha_t^* \otimes y_{t-1}') b_t^* + \delta_t$$

$$\widetilde{y_t} = \tilde{x}_t b_t^* + \delta_t. \tag{4.4}$$

The normality assumption of δ_t gives us a linear normal form for the measurement equation. The measurement equation (4.4) paired with the state equation of b_t^* in (4.3) forms a standard state model which can be fitted in the regular state space model frameworks to solve, for example, the method of Durbin and Koopman (2002).

4.3 Other Time Varying Coefficients

The remaining short terms coefficients and error covariance matrix can be modeled in state space form in a similar way as that in the applications of TVC-VAR in Primiceri (2005) and TVC-FAVAR in Korobilis (2009). Again, we follow the notations in Koop, Leon-Gonzalez and Strachan (2011), define

$$\Xi_t = \left(\alpha_t^*, \Phi_{1t}, \dots, \Phi_{pt}\right),$$

$$a_t = vec(\Xi).$$

We assume the state equation of a_t is a random walk:

$$a_t = a_{t-1} + \zeta_t, \tag{4.5}$$

where $\zeta_t \sim N(0, Q)$. In addition, let

$$z_t = \beta_t^{*\prime} y_{t-1}$$

$$Z_t = \left(z_t', \Delta y_{t-1}', \dots, \Delta y_{t-p}'\right)'.$$

and

$$x_t = Z_t' \otimes I_n.$$

Therefore, we can write the measurement model for a_t from (4.2) in its vector and matrix form as follows:

$$\Delta y_t = \Xi Z_t + \xi_t = (Z_t' \otimes I_n) vec(\Xi) + \delta_t$$

$$\Delta y_t = x_t a_t + \delta_t. \tag{4.6}$$

Again, thanks to the normal distribution of δ_t, conditioning on β_t^*, (4.6) is a linear normal form measurement equation. (4.5) along with (4.6) is in the standard state space form which can be handled in a similar way as (4.3) and (4.4).

As to the covariance matrix of error $\delta_t \sim N(0, \Omega_t)$, we utilize the treatments in Primiceri (2005). This is also the method used in Korobilis (2009) and Koop, Leon-Gonzalez and Strachan (2011). To be precise, we first conduct a triangular reduction on the covariance matrix Ω_t:

$$A_t \Omega_t A_t' = \Sigma_t {\Sigma_t}',$$

where $A_t$is the lower triangular matrix

$$A_t = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ a_{21,t} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{n1,t} & \cdots & a_{nn-1,t} & 1 \end{bmatrix},$$

and $\Sigma_t = diag(\sigma_{1t}, \dots, \sigma_{nt})$ is the diagonal matrix. Then, there exists $\varepsilon_t \sim N(0, I)$ such that $\delta_t = A_t^{-1} \Sigma_t \varepsilon_t$. Let

$$\hat{a}_t = \left(a_{21,t}, a_{31,t}, a_{32,t}, \dots, a_{n1,t}, \dots, a_{nn-1,t}\right)',$$

the row stack of the non-zero off-diagonal elements in A_t, Primiceri(2005) imposed a random walk state equation of $\hat{a}_t$:

$$\hat{a}_t = \hat{a}_{t-1} + \xi_t, \tag{4.7}$$

where $\xi_t \sim N(0, C)$ is independent over t and other previously defined error terms. Denote $n \times \frac{n(n-1)}{2}$ matrix

$$\widehat{Z}_t = \begin{bmatrix} 0 & 0 & \cdots\cdots\cdots & 0 \\ -\hat{y}_{1,t} & 0 & \cdots\cdots\cdots\cdots & 0 \\ 0 & -\hat{y}_{[1,2],t} & \ddots & \cdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots\cdots\cdots & 0 & -\hat{y}_{[1,\dots,n-1],t} \end{bmatrix},$$

where $\hat{y}_{[1,\dots,i],t} = [\hat{y}_{1,t}, \hat{y}_{2,t}, \dots, \hat{y}_{i,t}]$ and $\hat{y}_{i,t}$is the i-th element of vector

$$\hat{y}_t = \Delta y_t - (\alpha_t \beta_t' y_{t-1} + \Phi_{1t} \Delta y_{t-1} + \cdots + \Phi_{pt} \Delta y_{t-p}).$$

Then from (4.2), we have

$$A_t(\Delta y_t - \alpha_t \beta_t' y_{t-1} - \Phi_{1t} \Delta y_{t-1} - \cdots - \Phi_{pt} \Delta y_{t-p}) = A_t \hat{y}_t = \Sigma_t \varepsilon_t$$

The measurement equation for $\hat{a}_t$ is given when we apply that A_t is a lower triangular matrix with ones on the main diagonal:

$$\hat{y}_t = \widehat{Z}_t \hat{a}_t + \Sigma_t \varepsilon_t. \tag{4.8}$$

Hence (4.7) and (4.8) describe the dynamic of $\hat{a}_t$ in a standard state space model.

Last but not the least, we lay out the evolution of Σ_t. A Stochastic volatility framework is used on Σ_t. Let

$$\sigma_t = [\sigma_{1t}, \dots, \sigma_{nt}]',$$

$$h_{it} = ln(\sigma_{it}),$$

$$h_t = [h_{1t}, \dots, h_{nt}]'.$$

The state equation is given on $\mathrm{h_t}$,

$$h_t = h_{t-1} + \omega_t, \tag{4.9}$$

with $\omega_t \sim N(0, W)$ is independent over t and other previously defined error terms. Denote

$$y_t^* = A_t \hat{y}_t = \Sigma_t \varepsilon_t,$$

$$y_t^* = [y_{1t}^*, \dots, y_{nt}^*]',$$

$$y_t^{**} = [y_{1t}^{**}, \dots, y_{nt}^{**}]',$$

where $y_{it}^{**} = ln[(y_{it}^*)^2 + \bar{c}]$ with $\bar{c} = 0.001$ is the offset constant.

Then the measurement equation for h_t is

$$y_t^{**} = 2h_t + e_t, \tag{4.10}$$

where $e_t = [e_{1t}, \dots, e_{nt}]'$ with $e_{it} = ln(\epsilon_{it}^2)\ for\ i = 1, \dots, n$. (4.9) and (4.10) complete the state space form of h_t.

Note that different from the previous state space models, the system consisting of (4.9) and (4.10) has a linear but non-Gaussian state space form, because the innovations in the measurement equation (4.10) are distributed as $\ln(\chi^2(1))$. Kim, Shephard and Chib(1998) provided a practical way to approximate the $\ln(\chi^2(1))$ by a mixture of normal distributions. Hence, we could employ this approximation to transform the state space system in (4.9) and (4.10) into a Gaussian one. Since we have $\varepsilon_t \sim N(0, I)$, implying an identity variance covariance matrix of ε_t, the variance covariance matrix of e_t should also be diagonal. This allows one to use the same (independent) mixture of normal distribution approximation for any element of e_t (Primiceri, 2005). Kim, Shephard and Chib (1998) selected a mixture of seven normal distributions with component probabilities q_j, means $m_j - 1.2704$ and variance

$v_j^2, j = 1, \ldots, 7$. The constants $\{q_j, m_j, v_j^2\}$ listed in Table 4.1, from Kim, Shephard and Chib (1998), are chosen to match a number of moments of the $\ln(\chi^2(1))$.

ω	$q_j = \Pr(\omega = j)$	m_j	v_j^2
1	0.00730	-10.13	5.79596
2	0.10556	-3.97281	2.61369
3	0.00002	-8.56686	5.17950
4	0.04395	2.77786	0.16735
5	0.34001	0.61942	0.64009
6	0.24566	1.79518	0.34023
7	0.25750	-1.08819	1.26261

Table 4.1 Selection of the mixing distribution to be $\ln(\chi^2(1))$ (Source: Kim, Shephard and Chib, 1998.)

Let s_t denote the indicator variables selecting which normal approximation in Table 4.1 has been used for e_t at time point $t, t = 1, \ldots, T$. Conditional on $s^T = [s_1, \ldots, s_T]'$, other FECM coefficients and variance covariance matrices, the state space system of (4.9) and (4.10) would have an approximate linear and Gaussian state space form. At the same time, condition on h_t and y_t^{**} from (4.9) and (4.10), it is easy to draw the sample of new s^T.

4.4 Bayesian Inference

We estimate the TVC-FECM through Markov Chain Monte Carlo (MCMC) algorithm. In the previous two sections, we divided the time varying coefficients in three main blocks: the short term coefficients $(\alpha_t^*, \Phi_{1t}, \ldots, \Phi_{pt})$, the cointegration coefficients β_t^*, and the error covariance matrix Ω_t. Their time varying patterns are given in four state space models, specified in equations (4.3) – (4.10). We employ the Gibbs sampling approach to accomplish the model estimation.

We first list the prior distributions we use in the Gibbs samplers and their corresponding posterior distributions.

(1) The block for short term coefficients described in (4.5) and (4.6).

The initial value of the state a_t is given as $a_0 \sim N(0, 2V_{a_0})$, where

$V_{a_0} = \begin{bmatrix} (1-\rho^2)I_{nk} & 0 \\ 0 & I \end{bmatrix}, n = \dim(y_t)$ and k is the cointegration rank. This prior variance V_{a_0}together with the setting in (4.3) imposes a prior variance of $2k$ on each of the elements of the product $\alpha_0^* \beta_0^{*\prime}$. The posterior distribution of a_0 given data and other coefficients is therefore $a_0|data \sim N(\overline{\mu_{a_0}}, \ \overline{V_{a_0}})$, where $\overline{V_{a_0}} = \left(\sum_{i=1}^T x_t \Omega_t^{-1} x_t' + V_{a_0}\right)^{-1}$ and $\overline{\mu_{a_0}} = \ \overline{V_{a_0}} \cdot [\sum_{i=1}^T (\Delta y_t - x_t a_t) \Omega_t^{-1} x_t']$.

The error term ζ_t in (4.5) has prior normal distribution $\zeta_t \sim N(0, Q)$, where the prior of Q *is* Wishart $Q^{-1} \sim W(\underline{\nu}_Q, \underline{Q}^{-1})$, with $\underline{\nu}_Q = \dim(a_0) + 2$ and $\underline{Q} = 0.0001 \cdot I$. Here the constant 0.0001 is chosen following Primiceri (2005). It appears small but in reality allows for substantial evolution of the states. The posterior distribution of Q given data is therefore $Q^{-1}|data \sim W(\bar{\nu}_Q, \bar{Q}^{-1})$, where $\bar{\nu}_Q = T + \underline{\nu}_Q = T + \dim(a_0) + 2$ and $\bar{Q}^{-1} = \left[\underline{Q} + \sum_{t=1}^T (a_t - a_{t-1})(a_t - a_{t-1})'\right]^{-1}$.

(2) The block for cointegrating space coefficients in (4.3) and (4.4).

The initial value of the state in (4.3) is given by $b_0^* \sim N\left(0, \frac{1}{1-\rho^2} I_{nr}\right)$.

Error term η_t also has prior specified in (4.3) as $\eta_t \sim N(0, I_{nr})$.

The autoregression coefficient ρ can be set fixed in the prior or supposed uniformly distributed on [0.999, 1) as the simulation results in Koop, Leon-Gonzalez and Strachan (2011) suggested. If ρ is treated as unknown, we could add one more block into the MCMC algorithm to iteratively draw it. Because of the nonlinear way in which ρ involved in the initial condition for b_0^*, the posterior distribution of ρ is non-standard and we use a Metropolis–Hastings algorithm to draw it in iterations.

(3) The block of error covariance matrix parameters in (4.7) –(4.10).

Error term ξ_t in (4.7) has prior normal distribution $\zeta_t \sim N(0, C)$, where the prior of C *is* Wishart $C^{-1} \sim W(\underline{\nu}_C, \underline{C}^{-1})$ with $\underline{\nu}_C = \frac{n(n-1)}{2} + 2$ and $\underline{C} = 0.0001 \cdot I$. Here the constant 0.0001 is chosen due to the same reason as the short term coefficient block. The posterior distribution of C given data is therefore $C^{-1}|data \sim W(\bar{\nu}_C, \bar{C}^{-1})$, where $\bar{\nu}_C = T + \underline{\nu}_C = T + \frac{n(n-1)}{2} + 2$ and $\bar{C}^{-1} = \left[\underline{C} + \sum_{t=1}^{T}(\hat{a}_t - \hat{a}_{t-1})(\hat{a}_t - \hat{a}_{t-1})'\right]^{-1}$.

Error term ω_t in (4.9) has prior normal distribution $\omega_t \sim N(0, W)$, where the prior of W *is* Wishart $W^{-1} \sim W(\underline{\nu}_W, \underline{W}^{-1})$, with $\underline{\nu}_W = n + 2$ and $\underline{W} = 0.0001 \cdot I$, where constant 0.0001 is determined, again, to ensure properly variation of states from time to time. The posterior distribution of W given data is therefore $W^{-1}|data \sim W(\bar{\nu}_W, \bar{W}^{-1})$, where $\bar{\nu}_W = T + \underline{\nu}_W = T + n + 2$ and $\bar{W}^{-1} = \left[\underline{W} + \sum_{t=1}^{T}(h_t - h_{t-1})(h_t - h_{t-1})'\right]^{-1}$.

