STRESS TESTING BANK PROFITABILITY

Michael Duane, Til Schuermann[*], Peter Reynolds
Oliver Wyman

First Draft: July 2013
This Draft: September 22 2013

Abstract

A defining difference of macro-style stress testing is the explicit consideration of profitability dynamics in the stress scenario. Traditional stress testing had focused almost exclusively on losses only, but a complete assessment of capital adequacy under stress must take into account not just the balance sheet but also the income statement. For instance, in the 2013 US stress test, reduction in projected income for the 18 mandatory bank holding companies (BHCs) covered nearly 60% of projected stress losses. We describe and discuss a framework for modeling the major components of the income statement for BHCs using the U.S. regulatory reports as an empirical illustration. We review approaches taken by the industry and trace its remarkable development in the wake of the financial crisis. We find – perhaps unsurprisingly and in line with previous literature – that successfully modeling profitability requires a tailored BHC-specific approach to revenue segmentation and modeling. We argue that failure to pursue a relatively granular income source segmentation along different business activities, far more granular than reflected in typical regulatory reports, will obscure significant underlying differences in macro risk factor sensitivities.

Keywords: revenue dynamics, capital requirements, leverage, systemic risk
JEL Codes: G21, G28, G20.

---

[*] Oliver Wyman and Wharton Financial Institutions Center; (corresponding author: *til.schuermann@oliverwyman.com*). We would like to thank Cary Lin and Henry McLoughlin for outstanding analytical support and Dov Haselkorn, Umit Kaya, Andy McGee, an anonymous referee and the editor for comments and discussion. All remaining errors are ours, of course.

1. Introduction

Macro-conditional stress testing is entering its fourth year following the initial Supervisory Capital Assessment Program (SCAP)[1] in 2009, the banking stress test that effectively put a line under the U.S. banking crisis. Stress testing as a risk management tool is hardly new, but it had historically focused on losses only, given a single shock. The SCAP changed stress testing significantly by making the entire exercise dynamic and including profitability (pre-provision net revenue or PPNR in the nomenclature) under stress. If old-style stress testing focused mostly on the balance sheet, and there mostly on assets, new macro-conditional stress testing required balance sheet evolution over time as the scenario unfolds, with the income statement linking period t to $t+1$ balance sheets, all to arrive post-stress capital impacts, the "bottom line" of stress testing.

In this paper we focus on the pre-provision net revenue (PPNR) side of the stress testing problem. Loss generation and emergence has received considerably more attention in the literature; for a recent survey see Schuermann (2013) and papers in this volume. We follow the common definition where PPNR is the sum of three parts: net interest income (typically expressed in the form of net interest margin or NIM), non-interest income and non-interest expense.[2] For the typical commercial bank, NIM dominates PPNR as most income derives from lending activities. The macro stress scenario will typically specify both the (risk-free) yield curve and credit spreads, and both are needed to model the impact on NIM. Unfortunately the literature has found it difficult to find to find any meaningful relation between rising or falling rates on bank profitability, perhaps in part because of interest rate hedging strategies (English 2002, Purnanandam 2007).

The share of non-interest income has been rising steadily in the last two decades for the entire U.S. banking system. Stiroh (2004, 2006) shows that not only has the share of noninterest income been steadily rising in U.S. banks, from 25% in 1985 to 42% in 2004, but it is associated with greater volatility, higher beta and thus lower risk-adjusted returns, particularly among the larger banks that are more dependent on non-interest income activities. It is noteworthy that trading is by far the most volatile of the non-interest income categories, more than five times more volatile than the other categories (Stiroh 2004).

[1] See Board of Governors (2009a, b, c), as well as Hirtle, Schuermann and Stiroh (2009).
[2] See Board of Governors (2013a), p.9.

Our analysis will focus on U.S. bank holding companies (BHCs) for several reasons. First, between the SCAP and subsequent CCAR programs, we now have three rounds of public disclosures of stress test results, including for the most recent (2013) a parallel set of bank-own disclosed results. Second, by virtue of regulatory call reports (for banks) and FR Y-9C reports (consolidated reports for BHCs) going back to the 1980s, the U.S. has some of the richest data available to analysts and researchers of any banking system.

Our analytical approach builds on existing literature, including Guerreri and Welch (2012), who attempt to model PPNR at the level of NIM and non-interest income/expense, and Covas, Rump, and Zakrajsek (2012), who segment PPNR into six components. Both of these papers, however, focus mainly on the "price" component of PPNR – normalizing revenue as a percentage of assets – but do not explicitly address the "volume" component needed to produce a fully integrated and dynamic forecast of revenue under stress. In our illustrative analysis, we develop models of both "price" – defined as revenue as a % of assets – and "volume" – defined as asset volume – for a set of PPNR sub-components. Because of the requirements of CCAR to model the balance sheet and income statement dynamically over 9 quarters, quarter by quarter, one must develop models for both price (yield) and volume or balances. Importantly, it is not clear that the same risk factors drive both price and volume; models of PPNR directly would miss such differences.

We also draw on extensive insights gained over the years of working for banks to implement and execute CCAR-style stress tests, having by now worked for a majority of BHCs subject to mandatory stress testing.

The remainder of the paper proceeds as follows: Section 2 lays out the analytical framework, Section 3 provides an empirical illustration, and Section 4 concludes.

2. Analytical Framework

2.1. Overview

We begin the discussion with the definition of PPNR: net interest income (interest income less interest expense) plus non-interest income less non-interest expense. Non-interest income includes a very wide variety of activities. For the purposes of describing the breadth of non-interest income, we follow Stiroh (2004, 2006) by grouping into four broad categories on the Y-9C regulatory reports: fiduciary, service charges (dominated by checking account related fees, ATM, and so on), trading, fees and other income. This last category is a very wide collection and includes among other things loan commitment fees, investment banking, advisory, and underwriting fees and commissions, fees from safe deposit boxes, commissions, land rental fees, net income from securitizations, and bank card and credit card interchange fees. Clearly some of these line items are likely to be especially important for some institutions, but less important for others. Note that PPNR also includes income from mortgage servicing rights and losses from operational-risk events, mortgage put-back losses, and OREO (other real estate owned) expenses. However, PPNR excludes gains or losses from AFS/HTM (available for sale / held to maturity) securities, as well as trading and counterparty losses due to the stress scenario (direct impact of market stress shock), but it does include trading income.[3]

To be sure, this specific and somewhat rigid (but unambiguous) definition used by the Federal Reserve in its CCAR program could be reasonably considered as narrow.[4] It is particular to U.S. bank regulatory reporting and U.S. GAAP accounting. Indeed IFRS accounting also treats off-balance sheet exposures differently which likely will influence the modeling approach of balances which, in turn, generate income. Despite these limitations, we think that coverage is sufficiently broad and generic to allow for broader insights into the modeling of bank profitability under stress.

