data source would be accessible to central banks, like the Federal Reserve Board or the European Central Bank. By using these two data sources, and by building on existing accounting frameworks [3, 4, 9, 10], we partition the balance sheet of the banking sector to isolate the underlying (and unobserved) portfolios held by banks. This partitioning, when combined with balance sheet identities, implies a non-negative matrix factorization problem [6], which has been extensively studied in other domains. Solving the matrix factorization problem provides estimates of a bank’s underlying portfolio along different asset classes (equities, bonds, commodities, etc.), thus giving regulators meaningful information synthesized from data sources that are not typically integrated or studied together in this manner.

2 AN ACCOUNTING FRAMEWORK FOR BANKS

We follow the setup of previous works [3, 4, 9, 10]. Before introducing the accounting framework and factorization problem, we first introduce some notation. Let there be $n$ banks under consideration and $X$ be the vector of interbank debt (the total value of liabilities held by other banks). $\Pi_{ij}$ is the share of bank $i$’s liabilities held by bank $j$. $W_{i k}$ is the weight invested in each of the $k$ asset classes by bank $i$ ($\sum_k W_{i k} = 1$). $V_{i k}$ denotes the market value of bank $i$’s assets, including loans to firms and households as well as $k$ asset classes (equities, bonds, commodities, etc.). $E_i$ indicates the market value of bank $i$’s equity, and $D_i$ is the total value of liabilities of bank $i$ held by non-banks.

Table 1: Representation of the balance sheet of a bank $i$.

<table>
<thead>
<tr>
<th>Assets</th>
<th>Liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sum_k W_{i k} V_{i k}$</td>
<td>$e_i$</td>
</tr>
<tr>
<td>$\sum_j x_j \Pi_{i j}$</td>
<td>$x_i$</td>
</tr>
<tr>
<td>$d_i$</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Representation of the balance sheet of a bank $i$.  

We obtain the balance sheet identity by summing assets and liabilities respectively

$$\sum_j x_j \Pi_{i j} + \sum_k W_{i k} V_{i k} = e_i + x_i + d_i. \quad (1)$$
Written using matrix notation (with capital letters), the balance sheet identity is
\[ (W \odot V)u = E + X + D, \]
where \( u \) is a vector of ones and \( \odot \) is the Schur product (element-wise multiplication).

Thus, one can see that \( W \) and \( V \) are a function of the equity, interbank market, and household debt, which we ignore since it is roughly constant [9].

(3)

\[ (W \odot V)u = E + (I - \Pi)X. \]

### 2.1 Probability Factorization

To estimate \( W \), one could pose an optimization problem based on minimizing the Frobenius norm of the difference between a time-series of the combination of stock returns and interbank lending data (captured in \( Z = \{Z_1, \ldots, Z_T\} \)) and the estimated factors

\[
\min_{W,V} \|Z - WV\|_F^2
\]

subject to \( W \in \mathbb{R}^{n \times K} \)
\( V \in \mathbb{R}^{K \times T} \)
\( (W)_{ij} \geq 0 \) for all \( i, j \)
\( \sum_{j=1}^{K} (W)_{ij} = 1 \) for all \( i \).

The estimation approach we present alternates between optimizing with respect to \( W \) and \( V \). The algorithm solves for \( W \) using a projected gradient descent method that has been effective at balancing cost per iteration and convergence rate for similar problems posed in Nonnegative Matrix Factorization [7, 8]. After solving for \( W \), probability constraints are enforced without changing the overall quality of the solution since \( V \) can be rescaled without changing the objective function value. Specifically, we have \( WD \) and \( D^{-1}V \), where \( D \) is a diagonal matrix with positive entries, provides different solutions with identical objective function values.

#### 2.1.1 Solving for \( V \)

When holding \( W \) fixed, the remaining optimization problem written is exactly the usual least squares problem from linear regression.

Starting the objective function,
\[
O = \|Z - WV^T\|_F^2
\]
\[
= (Z - WV^T)^T(Z - WV^T)
\]
\[
= Z^TZ - VV^TZ - Z^TWV^T + WV^TW^T.
\]

Holding \( W \) fixed and differentiating with respect to \( V \) yields
\[
\frac{\partial O}{\partial V} = -2Z^TW + 2VV^TW.
\]

Setting the partial derivative equal to zero and solving for \( V \) yields
\[
V = Z^TW(W^TW)^{-1}.
\]

This is the optimal update for the problem \( \min_V \|Z - WV^T\|_F^2 \).

#### 2.1.2 Solving for \( W \)

We now turn our attention to solving for \( W \), holding \( V \) fixed.

A standard gradient descent algorithm would start with an initial condition \( W^{(0)} \) and constants \( \alpha_i \) and iterate

(1) For \( i = 1, 2, \ldots \)

(2) Set \( W^{(i+1)} = W^{(i)} - \alpha_i \Delta W \),

where the gradient of the objective function with respect to \( W \) is
\[
\Delta W = WV^TV - ZV.
\]

Due to the subtraction, the non-negativity of \( W \) cannot be guaranteed. Thus, the basic idea of projected gradient descent is to project elements in \( W \) to the feasible region using the projection function, which for our problem is defined as \( P(y) = \max(0, y) \). The basic algorithm is then

(1) For \( i = 1, 2, \ldots \)

(2) Set \( W^{(i+1)} = P(W^{(i)} - \alpha_i \Delta W) \).

The constants \( \alpha_i \) regulate the step size or amount of change in the estimate at each iteration, and converge to zero with \( i \). However, the exact specification of \( \alpha_i \) is a main challenge. If the step size is too small, the algorithm will not converge to a stationary point. If the step size is too large, then too many elements of \( W \) will be projected to zero and the quality of the estimate will suffer. To guarantee a sufficient decrease at each iteration and convergence to a stationary point, the “Armijo rule” developed in [1, 2] provides a sufficient condition for a given \( \alpha_i \) at each iteration
\[
\|Z - W^{(i+1)}V^T\|_F^2 - \|Z - W^{(i)}V^T\|_F^2 \leq \sigma(\Delta W^T \Delta W),
\]

(10)

where \( \sigma \in (0, 1) \) and \( \langle \cdot, \cdot \rangle \) is the sum of element wise products of two matrices. Thus, for a given \( \alpha_i \), one calculates \( W^{(i+1)} \) and checks whether (10) is satisfied. If the condition is satisfied, then the step size \( \alpha_i \) is appropriate to guarantee convergence to a stationary point.

### 3 EMPIRICAL RESULTS

We use data from e-MID interbank market and public stock returns data to form \( Z \) and subsequently estimate \( W \) over monthly intervals. Our preliminary results of one-way Granger causality of factorization-based variables to the St. Louis Fed Financial Stress Index [5] is an encouraging result. More work is needed to validate the model before using it to gain insight into the balance sheets of banks at a higher frequency than current disclosures allow, and thus better estimate systemic risk and monitor the financial ecosystem.

#### REFERENCES