Denote $a^T = [a_0', a_1', \dots, a_T']'$, $b^{*T} = [b_0^{*\prime}, b_1^{*\prime}, \dots, b_T^{*\prime}]'$, $A^T = [A_1', \dots, A_T']'$, $\Sigma^T = [\Sigma_1', \dots, \Sigma_T']'$, $s^T = [s_1, \dots, s_T]'$, and $y^T = [y_1', \dots, y_T']'$.

To summarize, the Gibbs sampler takes the form:

(i) Initialize $a^T, b^{*T}, A^T, \Sigma^T, s^T$ and ρ, Q, C, W;

(ii) Draw sample a^T from $p(a^T|y^T, b^{*T}, A^T, \Sigma^T, s^T, \rho, Q, C, W)$;

(iii) Draw sample Q from $p(Q|y^T, a^T, b^{*T}, A^T, \Sigma^T, s^T, \rho, C, W)$;

(iv) Draw sample b^{*T} from $p(b^{*T}|y^T, a^T, A^T, \Sigma^T, s^T, \rho, Q, C, W)$;

(v) Draw sample ρ from $\mathrm{p}(\rho|\mathrm{y}^\mathrm{T}, \mathrm{a}^\mathrm{T}, \mathrm{b}^{*\mathrm{T}}, \mathrm{A}^\mathrm{T}, \Sigma^\mathrm{T}, \mathrm{s}^\mathrm{T}, \mathrm{Q}, \mathrm{C}, \mathrm{W})$;

(vi) Draw sample A^T from $p(A^T|y^T, a^T, b^{*T}, \Sigma^T, s^T, Q, C, W)$;

(vii) Draw sample C from $p(C|y^T, a^T, b^{*T}, A^T, \Sigma^T, s^T, Q, W)$;

(viii) Draw sample s^T from $p(s^T|y^T, a^T, b^{*T}, A^T, \Sigma^T, Q, C, W)$;

(ix) Draw sample Σ^T from $p(\Sigma^T|y^T, a^T, b^{*T}, A^T, s^T, Q, C, W)$;

(x) Draw sample W from $p(W|y^T, a^T, b^{*T}, A^T, s^T, Q, C)$.

Specially, in the steps draw $a^T, b^{*T}, A^T, \Sigma^T$, we implement the method of Durbin and Koopman (2002) for the state space model simulation smoothing. Consider the state space model, for $t = 1, \dots, T$, the measurement equation is

$$y_t = Z_t \alpha_t + \varepsilon_t,\ \ \varepsilon_t \sim N(0, H_t), \tag{4.11}$$

the state equation is

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t,\ \ \eta_t \sim N(0, Q_t). \tag{4.12}$$

where $\alpha_1 \sim N(a_1, P_1)$, $Z_t, T_t, R_t, H_t, Q_t, a_1$ and P_1 are assumed to be known. Denote $w = [\varepsilon_1', \eta_1', \dots, \varepsilon_T', \eta_T']'$. $w \sim N(0, \Omega)$, where $\Omega = diag(H_1, Q_1, \dots, H_T, Q_T)$. The Kalman filter is given by, for $t = 1, \dots, T$,

$$v_t = y_t - Z_t a_t, \qquad F_t = Z_t P_t Z_t' + H_t, \qquad K_t = T_t P_t Z_t' F_t^{-1},$$

$$L_t = T_t - K_t Z_t, \;\; a_{t+1} = T_t a_t + K_t v_t, \;\; P_{t+1} = T_t P_t L_t' + R_t Q_t R_t'. \tag{4.13}$$

The disturbance smoothers are

$$\widehat{w_t} = \begin{bmatrix} H_t F_t^{-1} & -H_t K_t' \\ 0 & Q_t R_t \end{bmatrix} \begin{pmatrix} v_t \\ r_t \end{pmatrix}, \text{ for } t = 1, \dots, T, \tag{4.14}$$

where r_t is evaluated by the backwards recursion

$$r_{t-1} = Z_t' F_t^{-1} v_t + L_t' r_t, \quad \text{for } t = T, T-1, \dots, 1, \text{with } r_T = 0. \tag{4.15}$$

Then the smoother for state vector is computed as

$$\hat{\alpha}_{t+1} = T_t \hat{\alpha}_t + R_t Q_t R_t' r_t, \text{ for } t = 1, \dots, T \tag{4.16}$$

with the initialization $\hat{\alpha}_1 = a_1 + P_1 r_0$.

Durbin and Koopman (2002) used the following algorithm to generate the draws of the state vector $\alpha = [\alpha_1', \dots, \alpha_T']'$ from the conditional density $p(\alpha|y)$:

Step 1. Draw a random vector $w^+$from density $p(w)$ and use it to generate α^+ and y^+ by means of recursion (4.11) and (4.12) with w replace by w^+, where the recursion is initialized by the draw $\alpha_1^+ \sim N(a_1, P_1)$.

Step 2. Compute $\hat{\alpha} = E(\alpha|y)$ by means of standard filtering and smoothing using (4.13) – (4.16).

Step 3. Take $\tilde{\alpha} = \hat{\alpha} - \hat{\alpha}^{+} + \alpha^{+}$.

The final result $\tilde{\alpha}$ is a simulation smoother of the state space vector α. Durbin and Koopman (2002) showed that this simulation smoothing method is simple and computational efficient when comparing to other standard methods. In the empirical application of this thesis, we use this simulation smoother in our MCMC algorithm.

Chapter 5 Application to U.S. Bank Stress Testing

In this chapter, we apply the newly developed TVP-FECM to the U.S. bank stress testing analysis. Typically, we are interested in the projection of the pre-provision net revenue (PPNR). We use our model to look for potential variation of the relation between PPNR and the macroeconomic variables in different historical periods.

5.1 Pre-Provision Net Revenue

In the Federal Reserve published CCAR reports, the key feature to make the "object or not object" decision about a BHC's capital plan is the Tier 1 Common Capital Ratio (T1CR). The calculation of T1CR requires the projection of a group bank balance sheet data which reflect the profit, risk and regulatory capital of banks. As shown in Figure 5.1, this framework begins with a projection of the Pre-Provision Net Revenue (PPNR). Hence, as a first step experiment, we are of specially interest to apply our approach to analyzing the PPNR.

Generally speaking, the PPNR consists of three part, Net Interest Income, Non-interest Income and Non-interest Expense. Covas, Rump and Zakrajšek (2013) separated the PPNR into six components: (1) NII = net interest income; (2) TI = trading income; (3) ONII = noninterest income, excluding trading income; (4) CE = compensation expense; (5) FA = fixed assets expense; and (6) ONIE = other noninterest expense. Duane, Schuermann, Reynolds and Wyman (2013) used regression models on 10 components related to PPNR. Their results showed

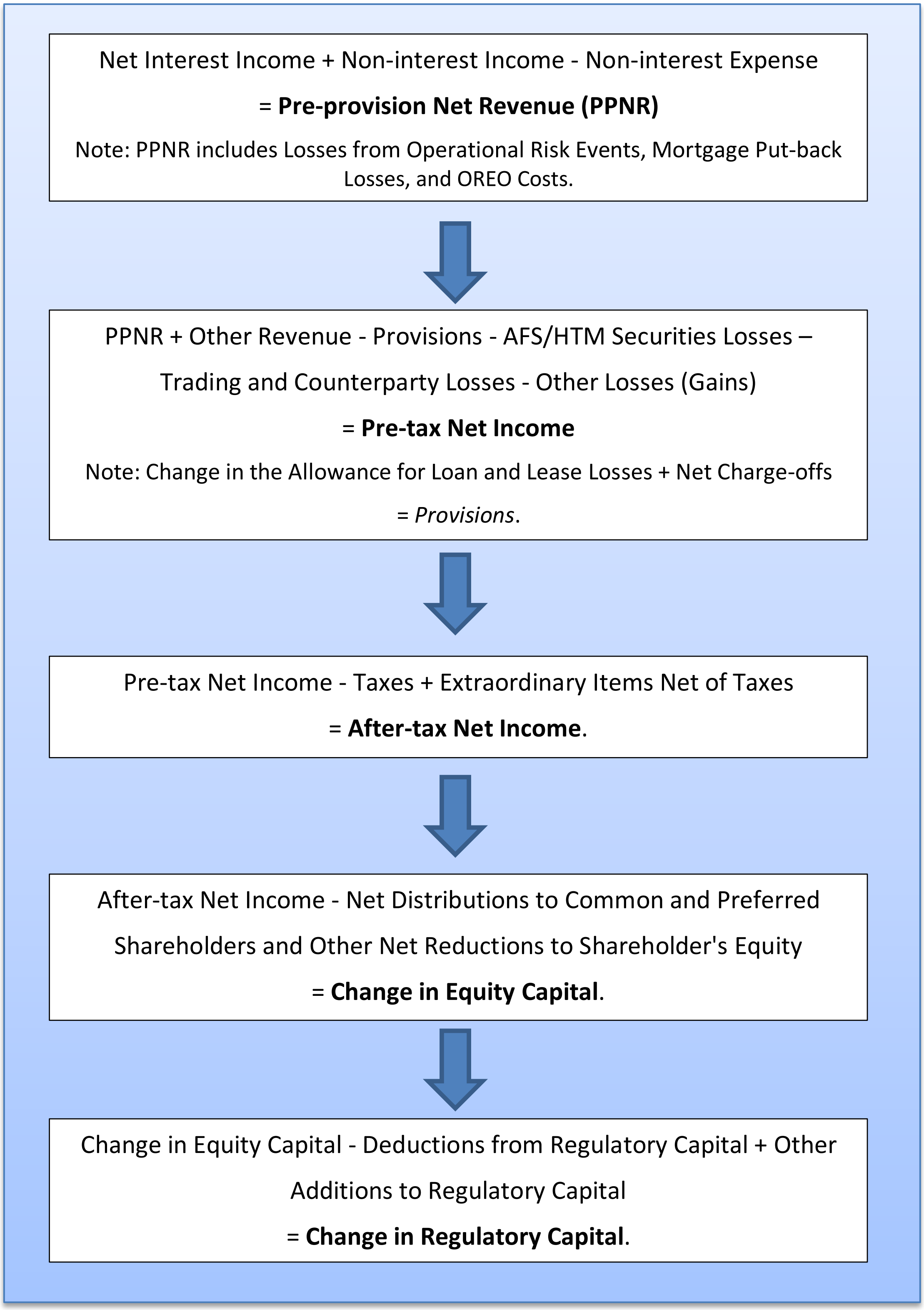

Figure 5.1 Projecting Net Income and Regulatory Capital (Source: Board of Governors of the Federal Reserve System, Comprehensive Capital Analysis and Review 2012: Methodology and Results for Stress Scenario Projections.)

that a single model attempting to directly project PPNR confounds and obscures the individual dynamics of each component and is therefore inferior in its prediction ability. This is consistent with the findings of Guerrieri and Welch (2012) that even the best performing model in their experiment could not beat a random walk at all horizons for Pre Provision Net Income modeling. Duane, Schuermann, Reynolds and Wyman (2013)'s empirical illustration suggested that the granularity and selection of PPNR components as response variables in the model substantially influence the projection results. Despite in the Federal Reserve's model description, the number of PPNR components used in separate models are as many as to 22 (eight components of interest income, five components of interest expense, five components of noninterest non-trading income, three components of noninterest expenses, and trading revenue), we adopt the division method similar to that in Covas, Rump and Zakrajšek (2013) due to the limitation of publicly available data.

5.2 Data

Our TVC-FECM is designed for large data environments. We use a dataset of 156 quarterly U.S. macroeconomic time series covering the period from the first quarter of 1990 (1990:Q1) to the fourth quarter of 2014 (2014:Q4) with 100 observations to extract factors. The variables are selected based on the dataset used in Bernanke, Boivin and Eliasz (2005), Stock and Watson (2009) and Korobilis (2009). All series are downloaded from St. Louis' FRED database except Commercial real estate price index (the Federal Reserve's Z1 releases), Total Gross Domestic Product for China (National Bureau of Statistics of China), Dow Jones U.S. Total Stock Market Index, NYSE Composite, Dow Jones Industrial Average (Yahoo Finance).

All series are seasonally adjusted either by FRED or by applying the X-13 Seasonal Adjustment Program developed by United States Census Bureau after seasonality test. Series recorded in monthly basis are used by quarterly average, except the stock index variables which are end of period closing. Table 5.1 is a summary of the selected macroeconomic variables and the full list can be found in the Appendix at the end of this chapter. When comparing with the 28 macroeconomic variables in the supervisory scenarios provided by the Federal Reserve, our data set includes all of these variables or their substitutions.

Category		Number of Macro Variables
1.	Real Activity Factor	33
2.	Unemployment and Employment	23
3.	Housing	11
4.	Price Index	17
5.	Interest Rate	12
6.	Stock Market	8
7.	International Factors	19
8.	Money Credit and Finance Factor	33
Total		156

Table 5.1 Summary of 156 Macroeconomic Variables.