Projecting PPNR is hard. Guerrieri and Welch (2012), in their study on linking macro variables to bank regulatory reports for purposes of stress testing, state rather soberly: “In the case of PPNR, even the best performing model does not beat a random walk at all horizons.” Large banks are complex institutions with a diverse set of revenue streams. Different revenue streams are sensitive to different macro factors, and the degree of sensitivity will also vary across banks. For instance, asset management revenue will be sensitive to stock market returns and volatility as well as interest rate movements. Mortgage origination fees are sensitive to HPI (home price index) and interest rates. Net interest income is itself a complex collection of revenue streams from different lending products: on the consumer side they include mortgages, auto loans and credit cards, and on the commercial side include CRE (commercial real estate), the very heterogeneous class of C&I (commercial & industrial) lending, SME (small and medium enterprise) lending, and so on. Many of these products have embedded optionality (e.g. mortgages) with corresponding complex behaviors driven by the underlying risk factors. This

[3] See Federal Reserve Board (2012a).
[4] For a non-US example, see Coffinet and Lin (2010).

hardly exhaustive list serves to show that net interest income is likely sensitive to a broad set of risk factors, from interest rates, especially the shape of the yield curve, to the shape of the term structure of credit spreads (across maturity *and* credit quality), and geographic variation of HPI and unemployment.

We argue that a model which aims to predict $ PPNR directly confounds price and volume effects, and creates challenges that we describe further in Section 2.2. Indeed, PPNR is crucially dependent on the size and composition of the balance sheet. CCAR-style stress testing may be thought of as an exercise of generating dynamic forecasts of balance sheets and the income statements which connect them, as illustrated in Figure 1 below, here for the case of two years.[5] In each period the relevant regulatory ratios (capital and, say, liquidity) need be satisfied.

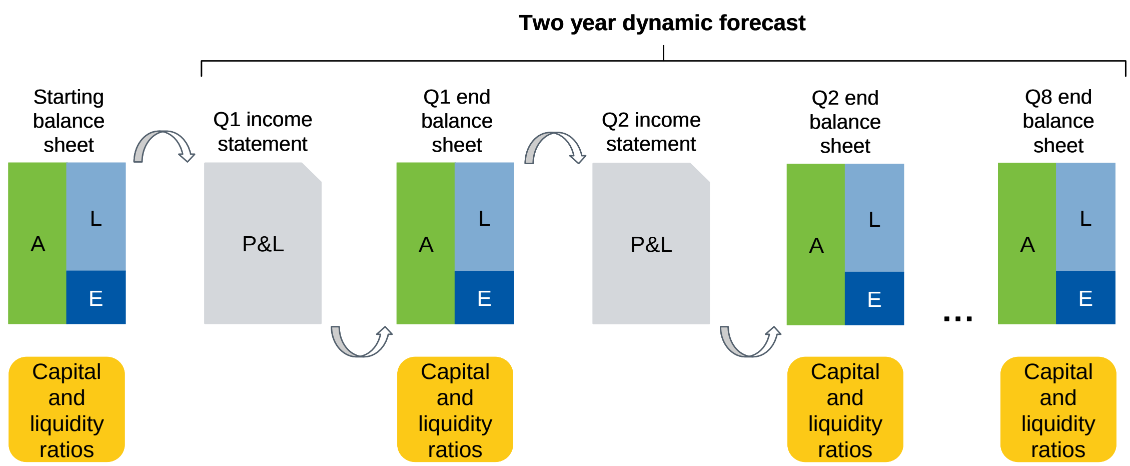


Figure 1: Stress testing balance sheet and income statement dynamics[6].

Determining post-stress capital adequacy requires modeling both the income statement and the balance sheet, both flows and stocks, over the course of the stress test horizon. The point of departure is the current balance sheet, at which point the bank meets the required capital (and, if included, liquidity) ratios. The starting balance sheet generates the first quarter's income and loss, which in turn determines the quarter-end balance sheet. The modeler is then faced with the

[5] The horizon is 9 quarters for the CCAR as it is based on Q3, not Q4, balance sheets.
[6] Note that the denominator of the regulatory ratios is typically risk-weighted assets (RWA), where the risk weights are determined by the prevailing regulatory capital regime, namely Basel I (in the U.S. cases of the SCAP and CCAR) and Basel II (in the European stress tests). The Basel III rules have recently been finalized in the U.S. with rolling implementation over the next five years (Federal Reserve Board, 2013b). The many subtleties of what this implies is beyond the scope of this paper, but suffice it to say that a bank may be forced to raise capital under one regime but not the other, and without considerable detail about the portfolio, there is no way to know which regime will result in a more favorable treatment.

problem of considering renewals for maturing assets, the nature and amount of new assets originated and/or sold during the quarter, changes to the pricing of services provided (both strategic and reactive changes), and any other capital depleting or conserving actions such as acquisitions or spin-offs, dividend changes or share (re-)purchase or issuance programs, including employee stock and stock option programs.

Each of the major dynamics that influence an iterative projection of the balance sheet and income statement should be considered and modeled independently. Developing a model that either explicitly or (worse) implicitly attempts to combine these dynamics into a single regression, such as an attempt to directly model $ PPNR, confounds and obscures these individual dynamics and creates practical challenges described in Section 2.2. Direct forecasts of $ PPNR can conflate balance-related trends (e.g. mortgage origination volume) and pricing-related trends (e.g. origination fee pricing schemes) that may not move in the same direction or be sensitive to the same macroeconomic conditions. Such a model will have inferior predictive power, may be less economically intuitive, and will be more difficult to explain and attribute than developing two separate models for balance-related trends (volume) and pricing-related trends (price). Further, for the many components of PPNR that have both balance-related and pricing-related components an attempt to model $ PPNR directly violates the iterative approach to balance sheet and income statement forecasting described in Figure 1above by mixing balance sheet and income statement elements.

