Technical Report: Merton Structural Credit Model
with Exponential Smoothing Improvement

Ruhani Rekhi

December 25, 2025

1 Introduction

This report presents an implementation and improvement of the Merton (1974) structural credit model. After implementing and diagnosing the baseline model on API-fetched 2020 data for five U.S. firms (Apple, JPMorgan Chase, Tesla, ExxonMobil, and Ford), we identified default probability (PD) instability as a primary weakness. To approach this issue we implement exponential smoothing as a minimal improvement that reduces PD volatility while preserving the model's ability to properly capture credit deterioration during crisis periods.

2 Model Formulation

The Merton model views a firm's equity as a European call option on its assets, with the strike price being the face value of its debt. The key equations and assumptions of the baseline Merton model are outlined below.

2.1 Baseline Merton Model

2.1.1 Key Assumptions

1. Firm asset value V_t follows geometric Brownian motion:
$$dV_t = \mu V_t \, dt + \sigma_V V_t \, dW_t \tag{1}$$
2. Firm has zero-coupon debt with face value D maturing at time T
3. Default occurs only at maturity T if $V_T < D$
4. Equity is a European call option on firm assets: $E = \max(V_T - D, 0)$
5. Markets are frictionless (no transaction costs), trading is continuous, no arbitrage(no riskless profit opportunities)

2.1.2 Key Equations/ Mathematical Formulation

Equity Value Under the Merton model, equity value is given by the Black-Scholes call option formula:

$$E = V \cdot \Phi(d_1) - D \cdot e^{-r(T-t)} \cdot \Phi(d_2) \tag{2}$$

where

$$d_1 = \frac{\ln(V/D) + (r + \sigma_V^2/2)(T-t)}{\sigma_V \sqrt{T-t}} \tag{3}$$

$$d_2 = d_1 - \sigma_V \sqrt{T-t} \tag{4}$$

and $\Phi(\cdot)$ denotes the standard normal cumulative distribution function.

Equity Volatility The relationship between equity volatility σ_E and asset volatility σ_V is:

$$\sigma_E \cdot E = \Phi(d_1) \cdot \sigma_V \cdot V \tag{5}$$

where $\Phi(d_1)$ is the option delta, representing the sensitivity of equity value to changes in asset value.

Default Probability The risk-neutral probability of default at maturity is:

$$\text{PD} = \Phi(-d_2) = \Phi\left(-\frac{\ln(V/D) + (r - \sigma_V^2/2)T}{\sigma_V \sqrt{T}}\right) \tag{6}$$

Distance-to-Default The distance-to-default metric is:

$$\text{DD} = \frac{\ln(V/D) + (r - \sigma_V^2/2)T}{\sigma_V \sqrt{T}} \tag{7}$$

2.1.3 Calibration Approach

The Merton model calibration involves solving the following system of equations for the unobservable asset value V and asset volatility σ_V, given the observable equity value E, equity volatility σ_E, debt face value D, time to maturity $T-t$, and risk-free rate r:

$$E = \text{BlackScholes}(V, D, T-t, r, \sigma_V) \tag{8}$$

$$\sigma_E \cdot E = \Phi(d_1) \cdot \sigma_V \cdot V \tag{9}$$

2.2 Improved Model

2.2.1 Modification

We apply post-processing exponential smoothing to default probabilities. For each firm's time series of raw PDs $\{\text{PD}_t^{\text{raw}}\}$ we applied the following smoothing formula:

$$\text{PD}_t^{\text{smooth}} = \alpha \cdot \text{PD}_t^{\text{raw}} + (1-\alpha) \cdot \text{PD}_{t-1}^{\text{smooth}} \tag{10}$$

where $\alpha = 0.1$ is the smoothing parameter.