As to the bank PPNR data, we use the FR Y-9C form, the Consolidated Financial Statements for Bank Holding Companies, published by the Federal Reserve to construct a balanced panel of the 4 largest U.S. BHCs, covering the period from 1990:Q1to 2014:Q4. The four selected BHCs, as listed in the top of Table 5.2, reported total consolidated assets over $1 trillion in quarterly average at the end of sample period. The total consolidated assets of these four largest BHCs count for 58.65% of total assets of the 31 bank holding companies participated in the CCAR 2015, and 45.26% of total assets of all U.S. BHCs that have reported in the FR Y-9C form. We consider the mergers of these four BHCs during the sample period. Mergers and acquisitions information is from the Federal Financial Institutions Examination Council's (FFIEC) web site, and listed in the table under Appendix of this chapter. Typically, for each of these four large banks, we aggregate all institutions acquired by which during our sampling period. Due to the limit of data availability, we only consider the entities that report the FR Y-9C

RSSD ID	Bank Holding Company	Ticker	Assets
1073757	Bank of America Corporation	BAC	2141.074
1951350	Citigroup Inc.	C	1898.981
1039502	JPMorgan Chase & Co.	JPM	2503.514
1120754	Wells Fargo & Co.	WFC	1656.513
31 CCAR 2015 Participating BHCs			13981.874
All U.S. BHCs reported in the FR Y-9C form			18119.503

Table 5.2 Assets of Bank Holding Companies in the U.S.: The RSSD ID is a unique identifier assigned to institutions by the Federal Reserve, used in the FR Y-9C form. Assets are referring to the quarterly average of total consolidated assets ($billions) reported at the end of 2014:Q4.

Components of PPNR	Regulatory Reports Mnemonics	Transformation
Net interest income	BHCK4074	$400 \times \frac{\text{Net interest income}}{\text{Consolidated assets}}$
Total noninterest income	BHCK4079	$400 \times \frac{\text{Total noninterest income}}{\text{Consolidated assets}}$
Compensation expense	BHCK4135	$400 \times \frac{\text{Compensation expense}}{\text{Consolidated assets}}$
Fixed assets expense	BHCK4217	$400 \times \frac{\text{Fixed assets expense}}{\text{Consolidated assets}}$
Total noninterest expense	BHCK4093	$400 \times \frac{\text{Total noninterest expense}}{\text{Consolidated assets}}$

Table 5.3 Components of Pre-Provision Net Revenue (PPNR): the regulatory reports mnemonics of the consolidated assets is BHCK3368.

form. Some of the famous mergers are omitted: for example, in 2008, Bank of America bought Merrill Lynch while JPMorgan Chase bought Bear Stearns and Washington Mutual – however, Merrill Lynch, Bear Stearns, and Washington Mutual did not file the FR Y-9C.

In terms of target variables, we consider the five components of PPNR as shown in Table 5.3. Comparing with the six components of PPNR in Covas, Rump and Zakrajšek (2013), we skip the trading income since the records of it in FR Y-9C can only date back to 1996:Q1.

5.3 Fitting TVC-FECM on PPNR Components

We start the model by estimating factors. ADF unit-root test shows that 148 out of 156 macroeconomic series are I(1), while the remaining 8 are considered to be stationary. The number of total nonstationary and stationary factors is determined to be 5 from estimation on differenced data. Information criterion for level nonstationary data in Bai (2004) identifies the number of I(1) factor to be 2 when the threshold of maximum number of factors is set to 5.

Rank	H_0:Unit Root	Test Statistics	Critical Value	p-Value
0	1	137.8119	40.1751	0.001
1	1	47.82	24.2747	0.001
2	0	4.8833	12.3206	0.6229
3	0	0.0079	4.1302	0.9423

Table 5.4 Johansen Trace Test of the First 4 Estimated Factors: assume no intercepts or trends in the cointegrating relations and no trends in the data.

We extract 5 factors from the level data via principle components estimation. The ADF unit-root test suggest the first 4 estimated factors are I(1) at 5% significance. Table 5.4 shows the result of Johansen trace test – cointegration rank is 2 among these first 4 factors, which suggests there are two non-cointegrated nonstationary trends in the extracted factors. Since the principle components estimator is consistently estimate the true factors up to a rotation, this finding is consistent with the result of factor number estimators.

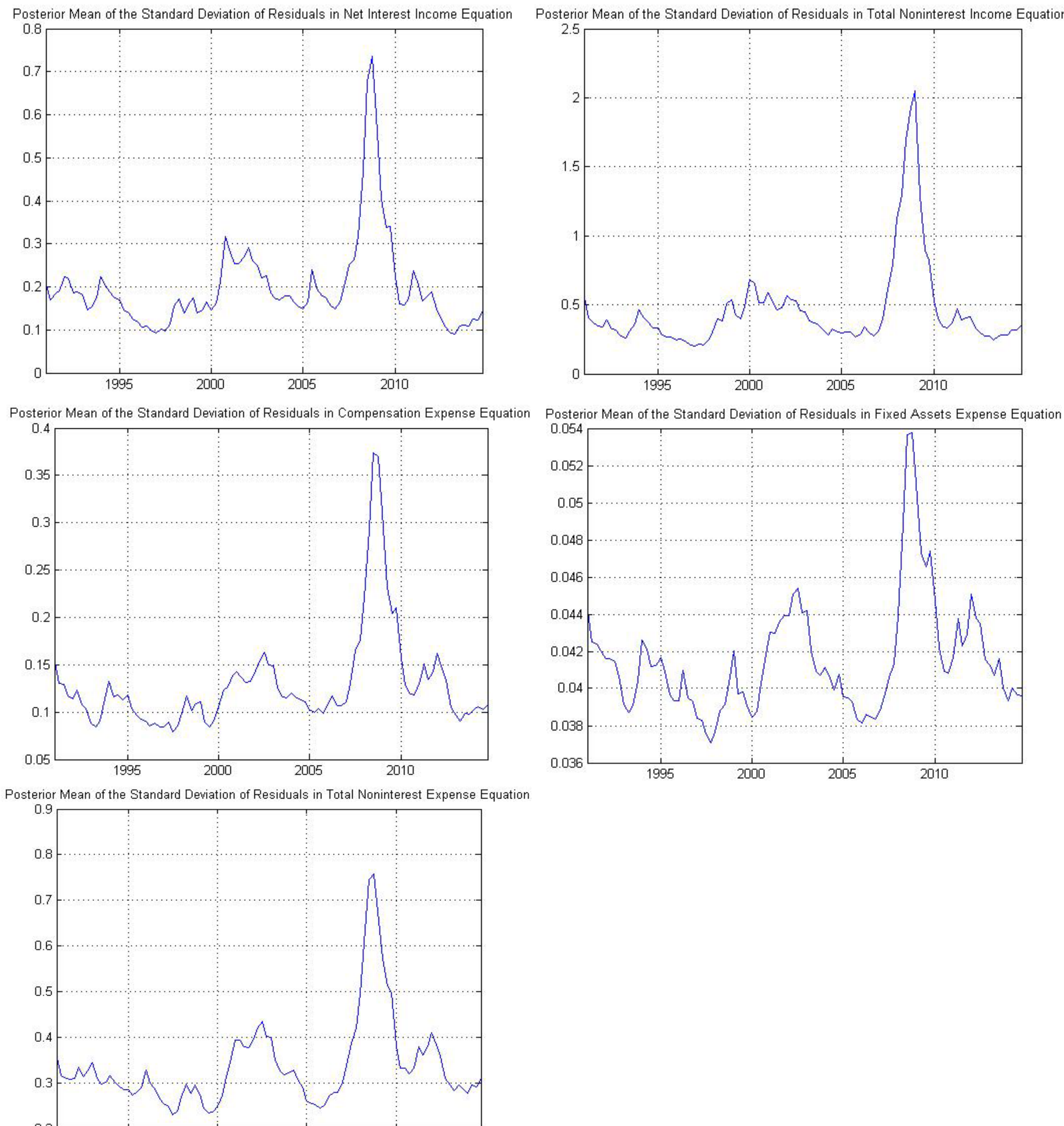

Figure 5.2 Time Varying Standard Deviations (std) of Errors in the Aggregated PPNR Components Equations: from left to right, top to bottom, the models are (1) Net interest income, (2) Total noninterest income, (3) Compensation expense, (4) Fixed assets expense, and (5) Total noninterest expense.

We aggregate the four selected banks data as a representative of the U.S. bank sector. Five PPNR components are fitted in bank industry level. Besides, we fit the (1) net interest income and (2) Total noninterest income of Citigroup Inc. and Wells Fargo & Co. to conduct a more institution-specific comparison. Comparing to Wells Fargo, Citigroup has a more diversified financial services to generate revenues, which may lead to the different reactions of the two banks interest income and noninterest income distribution. In addition, Citigroup has more geographically diversified business enabling which to raise cheap funds outside the U.S., and hence, it is expected to be less sensitive to U.S. economy changes. We fit the TVC-FECM using the same five estimated factors on all nine tested models. Each model is estimated via Gibbs sampling MCMC algorithm with 20,000 iterations after 10,000 burn-in.

Figure 5.2 presents the median posterior estimates of the standard deviation of the errors coming from the PPNR components equation in the TVP-FECM. At the aggregated bank industry-level, standard deviation of errors for all five PPNR components share a similar pattern: a clear spike is observed around 2008, reflecting the big variations in the 2007-2009 financial crisis and recession period; a moderate rise after 2000 could be explained by the market instability caused by the bursting of the dot com bubble.

Our main results are shown in the form of Impulse Response Functions (IRF). Impulse response functions quantify the reaction of a dynamic system in response to some external shocks which is known as impulse, at different horizon. Impulse response function is a powerful analytic tool in the empirical econometric application of macroeconomic models. We employ the IRF to illustrate the comparisons of relations between variables at different time. Condition on each time point, the TVC-FECM can be treated as a standard VECM. Hence we are able to transform it into a structural VAR form and obtain the IRF for desired horizons.

Figure 5.3 to Figure 5.11 show the impulse response for the target variables in the 9 models at time points of 1996:Q1, 2001:Q3, 2008:Q3, 2010:Q1. The four periods are chosen since they are in the four typical economy environment of U.S. in the past two decades. From boom to recession, the four time points represent (1) the 1990s boom (1996:Q1), (2) start to recover from Subprime mortgage crisis and Great Recession (2010:Q1), (3) the bursting of the dot com bubble and the end of 1990s boom (2001:Q3), and (4) Subprime mortgage crisis and Great Recession (2008:Q3).

The macroeconomic variables listed in Table 5.5 are the ones we investigate in the impulse response analysis. That is, when a standard shock occurs on these macroeconomic variables, we evaluate the response of interested PPNR components.

ID	Description
GDP	Real Gross Domestic Product, 3 Decimal
DPI	Real Disposable Personal Income
IP	Industrial Production: Total index
Unemployment Rate	Civilian Unemployment Rate
Home Price Index	S&P/Case-Shiller U.S. National Home Price Index
CPI	Consumer Price Index: Total All Items
3-Month	3-Month Treasury Bill: Secondary Market Rate
5-Year	5-Year Treasury Constant Maturity Rate
BAA	Moody's Seasoned Baa Corporate Bond Yield
Fed Fund	Effective Federal Funds Rate
VIX	CBOE Volatility Index: VIX©
Dow Jones	Dow Jones U.S. Total Stock Market Index
EX Rate-Yen	Japan / U.S. Foreign Exchange Rate
Real Estate	Real Estate Loans, All Commercial Banks
Bank Loans	Commercial and Industrial Loans, All Commercial Banks
Bank Credit	Bank Credit at All Commercial Banks
Monetary Base	Monetary Base; Total, Billions of Dollars
M1	M1 Money Stock