2.2. Producing a PPNR forecast

Operationally developing a PPNR framework and producing a PPNR forecast is complicated, and involves tracking and synchronizing a number of moving parts. In Figure 2 below, we have illustrated the high-level components that would exist and interact within an institution's PPNR framework, as well as the linkage points between the PPNR components and the rest of the stress test architecture. We expand upon these interactions later in this section.

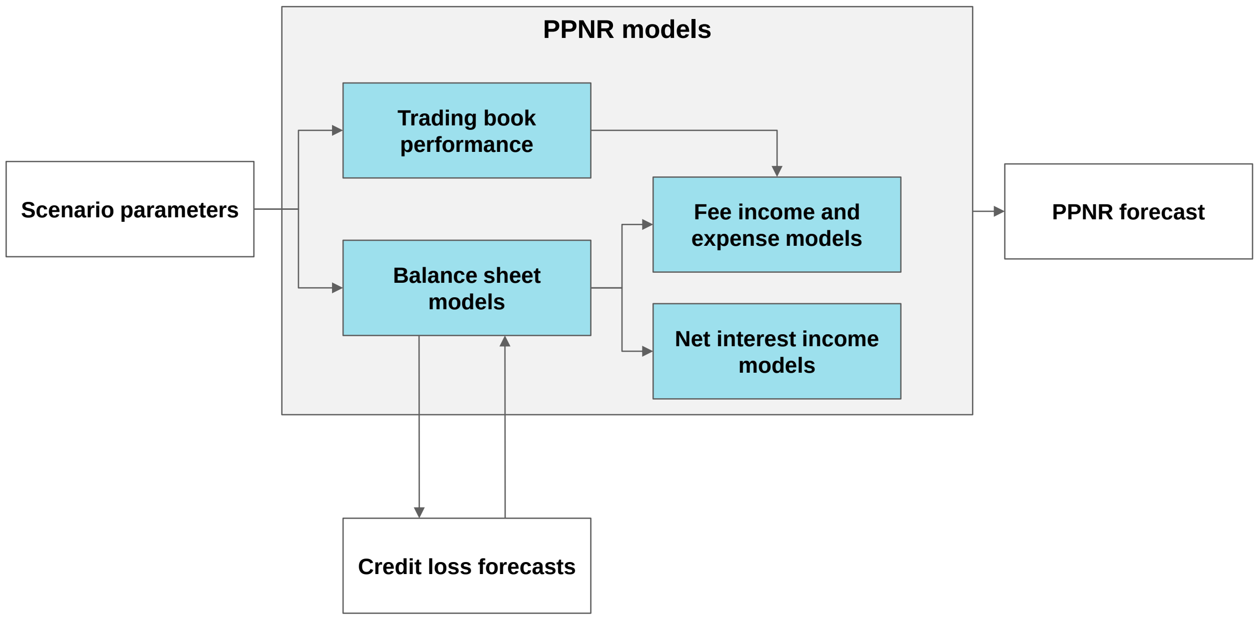


Figure 2: PPNR Forecasting schematic

Producing a PPNR forecast that satisfies regulator expectations requires a well-developed analytical infrastructure and set of processes. Where possible, institutions have tried to make use of existing infrastructure and resort to building new models only where needed. In practice, institutions will typically use their ALM (asset-liability management) infrastructure to produce net interest income forecasts. If the institution has a well-functioning and integrated ALM process, it is the natural environment to use for generating net interest income, conditional on the relevant interest rate (risk-free and credit sensitive) factors. Producing stressed net interest income still requires a set of inputs beyond the macro risk factors – namely balance forecasts and pricing/spread forecasts on new origination – but the analytical machinery to translate these inputs into net interest income is largely in place.

For the remaining components of PPNR – balances and non-interest income/expense "price" forecasts – similar machinery typically does not exist. As a result, many institutions must build bespoke stress models to forecast these PPNR components. The sophistication of these models varies across the industry. At one end of the spectrum some attempt direct regressions on P&L line items – essentially combining "price" and "volume" into a single model – while at the other end some disaggregate a P&L item into a number of discrete business drivers and model each separately. This is also a model maturation process; with four rounds of stress testing behind us in the U.S., industry practice is clearly moving towards the disaggregated approach. The degree to which disaggregation and greater granularity will find greater dependence on easily observable macro risk factors, the problems pointed out by Guerrieri and Welch (2012), Covas, Rump, and Zakrajsek (2012) and also Pritsker (2012), remains an empirical matter. But as we point out below, one ignores granularity at one's peril.

Operationally developing and executing a robust PPNR framework poses a set of challenges, which we have broadly grouped into (1) technical challenges and (2) consistency challenges. We describe each below.

Technical challenges relate mainly to the analytical issues that arise in building PPNR models. These are not different in kind from challenges that arise in building any other model. But in some cases the magnitude of the challenge is exacerbated in the PPNR context by the lack of historical focus on quantitative revenue forecasts as opposed to say, credit loss or migration. Specific technical challenges include:

- Model granularity: institutions must determine which business areas and balance sheet/P&L line items to model. There is a natural tension between the desire for sharper resolution on stress performance and the marginal cost/benefit of building and maintaining an additional model. In the extreme, an institution might separately model each balance sheet/P&L line item: such an approach is cumbersome to build and maintain and potentially wastes resources analyzing areas with little macro sensitivity. Striking the right balance requires business intuition and expert judgment, and should be focused on identifying the largest and/or most volatile line items and concentrating efforts on modeling volumes (e.g. balances) and profitability (e.g. spreads, margins) for these areas.
- Data availability: in developing models, institutions are often faced with a choice between internal and external data. Internal data is typically more granular and allows banks to model the response of their own specific portfolios to stress, but may be a shorter, often much shorter time series. Typically, institutions have not invested as heavily in capturing detailed revenue-related information as they have for credit risk management purposes. Moreover, the internal data will almost certainly include non-economic events that confound analysis (e.g. shift in strategy, M&A). External sources have the opposite strengths and weaknesses: they typically have a longer time series, but are not as reflective of an institution's own internal strategy and portfolio mix. Both data sources are viable, and institutions should consider the relevant trade-offs for each model, and firms are increasingly pursuing this path.
- Performance under stress conditions: PPNR models ought to be calibrated to perform well under extreme conditions. For obvious reasons this is hard, and a naïve adherence to

maximizing goodness of fit statistics may not always be the best strategy. It is not uncommon that expert judgment plays a significant role in the final model choice, though rarely in a formal Bayesian sense. Institutions that do not apply an expert-based review may find themselves with statistically accurate, but economically incoherent models. Some firms have explored forecast combinations or model averaging as described, for instance, by Timmermann (2006). Such challenges are increased, as best practices dictate that institutions use the same PPNR projection models for both baseline and more adverse scenarios, and that baseline results are related to strategic planning within the institution. As such, the models need not only to be appropriately conservative in the stress case, but also appropriately calibrated to reflect expected values.

Given the number of different components that make up a PPNR forecast – especially the use of the same balance forecast to produce multiple PPNR line items under our proposed "price"/"volume" framework – maintaining consistency can be a significant operational challenge that requires strong process discipline. Specific consistency challenges include:

- Consistency within the PPNR infrastructure: the components of a PPNR forecasts must be internally consistent. Applying a "price"/"volume" framework as we articulate in this paper certainly mitigates this problem significantly. When PPNR models are built to directly forecast P&L line items, these forecasts will embed assumptions about underlying balances and fees/margins associated with a specific P&L value, which may differ from assumptions used elsewhere. We allude to such an example in Section 2.1 above: directly forecasting mortgage origination fee revenue would embed assumptions on mortgage origination volume and fees charged per mortgage origination. The implicit assumptions on origination volume, in particular, may differ from those used to forecast balance sheet size and net interest income. Sub-dividing forecasts into balance ("volume") forecasts and margin/fee ("price) forecasts breaks the process into transparent steps and allows a single balance forecast to be consistently used in multiple places, but still requires disciplined tracking to ensure that common inputs are consistent across the entire PPNR infrastructure.
- Consistency with other stress components: as PPNR is not used in vacuum, it is equally important to be sure that the inputs and outputs used in PPNR forecasts are consistent with those used in other parts of the stress test. In particular, balance volumes must be reflective not only of new originations, but of losses projected as part of an institution's credit forecast.

Similarly, the forecast size and composition of the balance sheet at each point in the forecast must be linked to the credit risk RWA forecast to ensure consistency on this dimension. A pessimistic credit view expressed by the loss modeling teams cannot co-exist with a more optimistic balance sheet evolution and PPNR generation expressed by those respective modeling teams.

2.3. Challenges for Banks with Significant Trading Operations

Banks with a sizable trading book are faced with an additional set of challenges. First, macroprudential stress test in both the U.S. and Europe apply an instantaneous shock to their trading portfolios, including all derivative positions. For example, in CCAR-2013 the six BHCs with the largest trading books[7] were required to implement a scenario with well over 2500 Fed specified shocks, spanning the risk factors set for both domestic and overseas markets.[8] This compares with a far sparser representation of the state space for the nine quarter CCAR scenario: 14 domestic, and three each from four regions for a total of 26 risk factors.[9] We compare a few of these factor paths across the different macroprudential stress tests in Figure 3 and Figure 4. One is then faced with the problem of applying consistently the highly specific market shock scenario, which is instantaneous, with the much more thinly specific but gradually unfolding CCAR scenario. For example, the spot shock for U.S. equities in the 2013 instructions is given as -29.43%, while the peak to trough drop in the 2013 CCAR scenario is about 51%, spread out over five quarters, with the largest one-quarter drop of 20.3% (in the second quarter of the nine-quarter scenario).

The second challenge is the predominance of non-interest income, especially for the investment (i.e. non-universal) banks. For example, at year-end 2012, Goldman Sachs' and Morgan Stanley's non-interest income share (of net interest and non-interest income) was 84% and 98% respectively. While income from trading directly is highly correlated with balance sheet size, this is less clear for other sources of revenue such as fees from investment banking or other

[7] These BHCs are Bank of America, Citigroup, Goldman Sachs, JP Morgan Chase, Morgan Stanley and Wells Fargo.

[8] The four non-U.S. regions are: UK, Euro-zone, Japan, and Asia ex Japan. Common risk factor types are real GDP growth and inflation, and FX (to USD). In addition, the U.S. risk factor set includes nominal GDP, real disposable income, unemployment, inflation, the 3M and 10Y Treasury rate, the BBB corporate yield, mortgage rate, total stock market index, VIX, and residential and commercial real estate price index.

[9] See Board of Governors (2012b).

advisory services. As we show in the empirical illustration below, non-interest income is typically less correlated with macro risk factors than are interest income components.

3. Empirical Illustration

In this section we present an empirical illustration using FR Y-9C regulatory reports. The objective is to build simple, at most 2-factor, time series regressions to predict PPNR at the aggregate (industry) level. Specifically, we develop models of both “price” – defined as revenue as a % of assets – and “volume” – defined as asset volume – for a set of PPNR sub-components. Our sample is made up of all banks that produced a Y-9C regulatory report between 1994Q1 and 2012Q4, with all “intermediate”-level BHCs for which a parent BHC also produced a Y-9C report removed (to avoid a double-count). Taken together, this sample constitutes the large majority – if not all – of BHCs in the US.

We assemble a time series of total quarterly PPNR generated by US BHCs from 1994Q1 to 2012Q4. Table 1 below is a list of the PPNR components we consider (the dependent variables) along with the candidate macro-factors chosen by and large from the CCAR risk factor set. Note that this is not a panel regression model but rather an aggregate banking model.

