Electronic Companion to

Should Bank Stress Tests Be Fair?

Paul Glasserman and Mike Li
Columbia Business School

This Electronic Companion covers the following topics. Section EC.1 examines how changes in bank-specific parameters affect parameters in an aggregated model. We show that in many cases these parameter externalities favor FEO over a pooled model; Section EC.2 provides supporting analysis. Section EC.3 argues that the only parameter aggregation rules satisfying some simple conditions are convex combinations, again supporting the FEO model over the pooled model.

Sections EC.4– EC.7 investigate the empirical relevance of our theoretical results. Section EC.4 regresses the loss rates of loan portfolios (credit cards, first lien mortgages, commercial real estate, and commercial and industrial loans) on measures of portfolio quality (past-due rates and allowances for losses) and macroeconomic variables, where significant heterogeneity across banks in their estimated coefficients is documented. Section EC.5 extends the empirical results to nonlinear models. Section EC.6 provides some additional information on the data used in Sections EC.4 and EC.5. Section EC.7 discusses bank heterogeneity in revenue models.

We emphasize that our empirical investigation is limited by the information banks make public — the Federal Reserve has access to far more granular data in estimating models and forecasting losses. The Fed's models take a bottom-up approach to forecasting losses, based on detailed portfolio characteristics; we are limited to a top-down approach using limited accounting data and macro variables. We cannot claim to approximate the Fed's forecasts; our goal is to provide evidence of the potential importance of heterogeneity.

EC.1 Cross-Bank Parameter Externalities

As a consequence of aggregating bank-specific results into a single industry model, changes at one bank can affect loss forecasts at other banks, and the results are sometimes counterintuitive. In this section, we argue that these cross-bank externalities are generally more reasonable under FEO forecasts than under the pooled method.

For simplicity, we consider a setting with a single scalar feature x. More generally, we can think of this as a feature that is uncorrelated with all other features. We adopt the convention that this feature is nonnegative, and that higher values of x are associated with higher losses. Thus, for each bank s we assume $\mu_s \geq 0$ and $\beta_s \geq 0$. In reducing μ_s, a bank improves its portfolio quality; in reducing β_s, a bank improves its ability to manage portfolio risk; and in reducing α_s, a bank improves unobserved features to reduces its losses. We examine how these improvements — reductions in μ_s, α_s, and β_s — affect stress test results for bank s and other banks l.

We can write the FEO loss forecast (24) for bank l evaluated at $X_l = x$ as

$$\hat{Y}_{F,l}(x) = \hat{Y}_F(x) = \sum_s p_s(\alpha_s + \beta_s\mu_s) + \beta_F(x - \bar{\mu}), \tag{58}$$

with $\beta_F = \sum_i p_i\sigma_i^2\beta_i / \sum_i p_i\sigma_i^2$, as in (27). The forecast is the same for all banks l because FEO satisfies equal treatment. It is now easy to see that

$$\frac{\partial \hat{Y}_F(x)}{\partial \mu_s} = p_s\beta_s - p_s\beta_F \geq 0, \quad \text{if and only if } \beta_s \geq \beta_F; \tag{59}$$

$$\frac{\partial \hat{Y}_F(x)}{\partial \alpha_s} = p_s \geq 0; \tag{60}$$

and

$$\frac{\partial \hat{Y}_F(x)}{\partial \beta_s} = p_s\mu_s + (x - \bar{\mu})p_s\sigma_s^2 / \sum_i p_i\sigma_i^2 \geq 0, \quad \text{if } x > \bar{\mu}. \tag{61}$$

In (59) we see that if bank s has above-average (relative to β_F) sensitivity to feature x, then reducing its average exposure to that feature μ_s reduces loss forecasts for all banks. Equation (60) shows a similar overall benefit if bank s improves on the other dimensions captured by α_s. In (59) we see that an improvement in risk management at bank s, corresponding to a reduction in β_s, reduces loss forecasts at above-average levels of x. If x is part of the stress scenario, then large values of x are particularly relevant.

The directional effects in (59)–(61) are fairly simple and reasonable, considering that cross-bank effects are inevitable in an industry model. If the industry improves its performance (perhaps because of improvements at one bank) we generally expect loss forecasts to decrease. (A decrease in a forecast corresponds to a positive derivative because we are considering a decrease μ_s, α_s, or β_s.) Counterparts to (59)–(61) continue to hold if we replace β_F in (58) with any convex combination of the β_s, as in the WATE model. However, the pooled method behaves quite differently.

The pooled forecast $\hat{Y}_P(x)$ can be written in the same form as (58) but with β_F replaced by β_{Pool} in (11). We now get

$$\frac{\partial \hat{Y}_P(x)}{\partial \mu_s} = p_s(\beta_s - \beta_{Pool}) + (x - \bar{\mu})\frac{\partial \beta_{Pool}}{\partial \mu_s}.$$

The sign of the last term is not determined by a simple condition, so the overall directional effect is difficult to predict. The sign of

$$\frac{\partial \hat{Y}_P(x)}{\partial \alpha_s} = p_s + p_s \frac{(\mu_s - \bar{\mu})(x - \bar{\mu})}{\sum_s p_s \sigma_s^2 + \mathsf{var}(\mu_S)} \beta_s,$$

depends on the magnitudes of μ_s and x, relative to $\bar{\mu}$. For the sensitivity to β_s, we can write

$$\frac{\partial \hat{Y}_P(x)}{\partial \beta_s} = p_s \mu_s + (x - \bar{\mu}) \frac{\partial \beta_{Pool}}{\partial \beta_s}, \quad \frac{\partial \beta_{Pool}}{\partial \beta_s} = \frac{p_s(\sigma_s^2 + \mu_s(\mu_s - \bar{\mu}))}{\sum_i p_i \sigma_i^2 + \mathsf{var}(\mu_S)}.$$

Among the most troubling aspects of the pooled model is that the last term could be negative: a reduction in β_s could produce an increase in β_{Pool}. In particular, $\sigma_s^2 + \mu_s(\mu_s - \bar{\mu})$ is negative for a bank with below-average exposure to feature x (so $\mu_s < \bar{\mu}$) and low variability σ_s^2 in this exposure. Under the pooled model, it is therefore possible for an improvement in risk management at one bank (a reduction in β_s) to produce an *increase* in loss forecasts at all banks.

The top panel of Table EC.1.1 shows sufficient conditions for positive sensitivities of $\hat{Y}_F(x)$ and $\hat{Y}_P(x)$. The middle and bottom panels show corresponding results for the expected forecasts $\mathsf{E}[\hat{Y}_l] = \mathsf{E}[\hat{Y}(X_l)]$ and for the bias $\mathsf{E}[\hat{Y}(X_l) - Y_l]$. Supporting details for the second and third cases are provided in Section EC.2. We have tried to provide simple sufficient conditions, and in most cases the conditions are not necessary. All of the conditions for FEO extend to WATE with β_F replaced by the weighted average coefficient.

Some counterintuitive and undesirable cases can arise at empirically plausible parameter values. For example, in equation (64) we derive an expression for $\partial \mathsf{E}[\hat{Y}_P(X_l)]/\partial \alpha_s$. Using estimated parameters for the credit card data in Section EC.4, we find that this derivative is negative when l is Citigroup and s is JPMorgan Chase. In other words, an improvement at JPMorgan Chase would result in a higher expected loss forecast at Citigroup under the pooled model.

The bias sensitivities in Table EC.1.1 are more complicated than the other cases because the bias involves the difference between the predicted and actual loss rates. A reduction in the predicted loss rate can increase or decrease bias, depending on whether the initial forecast is too low or too high.

EC.2 Sensitivity Analysis

This section provides supporting details for Section EC.1, particularly the conclusions summarized in the middle and bottom panels of Table EC.1.1. We begin with an analysis of forecast bias that is of independent interest.