Table 5.5 Selected Macroeconomic Indicator tested in Impulse Response Function

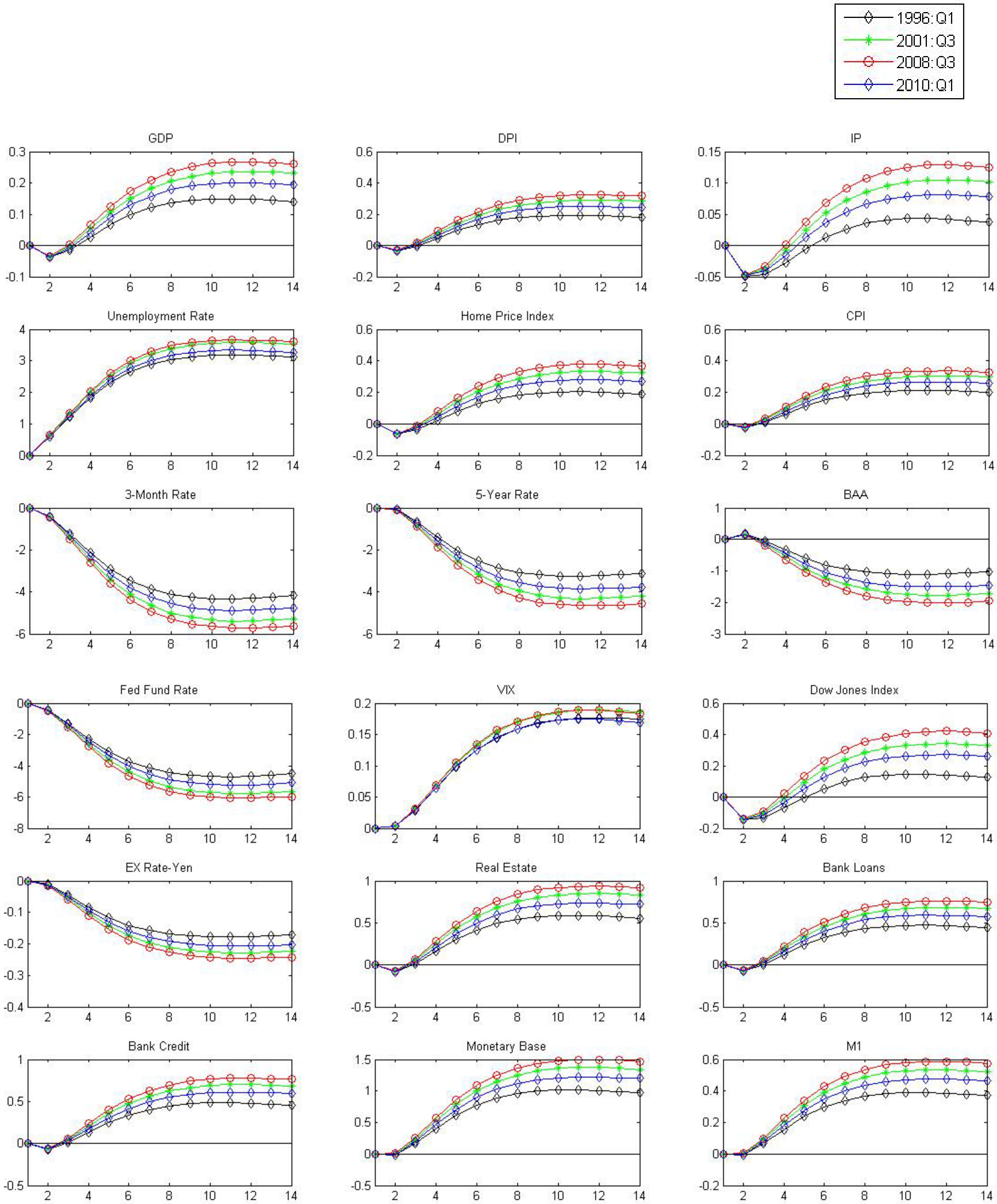

Figure 5.3 Impulse Responses for aggregated Net Interest Income of U.S Banks to the Selected Variables on the Periods 1996:Q1, 2001:Q3, 2008:Q3, 2010:Q1

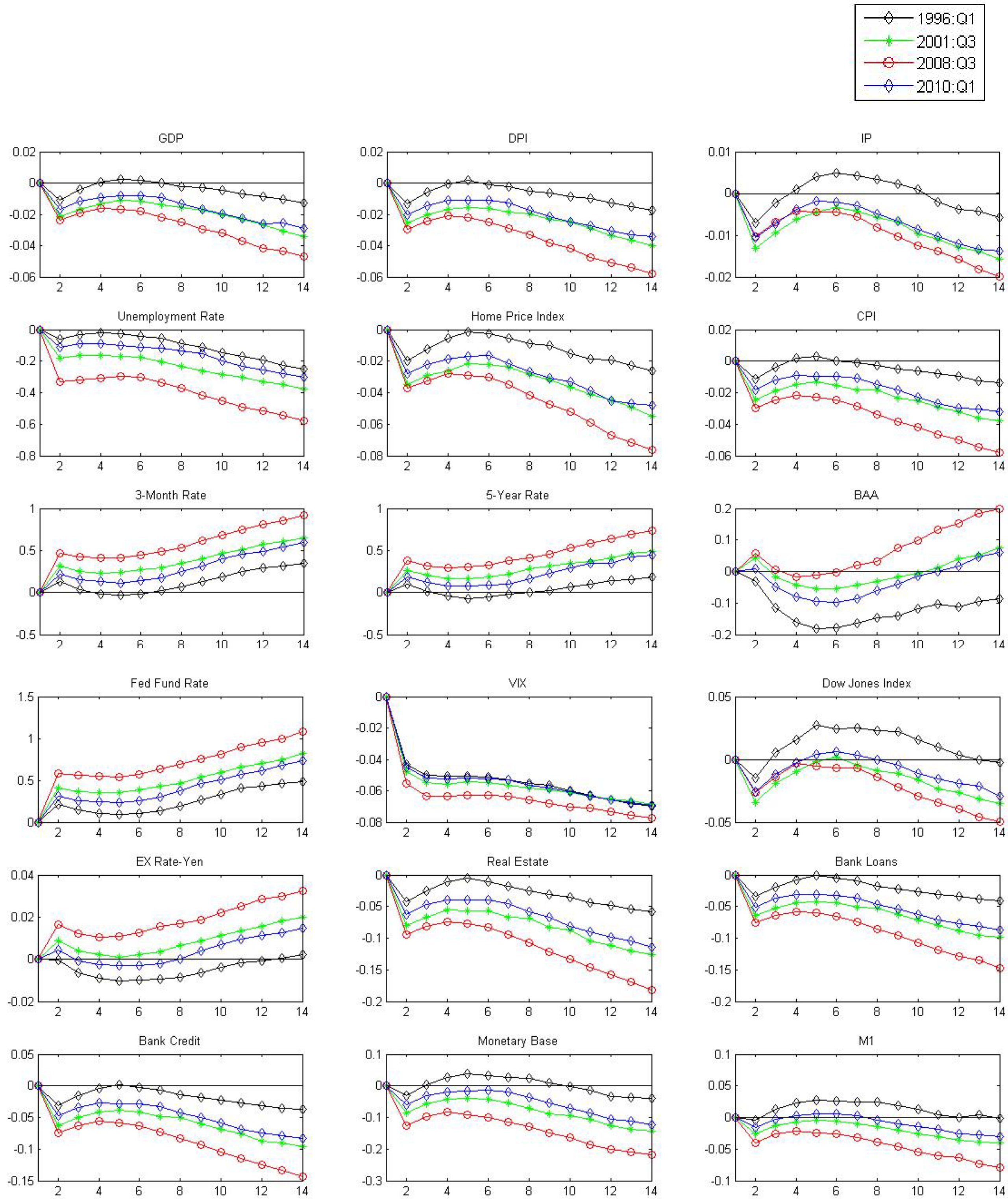

Figure 5.4 Impulse Responses for aggregated Total Noninterest Income of U.S Banks to the Selected Variables on the Periods 1996:Q1, 2001:Q3, 2008:Q3, 2010:Q1

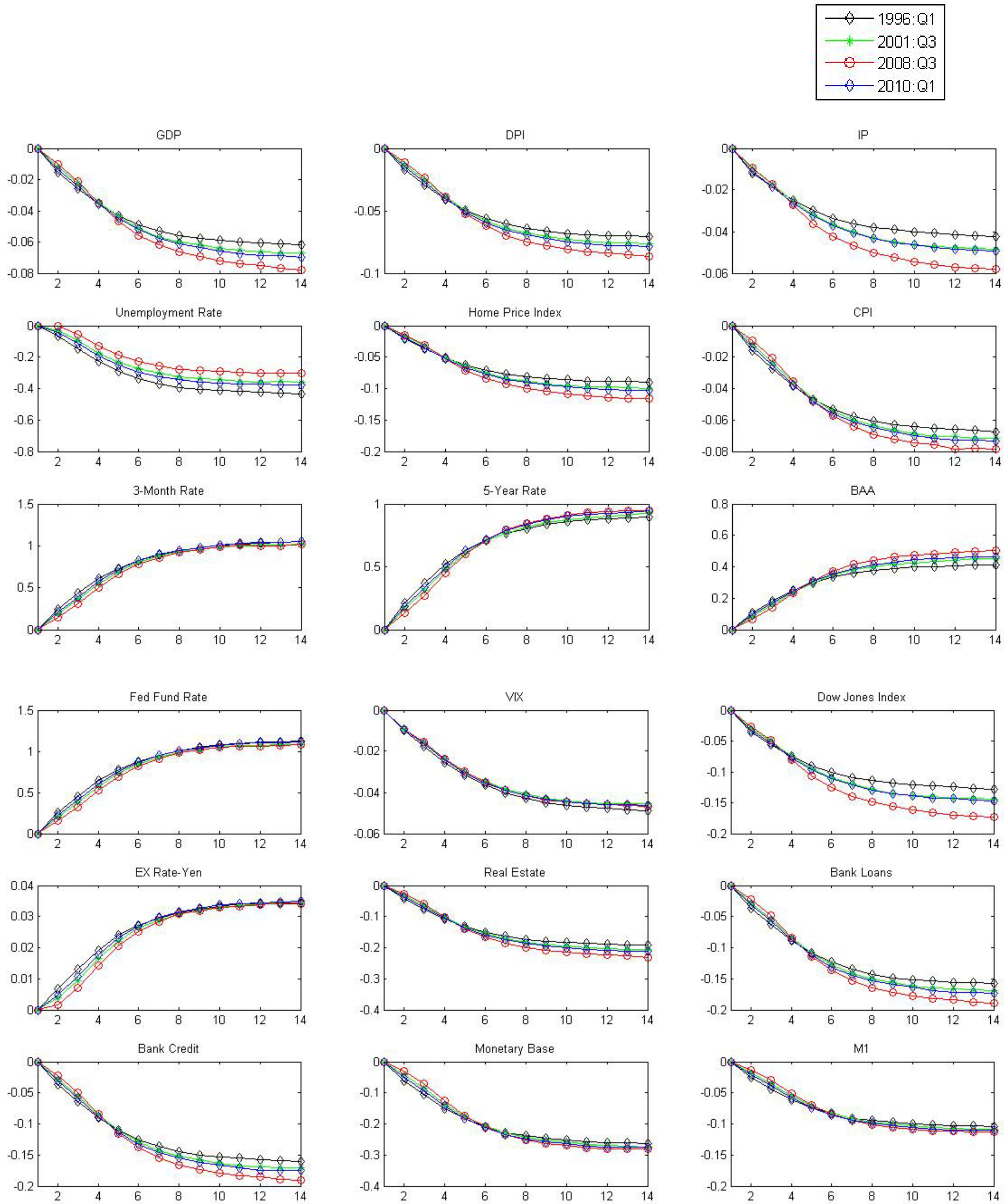

Figure 5.5 Impulse Responses for aggregated Compensation Expense of U.S Banks to the Selected Variables on the Periods 1996:Q1, 2001:Q3, 2008:Q3, 2010:Q1

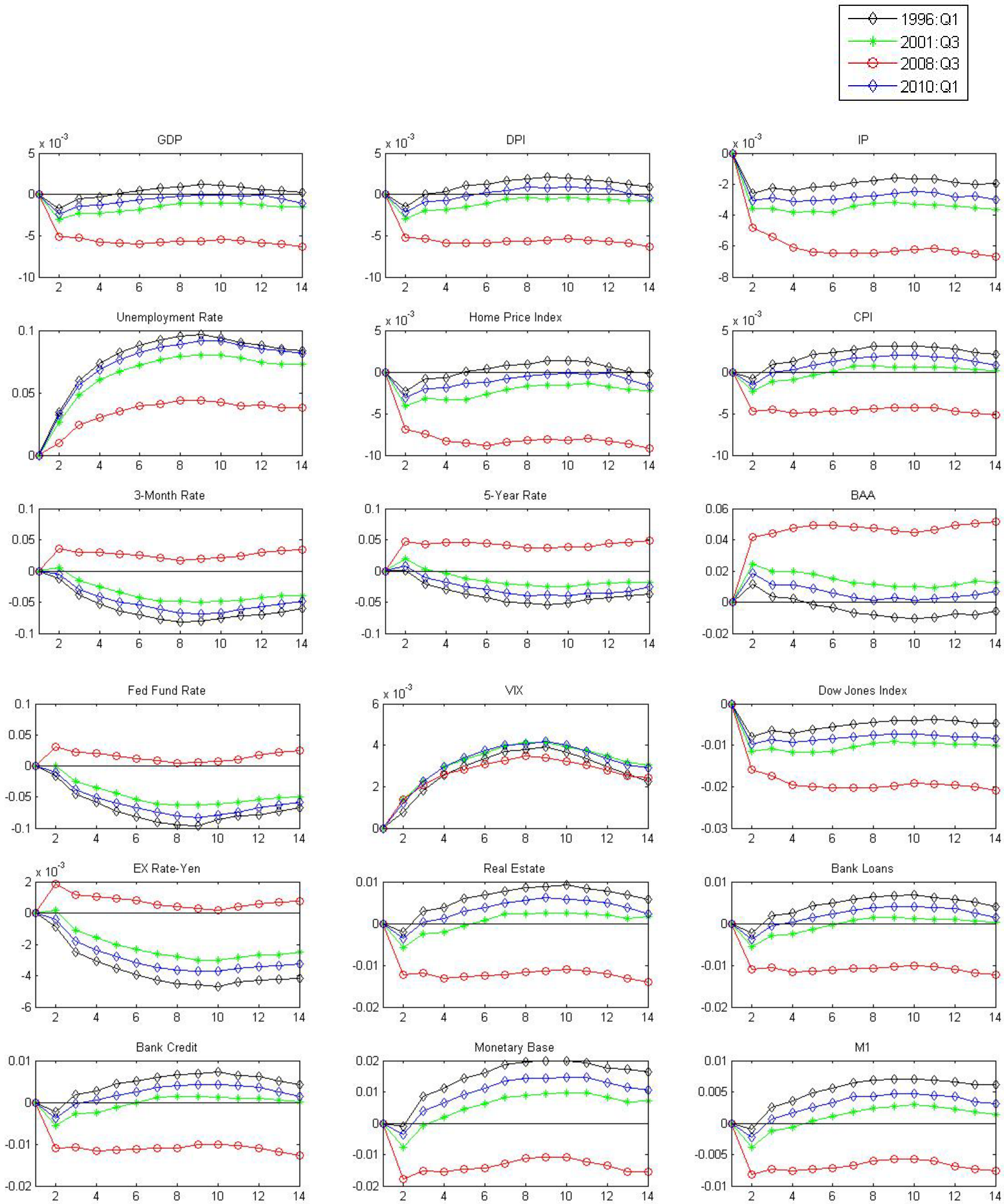

Figure 5.6 Impulse Responses for aggregated Fixed Assets Expense of U.S Banks to the Selected Variables on the Periods 1996:Q1, 2001:Q3, 2008:Q3, 2010:Q1