Table 1: PPNR component variables and macro risk factors considered

PPNR component variable	Macro risk factor considered
Total assets	All domestic CCAR risk factors
Total liabilities	All domestic CCAR risk factors
Total interest income (% total assets)	GDP growth (real and nominal), 3M Treasury Yield, 10Y Treasury Yield, BBB Corporate Yield, BBB Corporate Spread (vs. 10Y Treasury)
Total interest expense (% total liabilities)	GDP growth (real and nominal), 3M Treasury Yield, 10Y Treasury Yield, BBB Corporate Yield, BBB Corporate Spread (vs. 10Y Treasury)
Fiduciary revenue (% total assets)	GDP growth (real and nominal), unemployment, BBB Corporate Yield, BBB Corporate Spread (vs. 10Y Treasury), Dow Jones Total Stock Market Index, Market Volatility Index (VIX)
Service charges (% total assets)	All domestic CCAR risk factors
Trading revenue (% total assets)	GDP growth (real and nominal), 3M Treasury Yield, 10Y Treasury Yield, BBB Corporate Yield, BBB Corporate Spread (vs. 10Y Treasury), Dow Jones

	Total Stock Market Index, Market Volatility Index (VIX)
Other non-interest income (% total assets)	All domestic CCAR risk factors
Salary expense(% total assets)	GDP growth (real and nominal), disposable income growth (real and nominal), unemployment rate, CPI inflation rate
Total non-interest expense (% total assets)	All domestic CCAR risk factors

For the explanatory variables, we also considered lags of up to 4 quarters, to reflect potential timing differences between the behavior of the explanatory and response variables. We also consider both the level and the change (where appropriate).

Despite the application of our framework and the use of more granular response variables, our analysis did not produce very strong relationships for most of the Y-9C categories, consistent with the existing literature. We have included in Table 2 the strongest relationships for each of the response variables. Goodness of fit was generally better for interest-income, ranging from interest expense (as a % of total assets) at R^2 = 51% to trading revenue (% of total assets) at R^2 = 31% to service charge-related revenue (as % of total assets) at R^2 = 12%.

Table 2: Top regressions – Y-9C variables vs. domestic CCAR macroeconomic factors

	Dependent variable (% change in…)	Coefficient 1	Value (p-value)	Coefficient 2	Value (p-value)	R^2
(1)	Total interest expense (% of total liabilities)	% change in 3-month treasury yield	0.0047 (<0.0001)	change in nominal GDP (2 qtr lag)	0.0002 (0.004)	0.513
(2)	Total interest income (% of total assets)	% change in 3-month treasury yield	0.0040 (<0.0001)	change in real GDP (2 qtr lag)	0.0003 (0.019)	0.271
(3)	Salary expense (% total assets)	CPI inflation (1 qtr lag)	-0.0002 (<0.0001)	change in nominal GDP (2 qtr lag)	9.2×10^{-5} (0.0396)	0.225
(4)	Other non-interest expense (% total assets)	% change in CRE Price Index (1 qtr lag)	-0.0214 (0.0022)	real GDP growth (4 qtr lag)	0.0003 (0.0039)	0.159
(5)	Fiduciary revenue (% total assets)	VIX Level	-4.6×10^{-6} (0.02)	US real GDP growth (2 qtr lag)	2.7×10^{-5} (0.001)	0.183
(6)	Other non-interest income (% total assets)	US real GDP growth	-0.0002 (0.0073)	% change in Dow Jones Total Stock Market Index (1 qtr lag)	0.0045 (0.0093)	0.139
(7)	Service charges (% total assets)	% change in VIX (1 qtr lag)	-7.9×10^{-6} (0.0024)	US nominal GDP growth (1 qtr lag)	-2.4×10^{-5} (0.0257)	0.122
(8)	Trading revenue (% total assets)	% change in BBB corporate yields (1 qtr lag)	0.0031 (<0.0001)	% change in mortgage rates (1 qtr lag)	-0.0037 (<0.0001)	0.311
(9)	Total assets	nominal disposable income growth (1 qtr lag)	0.0051 (0.0015)	VIX (4 qtr lag)	0.0013 (0.0119)	0.152
(10)	Total liabilities	nominal disposable income growth (1 qtr lag)	0.005 (0.0021)	VIX (4 qtr lag)	0.0013 (0.0144)	0.144

Given the level of resolution in our dataset, the lack of strong macro-sensitive relationships for many PPNR components may not be surprising and underscores the challenge involved in modeling PPNR. The challenges in using industry-level Y-9C data are twofold: first, the Y-9C categories are not sufficiently granular to capture the diversity of business activities and related business drivers at the level of an individual firm. As an example, Tables 3 – 6 below provide a high-level overview of the differences between BHC quarterly P&L disclosures on the SEC 10-K and FR Y-9C disclosures. For the purposes of illustration, we have selected two large CCAR banks, American Express and Bank of New York Mellon.[10] They are admittedly non-

[10] Note: this analysis is based entirely on publically available data.

representative of a traditional commercial bank, but that choice is deliberate to make the point of income source heterogeneity. American Express represents one extreme case: nearly all of its revenue is classified in a single Y-9C category. The Y-9C data captures several different categories of non-interest income that are separately reported in American Express' own regulatory filings. Similarly, the majority of BNY Mellon's non-interest income falls in Fiduciary Activity, which likely includes at minimum its asset servicing and asset management business lines. Further, it is reasonable to expect that for the largest of these categories, a PPNR model may attempt further granularity than what is captured even in the institution's own public filings.

Table 3: American Express non-interest income: Y-9C breakdown (FY 2012)

Category	$MM	% Total
NII: Fiduciary Activities	0	0.0%
NII: Services Charges on Deps	0	0.0%
NII: Tot Trading Revenue	79	0.3%
NII: Total Oth Svc Chgs & NII	26,757	99.7%
Total	**26,836**	**100.0%**

Table 4: American Express non-interest income: SEC 10-K report (FY 2012)

Category	$MM	% Total
Discount revenue	17,739	65.8%
Net card fees	2,506	9.3%
Travel commissions and fees	1,940	7.2%
Other commissions and fees	2,317	8.6%
Other	2,452	9.1%
Total	**26,954**	**100.0%**

Table 5: Bank of New York Mellon non-interest income: Y-9C breakdown (FY 2012)

Category	$MM	% Total
NII: Fiduciary Activities	8,062	70.2%
NII: Services Charges on Deps	434	3.8%
NII: Tot Trading Revenue	692	6.0%
NII: Total Oth Svc Chgs & NII	2,302	20.0%
Total	**11,490**	**100.0%**