2.2.2 Justification

This improvement is justified on two grounds. Theoretically speaking, the baseline model assumes you can trade infinitely with no transaction costs. Whereas in reality, trading is discrete and every trade has costs leading to noise in observed prices that isn't properly captured in the model. Additionally, daily equity volatility contains high-frequency noise (changes in company value, as well as market microstructure noise) which is further amplified through the nonlinear Black-Scholes model into asset volatility thus into PD estimates. Practically speaking, PD estimates play a key role in credit risk management for portfolio monitoring, capital alllocation, and risk limits. When daily PD estimates oscillate wildly (for example 5-20 percent points as seen in the baseline model), they become operationally unusable. This is why smoothing techniques are applied in practice to reduce noise while preserving signal.

2.2.3 Mathematical Formulation of Improvement

The exponential weighted moving average (EWMA), the smoothing formula we are using, has the following properties.

- Parameter α controls reactivity: lower α yields more smoothing
- No look-ahead bias (uses only past observations)
- Preserves long-term trends while filtering short-term noise
- Weights decay exponentially: observation k periods ago has weight $\alpha(1-\alpha)^k$

Diving into the math, property 1 of the EWMA smoothing formula which is equation (10) can be demonstrated by unrolling the recursion:

$$\mathrm{PD}_t^{\text{smooth}} = \alpha \sum_{k=0}^{\infty} (1-\alpha)^k \cdot \mathrm{PD}_{t-k}^{\text{raw}} \tag{11}$$

This shows how α controls the weight decay rate, thus the reactivity. For example, with $\alpha = 0.9$ the decay factor $(1-\alpha) = 0.1$ which means today would be 90%, yesterday 9%, two days ago 0.9%, etc. Showing that the higher alpha has high reactivity to recent changes since recent data dominates. Conversely, with $\alpha = 0.1$ the decay factor $(1-\alpha) = 0.9$ which means today would be 10%, yesterday 9%, two days ago 8.1%, etc. This shows that lower alpha has low reactivity since weights are more evenly spread out over time.

Property 2 is evident since the formula only uses $\mathrm{PD}_t^{\text{raw}}$ and $\mathrm{PD}_{t-1}^{\text{smooth}}$ which is computed from past raw PDs. Additionally the code implementation processes data chronologically to avoid look-ahead bias.

Property 3 is evident since the EWMA formula averages over past raw PDs, thus filtering out high-frequency noise while preserving long-term trends.

Property 4 is shown in equation (11) where the weight for observation k periods ago is $\alpha(1-\alpha)^k$, demonstrating the exponential decay.

2.2.4 Calibration Changes

Smoothing is applied post-calibration. The algorithm is:

1. Calibrate (V, σ_V) using baseline approach
2. Compute PD_t^{raw} from equation (6)
3. Apply EWMA smoothing to obtain $\text{PD}_t^{\text{smooth}}$

Note that asset values V, asset volatilities σ_V, and distance-to-default remain unchanged from the baseline model.

3 Calibration Methodology

3.1 Numerical Method

We solve the system of two non-linear equations for two unknowns using scipy's `fsolve` optimizer, which implements a modified Powell hybrid method—a quasi-Newton approach with adaptive damping.

System of Equations

$$F_1(V, \sigma_V) = \text{BlackScholes}(V, D, T, r, \sigma_V) - E = 0 \tag{12}$$

$$F_2(V, \sigma_V) = \Phi(d_1) \cdot \sigma_V \cdot V - \sigma_E \cdot E = 0 \tag{13}$$

Algorithm We use scipy's `fsolve` optimizer with the following parameters:

1. **Initialize:**

$$V_0 = E + D \quad \text{(balance sheet approximation)} \tag{14}$$

$$\sigma_{V,0} = \sigma_E \cdot \frac{E}{E + D} \quad \text{(de-levered volatility)} \tag{15}$$

2. **Solve:** Call `fsolve(equations, [V_0, σ_V,0], xtol=1e-6)`

   `fsolve` implements a modified Powell hybrid method that approximates the Jacobian via finite differences, uses a trust region approach with adaptive damping, and converges when the relative error in the solution update (Δx) is $||\Delta x|| < 10^{-6}$.