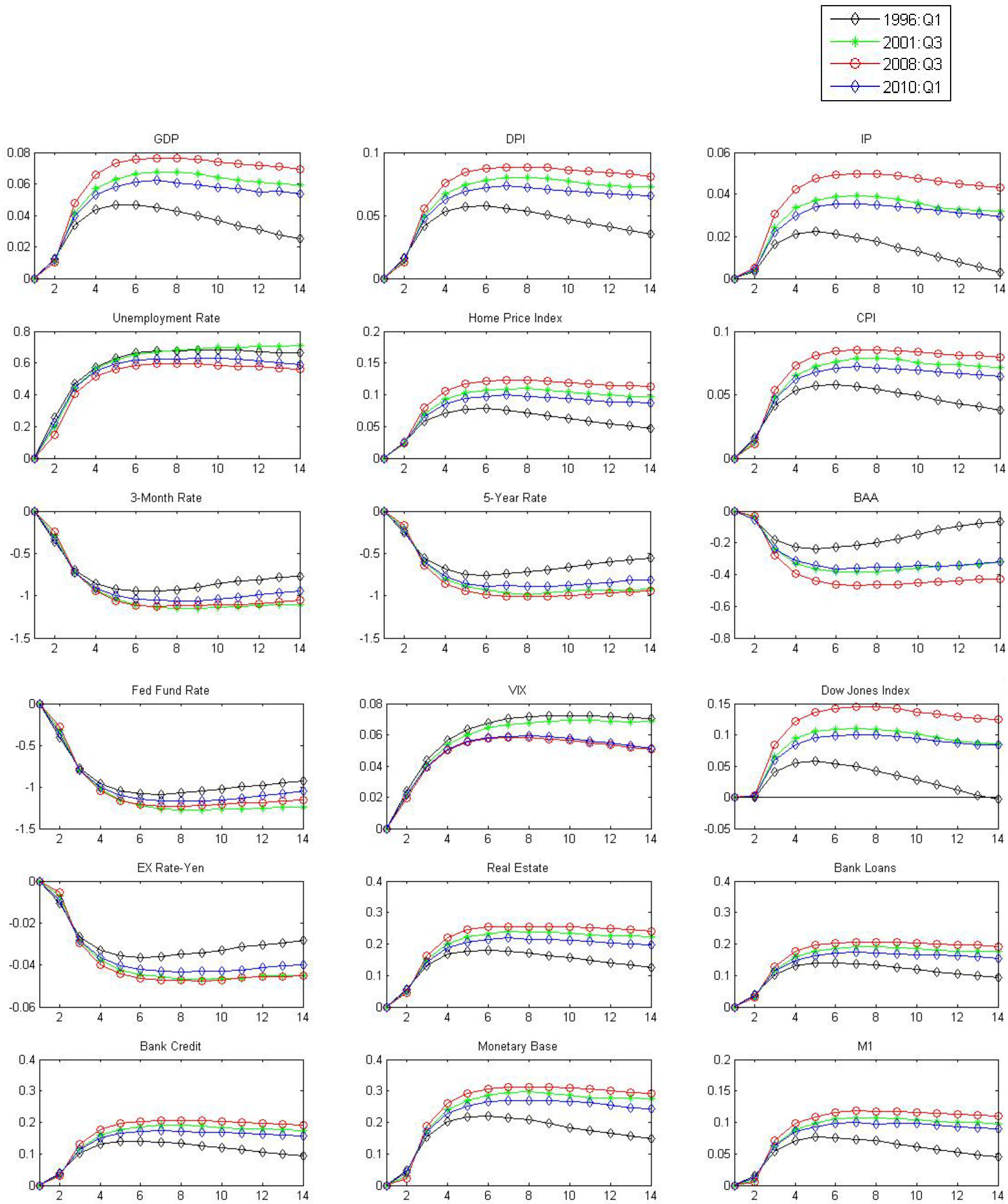

Figure 5.7 Impulse Responses for aggregated Total Noninterest Expense of U.S Banks to the Selected Variables on the Periods 1996:Q1, 2001:Q3, 2008:Q3, 2010:Q1

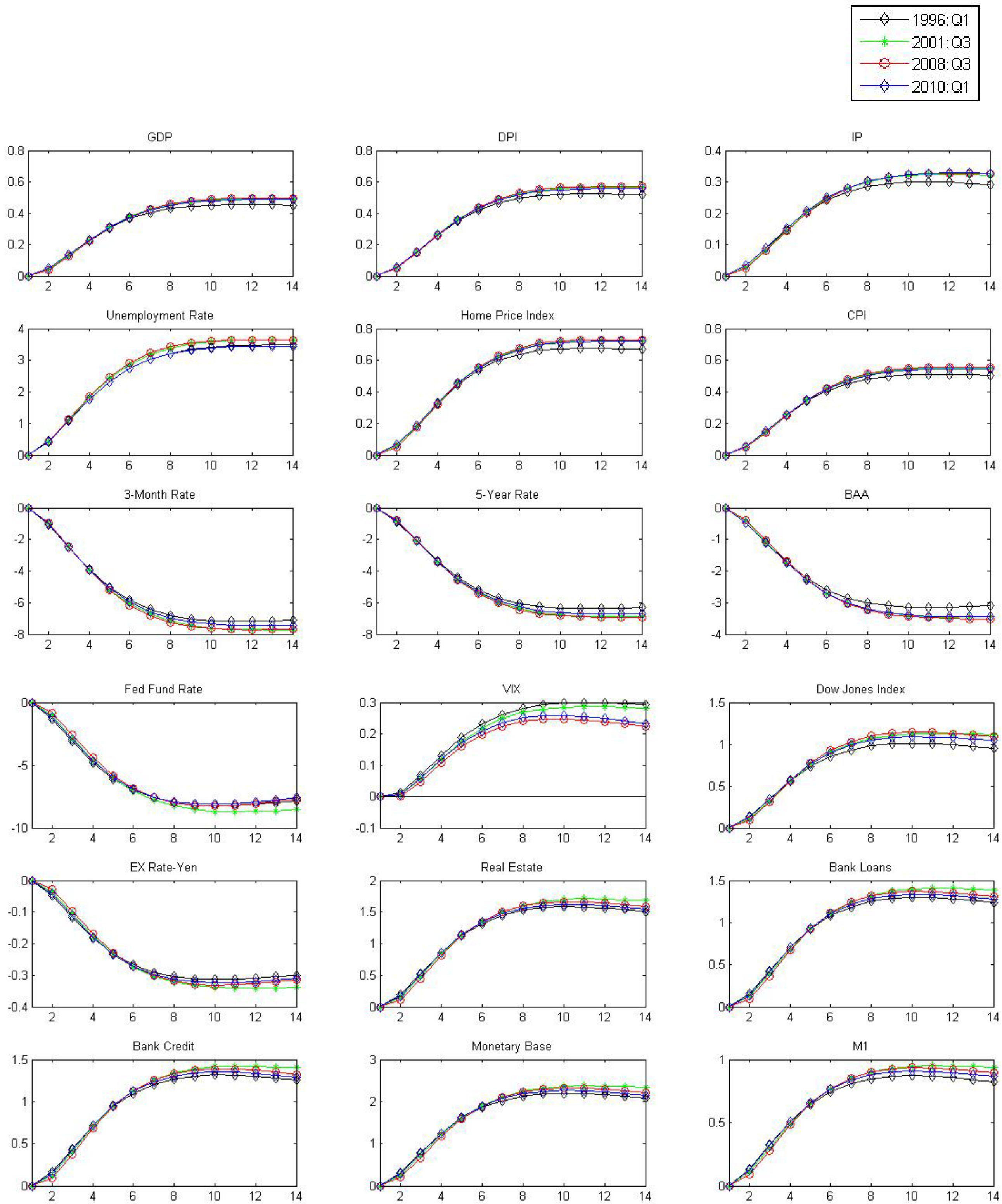

Figure 5.8 Impulse Responses for Net Interest Income of Citigroup to the Selected Variables on the Periods 1996:Q1, 2001:Q3, 2008:Q3, 2010:Q1

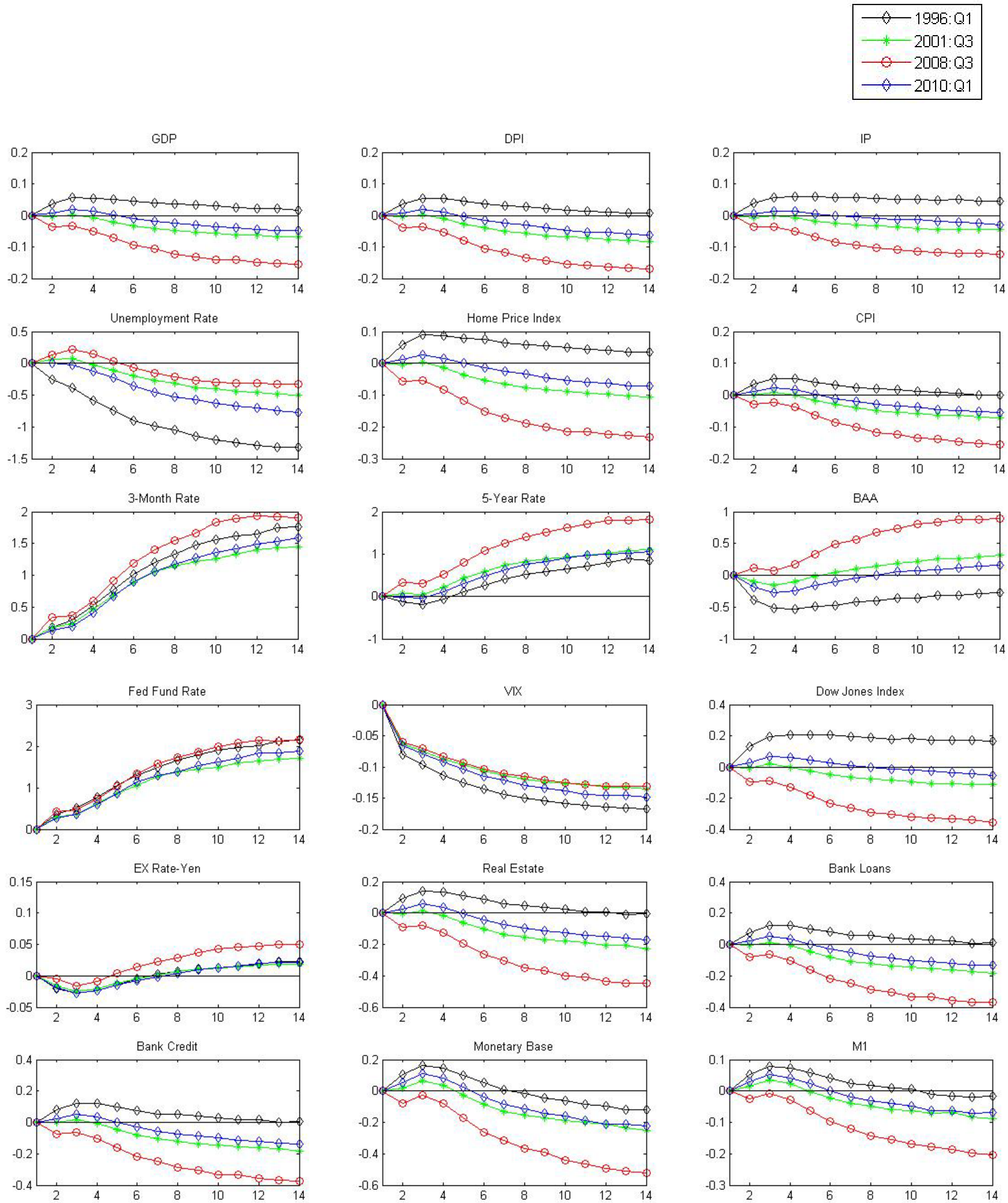

Figure 5.9 Impulse Responses for Total Noninterest Income of Citigroup to the Selected Variables on the Periods 1996:Q1, 2001:Q3, 2008:Q3, 2010:Q1