Table 6: Bank of New York Mellon non-interest income: SEC 10-K report (FY 2012)

Category	$MM	% Total
Asset servicing	3,780	33.7%
Issuer services	1,052	9.4%
Clearing services	1,193	10.6%
Treasury services	549	4.9%
Investment management and performance fees	3,174	28.3%
Foreign exchange and other trading revenue	692	6.2%
Distribution and servicing	192	1.7%
Financing-related fees	172	1.5%
Investment and other income	427	3.8%
Total	**11,231**	**100.0%**

The examples highlight that even for an individual firm, the Y-9C categories combine a set of heterogeneous business lines. For more complex institutions, multiple line-items known to move in opposing directions in response to specific macroeconomic variables are all captured in common Y-9C categories. For example, mortgage prepayment behavior is known to differ materially in reaction to interest rate changes dependent on interest rates at which the mortgage was originated. Similarly, different product types – such as those with adjustable rather than fixed rates – also react in opposing directions to rate changes, but are often captured in common reporting lines. This problem is further compounded by the aggregation to the industry level, which further obscures important underlying heterogeneity.

The benefits to modeling of data granularity have been implicitly recognized by the Federal Reserve in the construction of its FR Y-14 forms, which form the basis of the quantitative CCAR submission (both historical and forecast data in the FR Y-14Q/M and -14A forms, respectively) and provide underlying data the Fed uses for its own CCAR models. The PPNR-related information in the Y-14 forms represents a significant departure from the Y-9C form. It certainly suggests that in order to effectively model revenues (even at the industry level without attempting to account for firm-specific nuance), the Federal Reserve has concluded that one must (a) align information and modeling along logical business segments, (b) disaggregate revenue information more finely than do Y-9C data, and (c) capture additional categories of information not contained in Y-9C data. Taking each of these in turn:

- The Y-14 forms request historical information for non-interest income to be broken down across 7 major business areas: Retail and Small Business, Commercial Lending, Investment

Banking, Merchant Banking / Private Equity, Sales and Trading, Investment Management, and Investment Services. This breakdown represents a significant departure from the Y-9C forms, which treat each bank as a consolidated enterprise and may not differentiate like revenue categories (e.g. "income from fiduciary activities") along the dimension of business area. The request for business area-specific information implicitly acknowledges that in order to effectively model revenues, modelers must approach a bank's revenue streams from a business segmentation perspective.

- Within each business segment, the FR Y-14 forms require a set of granular, business-specific information that exceeds the depth of information required in the Y-9C. As compared to the four major categories of non-interest income that the Y-9C uses to comprehensively cover an institution, the Y-14 forms request up to 15 separate line items for each business area listed above. These line items are tailored to the relevant sub-segments of each business area, and further imply that granular, business-specific information is needed to effectively model the fee revenue component of PPNR.
- Finally, the Y-14 forms request a set of qualitatively different information from the Y-9C: a set of "PPNR metrics" segmented by business area. Within our framework these metrics generally fall into the "volume" category, and represent additional specific volume information beyond simple asset balances. For example, within the Investment Management business area, the Y-14 forms request information on total assets under management (AuM), split by equities and fixed income, as well as net inflows and outflows. Similarly, total open credit card accounts and advisory deal volumes are requested for the Retail/Small Business and Investment Banking business areas, respectively. The presence of this information on the Y-14 forms strongly suggests the need for information beyond merely aggregate revenue and balance sheet information to effectively model PPNR, and further suggests that regulators are aware of the limitation in Y-9C PPNR data to at least some degree.

Taken together the results of our illustrative regression analysis, heterogeneity in other firm-specific regulatory disclosures, and information requested by the Fed to support its own CCAR evaluation clearly indicates that in order to develop meaningful PPNR model linked to macroeconomic factors, it is necessary to model different revenue streams at a granular level, and to do so using the reporting taxonomy the firm has developed over decades, which is likely different from the FR Y-9C reports.

To be sure, none of the call for granular data is an absolute guarantee of high-quality PPNR models. Revenue streams may be truly idiosyncratic, and data may be too “noisy” (e.g. influence by deliberate business strategies, M&A activity, etc.) to tease apart underlying macro trends. Alternatively certain business areas may truly be acyclical, and the search for macroeconomic drivers fruitless. But for many business areas our experience indicates– and appear to be implicitly supported by regulatory request for granular information in the CCAR context – that segmenting revenue streams according to a bank’s individual business areas and analyzing those streams at a more granular level will stand a good chance of highlighting the heterogeneous economic sensitivities of individual businesses.

4. Conclusion

CCAR-style macro-conditional stress testing is here to stay, and the revenue offset to credit losses will remain a critical part of an institution's capital adequacy assessment. As stress testing practices become a more routine part of bank risk management, we expect PPNR modeling practices to continue to mature towards disaggregated modeling approaches along the lines of the framework we propose. For now the degree of granularity is limited in part by data availability, but we see it as possible that CCAR will provide the motivation for institutions to upgrade their revenue-related data capture much the way Basel II provided this imperative for credit and operational risk.

PPNR modeling is not (yet) as sophisticated as loss modeling, but the industry has made significant leaps since the results of the 2012 CCAR were released. Revenue projections, such as it was, had been closely tied, if not identical to the budgeting process, and it had relied heavily on expert judgment and thus was rarely tied, even informally, to any specific state of the economy. It is this formalization of the revenue projection process and its linkage to the annual budget cycle that has been a significant cultural shift in banks.

References

Board of Governors of the Federal Reserve System, 2009a. "Agency Joint Stress Testing Program FAQs." Available at http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20090225a1.pdf

_____, 2009b "The Supervisory Capital Assessment Program: Design and Implementation." April 24, 2009. http://www.federalreserve.gov/bankinforeg/bcreg20090424a1.pdf.

_____, 2009c. "The Supervisory Capital Assessment Program: Overview of Results." May 7, 2009. http://www.federalreserve.gov/bankinforeg/bcreg20090507a1.pdf.

_____, 2012a. "2013 Supervisory Scenarios for Annual Stress Tests Required under the Dodd-Frank Act Stress Testing Rules and the Capital Plan Rule." November 9, 2012. Available at http://www.federalreserve.gov/bankinforeg/bcreg20121115a1.pdf.