3. **Validate:** After convergence, check the convergence flag utilizing ier (integer flag that indicates sucess or failure of solver) `ier == 1`, the economic validity (asset value should be larger than equity value by a small margin) $V \geq 1.01E$, and reasonable range $0.01\% \leq \sigma_V \leq 200\%$.

3.2 Edge Case Handling

1. **Invalid inputs**: ($E \leq 0$, $\sigma_E \leq 0$, $D < 0$): reject observation
2. **Jacobian singularity**: reject observation (caught through code exception handling)
3. **Non-convergence**: reject observation (caught through `fsolve` ier flag)
4. **Economic Validity**: $V > 1.01E$ (covered in validation above)

4 Empirical Setup

Sample Description

- **Firms:** Apple (AAPL), JPMorgan Chase (JPM), Tesla (TSLA), ExxonMobil (XOM), Ford (F)
- **Period:** January 2 – December 30, 2020 (252 trading days)
- **Total observations:** 1,260

4.1 Implementation Details

4.1.1 Time to Maturity

We assume $T = 1.0$ year for all observations. This is justified as it is a standard horizon for credit risk assessment as well as consistent with industry practice.

4.1.2 Market Capitalization Scaling

The dataset provides equity as price per share, but the Merton model requires total market capitalization. We convert:

$$E = \text{price per share} \times \text{shares outstanding}/10^6 \quad \text{(millions)} \tag{16}$$

Table 1 shows shares outstanding adjusted for 2020 stock splits:

Table 1: Shares Outstanding (Post-2020 Stock Splits)

Firm	**Shares (billions)**	**Notes**
AAPL	17.00	Post 4-for-1 split (August 2020)
JPM	3.07	No split
TSLA	0.96	Post 5-for-1 split (August 2020)
XOM	4.23	No split
Ford	3.92	No split

Source: Official 10-K filings, 8-K filings, Press Releases, split-adjusted for consistency with Yahoo Finance price data.

4.1.3 Quarterly Debt Alignment

Debt is reported quarterly while equity prices are daily to solve this issue we use **forward-fill**. This entails using the most recent quarterly debt value for each date. The code implementation uses `pd.merge_asof()` with `direction='backward'`.

Assumption: Debt changes slowly relative to daily equity movements, allowing quarterly observations to take place for daily values. While this generally holds for stable periods, this assumption may be violated during crisis periods such as the March 2020 COVID crisis which is captured in our data. However, most of the variation we see in debt-to-equity ratios stems from equity price changes rather than debt changes which in turn limits the impact of this approximation.

4.1.4 Additional Assumptions

- Dividends ignored for simplicity
- Zero recovery rate at default (standard in Merton model)
- Risk-free rate approximated using daily 10 year Treasury yield (from FRED)

5 Baseline Model Diagnosis

5.1 Identified Weaknesses

The primary weakness identified in the baseline Merton model is default probability insability. All five firms show excessive PD volatility displayed in key metrics below:

Table 2 summarizes the stability metrics for the baseline model:

Table 2: Baseline Model PD Stability Metrics

Firm	**Mean PD**	**Std Dev**	**CV**
AAPL	0.10%	0.30%	3.03
JPM	3.90%	8.64%	2.22
TSLA	3.53%	6.07%	1.72
XOM	0.70%	1.75%	2.52
Ford	6.38%	10.78%	1.69

High Coefficient of Variation: All firms exhibit CV between 1.69 and 3.03, indicating that standard deviation is 1.7 to 3 times the mean PD. While no universal CV threshold exists in credit risk practice, CV values exceeding 1.0 indicate standard deviation exceeds the mean, representing a high signal-to-noise ratio in this context making it unsuitable for operational risk management.

Universal Instability Pattern The instability problem in Table 2 affects all firms regardless of risk level or leverage. Both the highest-risk firm (Ford, 6.38% mean PD) and the lowest-risk firm (AAPL, 0.10% mean PD) exhibit CV values substantially exceeding 1.0. This suggests an issue with the model structure rather than firm-specific calibration problems. Notably, AAPL produces the highest relative volatility ($CV = 3.03$) despite having minimal default risk, exemplifying perhaps that the baseline model's instability worsens for low-risk firms where small changes create large percentage swings. The consistency of $CV > 1.0$ across all firms confirm that the baseline implementation produces estimates where noise dominates signal.