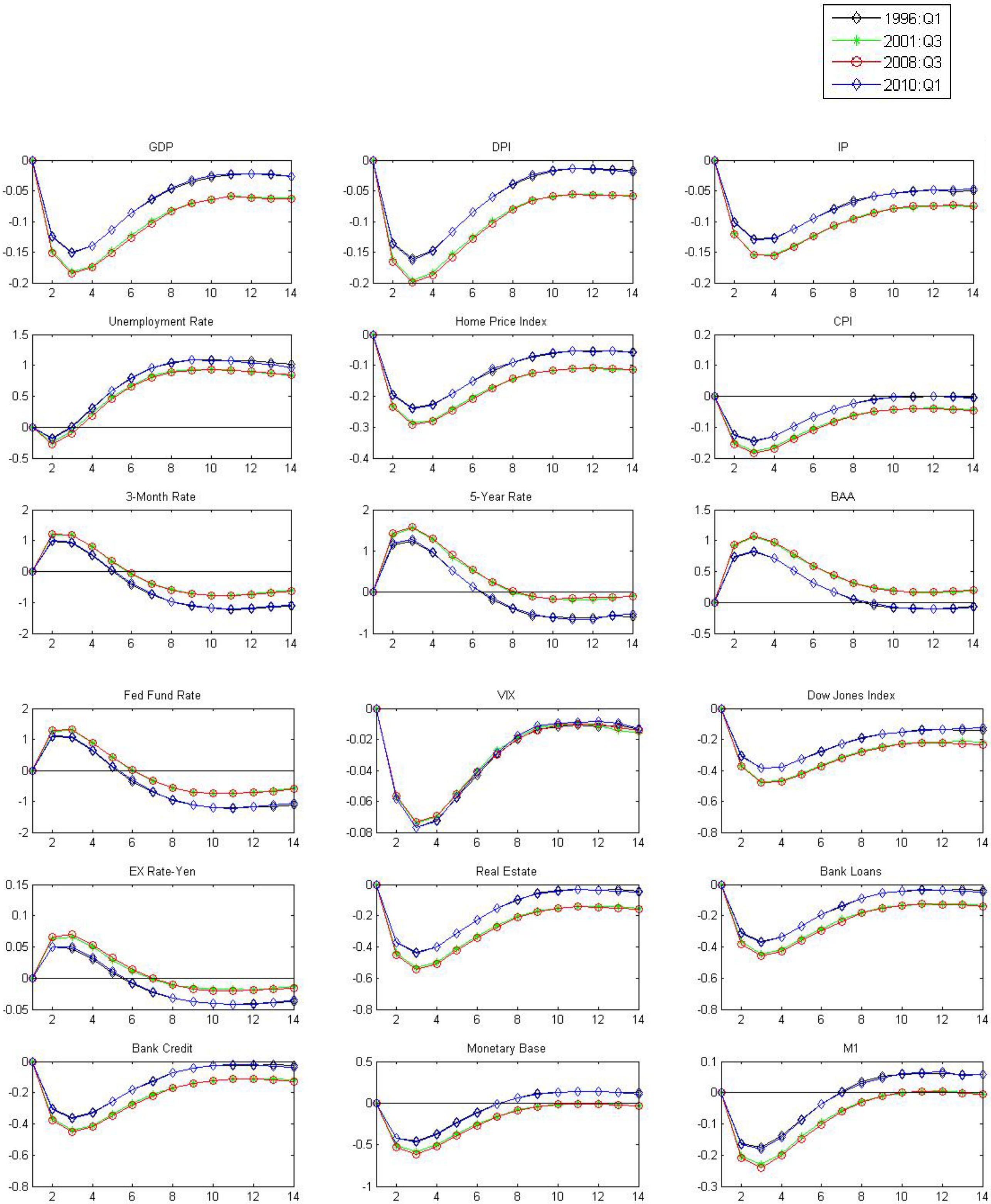

Figure 5.10 Impulse Responses for Net Interest Income of Wells Fargo to the Selected Variables on the Periods 1996:Q1, 2001:Q3, 2008:Q3, 2010:Q1

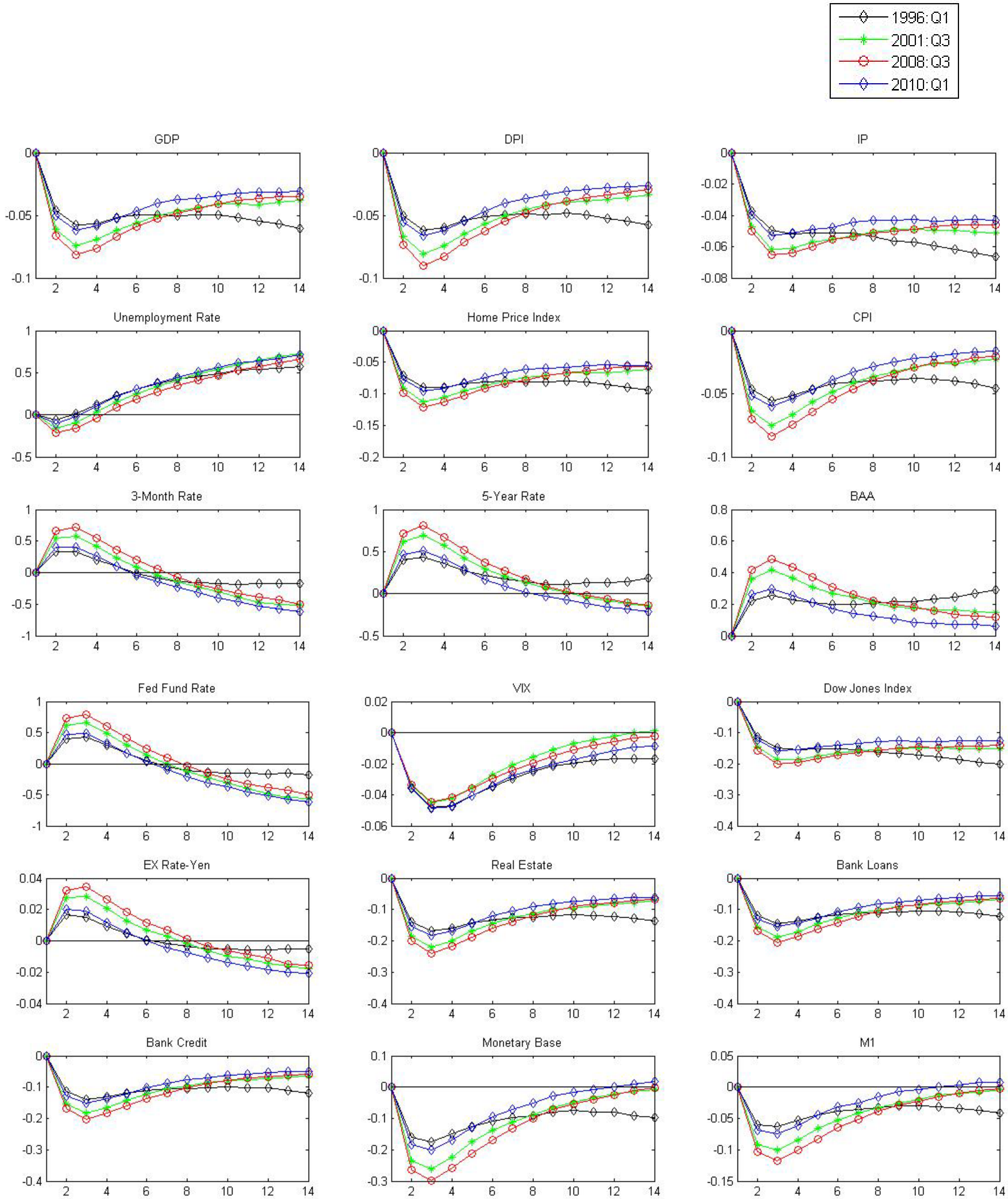

Figure 5.11 Impulse Responses for Total Noninterest Income of Wells Fargo to the Selected Variables on the Periods 1996:Q1, 2001:Q3, 2008:Q3, 2010:Q1

Overall, all plots have the four lines representing the four periods sorted in the order of economy conditions. Among the aggregated PPNR components, the compensation expense has the IRF more concentrated than others, implying that the changes in compensation expense responding to the macroeconomic variables shock are relatively invariant under different economic environments. In contrast, the fixed assets expense has the red lines, representing the period of 2008:Q3, stand apart from other time periods lines. Although the scale is relatively small, this is an interesting finding that the fixed assets expense would have very different, for some macroeconomic metrics even the inversed, reactions to system shocks under the stressed environment.

As to the scale of the responses, interest rates receive most significant response from all components considered, even in the noninterest categories. Net interest income does more sensitive to interest rates variables than other noninterest components. The Unemployment rate following the interest rates variables, has the second large influences on PPNR components.

Another surprising finding from the impulse response is that when comparing Figure 5.3, Figure 5.8 and Figure 5.10, the three net interest income groups, the response to interest rates shocks are negative in long term. While Wells Fargo does show a short increase in the net interest income when the interest rates were raised, the net interest income of the aggregated four banks and that of the Citigroup would go down when the rates were increased.

5.4 Appendix

Table 5.6 Macroeconomic Variables used in Factor Estimation: The series are left in untransformed if indicated by Tcode 1. Others with Tcode 2 are the logarithm.

Series ID	Description	Tcode
	1. Real Activity Factor	
GDPC96	Real Gross Domestic Product, 3 Decimal	2
DPIC96	Real Disposable Personal Income	2
CBI	Change in Private Inventories	1
FINSLC96	Real Final Sales of Domestic Product	2
GSAVE	Gross Saving	2
FPI	Fixed Private Investment	2
PRFI	Private Residential Fixed Investment	2
SLINV	State & Local Government Gross Investment	2
EXPGSC96	Real Exports of Goods & Services, 3 Decima	2
IMPGSC96	Real Imports of Goods & Services, 3 Decimal	2
CIVA	Corporate Inventory Valuation Adjustment	1
CP	Corporate Profits After Tax (without IVA and CCAdj)	2
CNCF	Corporate Net Cash Flow with IVA	2
DPCERA3Q086SBEA	Real personal consumption expenditures	2
DSERRA3Q086SBEA	Real personal consumption expenditures: Services	2
DDURRA3Q086SBEA	Real personal consumption expenditures: Durable goods	2
DNDGRA3Q086SBEA	Real personal consumption expenditures: Nondurable goods	2
IPB50001SQ	Industrial Production: Total index	2
IPGMFSQ	Industrial Production: Manufacturing (NAICS)	2
IPB51100SQ	Industrial Production: Durable consumer goods	2
IPB51200SQ	Industrial Production: Nondurable consumer goods	2
IPB54100SQ	Industrial Production: Construction supplies	2
NAPM	ISM Manufacturing: PMI Composite Index	2
NAPMNOI	ISM Manufacturing: New Orders Index	2
NAPMII	ISM Manufacturing: Inventories Index	2
NAPMSDI	ISM Manufacturing: Supplier Deliveries Index	2
HOABS	Business Sector: Hours of All Persons	2
RCPHBS	Business Sector: Real Compensation Per Hour	2
ULCBS	Business Sector: Unit Labor Cost	2
COMPNFB	Nonfarm Business Sector: Compensation Per Hour	2
ULCNFB	Nonfarm Business Sector: Unit Labor Cost	2
HOANBS	Nonfarm Business Sector: Hours of All Persons	2
COMPRNFB	Nonfarm Business Sector: Real Compensation Per Hour	2

2. Unemployment and Employment		
UNRATE	Civilian Unemployment Rate	1
UEMPLT5	Number of Civilians Unemployed for Less Than 5 Weeks	2
UEMP5TO14	Number of Civilians Unemployed for 5 to 14 Weeks	2
UEMP15OV	Number of Civilians Unemployed for 15 Weeks and Over	2
UEMP15T26	Number of Civilians Unemployed for 15 to 26 Weeks	2
UEMP27OV	Number of Civilians Unemployed for 27 Weeks and Over	2
NDMANEMP	All Employees: Nondurable goods	2
MANEMP	All Employees: Manufacturing	2
SRVPRD	All Employees: Service-Providing Industries	2
USTPU	All Employees: Trade, Transportation & Utilities	2
USWTRADE	All Employees: Wholesale Trade	2
USFIRE	All Employees: Financial Activities	2
USEHS	All Employees: Education & Health Services	2
USPBS	All Employees: Professional & Business Services	2
USINFO	All Employees: Information Services	2
USSERV	All Employees: Other Services	2
USGOVT	All Employees: Government	2
USLAH	All Employees: Leisure & Hospitality	2
CES2000000008	Average Hourly Earnings: Construction	2
CES3000000008	Average Hourly Earnings: Manufacturing	2
AHETPI	Average Hourly Earnings of: Total Private	2
AWOTMAN	Average Weekly Overtime Hours: Manufacturing	1
AWHMAN	Average Weekly Hours: Manufacturing	2
3. Housing		
USSTHPI	All-Transactions House Price Index for the United States	2
CSUSHPISA	S&P/Case-Shiller U.S. National Home Price Index©	2
CREPI	Commercial real estate price index	2
HOUST	Total: New Privately Owned Housing Units Started	2
HOUSTNE	Housing Starts in Northeast Census Region	2
HOUSTMW	Housing Starts in Midwest Census Region	2
HOUSTS	Housing Starts in South Census Region	2
HOUSTW	Housing Starts in West Census Region	2
HOUST1F	Privately Owned Housing Starts: 1-Unit Structures	2
PERMIT	New Private Housing Units Authorized by Building Permits	2
PERMIT1	New Private Housing Units Authorized by Building Permits: 1-Unit Structures	2
4. Price Index		
CPALCY01USQ661N	Consumer Price Index: Total All Items for the United States	2
CPIAUCSL	Consumer Price Index for All Urban Consumers: All Items	2