_____, 2012b. "Comprehensive Capital Analysis and Review 2013 Summary Instructions and Guidance." November 9, 2012. Available at http://www.federalreserve.gov/newsevents/press/bcreg/bcreg20121109b1.pdf

_____, 2013a. "Dodd-Frank Act Stress Test 2013: Supervisory Stress Test Methodology and Results." March 7, 2013. http://www.federalreserve.gov/newsevents/press/bcreg/dfast_2013_results_20130314.pdf.

_____, 2013b. "Basel III Final Rule." Available at http://www.federalreserve.gov/bcreg20130702a.pdf.

Covas, Francisco, Ben Rump and Egon Zakrajsek, 2012, "Stress-Testing U.S. Bank Holding Companies: A Dynamic Panel Quantile Regression Approach." Available at http://www.frbsf.org/economic-research/events/2012/september/systemic-risk-forecast/zakrajsek.pdf.

Coffinet, Jérôme and Surong Lin, 2010, "Stress testing banks' profitability, the case of French banks." Banque de France Working Paper No. 306.

English, William B., 2002. "Interest Rate Risk and Bank Net Interest Margins." *BIS Quarterly Review*, December, 67-82.

Guerrieri, Luca and Michelle Welch, 2012, "Can macro-variables used in stress-testing forecast the performance of banks?" FEDS 2012-49.

Hirtle, Beverly, Til Schuermann and Kevin Stiroh, 2009. "Macroprudential Supervision of Financial Institutions: Lessons from the SCAP." Federal Reserve Bank of New York Staff Report No. 409.

Pritsker, Matthew, 2012," Enhanced Stress Testing and Financial Stability." Federal Reserve Bank of Boston Working Paper. Available at SSRN: http://ssrn.com/abstract=2082994.

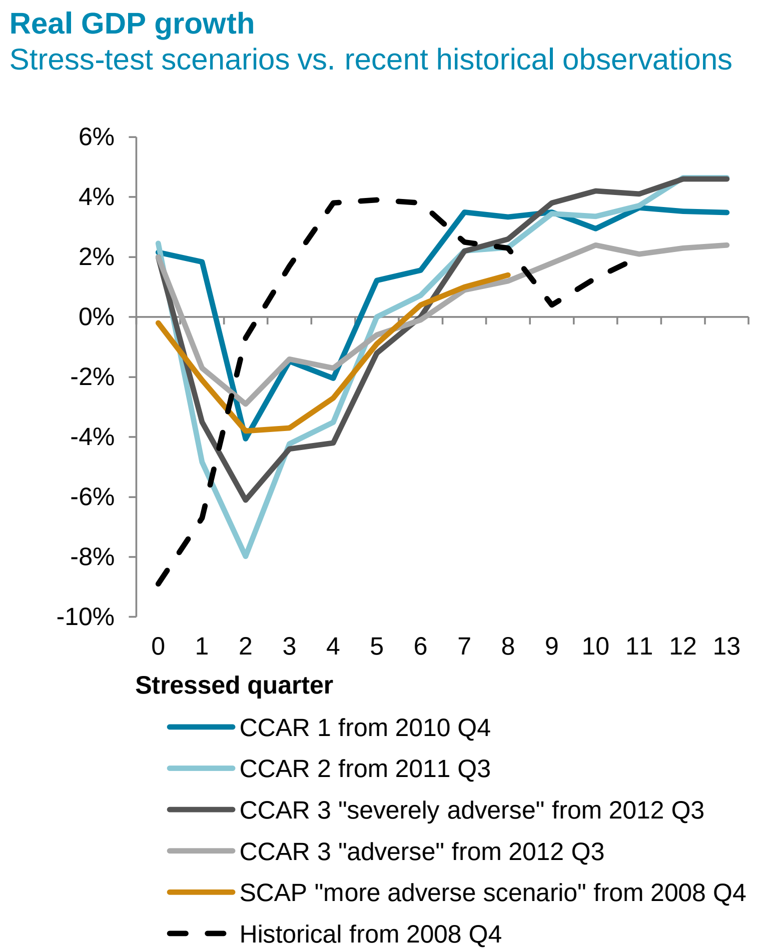
Purnanandam, Amiyatosh, 2007. "Interest Rate Risk Management at Commercial Banks: An Empirical Investigation." *Journal of Monetary Economics* 54, 1769-1808.

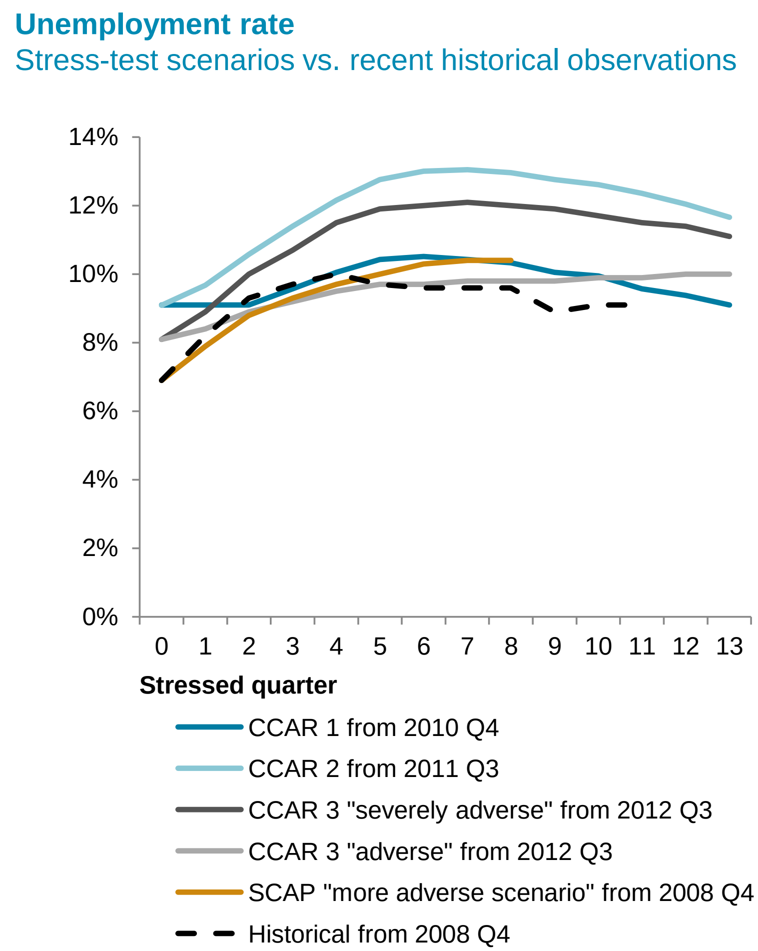
Schuermann, Til, 2013, "Stress Testing Banks." Forthcoming, *International Journal of Forecasting*. A working paper is available as Wharton Financial Institutions Center Working Paper 12-08.