$\hat{Y}(x)$	FEO	Pool
$\mu_s \downarrow$	$\downarrow$ iff $\beta_s > \beta_F$	no simple rule
$\alpha_s \downarrow$	$\downarrow$	$\downarrow$ if $(\mu_s - \bar{\mu})(x - \bar{\mu}) > 0$
$\beta_s \downarrow$	$\downarrow$ if $x > \bar{\mu}$	$\downarrow$ if $[\sigma_s^2 + \mu_s(\mu_s - \bar{\mu})](x - \bar{\mu}) > 0$
$\mathsf{E}[\hat{Y}(X_l)]$		
$\mu_s \downarrow$	$l = s$: $\downarrow$ $l \neq s$: $\downarrow$ iff $\beta_s > \beta_F$	no simple rule
$\alpha_s \downarrow$	$\downarrow$	$l = s$: $\downarrow$ $l \neq s$: $\downarrow$ if $(\mu_s - \bar{\mu})(\mu_l - \bar{\mu}) > 0$
$\beta_s \downarrow$	$\downarrow$ if $\mu_s + \mu_l > \bar{\mu}$	$\downarrow$ if $[\sigma_s^2 + \mu_s(\mu_s - \bar{\mu})](\mu_l - \bar{\mu}) > 0$ or if μ_s sufficiently large
$\mathsf{bias}(l)$		
$\mu_s \downarrow$	$l = s$: $\downarrow$ iff $\beta_s < \beta_F$ $l \neq s$: $\downarrow$ iff $\beta_s > \beta_F$	no simple rule
$\alpha_s \downarrow$	$l = s$: $\uparrow$ $l \neq s$: $\downarrow$	$l = s$: no simple rule $l \neq s$: $\downarrow$ if $(\mu_s - \bar{\mu})(\mu_l - \bar{\mu}) > 0$
$\beta_s \downarrow$	$l = s$: $\uparrow$ if $\mu_s < \bar{\mu}$ $l \neq s$: $\downarrow$ if $\mu_s + \mu_l > \bar{\mu}$	no simple rule $\downarrow$ if $[\sigma_s^2 + \mu_s(\mu_s - \bar{\mu})](\mu_l - \bar{\mu}) > 0$

Table EC.1.1: Sensitivity of results for bank l in response to a decrease in parameter μ_s, α_s, or β_s for bank s. Sensitivities shown are for predicted loss $\hat{Y}_l(x)$ (top), mean predicted loss $\mathsf{E}[\hat{Y}(X_l)]$ (middle), and the bias $\mathsf{E}[\hat{Y}(X_l) - Y_l]$.

EC.2.1 Forecast Bias

If losses at different banks are described by different models, then forecast bias becomes inevitable when we apply a single model to all banks. But the distribution of bias across banks may differ under different choices of the single model.

Let $\hat{Y}_s$ be any of the forecasts for bank s in Table 4.1, and, as in (1), let Y_s denote the actual loss rate for bank s. Both $\hat{Y}_s$ and Y_s are evaluated at X_s. Define the forecast bias for bank s to be

$$\mathsf{bias}(s) = \mathsf{E}[\hat{Y}_s - Y_s]. \tag{62}$$

The expectation integrates over the distribution of the error ϵ_s in (1) and the features X_s.

Proposition EC.2.1. *For each forecast in Table 4.1, the bias is as follows.*

(i) Pooled: $\mathsf{bias}(s) = \mathsf{E}[Y_S] - \mathsf{E}[Y_s] + \beta_{Pool}^\top(\mu_s - \bar{\mu})$;

(ii) PTF in (18): $\mathsf{bias}(s) = \mathsf{E}[Y_S] - \mathsf{E}[Y_s]$;

(iii) Conditional expectation: $\mathsf{bias}(s) = \mathsf{E}[\hat{Y}_C(X_s)] - \mathsf{E}[Y_s]$;

(iv) FEO: $\mathsf{bias}(s) = \mathsf{E}[Y_S] - \mathsf{E}[Y_s] + \beta_F^\top(\mu_s - \bar{\mu})$;

(v) *SEO:* $\text{bias}(s) = \mathsf{E}[Y_S] - \mathsf{E}[Y_s]$.

Proof. For (i), we have, using the definition of α_{Pool} in (12),

$$\begin{aligned} \mathsf{E}[\hat{Y}_s - Y_s] &= \mathsf{E}[\alpha_{Pool} + \beta_{Pool}^\top X_s - Y_s] \\ &= (\mathsf{E}[Y_S] - \beta_{Pool}^\top \bar{\mu}) + \beta_{Pool}^\top \mu_s - \mathsf{E}[Y_s] \\ &= \mathsf{E}[Y_S] - \mathsf{E}[Y_s] + \beta_{Pool}^\top (\mu_s - \bar{\mu}). \end{aligned}$$

For the PTF forecast, (17) and (18) yield

$$\mathsf{E}[\hat{Y}_s] = \bar{\alpha}^o = \sum_s p_s \alpha_s^o = \sum_s p_s \mathsf{E}[Y_s] = \mathsf{E}[Y_S],$$

and the bias in (ii) follows. The expression in (iii) holds by definition. The argument for (iv) is the same as the argument for (i). The bias in (v) follows from (iv) because we see from (38) that the SEO forecast for bank s subtracts $\beta_F^\top(\mu_s - \bar{\mu})$ from the FEO forecast. □

In every case of Proposition EC.2.1, the average bias $\sum_s p_s \text{bias}(s)$ is zero, but the methods differ in how they distribute bias across banks. We saw previously that the PTF and SEO methods go the farthest in equalizing differences; we now see that the bias for each of these methods is the difference $\mathsf{E}[Y_S] - \mathsf{E}[Y_s]$ between the average loss rate for all banks and the average for an individual bank.

Using the relationship $\beta_{Pool} = \beta_F + \Lambda\delta$ from (30), we see that the difference between the expressions in (i) and (iv) is

$$\text{bias}_{Pool}(s) - \text{bias}_{FEO}(s) = \delta^\top \Lambda^\top (\mu_s - \bar{\mu}).$$

In light of the discussion in Section 4.2, this difference is the expected disparate impact on bank s of using the pooled model.

EC.2.2 Improvement in Intercept α_s

By taking the expectation of (58), we get

$$\mathsf{E}[\hat{Y}_F(X_l)] = \sum_s ps(\alpha_s + \beta_s \mu_s) + \beta_F(\mu_l - \bar{\mu}), \tag{63}$$

and the same holds for the expected pooled forecast with β_F replaced by β_{Pool}. It follows that, for any banks s and l,

$$\frac{\partial \mathsf{E}[\hat{Y}_F(X_l)]}{\partial \alpha_s} = p_s > 0.$$

In other words, all expected forecasts decrease following a reduction in α_s.

In contrast, for the pooled model we get

$$\frac{\partial \mathsf{E}[\hat{Y}_P(X_l)]}{\partial \alpha_s} = p_s - \frac{\partial \mathsf{cov}(\alpha_S, \mu_S)/\partial \alpha_s}{\sum_t p_t \sigma_t^2 + \mathsf{var}(\mu_S)} \beta_s(\bar{\mu} - \mu_l) = p_s + p_s \frac{(\mu_s - \bar{\mu})(\mu_l - \bar{\mu})}{\sum_s p_s \sigma_s^2 + \mathsf{var}(\mu_S)} \beta_s. \tag{64}$$

Bank s benefits from its reduction of α_s, in the sense that the derivative with $l = s$ is positive. For $l \neq s$, the sign of (64) does not admit a simple description. In particular, it may be negative when μ_s and μ_l are on opposite sides of $\bar{\mu}$, meaning that one bank's loans are riskier than average and the other bank's loans are less risky than average.

For the bias under FEO we have

$$\frac{\partial \mathsf{bias}_F(l)}{\partial \alpha_s} = p_s - \mathbf{1}\{l = s\}$$

It is then immediate that

$$\frac{\partial \mathsf{bias}_F(l)}{\partial \alpha_s} > 0 \text{ if } l \neq s \quad \text{and} \quad \frac{\partial \mathsf{bias}_F(s)}{\partial \alpha_s} < 0.$$

The direction of change makes sense. If the bias for a bank is positive, meaning that the industry model overestimates its losses, then improvements at other banks will reduce loss forecasts and thus reduce the bias. The bank's own improvements will increase the bias by reducing the bank's own losses by more than they reduce the model's forecasts. The situation is reversed for a bank with a negative bias.

However, for the pooled regression method,

$$\frac{\partial \mathsf{bias}_P(l)}{\partial \alpha_s} = p_s + p_s \frac{(\mu_s - \bar{\mu})(\mu_l - \bar{\mu})}{\sum_s p_s \sigma_s^2 + \mathsf{var}(\mu_S)} \beta_s - \mathbf{1}\{l = s\},$$

and the direction of change is unclear.