Root Cause Analysis The baseline Merton model amplifies equity volatility through non-linear transformations and leverage effects, leading to unstable PD estimates !

5.2 Examples

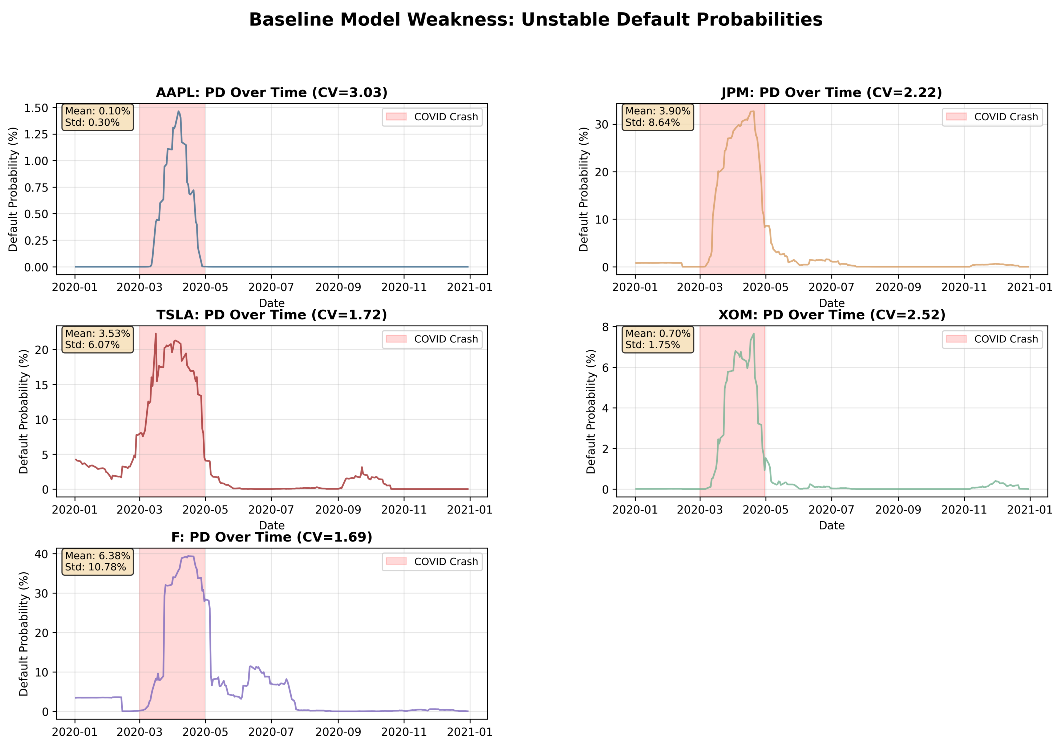

Figure 1: Baseline Model Default Probabilities Across 2020.

All five firms exhibit extreme PD spikes during the COVID-19 crisis (March–April 2020, shaded region), with peak increases of 6–14 times their annual mean values. For example we see that Ford reaches 40% PD (6.3× its mean), JPMorgan peaks at 32% (8.2× its mean), and even low-risk AAPL spikes to 1.4% (14× its mean). Additionally, these spikes reverse rapidly almost within 2-4 months which is inconsistent with typical credit cycles where rating changes occur over quarters and defaults unfold over years!