Stiroh, Kevin, 2004, “Diversification in Banking: Is Noninterest Income the Answer?” *Journal of Money, Credit and Banking* 36:5, 853-882.

_____, 2006, “A Portfolio View of Banking with Interest and Noninterest Activities”, *Journal of Money, Credit, and Banking* 38:5, 1351-1361.

Timmermann, Allan, 2006, “Forecast Combinations. “In Elliott, G., C.W.J. Granger and A. Timmermann (Eds.), *Handbook of Economic Forecasting* (pp. 135-196). Amsterdam: North Holland.

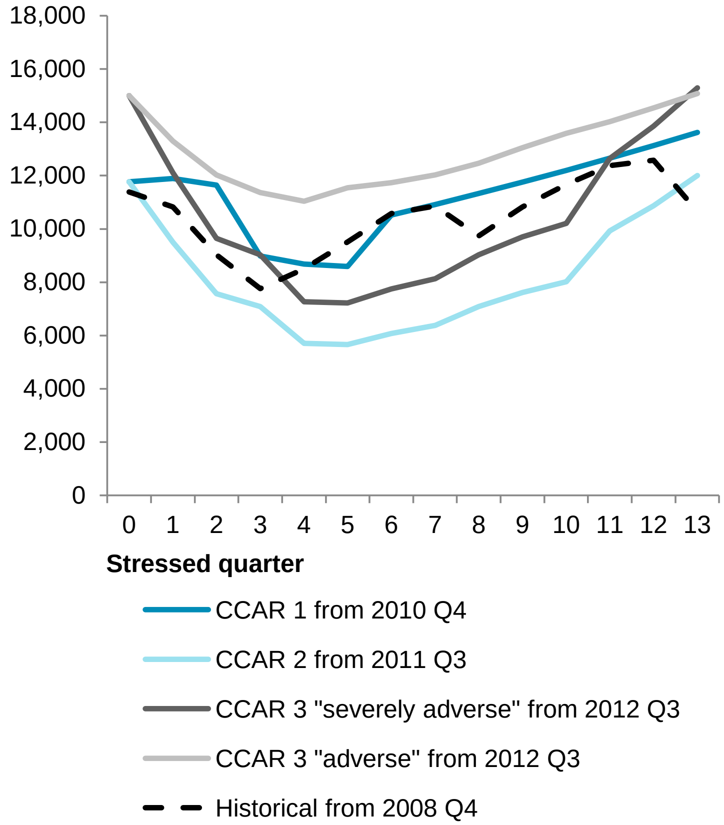




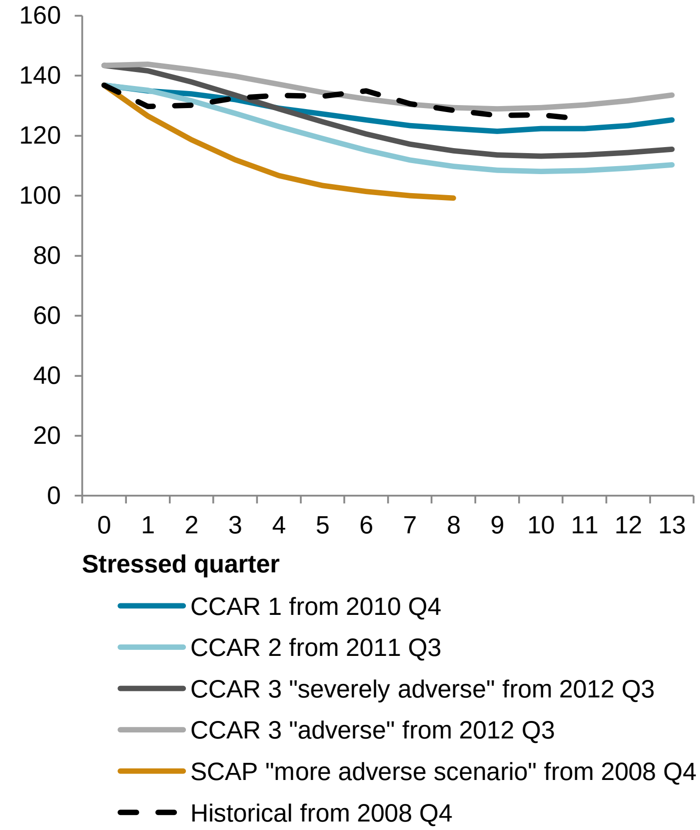
Source: Fed, The Supervisory Capital Assessment Program: Design and Implementation, 24 April 2009; Fed, Comprehensive Capital Analysis and Review: Objectives and Overview, 18 March, 2011; Fed, “Comprehensive Capital Review” document and “Capital Plan review” 22 November 2011; Fed, “2013 Supervisory Scenarios” 15 November 2012; Datastream

Figure 3: U.S. real GDP and unemployment scenarios compared

Dow Jones total stock market index level
Stress-test scenarios vs. recent historical observations



House Price index
Stress-test scenarios vs. recent historical observations



Source: Fed, The Supervisory Capital Assessment Program: Design and Implementation, 24 April 2009; Fed, Comprehensive Capital Analysis and Review: Objectives and Overview, 18 March, 2011; Fed, “Comprehensive Capital Review” document and “Capital Plan review” 22 November 2011; Fed, “2013 Supervisory Scenarios” 15 November 2012; Datastream

Figure 4: U.S. equity and house price indices compared