EC.2.3 Improvement in Loan Quality

Now suppose bank s improves the quality of its loan portfolio, resulting in a smaller μ_s. This has no effect on β_F, which makes sense — changing one bank's loan quality should not change the sensitivity of losses to loan quality. However, it is evident from (11) that β_{Pool} does change with μ_s.

Under FEO, the mean the mean predicted loss rate satisfies

$$\frac{\partial \mathsf{E}\hat{Y}_F(X_l)}{\partial \mu_s} = p_s \beta_s + \beta_F(\mathbf{1}\{l = s\} - p_s),$$

which is always positive if $l = s$. This means that an improvement in bank l's loan quality (a reduction in μ_l) reduces bank l's mean predicted losses. In the pooled model,

$$\frac{\partial \mathsf{E}\hat{Y}_P(X_l)}{\partial \mu_s} = p_s \beta_s + \beta_{Pool}(\mathbf{1}\{l = s\} - p_s) + (\mu_s - \bar{\mu}) \frac{\partial \beta_{Pool}}{\partial \mu_s};$$

this expression could be negative, even with $l = s$, meaning that a bank could be penalized (through a higher mean predicted loss rate) as a result of improving its loan quality.

The sensitivity of the bias under FEO is given by

$$\frac{\partial \mathsf{bias}_F(l)}{\partial \mu_s} = (\mathbf{1}\{l = s\} - p_s)(\beta_F - \beta_s);$$

in particular, the bias for bank l moves in opposite directions with respect to changes in μ_l and μ_s, $s \neq l$. Suppose industry model overestimates bank l's losses, in the sense that the bias is positive, and suppose the industry model overestimates bank l's sensitivity to loan quality, in the sense that $\beta_F > \beta_l$. Then bank l will benefit (in the sense of reducing the bias) from improving its loan quality by reducing μ_l.

For the pooled regression,

$$\frac{\partial \mathsf{bias}_P(l)}{\partial \mu_s} = (\beta_{Pool} - \beta_s)(\mathbf{1}\{l = s\} - p_s) + (\mu_s - \bar{\mu})\frac{\partial \beta_{Pool}}{\partial \mu_s}.$$

The sign of this expression does not admit a simple condition.

EC.2.4 Improvement in Loan Management

Now suppose bank s improves its abilities in loan management, resulting in a reduction in β_s. The mean predicted loss rate under FEO satisfies

$$\frac{\partial \mathsf{E}[\hat{Y}_F(X_l)]}{\partial \beta_s} = p_s(\mu_s + \mu_l - \bar{\mu}),$$

and is positive if $\mu_s + \mu_l > \bar{\mu}$. In the pooled model

$$\frac{\partial \mathsf{E}[\hat{Y}_P(X_l)]}{\partial \beta_s} = p_s\mu_s + \frac{p_s(\sigma_s^2 + \mu_s(\mu_s - \bar{\mu}))}{\sum_i p_i(\sigma_i^2 + \mu_i(\mu_i - \bar{\mu}))}(\mu_l - \bar{\mu}),$$

so $[\sigma_s^2 + \mu_s(\mu_s - \bar{\mu})](\mu_l - \bar{\mu}) > 0$ is a sufficient condition for the sensitivity to be positive. Regardless of the value of μ_l, the sensitivity is positive for all sufficiently large μ_s.

For $l \neq s$, the sensitivity of the bias for bank l with respect to β_s equals the sensitivity of the mean predicted loss because the actual expected loss $\mathsf{E}[Y_l]$ is unaffected by β_s. We therefore focus on the case $l = s$. Under FEO,

$$\frac{\partial \mathsf{bias}_F(s)}{\partial \beta_s} = (p_s - 1)\mu_s + p_s(\mu_s - \bar{\mu}),$$

which is guaranteed to be negative if $\mu_s < \bar{\mu}$. Under the pooled model, the sign of

$$\frac{\partial \mathsf{bias}_P(s)}{\partial \beta_s} = (p_s - 1)\mu_s + \frac{p_s(\sigma_s^2 + \mu_s(\mu_s - \bar{\mu}))}{\sum_i p_i(\sigma_i^2 + \mu_i(\mu_i - \bar{\mu}))}(\mu_s - \bar{\mu})$$

does not admit a simple characterization.

EC.3 Convex Combinations of Coefficients

Equation (13) aggregates the individual scalar slopes β_s into a single value. We can generalize this perspective and ask what properties we would like in an aggregation function, meaning a function $f : \mathbb{R}^{\bar{S}} \to \mathbb{R}$,

$$\beta_* = f(\beta_1, \ldots, \beta_{\bar{S}}),$$

that combines bank-specific coefficients β_s into an "industry" parameter β_*.

We consider the following properties:

(i) $f(kb_1, \ldots, kb_{\bar{S}}) = kf(b_1, \ldots, b_{\bar{S}})$, for all k, $b_1, \ldots, b_{\bar{S}} \in \mathbb{R}$;

(ii) $f(b, \ldots, b) = b$, for at least one nonzero $b \in \mathbb{R}$;

(iii) $b_s > 0$, for all s, implies $f(b_1, \ldots, b_{\bar{S}}) \geq 0$;

(iv) f is differentiable at zero.

Property (i) is needed for the aggregation to perform sensibly under a change of units in the measurement of X_s: if we divide each X_s by k, each β_s increases by a factor of k, and it is natural to require that β_* scale accordingly. Properties (ii) and (iii) are also very modest requirements. Property (iv) is harder to motivate but not unreasonable. These properties constrain the aggregation function as follows:

Proposition EC.3.1. *If (i)–(iv) hold, then $f(\beta_1, \ldots, \beta_{\bar{S}})$ is a convex combination of its arguments.*

Proof. Fix $\beta \in \mathbb{R}^{\bar{S}}$. Let $g(t) = f(t\beta)$. By condition (iv), $g'(0) = \beta^\top f'(0)$. Condition (i) and (ii) imply $g(t) = tf(\beta)$, so $g'(t) = f(\beta)$ for any t. Thus, $f(\beta) = g'(0) = \beta^\top f'(0) = \sum_{i=1}^{\bar{S}} f_i'(0)\beta_i$. Condition (ii) now implies $\sum_{i=1}^{\bar{S}} f_i'(0) = 1$, and condition (iii) implies $f_i'(0) \geq 0$, for all i. Thus, $f(\beta) = \sum_{i=1}^{\bar{S}} f_i'(0)\beta_i$ is a convex combination of the components of β. □

The scalar FEO coefficient in (27) is a convex combination of the bank-specific coefficients β_s, but the pooled coefficient (13) is generally not. This property of the FEO model extends to the multivariate case under additional conditions. If all the bank-specific covariance matrices Σ_s, $s = 1, \ldots, \bar{S}$, coincide, then in (25) we get $\beta_F = \mathsf{E}[\beta_S] = \sum_s p_s \beta_s$. If all Σ_s are diagonal (but not necessarily identical), then the representation of the scalar FEO coefficient in (27) applies to each coordinate of β_F. If all Σ_s have the same eigenvectors, then we can transform the original features X_s into uncorrelated features using principal components. Using these transformed features, each coordinate of β_F is a convex combination of bank-specific coefficients.

EC.4 Empirical Evidence

In this section, we document empirical evidence of heterogeneity in bank-specific models of loss rates, and we examine the implications of this heterogeneity for the choice of an industry-wide model. We find strong evidence of statistically significant differences in model parameters across banks. These differences can lead to material differences between pooled and FEO coefficients in an industry model.

We must emphasize, however, that our investigation is constrained by the very limited information made publicly available by banks about the risk characteristics and losses in their loan portfolios. The Federal Reserve has far more granular information about banks' loans and losses. Our results can therefore provide only a rough indication of the impact of bank heterogeneity in the Fed's stress tests.

EC.4.1 Data

We use two types of data and data sources: historical macroeconomic data and loan information for individual banks.

EC.4.1.1 Macroeconomic Data

We use data on seven of the macro variables used in the Federal Reserve's stress tests: real disposable income growth, real GDP growth, house price index level, inflation rate, unemployment rate, Dow Jones total stock index level, and the Treasury spread. The Federal Reserve provides historical data on its website for all variables used in forming stress scenarios, including these. We use the values reported by the Fed for these variables in the June 2020 stress test; these values run from 1990 through 2019.