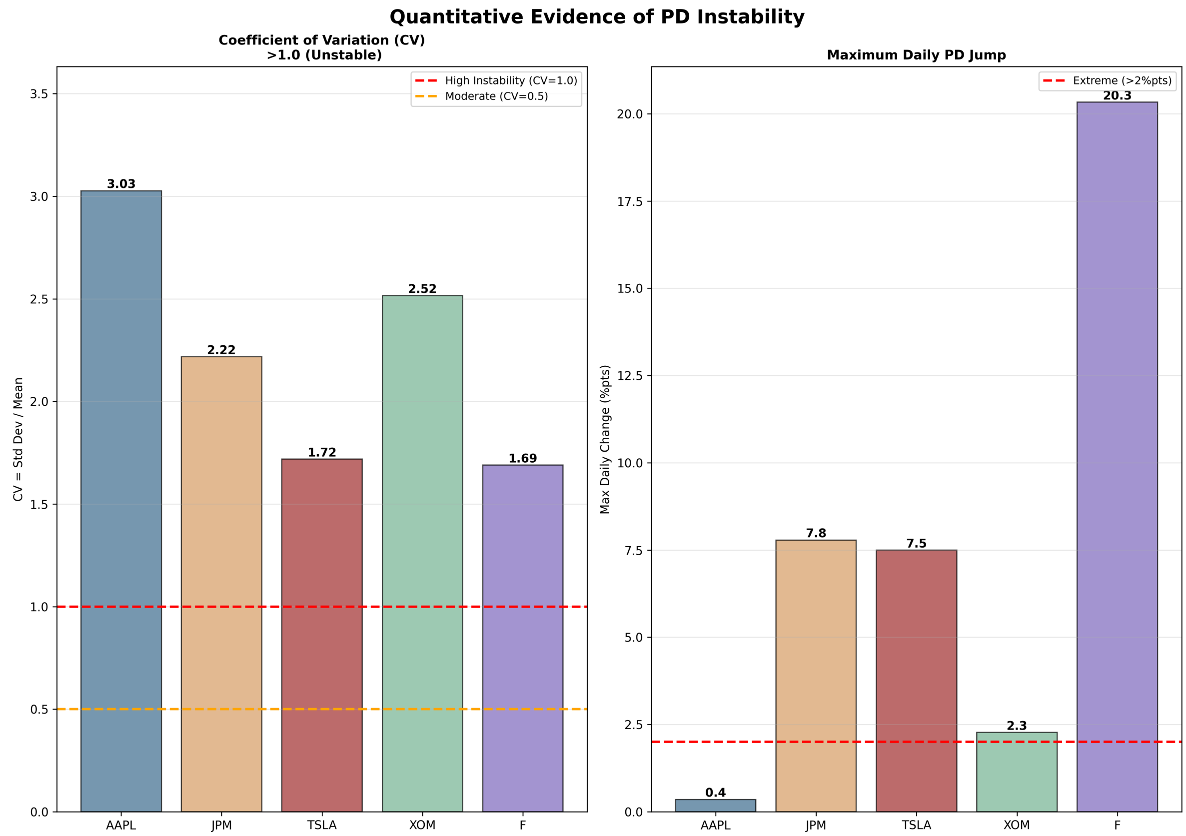

Figure 2: Quantitative Evidence of PD Instability in Baseline Model.

Figure 2 provides quantitative evidence of the instability problem through two metrics. The left panel shows that all firms exceed a CV of 1.0, indicating standard deviation exceeds the mean for all cases (operationally unacceptable, explained in further detail above). The right panel reveals extreme single-day PD movements, with four out of five firms exceeding the 2 percentage point reference line for operational viability. Ford's 20.3 percentage point single-day jump is almost 10 times this reference magnitude, while JPMorgan and Tesla show 7–8 percentage point jumps. These sort of extreme movements exemplify crisis amplification. Moreover, the contrasting patterns between panels reveal an important distinction between relative and absolute instability. Low-risk firms (AAPL, XOM) suffer from high relative volatility (CV $>$ 2.5) where noise dominates signal whereas high-risk firms (Ford, JPM, TSLA) suffer from large absolute movements ranging from 7.5 to 20.3 percentage points. Both forms of instability are problematic for risk management as they lead to unreliable PD estimates that cannot be used for portfolio monitoring or capital allocation. Even the best-performing case (XOM with CV = 2.52 and max jump = 2.3%pts) exceeds our reference thresholds, indicating no firm in the sample produces baseline estimates suitable for risk management without smoothing.