Appendix: Mathematical Derivations

1 Fixed Effects and Difference-in-Differences

1.1 Fixed Effects Estimator

Consider the baseline panel model

$$ret_{it} = \alpha_i + X_t'\beta + \varepsilon_{it}, \tag{1}$$

where ret_{it} is the weekly municipal bond return for issuer i at time t, α_i is an issuer-specific effect, and X_t is a vector of time-varying common regressors (e.g., changes in muni volatility and liquidity).

The within transformation subtracts issuer means:

$$\tilde{ret}_{it} = ret_{it} - \bar{ret}_i, \tag{2}$$
$$\tilde{X}_t = X_t - \bar{X}, \tag{3}$$
$$\tilde{\varepsilon}_{it} = \varepsilon_{it} - \bar{\varepsilon}_i, \tag{4}$$

where $\bar{ret}_i = \frac{1}{T_i}\sum_t ret_{it}$ and $\bar{X} = \frac{1}{T}\sum_t X_t$. The issuer fixed effect α_i drops out since it is constant over time for each i:

$$\tilde{ret}_{it} = \tilde{X}_t'\beta + \tilde{\varepsilon}_{it}. \tag{5}$$

Stacking observations across i and t, the fixed effects estimator is

$$\hat{\beta}_{FE} = \left(\tilde{X}'\tilde{X}\right)^{-1}\tilde{X}'\tilde{ret}, \tag{6}$$

where $\tilde{X}$ and $\tilde{ret}$ collect all within-transformed observations.

In the dynamic specification with lags of X_t,

$$ret_{it} = \alpha_i + \sum_{k=0}^{p}\beta_{1k}\Delta VIX_{t-k} + \sum_{k=0}^{p}\beta_{2k}\Delta LIQ_{t-k} + u_{it}, \tag{7}$$

the same within-transformation applies; the regressors simply include current and lagged values of the common shocks.

1.2 Difference-in-Differences Estimator

Let D_i be a treatment indicator equal to 1 for city issuers and 0 for county issuers. Let $Post_t$ be a post-treatment indicator equal to 1 after the COVID onset date (March 1, 2020) and 0 before. The standard DiD regression is

$$ret_{it} = \alpha_i + \gamma Post_t + \delta(Post_t \cdot D_i) + X_t'\theta + \varepsilon_{it}. \tag{8}$$

The coefficient of interest is δ, which measures the additional change in returns for treated (city) issuers relative to control (county) issuers after COVID, conditional on common shocks X_t.

Under the parallel trends assumption (in the absence of treatment, city and county returns would have evolved in parallel), the DiD estimator can be expressed in terms of group-time means:

$$\begin{aligned}\hat{\delta}_{DiD} &= \left[\bar{ret}^{\text{City}}_{\text{Post}} - \bar{ret}^{\text{City}}_{\text{Pre}}\right] - \left[\bar{ret}^{\text{County}}_{\text{Post}} - \bar{ret}^{\text{County}}_{\text{Pre}}\right] && (9)\\ &= \Delta\bar{ret}^{\text{City}} - \Delta\bar{ret}^{\text{County}}, && (10)\end{aligned}$$

where $\bar{ret}^{g}_{\text{Post}}$ denotes the mean after COVID for group g, and similarly for $\bar{ret}^{g}_{\text{Pre}}$.

The regression estimator of δ is numerically equivalent to this difference-in-differences when no additional regressors are included and is consistent under the same identification assumptions when controls are present.

2 VAR Representation and Impulse Responses

2.1 VAR(p) and Companion Form

Consider the VAR(p) model

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma_u), \tag{11}$$

where Y_t is a $K \times 1$ vector of variables, here $Y_t = (\Delta VIX_t, \Delta LIQ_t, ret_t)'$.

Define the $Kp \times 1$ companion state vector

$$Z_t = \begin{pmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{pmatrix}. \tag{12}$$

Then the VAR can be written in companion form as

$$Z_t = \mathcal{A} Z_{t-1} + \mathcal{U}_t, \tag{13}$$

where

$$\mathcal{A} = \begin{pmatrix} A_1 & A_2 & \cdots & A_{p-1} & A_p \\ I_K & 0 & \cdots & 0 & 0 \\ 0 & I_K & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & I_K & 0 \end{pmatrix}, \quad \mathcal{U}_t = \begin{pmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{14}$$

Under stability (all eigenvalues of $\mathcal{A}$ inside the unit circle), the system admits a finite-variance solution.

2.2 Moving Average Representation and IRFs

The VAR has an infinite-order moving-average representation:

$$Y_t = \sum_{j=0}^{\infty} C_j u_{t-j}, \tag{15}$$

where $C_0 = I_K$ and subsequent C_j are obtained recursively:

$$C_j = A_1 C_{j-1} + A_2 C_{j-2} + \cdots + A_p C_{j-p}, \quad j \geq 1, \tag{16}$$

with $C_j = 0$ for $j < 0$.