We aggregate these variables into a single macro variable by taking the first principal component of their correlation matrix. Table EC.4.1 shows the corresponding loadings. We see that an increase in the principal component corresponds to decreases in income growth and GDP growth and an increase in unemployment, suggesting that this composite variable serves as a reasonable measure of overall economic conditions. Figure EC.4.1 plots the level of this variable over time and shows a sharp climb around 2008 and 2020.[5]

[5] The loadings in Table EC.4.1 are calculated using data through 2019, and we use these loadings to extend PC1 through the end of 2021. When we include the COVID period in the calculation of the principal components, PC1 becomes harder to interpret. For example, the coefficients for income growth and unemployment have the same sign.

Macro Factor	PC1 Loading
Real disposable income growth	-0.229
Real GDP growth	-0.525
Change House Price Index	-0.467
CPI inflation rate	-0.079
Change unemployment	0.529
Change Dow	-0.293
Change Treasury Spread	0.287

Table EC.4.1: Loadings of first principal component on macro variables.

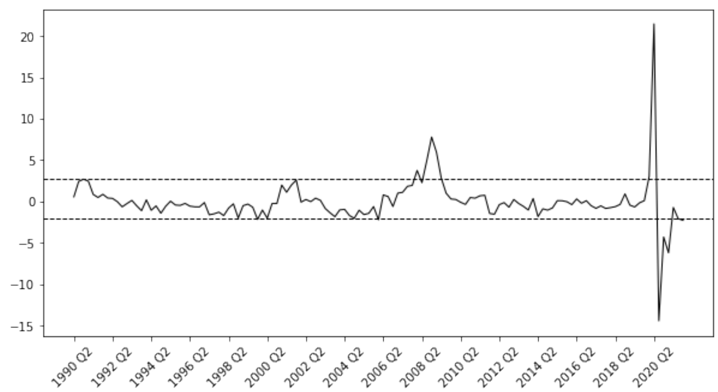

Figure EC.4.1: First principal component (PC1) of macro variables from 1990 Q2 to 2021 Q4. The dashed lines correspond to the 5th and 95th percentiles of PC1.

EC.4.1.2 Loan Information for Individual Banks.

Bank holding companies publicly report financial information quarterly through the Federal Reserve's form Y-9C. We use these filings to collect information on four loan types that are treated separately in the Fed's stress tests: credit cards, first lien mortgages, commercial real estate loans, and commercial and industrial loans. For each category, each bank, and each quarter, we collect loan balances, charge-offs, recoveries, and total amounts past due serving as our proxy measures of loan portfolio risk.

We collect this data from 2001 to 2021 for the thirty-five largest banks by total assets (as of December 2021). The banks are listed in Table EC.6.1. The stress test focuses on adverse economic conditions; we weight each observation by the level of stress in each quarter and each bank's load balance: for each bank s, each quarter t, and loan category p, we weight the observations by

$$w_{s,t}^{p} = e^{\lambda MacroPC_t} \times Loan_{s,t}^{p}, \tag{65}$$

where $Loan_{s,t}^{p}$ is the size of the loan portfolio of type p for bank s in quarter t. We choose λ so that for the same loan level, the worst economic quarter (as measured by $MacroPC_t$) is given twice the weight as the best quarter.

We would prefer to conduct our analysis using data from stress periods only, but that would leave us with too few observations. Weighting by the level of stress in a quarter allows us to approximate the effect of conditioning on stress while making greater use of the available data. This approach relies on the assumption that data from non-stressful periods is relevant to forecasting losses in periods of stress.

We merger-adjust all bank data. For example, Truist Financial, one of the banks in Table EC.6.1, was formed from the 2019 merger of BB&T and SunTrust, so our data for Truist in earlier years combines data from those two banks. We repeat this process as we work backwards in time. We obtain information on mergers and acquisitions from the Federal Financial Institutions Examination Council website. (We have also run our analysis without merger-adjusting the data; doing so does not change our conclusions and generally increases heterogeneity across banks.)

In each loan category, we calculate a loss rate (net charge-off rate) for each bank s and each quarter t as the ratio

$$LossRate_{s,t} = \frac{Charge\text{-}offs_{s,t} - Recoveries_{s,t}}{Total\ Loans\ in\ Category_{s,t-1}}. \tag{66}$$

This measure is commonly used in stress testing; see, for example, Guerrieri and Welch [3], Hirtle et al. [6], and Kapinos and Mitnik [7]. We similarly normalize the amounts past due to get a $PastDueRate_{s,t}$ for each bank-quarter. We remove values less than -50% or greater than 50% of *LossRate* and values greater than 20% of *PastDueRate*. We winsorize *PastDueRate* at the upper and lower 5% levels. To attain a mostly balanced panel for more reliable estimates, in each loan category we include only banks with at least 18 years (72 quarters) of history from 2001 Q1 to 2021 Q4.

Table EC.4.2 shows descriptive statistics for these variables. Loss rates and past due rates are shown by loan category — credit cards (CC), first liens (FL), commercial real estate (CRE), and commercial and industrial (CI). Columns 2–4 of the table summarize time-averaged values across banks. Columns 5–8 summarize observations across all banks and quarters.

	bank averages			all observations			
	min	mean	max	lower 5%	mean	upper 5%	std
Loss Rate: CC	-0.20	2.50	3.36	0.31	3.00	5.88	2.01
Loss Rate: FL	0.01	0.22	0.66	-0.01	0.27	1.03	0.51
Loss Rate: CRE	0.06	0.19	0.46	-0.04	0.16	1.00	0.41
Loss Rate: CI	0.05	0.41	1.00	0.00	0.42	1.56	0.52
Past Due Rate: CC	1.10	2.90	4.23	1.06	3.30	5.74	1.53
Past Due Rate: FL	0.71	4.01	8.02	0.45	6.20	11.80	4.60
Past Due Rate: CRE	1.07	2.15	3.80	0.36	2.12	6.52	1.88
Past Due Rate: CI	0.05	1.65	2.85	0.05	1.74	3.97	1.23

Table EC.4.2: Descriptive statistics in percent. Columns 2–4 are calculated from banks' time averages, and columns 5–8 are calculated from all observations, with mean and standard deviation are stressed time and loan balance weighted.

Figure EC.4.2 plots the mean past due rate (±1.96 standard errors) for each bank in each loan category. The banks are identified by their stock tickers. The figure illustrates substantial heterogeneity across banks in their loan portfolios. For example, Bank of America (BAC) has among the highest past due rates for credit card loans, but in the commercial real estate category it has among the lowest. This type of pattern is consistent with the idea that banks have different areas of specialization and may target different markets.

The widths of the bars in Figure EC.4.2 show differences across loan categories and banks in the volatility of their past due rates. We again observe significant heterogeneity among different banks.