CPILEGSL	Consumer Price Index for All Urban Consumers: All Items Less Energy	2
CPIULFSL	Consumer Price Index for All Urban Consumers: All Items Less Food	2
CPIENGSL	Consumer Price Index for All Urban Consumers: Energy	2
CPIUFDSL	Consumer Price Index for All Urban Consumers: Food	2
PPIACO	Producer Price Index for All Commodities	2
PPICRM	Producer Price Index for Crude Materials	2
PPIFCF	Producer Price Index for Finished Consumer Foods	2
PPIFCG	Producer Price Index for Finished Consumer Goods	2
PFCGEF	Producer Price Index for Finished Consumer Goods Excluding Foods	2
PPIFGS	Producer Price Index by Commodity for Finished Goods	2
PPICPE	Producer Price Index for Finished Goods: Capital Equipment	2
PPIENG	Producer Price Index for Fuels & Related Products & Power	2
PPIIDC	Producer Price Index for Industrial Commodities	2
PPIITM	Producer Price Index for Intermediate Materials: Supplies & Components	2
MCOILWTICO	Crude Oil Prices: West Texas Intermediate (WTI)	2
5. Interest Rate		
TB3MS	3-Month Treasury Bill: Secondary Market Rate	1
TB6MS	6-Month Treasury Bill: Secondary Market Rate	1
GS1	1-Year Treasury Constant Maturity Rate	1
GS3	3-Year Treasury Constant Maturity Rate	1
GS5	5-Year Treasury Constant Maturity Rate	1
GS10	10-Year Treasury Constant Maturity Rate	1
BAA	Moody's Seasoned Baa Corporate Bond Yield	1
AAA	Moody's Seasoned Aaa Corporate Bond Yield	1
MORTG	30-Year Conventional Mortgage Rate	1
MORTGAGE30US	30-Year Fixed Rate Mortgage Average in the United States	1
MPRIME	Bank Prime Loan Rate	1
FEDFUNDS	Effective Federal Funds Rate	1
6. Stock Market		
VIX	CBOE Volatility Index: VIX©	2
DWCF	Dow Jones U.S. Total Stock Market Index	2
NASDAQCOM	NASDAQ Composite Index	2
NYA	NYSE COMPOSITE (DJ) (^NYA)	2
DJI	Dow Jones Industrial Average	2
GSPC	S&P 500 Composite Stock Price Index	2

WILL5000IND	Wilshire 5000 Total Market Index	2
RU2000TR	Russell 2000® Total Market Index	2
7. International Factors		
NAEXKP01DEQ661S	Total Gross Domestic Product for Germany	2
NAEXKP01GBQ661S	Total Gross Domestic Product for the United Kingdom	2
NAEXKP01CAQ661S	Total Gross Domestic Product for Canada	2
NAEXKP01FRQ661S	Total Gross Domestic Product for France	2
NAEXKP01ITQ661S	Total Gross Domestic Product for Italy	2
LORSGPORJPQ661S	Gross Domestic Product: Original Series for Japan	2
CNGDP	Total Gross Domestic Product for China	2
DEUCPIALLQINMEI	Consumer Price Index of All Items in Germany	2
GBRCPIALLQINMEI	Consumer Price Index of All Items in the United Kingdom	2
CANCPIALLQINMEI	Consumer Price Index of All Items in Canada	2
FRACPIALLQINMEI	Consumer Price Index of All Items in France	2
ITACPIALLQINMEI	Consumer Price Index of All Items in Italy	2
JPNCPIALLMINMEI	Consumer Price Index of All Items in Japan	2
CHNCPIALLQINMEI	Consumer Price Index of All Items in China	2
EXSZUS	Switzerland / U.S. Foreign Exchange Rate	2
EXUKUS	U.K. / U.S. Foreign Exchange Rate	2
EXCAUS	Canada / U.S. Foreign Exchange Rate	2
EXJPUS	Japan / U.S. Foreign Exchange Rate	2
EXCHUS	China / U.S. Foreign Exchange Rate	2
8. Money Credit and Finance Factor		
REALLN	Real Estate Loans, All Commercial Banks	2
NONREVSL	Total Nonrevolving Credit Owned and Securitized	2
USGSEC	Treasury and Agency Securities at All Commercial Banks	2
OTHSEC	Other Securities at All Commercial Banks	2
TOTALSL	Total Consumer Credit Owned and Securitized	2
BUSLOANS	Commercial and Industrial Loans, All Commercial Bank	2
CONSUMER	Consumer Loans at All Commercial Banks	2
LOANS	Loans and Leases in Bank Credit, All Commercial Banks	2
LOANINV	Bank Credit at All Commercial Banks	2
INVEST	Securities in Bank Credit at All Commercial Banks	2
BOGMBASE	Monetary Base; Total	2
REQRESNS	Required Reserves of Depository Institutions	2
TOTRESNSW	Reserves Of Depository Institutions	2
M1SL	M1 Money Stock, Billions of Dollars	2
CURRSL	Currency Component of M1	2
DEMDEPSL	Demand Deposits at Commercial Banks	2
TCDSL	Total Checkable Deposits	2

M2SL	M2 Money Stock, Billions of Dollars	2
M2OWN	M2 Own Rate	2
M2MSL	M2 Less Small Time Deposits	2
M2MOWN	M2 Minus Own Rate	2
MZMSL	MZM Money Stock	2
SVSTCBSL	Savings and Small Time Deposits at Commercial Banks	2
SVSTSL	Savings and Small Time Deposits - Total	2
SVGCBSL	Savings Deposits at Commercial Banks	2
SVGTI	Savings Deposits at Thrift Institutions	2
SAVINGSL	Savings Deposits - Total	2
STDCBSL	Small Time Deposits at Commercial Banks	2
STDTI	Small Time Deposits at Thrift Institutions	2
STDSL	Small Time Deposits - Total	2
USGVDDNS	U.S. Government Demand Deposits and Note Balances - Total	2
USGDCB	U.S. Government Demand Deposits at Commercial Banks	2
CURRCIR	Currency in Circulation	2

Table 5.7 Mergers and Acquisition Information (Source: FFIEC).

RSSD ID	Institutions Acquired by	Acquisition Date
	Bank of America Corporation	
1073364	MNC FINANCIAL, INC.	10/1/1993
1831715	C&S/SOVRAN CORPORATION	12/1/1993
1250802	NATIONSBANK TEXAS CORPORATION	12/1/1993
2122054	NATIONS FINANCIAL HOLDINGS CORPORATION	12/31/1995
1079638	BANK SOUTH CORPORATION	1/9/1996
1026016	BANKAMERICA CORPORATION	9/30/1998
2372064	NATIONSCREDIT CORPORATION	1/1/1999
1113514	FLEETBOSTON FINANCIAL CORPORATION	4/1/2004
1871159	MBNA CORPORATION	1/1/2006
1246140	MERRILL LYNCH & CO., INC.	10/1/2013
	Citigroup Inc.	
1246092	TRAVELERS CORPORATION	12/31/1993
1042351	CITICORP	8/1/2005
2879844	CITIGROUP HOLDINGS COMPANY	8/1/2005
3158452	CITIBANK (WEST) BANCORP INC.	10/1/2006
3158395	CITIBANK (WEST) HOLDINGS INC.	10/1/2006
3609114	CITIGROUP JAPAN INVESTMENTS LLC	7/5/2007
3367236	CITIGROUP FUNDING INC.	1/1/2013
1277881	ASSOCIATED MADISON COMPANIES, INC.	11/1/2013
	JPMorgan Chase & Co.	
1035296	MANUFACTURERS HANOVER CORPORATION	12/31/1991
1040795	CHASE MANHATTAN CORPORATION	3/31/1996
1037115	J.P. MORGAN & CO. INCORPORATED	12/31/2000
1068294	BANK ONE CORPORATION	7/1/2004
2881511	HAMBRECHT & QUIST GROUP	12/21/2007
2282378	JPMORGAN MEZZANINE CORPORATION	10/1/2009
3367236	CITIGROUP FUNDING INC.	1/1/2013
1277881	ASSOCIATED MADISON COMPANIES, INC.	11/1/2013
	Wells Fargo & Co.	
1061709	FIRST BELLEVUE BANCSHARES CO.	1/1/1990
1199581	FIRST INTERSTATE CORPORATION OF WISCONSIN	4/30/1990
1049042	UNITED BANKS OF COLORADO, INC.	4/19/1991
1065958	UNITED BANCSHARES, INC.	10/3/1992
1143463	MERCHANTS AND MINERS BANCSHARES, INC.	2/1/1993
1208429	FINANCIAL CONCEPTS BANCORP, INC.	4/1/1993
1124154	M & D HOLDING COMPANY	10/1/1993
1067840	RALSTON BANCSHARES, INC.	10/7/1993

1124248	WINNER BANSHARES, INC.	12/10/1993
1870590	NORWEST COLORADO, INC.	8/8/1997
1120763	NORWEST AGRICULTURAL CREDIT, INC.	12/23/1999
1847912	NORWEST HOLDING COMPANY	12/31/1999
2058528	INDEPENDENT BANCORP OF ARIZONA, INC.	9/1/2000
1125834	IRENE BANCORPORATION, INC.	9/1/2000
1121827	LINDEBERG FINANCIAL CORPORATION	9/1/2000
1427127	NORWEST AMG, INC.	9/1/2000
1927131	VICTORIA FINANCIAL SERVICES, INC.	9/1/2000
1124079	INTERNATIONAL BANCORPORATION, INC.	4/13/2001
1207093	FARMERS NATIONAL BANCORP, INC.	9/1/2001
1246627	FIRST NATIONAL BANKSHARES, INC., THE	9/1/2001
1053076	PACKERS MANAGEMENT COMPANY, INC.	9/1/2001
1120839	WELLS FARGO AUDIT SERVICES, INC.	12/10/2001
1021981	BUFFALO NATIONAL BANCSHARES, INC.	12/13/2002
1105939	TEXAS BANCSHARES, INC.	12/13/2002
1073551	WACHOVIA CORPORATION	12/31/2008
3297007	IJL 2004, LLC	2/18/2011
1109946	CENTURY BANCSHARES, INC.	8/1/2011
1136531	GREATER BAY BANCORP	8/1/2011
3084953	PLACER SIERRA BANCSHARES	8/1/2011
2537993	SIGNET STUDENT LOAN CORPORATION	10/3/2011
3648609	EDWARDS DEVELOPMENT CORPORATION	12/1/2011
2940986	NERO LIMITED, LLC	12/15/2011
1074893	CENTRAL FIDELITY PROPERTIES, INC.	2/1/2012
1145805	FAIRFAX CORPORATION, THE	2/1/2012
794578	CAPITOL FINANCE GROUP, INC.	5/1/2012
3597453	WACHOVIA CAPITAL INVESTORS, INC.	5/2/2012
2747372	FPFC MANAGEMENT LLC	8/1/2012

Chapter 6 Discussion and Future Work

The FECM is a promising tool that can simultaneously handle the long term and short term equilibrium information in a high dimensional nonstationary data. It is an important extension of the classical ECM and FAVAR model for both forecasting and structural analysis. Further extending the FECM to TVC-FECM, we are able to conduct multi-dimensional dissection of the time series data structure. In our empirical analysis, the TVC-FECM reveals the varying importance of the observed variables to the target variables, as well as the time evolution of their relations. It suggests that the TVC-FECM could be very useful for empirical analysis, especially when considering the models under different scenarios.

In this thesis, the macroeconomic variable data set we used is a simple merger of the data sets in Bernanke, Boivin and Eliasz (2005), Stock and Watson (2009), and Korobilis (2009). The factors estimated actually are quite sensitive to the variables included in the data set. The three previous works all had a variable refining process – not all macroeconomic variables they collected were used in their factor estimation. To obtain estimated factors that would capture the characteristic of the economy environment more accurately, one of our future studies is to improve the observed data set based on macroeconomic knowledge. Another related issue is that we used only publicly available data from FR Y-9C in our PPNR modeling, while in the real stress testing world regulators and each bank holding company has access to a much greater collection of information. An improvement of model performance is expected if bank-specific granular information could be introduced in defining the components of PPNR.