For structural analysis, we assume that reduced-form residuals u_t are related to orthogonal structural shocks e_t by

$$u_t = P e_t, \tag{17}$$

where P is a $K \times K$ matrix satisfying $\Sigma_u = PP'$. A common choice is the lower-triangular Cholesky factor of Σ_u, which imposes the contemporaneous causal ordering.

The structural MA representation is then

$$Y_t = \sum_{j=0}^{\infty} \Psi_j e_{t-j}, \quad \Psi_j = C_j P, \tag{18}$$

and the impulse response at horizon h to a unit structural shock in component k is given by the kth column of Ψ_h.

2.3 Forecast Error Variance Decomposition

The h-step-ahead forecast error is

$$Y_{t+h} - \mathbb{E}_t[Y_{t+h}] = \sum_{j=0}^{h-1} \Psi_j e_{t+h-j}. \tag{19}$$

The forecast error variance of the ith variable at horizon h is

$$\mathrm{Var}\left(Y_{i,t+h} - \mathbb{E}_t[Y_{i,t+h}]\right) = \sum_{j=0}^{h-1} \sum_{k=1}^{K} \Psi_{j,ik}^2, \tag{20}$$

where $\Psi_{j,ik}$ is the (i,k) element of Ψ_j. The contribution of shock k to this variance is

$$\frac{\sum_{j=0}^{h-1} \Psi_{j,ik}^2}{\sum_{j=0}^{h-1} \sum_{l=1}^{K} \Psi_{j,il}^2}. \tag{21}$$

This yields the forecast error variance decomposition (FEVD) used in the main text.

3 Bayesian VAR with Normal–Inverse-Wishart Prior

3.1 Model in Matrix Form

Consider again the VAR(p):

$$Y_t = A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + u_t, \tag{22}$$

and define the stacked regression form:

$$Y = XB + U, \tag{23}$$

where

- Y is an $(T-p) \times K$ matrix of dependent variables,
- X is an $(T-p) \times m$ matrix of regressors ($m = 1 + Kp$ including intercept),
- B is an $m \times K$ matrix of coefficients,
- U is $(T-p) \times K$ with rows u_t'.

Assume $u_t \sim \mathcal{N}(0, \Sigma)$ independently over t.

3.2 Normal–Inverse-Wishart Prior and Posterior

The conjugate prior is:

$$B \mid \Sigma \sim MN(B_0, \Sigma, \Omega_0), \tag{24}$$

$$\Sigma \sim IW(S_0, \nu_0), \tag{25}$$

where MN is the matrix-normal distribution with mean B_0, row covariance Ω_0, and column covariance Σ; and IW is the inverse-Wishart distribution.

Given data (Y, X), the posterior is also Normal–Inverse-Wishart:

$$\Omega_n^{-1} = \Omega_0^{-1} + X'X, \tag{26}$$

$$B_n = \Omega_n \left(\Omega_0^{-1} B_0 + X'Y\right), \tag{27}$$

$$\nu_n = \nu_0 + T - p, \tag{28}$$

$$S_n = S_0 + (Y - XB_n)'(Y - XB_n) + (B_n - B_0)'\Omega_0^{-1}(B_n - B_0). \tag{29}$$

Thus,

$$B \mid \Sigma, Y, X \sim MN(B_n, \Sigma, \Omega_n), \tag{30}$$

$$\Sigma \mid Y, X \sim IW(S_n, \nu_n). \tag{31}$$

In the paper, I focus on the posterior means $\mathbb{E}[B \mid Y, X] = B_n$ and $\mathbb{E}[\Sigma \mid Y, X] = S_n/(\nu_n - K - 1)$ for computing Bayesian impulse responses.

4 Historical Decomposition

Using the structural MA representation

$$Y_t = \sum_{j=0}^{\infty} \Psi_j e_{t-j}, \tag{32}$$

the contribution of shock type k to variable i at time t is

$$\text{Contrib}_{i,k}(t) = \sum_{j=0}^{H} \Psi_{j,ik} e_{k,t-j}, \tag{33}$$

where H is a finite truncation horizon chosen such that impulse responses have effectively decayed to zero. In practice, a moderate H (e.g. 12 weeks) suffices for the VAR estimated on municipal bond data.

Let i^* index the return variable (e.g. ret_t). The historical decomposition of returns is:

$$ret_t \approx \sum_{k=1}^{K} \text{Contrib}_{i^*,k}(t), \tag{34}$$

where each term corresponds to a shock category:

- $k = 1$: volatility shocks (ΔVIX),
- $k = 2$: liquidity shocks (ΔLIQ),
- $k = 3$: own (return) shocks.

The approximation error arises from truncation of the infinite MA and is typically small.

The stacked contribution series plotted in the main text (Figure 9) are obtained by computing $\text{Contrib}_{i^*,k}(t)$ for each t and k, then summing over k to compare the reconstructed series with the actual returns.