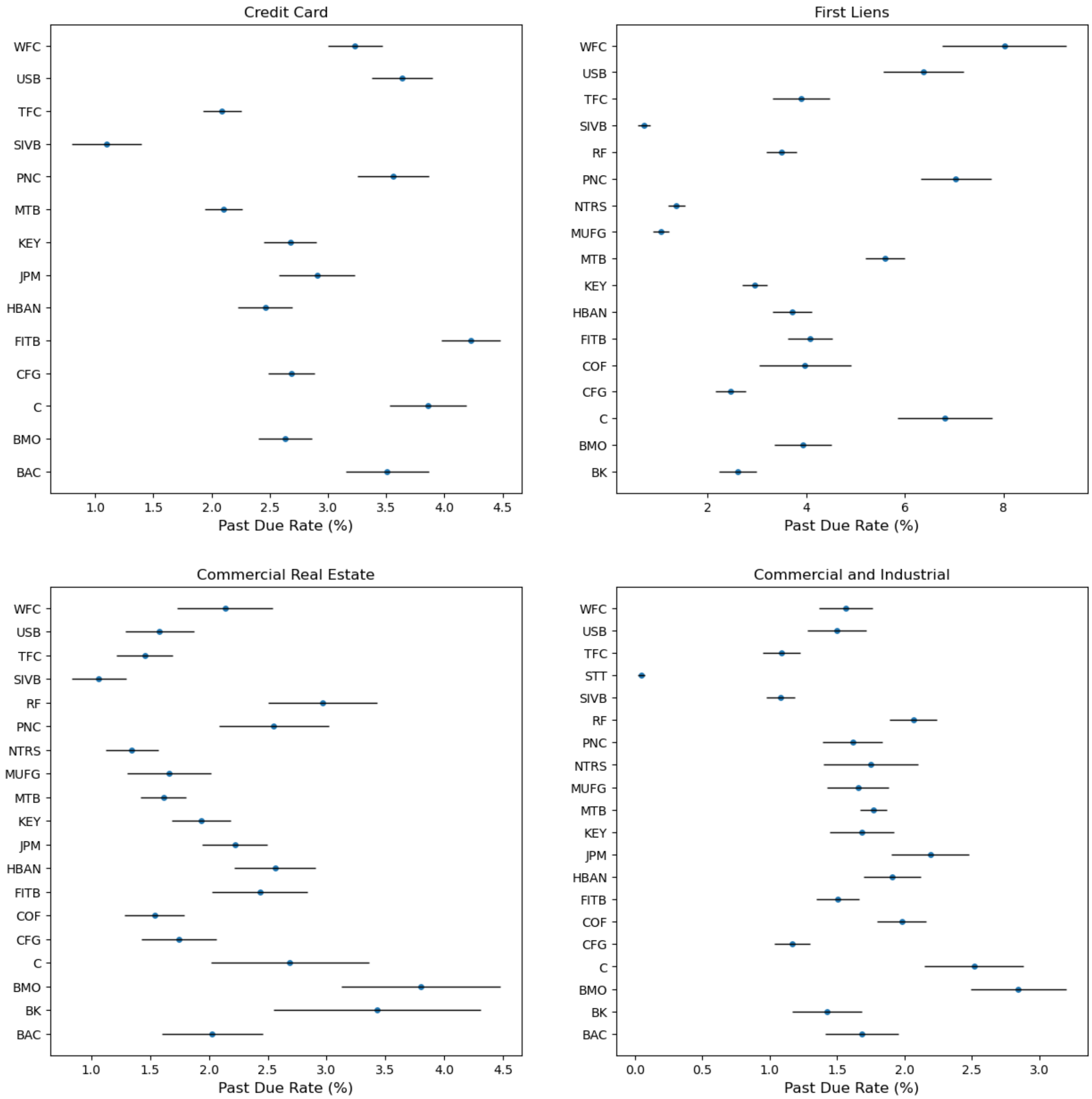

Figure EC.4.2: Past due rates (winsorized) by bank and loan category. The dots show mean values and each horizontal bar corresponds to ±1.96 standard errors.

EC.4.2 Heterogeneity in Slopes and Intercepts

We use the bank data to approximate our theoretical framework through the specification

$$LossRate_{s,t} = \alpha_s + \beta_s PastDueRate_{s,t-l} + \gamma_s MacroPC_{t-l} + \epsilon_{s,t}, \tag{67}$$

for bank s in quarter t, where $MacroPC$ is the principal component of the macro variables introduced in Section EC.4.1.1. (In Section EC.4.3, we also include allowances in (67) as a robustness check.) The lag l is four quarters to mimic the stress testing's forward-looking framework. We estimate separate coefficients for each of the four loan categories, for each bank, and the observations are loan balance and stress weighted using (65). Because these are bank-specific regressions, we do not add bank-specific controls.

For each loan category, we want to test for heterogeneity in parameters across banks. When we test for heterogeneity, the null hypothesis states that slopes for all banks are equal,

$$H_0 : \beta_1 = \cdots = \beta_{\bar{S}}, \tag{68}$$

or that the intercepts are equal,

$$H_0 : \alpha_1 = \cdots = \alpha_{\bar{S}}. \tag{69}$$

The alternative hypothesis in each case states that the indicated parameters are not identical across banks. We will run these tests with different subsets of the variables in (67) included and interpret the coefficients in (68) accordingly.

To test these hypotheses for a particular loan category, let $\mathbf{X}_s$ be the n_s by k data matrix for bank s, where n_s is the number of observations for bank s in the loan category, and k =1 or 2 is the number of variables included on the right side of (67). Let $\tilde{\mathbf{X}}_s = (\mathbf{1}, \mathbf{X}_s)$ be $\mathbf{X}_s$ concatenated with a column of 1s, and let $\mathbf{X}^*$ be the diagonal block matrix $\mathbf{X}^* = \text{diag}(\tilde{\mathbf{X}}_1, .., \tilde{\mathbf{X}}_{\bar{S}})$. Let $\theta^* = (\alpha_1, \beta_1^\top, ..., \alpha_S, \beta_S^\top)^\top, \epsilon^* = (\epsilon_1, ..., \epsilon_S)$, where ϵ_s is a column vector of length n_s. Our unrestricted model can be written as

$$Y = \mathbf{X}^* \theta^* + \epsilon^*, \tag{70}$$

and the restrictions in (68) and (69) impose linear constraints on the parameter θ^*.

We apply the Wald test to test linear constraints on θ^* in (70) under various assumptions on the error covariance matrix. We consider (i) bank-clustered errors, which allows correlation in errors across quarters for each bank, but no correlation across banks; (ii) time-clustered errors, which allows correlation in errors across banks in each quarter, but no correlation across time.

Table EC.4.3 reports p-values for the tests when different subsets of variables are included on the right side of (67), for a forecast horizon of one year. All tests indicate strong evidence of heterogeneity in the intercepts, the coefficients for past due rates, and macro variables.

Next we examine the impact of heterogeneity. Table EC.4.4 compares pooled and FEO coefficients for *PastDueRate* using a one-year lag when *MacroPC* is and is not included in (67). We estimate β_{Pool} in a pooled panel regression, and β_F in a panel regression with bank fixed effects included. Both regressions are weighted using (65).

In Table EC.4.4, the columns labeled "diff" show the difference in estimates $\beta_F - \beta_{Pool}$, serving as a measure of the impact of addressing heterogeneity in choosing an industry model. The table also shows p-values for tests of $H_0 : \beta_{Pool} = \beta_F$ (or $\gamma_{Pool} = \gamma_F$ in the case of the macro variable). To calculate these p-values, we estimate the pooled and FEO models simultaneously, as follows. Let $\tilde{\mathbf{X}}_{Pool} = (\mathbf{1}, \mathbf{X})$ be $\mathbf{X}$ concatenated with a column of 1s, and let $\tilde{\mathbf{X}}_F = (\mathbf{U}, \mathbf{X})$ be $\mathbf{X}$ concatenated with columns corresponding to centered bank identity variables $\mathbf{U}$. Let $\mathbf{X}_*$ be the diagonal block matrix $\mathbf{X}_* = \text{diag}(\tilde{\mathbf{X}}_{Pool}, \tilde{\mathbf{X}}_F)$ and $\theta_* = (\alpha_{Pool}, \beta_{Pool}^\top, \delta_1, \delta_2, \ldots, \delta_{\bar{S}}, \beta_F^\top)^\top$. Then we have

$$Y = \mathbf{X}_* \theta_* + \epsilon_*,$$

and testing H_0 is equivalent to testing linear constraints on the parameters θ_*, for which we apply the Wald test. The macro variable captures common variability over time, so we cluster errors by bank.

The results in Table EC.4.4 show that the differences between the pooled and FEO estimates are significant in three of the four loan categories. Moreover, the differences can be material. For example, for first lien loans in the top panel of Table EC.4.4, an absolute difference of 0.015 translates to a relative difference of 24% ($= |\beta_{Pool} - \beta_F|/\beta_F$), which can have a large relative impact on predicted losses. From Table EC.4.2, we see that the average bank has a past due rate of 4.01% on FL loans. The difference $0.015 \times 4.01\% = 0.060\%$ is 27% of the average FL loss rate of 0.22% in Table EC.4.2. The additional capital required to offset the higher predicted loss rate would be 27% of the capital required to offset the average loss rate.

To further analyze the differences in forecasts, we consider the relative prediction differences given by $|(\hat{Y}_{Pool}(x) - \hat{Y}_F(x))/Y(x)|$, where the denominator is the observed loss rate. Table EC.4.5 reports the mean and median of these relative differences in each of the four loan categories, for all banks in all quarters. In each loan category, the mean relative difference is large; moreover, in each case the mean is appreciably larger than the median, reflecting the presence of some very large relative prediction differences, which could be particularly important. To illustrate the differences, Figure EC.4.3 plots histograms of the FEO and pooled fitted values for Citigroup's first lien loans. The comparison shows, in particular, that the frequency of the largest and smallest predicted loss rates differ between the two methods.