References

Ahn, S.C., and A.R. Horenstein. (2013), Eigenvalue Ratio Test for the Number of Factors, *Econometrica*, 81(3), 1203–1227.

Bai, J. (2003), Inferential Theory for Factor Models of Large Dimensions, *Econometrica*, 71, 135-172.

Bai, J. (2004), Estimating Cross-section Common Stochastic Trends in Nonstationary Panel Data, *Journal of Econometrics*, 122, 137-183.

Bai, J., and S. Ng. (2002), Determining the Number of Factors in Approximate Factor Models, *Econometrica*, 70, 191-221.

Bai, J., and S. Ng. (2004), A PANIC Attack on Unit Roots and Cointegration, *Econometrica*, 72, 1127-1177.

Bai, J., and S. Ng. (2006a), Confidence Intervals for Diffusion Index Forecasts and Inference for Factor-Augmented Regressions, *Econometrica*, 74, 1133-1150.

Bai, J., and S. Ng. (2008), Large Dimensional Factor Analysis, *Foundations and Trends in Econometrics*, 3(2): 89-163.

Bai, J. and S. Ng. (2010), Instrumental Variable Estimation in a Data-Rich Environment, *Econometric Theory*, 26, 1577–1606.

Baltagi, B. H., Feng, Q., and C. Kao. (2015), Identification and Estimation of a Large Factor Model with Structural Instability, *Working Paper*, Syracuse University.

Banerjee, A., and M. Marcellino. (2009), Factor-Augmented Error Correction Models, in *The Methodology and Practice of Econometrics – A Festschrift for David Hendry*, ed. By J.L. Castle and N. Shephard, N. Oxford: Oxford University Press, 227-254.

Banerjee, A., Marcellino, M., and I. Masten. (2007), Forecasting Macroeconomic Variables Using Diffusion Indexes in Short Samples With Structural Change, *Forecasting in the Presence of Structural Breaks and Model Uncertainty*, ed. by D. Rapach and M. Wohar, Elsevier.

Banerjee, A.,Marcellino. M., and I. Masten. (2014a), Forecasting with Factor-augmented Error Correction Models, *International Journal of Forecasting*, 30(3), 589-612.

Banerjee, A., Marcellino. M., and I. Masten. (2014b), Structural FECM: Cointegration in large-scale structural FAVAR models, Discussion paper, No. 9858.

Banerjee, A., Marcellino. M., and I. Masten. (2015), An Overview of the Factor-augmented Error-Correction Model, University of Birmingham, Department of Economics Discussion paper, 15-03.

Bates, B., Plagborg-Møller, M., Stock, J.H., and M.W. Watson. (2013), Consistent Factor Estimation in Dynamic Factor Models with Structural Instability, *Journal of Econometrics*, 177, 289-304.

Bellotti, T. and J. Crook. (2009), Credit Scoring with Macroeconomic Variables Using Survival Analysis, *Journal of the Operations Research Society*, 60 (12), 1699–1707.

Bernanke, B.S., Boivin, J., and P. Eliasz. (2005), Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach, *Quarterly Journal of Economics*, 120, 387-422.

Beyer, A., Farmer, R., Henry, J. and M. Marcellino. (2005), Factor Analysis in a New-Keynesian Model. Working Paper Series #510, European Central Bank

Beyer, A., Farmer, R., Henry, J. and M. Marcellino. (2008), Factor Analysis in a Model with Rational Expectations, *Econometrics Journal*, 11, 271–286.

Bierens, H.J., and L. F. Martins, L.F. (2010), Time Varying Vointegration, *Econometric Theory*, 26 (05), 1453–1490.

Boivin, J., and S. Ng. (2005), Understanding and Comparing Factor-Based Forecasts, *International Journal of Central Banking*, 1, 117-151.

Board of Governors of the Federal Reserve System. (2012), Comprehensive Capital Analysis and Review 2012: Methodology and Results for Stress Scenario Projections, http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20120313a1.pdf

Board of Governors of the Federal Reserve System. (2014), 2015 Supervisory Scenarios for Annual Stress Tests Required under the Dodd-Frank Act Stress Testing Rules and the Capital Plan Rule, http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20141023a1.pdf

Board of Governors of the Federal Reserve System. (2014), Comprehensive Capital Analysis and Review 2015 Summary Instructions and Guidance, http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20141017a1.pdf

Board of Governors of the Federal Reserve System. (2015), Comprehensive Capital Analysis and Review 2015: Assessment Framework and Results, http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20150311a1.pdf

Board of Governors of the Federal Reserve System. (2015), Dodd-Frank Act Stress Test 2015: Supervisory Stress Test Methodology and Results, Appendix B: Models to Project Net

Income and Stressed Capital, http://www.federalreserve.gov/bankinforeg/stress-tests/2015-Appendix-B.htm

Breitung, J., and S. Eickmeier (2011): Testing for Structural Breaks in Dynamic Factor Models, *Journal of Econometrics*, 163, 71-84.

Cattell, R. B. (1966), The Scree Test for the Number of Factors, *Multivariate Behavioral Research*, 1, 245-76.

Chen, L., Dolado, J., and J. Gonzalo. (2014), Detecting Big Structural Breaks in Large Factor Models, *Journal of Econometrics*, 180, 30-48.

Cheng, X., Liao, Z., and F. Schorfheide. (2014), Shrinkage Estimation of High Dimensional Factor Models with Structural Instabilities, *NBER Working Paper*, No. 19792.

Chikuse, Y. (2003), Statistics on Special Manifolds, *Lecture Notes in Statistics*, vol. 174, Springer-Verlag, New York.

Cogley, T., and T. Sargent. (2005), Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US, *Review of Economic Dynamics*, 8, 262–302.

Corradi, V., and N. R. Swanson. (2014), Testing for Structural Stability of Factor Augmented Forecasting Models, *Journal of Econometrics*, 182, 100-118.

Covas, F., Rump, B., and E. Zakrajšek. (2014), Stress-testing US bank holding companies: A dynamic panel quantile regression approach, *International Journal of Forecasting*, 30(3), 691-713.

Doz, C., Giannone, D. and L. Reichlin. (2012), A Quasi Maximum Likelihood Approach for Large Approximate Dynamic Factor Models, *The Review of Economics and Statistics*, 94(4), 1014-1024.

Durbin, J., and S. Koopman. (2002), A Simple and Efficient Simulation Smoother for State Space Time Series Analysis, *Biometrika*, 89, 603–616.

Eickmeier, S., and C. Ziegler. (2008), How Successful are Dynamic Factor Models at Forecasting Output and Inflation? A Meta-Analytic Approach, *Journal of Forecasting*, 27(3), 237-265.

Engle, R.F., and C. W. J. Granger. (1987), Co-Integration and Error Correction: Representation, Estimation, and Testing, *Econometrica*, 55(2), 251-276.

Engle, R.F., and M.W. Watson. (1981), A One-Factor Multivariate Time Series Model of Metropolitan Wage Rates, *Journal of the American Statistical Association*, 76, 774-781.

Engle, R.F., and M.W. Watson. (1983), Alternative Algorithms for Estimation of Dynamic MIMIC, Factor, and Time Varying Coefficient Regression Models,*Journal of Econometrics*, Vol. 23, pp. 385-400.

Favero, C. A., Marcellino, M. and F. Neglia. (2005), Principal Components at Work: The Empirical Analysis of Monetary Policy With Large Data Sets, *Journal of Applied Econometrics*, 20, 603–620.

Geweke, J. (1977), The Dynamic Factor Analysis of Economic Time Series, in *Latent Variables in Socio-Economic Models*, ed. by D.J. Aigner and A.S. Goldberger, North-Holland.

Giannone, D., Reichlin, L., and L. Sala. (2004), Monetary Policy in Real Time, *NBER Macroeconomics Annual*, 2004, 161-200.

Giannone, D., Reichlin, L,. and D. Small. (2008), Nowcasting: The Real-Time Informational Content of Macroeconomic Data, *Journal of Monetary Economics*, 55, 665-676.

Guerrieri, L., and M. Welch. (2012), Can Macro Variables Used in Stress Testing Forecast the Performance of Banks?, *Finance and Economics Discussion Series Paper*, 2012-49, Federal Reserve Board.

Han, X., and A. Inoue. (2011), Tests for Parameter Instability in Dynamic Factor Models, North Carolina State University, Working Paper.

Johansen, S. (1995), Likelihood-based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press, Oxford and New York.

Kapetanios, G., and M. Marcellino. (2010), Factor-GMM Estimation with Large Sets of Possibly Weak Instruments, *Journal Computational Statistics & Data Analysis*, 54(11), 2655-2675.

Kim, S., Shephard, N., and S. Chib. (1998), Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models, *Review of Economic Studies*, 65, 361–393.

Koop, G., and D. Korobilis. (2014), A New Index of Financial Conditions, *European Economic Review*, 71, 101-116.

Koop, G., León-González, R., and R. Strachan. (2008), Bayesian Inference in a Cointegrating Panel Data Model, *Advances in Econometrics*, 23, 433–469.

Koop, G., León-González, R., and R. Strachan. (2010), Bayesian Inference in the Time Varying Cointegration Model, Working Paper.

Koop, G., León-González, R., and R. Strachan. (2011), Bayesian Inference in a Time Varying Cointegration Model, *The Journalof Econometrics*, 165, 210–20.

Korobilis, D. (2012), Assessing the Transmission of Monetary Policy Using Time-varying Parameter Dynamic Factor Models, *Oxford Bulletin of Economics and Statistics*, 75(2), 157–179.

Li, G., Song, H., and S. F. Witt. (2005), Recent developments in econometric modeling and forecasting, *Journal of Travel Research*, 44, 82–99.

Lütkepohl, H. (2014). Structural Vector Autoregressive Analysis in a Data Rich Environment: A Survey. *DIW Discussion paper*, 1351.

Nakajima, J., Kasuya, M., and T. Watanabe. (2011), Bayesian analysis of Time-Varying Parameter Vector Autoregressive Model for the Japanese Economy and Monetary Policy, *Journal of the Japanese and International Economies*, 25(3), 225–245.

Onatski, A. (2008), The Tracy-Widom Limit for the Largest Eigenvalues of Singular Complex Wishart Matrices, *Annals of Applied Probability*, 18, 470-490.

Onatski, A. (2009), Testing Hypotheses about the Number of Factors in Large Factor Models, *Econometrica*, 77, 1447-1479.

Primiceri, G. (2005), Time Varying Structural Vector Autoregressions and Monetary Policy, *Review of Economic Studies*, 72, 821–852.

Quah, D., and T.J. Sargent. (1993), A Dynamic Index Model for Large Cross Sections (with discussion), in *Business Cycles, Indicators, and Forecasting*, ed. by J.H.Stock and M.W. Watson, University of Chicago Press for the NBER,285-310.

Sargent, T.J. (1989), Two Models of Measurements and the Investment Accelerator, *Journal of Political Economy*, 97, 251–287.

Sargent, T.J., and C.A. Sims. (1977), Business Cycle Modeling Without Pretending to Have Too Much A-Priori Economic Theory, in *New Methods in Business Cycle Research*, ed. by C. Sims et al., Federal Reserve Bank of Minneapolis.

Simons, D., and F. Rolwes. (2009), Macroeconomic Default Modeling and Stress Testing, *International Journal of Central Banking*, 5(3), 177-204.

Strachan, R., and B. Inder. (2004), Bayesian Analysis of the Error Correction Model, *Journal of Econometrics*, 123, 307-325.

Stock, J.H., and M.W. Watson. (1989), New Indexes of Coincident and Leading Economic Indicators, *NBER Macroeconomics Annual*, 1989, 351-393.

Stock, J.H., and M.W. Watson. (1999), Forecasting Inflation, *Journal of Monetary Economics*, 44, 293-335.

Stock, J.H., and M.W. Watson. (2002a), Macroeconomic Forecasting Using Diffusion Indexes, *Journal of Business and Economic Statistics*, 20, 147-162.

Stock, J.H., and M.W. Watson. (2002b), Forecasting Using Principal Components from a Large Number of Predictors, *Journal of the American Statistical Association*, 97, 1167-1179.

Stock, J.H., and M.W. Watson. (2005), Implications of Dynamic Factor Models for VAR Analysis, manuscript.

Stock, J.H., and M.W. Watson. (2006), Forecasting with Many Predictors, ch. 6 in *Handbook of Economic Forecasting*, ed. by Elliott, G., Granger, C. W.J., Timmermann, A. G., Elsevier, 515-554.

Stock, J.H., and M.W. Watson. (2009), Forecasting in Dynamic Factor Models Subject to Structural Instability, in *The Methodology and Practice of Econometrics - A Festschrift for David Hendry*, ed. By J.L. Castle and N. Shephard, N. Oxford: Oxford University Press, Ch7.

Stock, J.H., and M.W. Watson. (2010), Dynamic Factor Models, in *Oxford Handbook of Economic Forecasting*, ed. by Clements, M. P., Henry, D. F. Oxford: Oxford University Press.

Tsay, R. S. (2010), Analysis of Financial Time Series. 3rd ed. New York: Wiley, 2010.

Watson, M.W. (2004), Comment on Giannone, Reichlin, and Sala, *NBER Macroeconomics Annual*, 2004, 216-221.