Our main results use data through 2021. As a robustness check, we run our analysis using data through 2019. This truncation serves two purposes. It ensures that our conclusions are not driven by a few extreme values during the COVID period 2020–2021, and it accounts for a change in how banks measure allowances (the Current Expected Credit Losses methodology) beginning at the end of 2019. We also consider including banks' allowances for losses as another proxy for the portfolio risk. Because allowances are not consistently reported separately by loan category, we use banks' total allowances across all loan types. That is, for each bank-quarter we calculate $AllowanceRate_{s,t}$ using (66), but normalizing by the total loans in all categories. We repeat our tests with pre-COVID data and the addition of allowance rates. The results,

Covariance Estimation	α CC	FL	CRE	CI	β_{PDR} CC	FL	CRE	CI	γ CC	FL	CRE	CI
bank clustered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
time clustered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
bank clustered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
time clustered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table EC.4.3: P-values for heterogeneity tests. In each loan category, the first two rows are for a model with *PastDueRate* only, and the last two rows are for a model with *PastDueRate* and *MacroPC*. The two rows for each model show results under alternative assumptions on the error covariance matrix.

Loan Type	*Past Due Rate* β_{Pool}	β_F	diff	p-value	*Macro PC* γ_{Pool}	γ_F	diff	p-value
CC	0.782	0.833	-0.050	0.009***				
FL	0.047	0.062	-0.015	0.001***				
CRE	0.131	0.129	0.002	0.464				
CI	0.208	0.219	-0.011	0.040**				
CC	0.774	0.823	-0.049	0.009***	0.040	0.037	0.003	0.000***
FL	0.046	0.061	-0.015	0.001***	0.018	0.019	-0.001	0.176
CRE	0.131	0.129	0.002	0.485	0.011	0.011	0.000	0.813
CI	0.205	0.216	-0.010	0.044**	0.022	0.022	0.000	0.127

*p<0.1; **p<0.05; ***p<0.01

Table EC.4.4: Comparison of coefficients for one-year forecasts. We regress *LossRate* on (i) *PastDueRate* and (ii) *PastDueRate* and *MacroPC*. Difference is calculated as $\beta_{Pool} - \beta_F$. p-values test $H_0 : \beta_{Pool} = \beta_F$ for *PastDueRate* or $H_0 : \gamma_{Pool} = \gamma_F$ for *MacroPC*.

reported in Section EC.4.3, are similar to those reported in this section.

Our results document evidence of bank heterogeneity and its potential impact on loss forecasts. We have not sought to identify the drivers of heterogeneity; that would require a very different investigation, particularly since some of the most interesting potential drivers are difficult to measure. Guerrieri and Harkrader [2], for example, find that bank-specific factors account for a sizable fraction of the variation in bank performance. But they measure the bank-specific component as the residual in a regression that removes the effect of macroeconomic and banking-wide factors; they do not identify specific bank features that influence performance. Some examples of bank features used as controls in stress testing models can be found in Hirtle et al. [6], Kapinos and Mittnic [7], and Kupiec [8]. These are balance sheet features, and they are usually found to be more relevant to forecasting revenues than losses. We discuss revenue models in Section EC.7.

	CC	FL	CRE	CI
mean	6.1	510.1	52.0	24.0
median	2.4	89.5	4.2	3.7

Table EC.4.5: Mean and median of relative prediction differences between the pooled and the FEO estimates (in %).

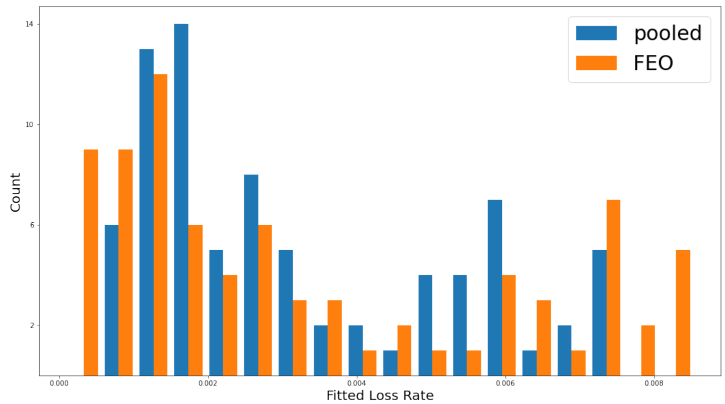

Figure EC.4.3: Pooled and FEO predicted loss rates for Citigroup's first lien loans.

EC.4.3 Robustness Checks

We repeat the analysis of Section EC.4, limiting the data to 2001–2019. This serves two purposes. It addresses the possibility that our results are driven by a few extreme values during the COVID period 2020–2021. It also accounts for a change in how banks measure allowances (the Current Expected Credit Losses methodology) that began to take effect at the end of 2019. We also consider adding allowance rate as another proxy for portfolio risks. Tables EC.4.6 and EC.4.7 report the results for heterogeneity tests and differences of the parameter estimates under this setting. The evidence for heterogeneity and its impact is generally at least as strong using the pre-COVID data as using data through 2021.

Covariance Estimation	α CC	FL	CRE	CI	β_{PDR} CC	FL	CRE	CI	β_{AR} CC	FL	CRE	CI	γ CC	FL	CRE	CI
bank clustered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00								
time clustered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00								
bank clustered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
time clustered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
bank clustered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00					0.00	0.00	0.00	0.00
time clustered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00					0.00	0.00	0.00	0.00
bank clustered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
time clustered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table EC.4.6: Heterogeneity tests using pre-COVID data with allowance rate as an additional proxy for banks' portfolio risks.

Loan	*Past Due Rate*				*Allowance Rate*				*Macro PC*			
Type	β_{Pool}	β_F	diff	p-value	β_{Pool}	β_F	diff	p-value	γ_{Pool}	γ_F	diff	p-value
CC	0.781	0.838	-0.057	0.006***								
FL	0.045	0.061	-0.016	0.002***								
CRE	0.135	0.133	0.002	0.449								
CI	0.207	0.217	-0.011	0.069*								
CC	0.794	0.852	-0.058	0.109	-0.029	-0.032	0.003	0.938				
FL	0.010	0.021	-0.011	0.306	0.335	0.307	0.028	0.424				
CRE	0.095	0.068	0.026	0.052*	0.127	0.191	-0.064	0.090*				
CI	0.178	0.196	-0.018	0.000***	0.064	0.054	0.010	0.514				
CC	0.595	0.633	-0.039	0.001***					0.417	0.397	0.020	0.000***
FL	0.037	0.052	-0.015	0.001***					0.110	0.106	0.003	0.329
CRE	0.129	0.126	0.002	0.477					0.057	0.057	0.000	0.601
CI	0.155	0.155	-0.001	0.886					0.190	0.190	0.000	0.981
CC	0.577	0.618	-0.041	0.046**	0.037	0.032	0.005	0.846	0.420	0.399	0.021	0.000***
FL	0.006	0.018	-0.012	0.146	0.315	0.268	0.047	0.041**	0.095	0.094	0.001	0.768
CRE	0.097	0.078	0.019	0.182	0.102	0.147	-0.045	0.265	0.052	0.050	0.002	0.267
CI	0.140	0.153	-0.012	0.035**	0.032	0.007	0.025	0.105	0.189	0.189	-0.001	0.583

*p<0.1; **p<0.05; ***p<0.01

Table EC.4.7: Comparison of pooled and FEO coefficients using pre-COVID data with allowance rate as an additional proxy for banks' portfolio risks.

EC.5 Nonlinear Models

Building on Section 5, we consider generalized additive models (GAMs) in which the effect of the past due rates and the macro variable are not restricted to be linear. More specifically, we consider the specifications

$$Y_{s,t}^{Pool} = f_0^P + f_1^P(PDR_{s,t-l}) + f_2^P(MacroPC_{t-l}) + f_3^P(PDR_{s,t-l} \times MacroPC_{t-l}) + \epsilon_{s,t}^P \quad (71)$$

and

$$Y_{s,t}^F = f_0^F + f_1^F(PDR_{s,t-l}) + f_2^F(MacroPC_{t-l}) + f_3^F(PDR_{s,t-l} \times MacroPC_{t-l}) + f_4^F(s) + \epsilon_{s,t}^F, \quad (72)$$

in which $f_i^{P/F}$, $i = 1, 2, 3$, are (possibly nonlinear) centered functions of *PastDueRate*, *MacroPC*, and their interaction *PastDueRate*×*MacroPC*, respectively, and f_4^F measures centered bank fixed effects. The pooled model (71) omits f_4^P; the FEO model (72) estimates f_4^F but discards it in forecasting loss rates to satisfy equal treatment. We consider the modeling of loss rates four quarters ahead, so $l = 4$ in both cases.

We use the R package `gam` (Hastie [5]) to fit (71) and (72), taking the $f_i^{P/F}$, $i = 1, 2, 3$, to be smoothing splines with 4 degrees of freedom. We choose `gam` because it is a direct implementation of the backfitting algorithm in Hastie and Tibshirani [4], which underpins the framework in Section 5. In particular, Proposition 5.1 applies to FEO forecasts based on (72).

Model (71) is nested within model (72), so we can use an F-test to compare the two. In all four loan categories, the test rejects (with p-values smaller than 0.01) the restriction to equal bank fixed effects ($f_4^P \equiv 0$) imposed in the pooled model (71). Figure EC.5.1 plots the centered bank fixed effects f_4^F for all four loan categories, expressed in percent. We observe significant variability within each loan type. For example, for credit card loans, JPM's fixed effect is three percentage points larger than WFC's. We also observe variability across loan types for individual banks. For example, CFG has the highest fixed effect for first lien loans, but one of the lowest for credit card loans. These observations again reflect the notable heterogeneity among the bank holding companies.

Table EC.5.1 reports the mean and median of the relative prediction differences between the FEO and the pooled predictions, given, as in Section EC.4, by $|(\hat{Y}_{Pool}(x) - \hat{Y}_F(x))/Y(x)|$). These summary statistics show that the relative prediction differences can indeed be very large. To further illustrate this point, Figure EC.5.2 contrasts the prediction distributions for Citigroup's first lien loans using the pooled and FEO methods. As in Proposition 5.1, the pooled method may yield smaller prediction errors overall, but it does so by implicitly misdirecting legitimate information.

	CC	FL	CRE	CI
mean	25.5	559.0	106.6	88.5
median	10.4	120.8	14.2	9.8

Table EC.5.1: Mean and median of relative prediction differences between the pooled and the FEO estimates (in %) for GAMs.

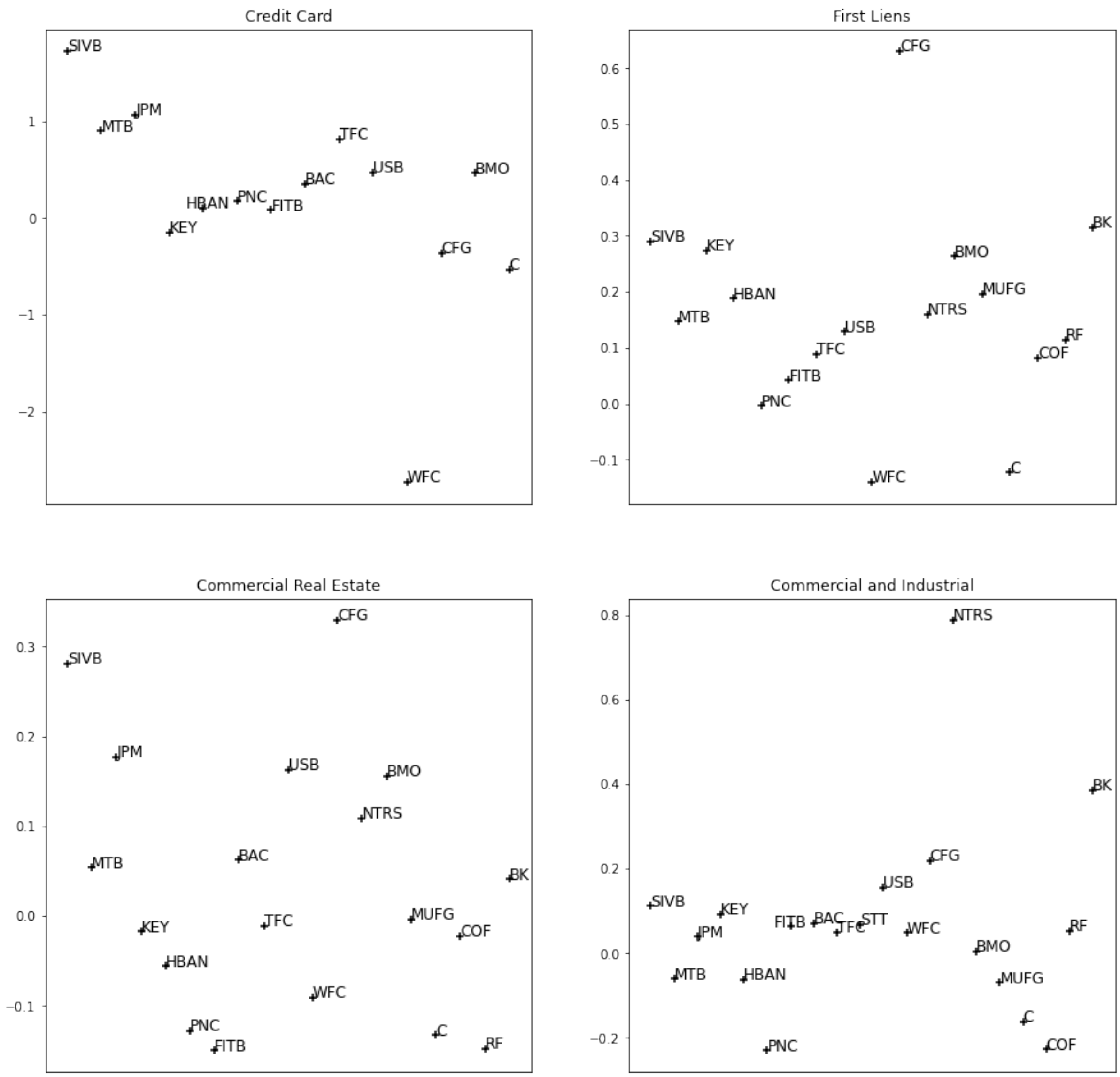

Figure EC.5.1: Banks' generalized fixed effects. Y-axis is in %.

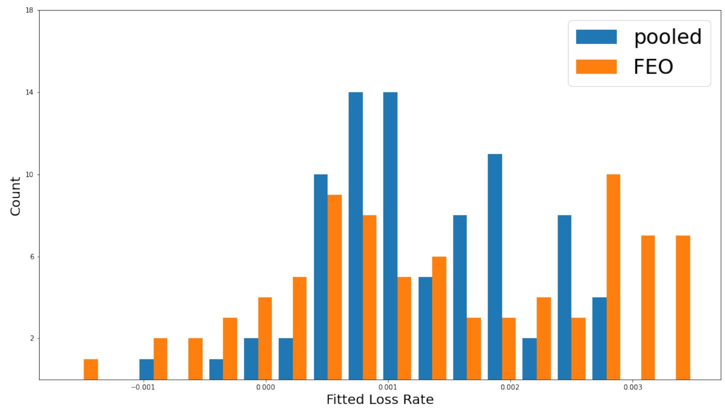

Figure EC.5.2: Pooled and FEO predicted loss rates for Citigroup's first lien loans.

EC.6 Additional Information on Empirical Analysis

Table EC.6.1 lists the bank holding companies included in our empirical analysis and the symbols we use to refer to them. The companies are listed in order of size by total assets.

Ticker	Bank Name
JPM	JPMORGAN CHASE & CO.
BAC	BANK OF AMERICA CORPORATION
C	CITIGROUP INC.
WFC	WELLS FARGO & COMPANY
GS	GOLDMAN SACHS GROUP, INC., THE
MS	MORGAN STANLEY
SCHW	CHARLES SCHWAB CORPORATION, THE
USB	U.S. BANCORP
PNC	PNC FINANCIAL SERVICES GROUP, INC., THE
TFC	TRUIST FINANCIAL CORPORATION
TD	TD GROUP US HOLDINGS LLC
BK	BANK OF NEW YORK MELLON CORPORATION, THE
COF	CAPITAL ONE FINANCIAL CORPORATION
STT	STATE STREET CORPORATION
HSBC	HSBC NORTH AMERICA HOLDINGS INC.
SIVB	SVB FINANCIAL GROUP
FITB	FIFTH THIRD BANCORP
USAA	UNITED SERVICES AUTOMOBILE ASSOCIATION
BMO	BMO FINANCIAL CORP.
CFG	CITIZENS FINANCIAL GROUP, INC.
AXP	AMERICAN EXPRESS COMPANY
KEY	KEYCORP
NTRS	NORTHERN TRUST CORPORATION
ALLY	ALLY FINANCIAL INC.
AMP	AMERIPRISE FINANCIAL, INC.
RY	RBC US GROUP HOLDINGS LLC
HBAN	HUNTINGTON BANCSHARES INCORPORATED
RF	REGIONS FINANCIAL CORPORATION
MUFG	MUFG AMERICAS HOLDINGS CORPORATION
BCS	BARCLAYS US LLC
SAN	SANTANDER HOLDINGS USA, INC.
MTB	M&T BANK CORPORATION
BNPQY	BNP PARIBAS USA, INC.
DB	DB USA CORPORATION
DFS	DISCOVER FINANCIAL SERVICES

Table EC.6.1: Symbols and names of included bank holding companies.

We construct the loss rates, past due rates, and allowance rates using the entries in FR Y-9C forms outlined in Table EC.6.2.

Variables	Loan Types	2007Q1 – Present	2003Q1 – 2006Q4
Loan Amount	CC	BHCKB538	BHCKB538
	FL	BHDM5367	BHDM5367
	CRE	Owned: BHCKF160 Other: BHCKF161	BHDM1480
	CI	BHCK1763	BHCK1763
	Total	BHCK2122	BHCK2122
Charge-Offs	CC	BHCKB514	BHCKB514
	FL	BHCKC234	BHCKC234
	CRE	Owned: BHCKC895 Other: BHCKC897	BHCK3590
	CI	BHCK4645	BHCK4645
Recoveries	CC	BHCKB515	BHCKB515
	FL	BHCKC217	BHCKC217
	CRE	Owned: BHCKC896 Other: BHCKC898	BHCK3591
	CI	BHCK4617	BHCK4617
Past Due: 30-89 days and accruing	CC	BHCKB575	BHCKB575
	FL	BHCKC236	BHCKC236
	CRE	Owned: BHCKF178 Other: BHCKF179	BHCK3502
	CI	BHCK1606	BHCK1606
Past Due: 90 days and accruing	CC	BHCKB576	BHCKB576
	FL	BHCKC237	BHCKC237
	CRE	Owned: BHCKF180 Other: BHCKF181	BHCK3503
	CI	BHCK1607	BHCK1607
Past Due: non-accrual	CC	BHCKB577	BHCKB577
	FL	BHCKC229	BHCKC229
	CRE	Owned: BHCKF182 Other: BHCKF183	BHCK3504
	CI	BHCK1608	BHCK1608

Table EC.6.2: Loan variables and FR Y-9C form correspondence.

EC.7 Revenue Models

The Federal Reserve's stress testing framework includes models of revenues as well as models of losses. We have focused on loan portfolio loss models because they fit most clearly within the Fed's policy of equal treatment and its preference for industry models. In this section, we show that the heterogeneity documented for loss models in Section EC.4 extends to revenue models, referred to in the Fed's framework as models of pre-provision net revenue or PPNR.

The Fed uses a suite of PPNR models to forecast difference sources of revenue. These models differ from the loss models in at least two important respects: they are typically autoregressive (AR) models, and, unlike the portfolio loss models, they do not rule out bank fixed effects; see [1]. This feature points to the presence of unmodeled bank heterogeneity in the revenue forecasts. Our goal in this section is to check for heterogeneity in a simple PPNR model and to compare coefficient estimates in the pooled and fixed-effect models.

We consider the modeling of trading revenue, which is one of the PPNR components in the Fed's framework. We compare AR models with fixed-effects,

$$Y^{FE}_{s,t} = \alpha_s + \rho_{FE} Y_{s,t-1} + \beta_{FE} X_{s,t} + \gamma_{FE}\, VIX_t + \epsilon^{FE}_{s,t} \tag{73}$$

or pooled without fixed effects,

$$Y^{P}_{s,t} = \alpha_P + \rho_P Y_{s,t-1} + \beta_P X_{s,t} + \gamma_P\, VIX_t + \epsilon^{P}_{s,t}. \tag{74}$$

In both models, Y is trading revenue normalized by total trading assets; this choice of normalization is consistent with [1]. For the AR term, we use either a one-quarter lag $Y_{s,t-1}$ or a four-quarter average lag, in which case we use $1/4\sum_{j=1}^{4} Y_{s,t-j}$ in place of $Y_{s,t-1}$ in (73) and (74) to capture average performance over the past year.

For $X_{s,t}$ we use the size of bank s in quarter t, as measured by the log of total assets. For the macro variable, we use the VIX_t, the market volatility index taken from the Federal Reserve's stress testing historical dataset. Market volatility is expected to have a direct impact on trading revenue, and indeed we observe a more significant effect of VIX_t than $MacroPC_t$ (from Section EC.4) in this setting.

Models (73) and (74) differ in their intercept terms: (73) captures banks' fixed effects, but (74) requires the same intercept across all banks. The fixed-effect coefficient estimates ρ_{FE}, β_{FE} and γ_{FE} are identical to those of FEO; the methods differ in their forecasts: FEO uses the average fixed effect, rather than bank-specific fixed effects in its forecasts.

As in Section EC.4, we use Y-9C financial reporting data for the top 35 banks by total asset size (as of year-end 2021), and we include only banks with at least 18 years of data to ensure our panel is mostly balanced. We fit the models using weighted least squares, weighting each observation by quarter stress and bank asset balance. As in (65), we choose the weights so that, for the same asset level, the quarters with the highest market volatility get twice the weight as the quarters with the lowest market volatility. As in the AR models in [6], we cluster standard errors by time.

Table EC.7.1 reports the results. The first two columns correspond to the one-quarter AR setting, and the last two columns are the one-year-average AR setting. The numbers in parentheses are standard errors.

As expected, both the lagged response and the *VIX* term are statistically significant. The volatility term is more significant in the one-year-average AR setting, presumably because the market environment changes less over one quarter, and its effect is partly captured by the lagged response.

The pooled and fixed-effect methods result in different estimates of ρ, and the differences in estimates in all settings are more than 20%. We tested the hypotheses that all banks share the same (i) intercept, (ii) AR coefficient term, (iii) *VIX* coefficient, or (iv) coefficients of total asset size, following the approach used in Section EC.4.2; all tests strongly reject that banks have identical model coefficients, with p-values less than 0.01, extending what we found for loss models. We have also examined the forecasts of trading interest income and trading interest expense (two other components of PPNR), and observed similar patterns across all three components.

	Normalized Trading Revenue			
AR	0.466***	0.588***	0.406***	0.596***
	(0.134)	(0.111)	(0.078)	(0.094)
VIX	-0.024**	-0.022*	-0.027***	-0.022***
	(0.011)	(0.012)	(0.008)	(0.008)
Log Assets	-0.003	-0.007***	0.000	-0.006***
	(0.004)	(0.001)	(0.003)	(0.002)
Bank FE	included	–	included	–

*p<0.1; **p<0.05; ***p<0.01

Table EC.7.1: Coefficient estimates for the AR models. The first two columns use a one-quarter lag, and the last two use a one-year average lag. Columns 1 and 3 correspond to AR models with bank fixed effects, and columns 2 and 4 are pooled AR models.

References

[1] Board of Governors (2021) Dodd-Frank Act Stress Test 2021: Supervisory Stress Test Methodology. Federal Reserve System, Washington, D.C.

[2] Guerrieri, L., and Harkrader, J. (2021) What drives bank performance?, Working paper 2021-009, Federal Reserve Board, Washington, D.C.

[3] Guerrieri, L., and Welch, M. (2012) Can macro variables used in stress test forecast the performance of banks? Working paper 2012-49, Federal Reserve Board, Washington, D.C.

[4] Hastie, T., and Tibshirani, R. (1986) Generalized additive models, *Statistical Science* 1(3), 297–318.

[5] Hastie, T. (2023) gam: Generalized Additive Models, R package version 1.22-1, https://CRAN.R-project.org/package=gam.

[6] Hirtle, B., Kovner, A., Vickery, J., and Bhanot, M. (2016) Assessing financial stability: The Capital and Loss Assessment under Stress Scenarios (CLASS) model. *Journal of Banking and Finance* 69, S35–S55.

[7] Kapinos, P., and Mitnik, O.A. (2016) A top-down approach to stress-testing banks. *Journal of Financial Services Research* 49, 229–264.

[8] Kupiec, P. (2020) Policy uncertainty and bank stress testing. *Journal of Financial Stability* 51, 100761.