

A Research Program for State-Dependent DTS Scaling in Corporate Credit Spreads: Integrating Structural Model Priors with Empirical Estimation

Bernd J. Wuebben
AllianceBernstein, New York
bernd.wuebben@alliancebernstein.com

December 3, 2025

Abstract

This document presents a comprehensive research program for enhancing the Duration-Times-Spread (DTS) framework by integrating theoretical predictions from structural credit models with empirical estimation. Recent theoretical work [Wuebben \[2025\]](#) demonstrates that proportional spread movements fail systematically across maturities even for investment-grade bonds, with 1-year bonds exhibiting 4–6× higher percentage spread sensitivity than 10-year bonds from the same issuer. These structural model predictions provide strong priors for the state-dependent elasticity $\lambda_{i,t}$ that we seek to estimate empirically.

The research program proceeds sequentially: (1) **Raw validation** testing structural model predictions through a three-pronged approach—bucket-level analysis, within-issuer tests, and sector interaction analysis—before any complex regression modeling; (2) **Establishing variation** documenting that DTS betas differ systematically across bonds; (3) **Theory testing** evaluating whether Merton predictions explain observed variation; (4) **Time-variation analysis** assessing stability of relationships; (5) **Robustness checks** examining tail behavior, shock decomposition, and liquidity effects; (6) **Production specification** selecting parsimonious model using hierarchical testing framework.

Key improvements over standard approaches: (i) separation of “is there variation?” (Stage A) from “does theory explain it?” (Stage B); (ii) three complementary Stage 0 methodologies—bucket-level for aggregate patterns, within-issuer for pure maturity effects holding credit quality constant, and sector interactions for formal inference about industry-specific deviations; (iii) hierarchical model selection guided by theory rather than atheoretical horse-racing; (iv) explicit treatment of unbalanced panels as methodological choice, not separate research stage.

The research delivers both academic and practical value: academically, we provide the first comprehensive empirical test of structural model spread dynamics using market-wide bond data with proper identification of within-issuer effects; practically, we determine whether simpler Merton-based adjustment factors suffice or whether complex empirical estimation with sector-specific refinements is necessary for production risk systems.

Structure: Sections 1–2 cover objectives and data construction (including issuer identification methodology critical for within-issuer analysis). Section 3 introduces structural model priors. Section 4 (Stage 0) conducts raw validation using three complementary approaches: bucket-level analysis, within-issuer tests, and sector interaction analysis. Sections 5–6 (Stages A–B) establish and explain cross-sectional variation. Sections 7–11 cover time-variation, robustness, production specification selection, and research task summary.

Contents

1 Objectives and High-Level Framework	7
1.1 Motivation and constraints	7
1.2 Core modeling equation and DTS anchor	7
1.3 Universe split: IG and HY	8
1.4 Integration with structural model theory	8
1.5 Sequential research design	9
2 Data Construction and Sample Filters	9
2.1 Sample period	9
2.2 Universe definition	10
2.3 Expected sample size	10
2.4 Core variables	10
2.5 Time frequency and return definitions	11
2.6 Panel structure and econometric approach	11
2.6.1 Unbalanced panel by design	11
2.7 Issuer identification methodology	12
2.7.1 Primary identification approach	12
2.7.2 Edge cases requiring special treatment	12
2.7.3 Validation procedure	12
2.7.4 Sample filters for within-issuer analysis	13
3 Structural Model Priors for $\lambda_{i,t}$	13
3.1 The Merton framework and spread elasticity	13
3.2 Adjustment factors from theory	14
3.2.1 Maturity adjustment factors	14
3.2.2 Credit quality adjustment factors	14
3.3 Regime-specific predictions	15
3.4 Empirical testing strategy	15
4 Stage 0: Raw Validation Using Three Complementary Approaches	16
4.1 Objective and overview	16
4.2 Bucket-Level Analysis	17
4.2.1 Methodology	17
4.2.2 Statistical tests for bucket-level analysis	18
4.2.3 Deliverables for bucket-level analysis	19
4.3 Within-Issuer Analysis	19
4.3.1 Theoretical motivation	19
4.3.2 Empirical specification	20
4.3.3 Pooled estimation across issuer-weeks	20
4.3.4 Sample construction	21
4.3.5 Concrete example: Illustrating within-issuer vs bucket-level	21
4.3.6 Diagnostic analyses	22
4.3.7 Deliverables for within-issuer analysis	23
4.4 Sector Interaction Analysis	23
4.4.1 Motivation: Descriptive vs inferential	23
4.4.2 The sector interaction specification	24

4.4.3	Statistical tests	24
4.4.4	Economic rationale for sector differences	25
4.4.5	Specification variants	25
4.4.6	Implementation considerations	26
4.4.7	Deliverables for sector interaction analysis	27
4.5	Synthesis and Integration	27
4.5.1	Decision tree for Stage 0 conclusion	27
4.5.2	Interpretation matrix	28
4.5.3	Resolving apparent contradictions	28
4.5.4	Complementary insights	29
4.6	Combined Deliverables for Stage 0	30
4.6.1	Tables	30
4.6.2	Figures	30
4.6.3	Written summary	31
4.7	Decision Point	31
5	Stage A: Establish Cross-Sectional Variation	33
5.1	Objective	33
5.2	Specification A.1: Bucket-level betas	33
5.3	Specification A.2: Continuous characteristics	33
5.4	Establishing the stylized fact	34
5.5	Deliverables for Stage A	34
5.6	Decision point	35
6	Stage B: Does Merton Explain the Variation?	35
6.1	Objective	35
6.2	Specification B.1: Merton as offset (constrained)	36
6.3	Specification B.2: Decomposed components	36
6.4	Specification B.3: With sector adjustments	36
6.5	Benchmarking against Stage A	37
6.6	Specification B.4: Unrestricted for comparison	37
6.7	Theory vs reality table	37
6.8	Deliverables for Stage B	38
6.9	Decision point	39
7	Stage C: Does Static Merton Suffice or Do We Need Time-Variation?	40
7.1	Objective and prerequisite	40
7.2	Rolling window stability test	40
7.3	Stability of sector effects	41
7.4	Visual stability assessment	41
7.5	Conditional on instability: Macro driver analysis	41
7.6	Maturity-specific time-variation	42
7.7	Sector-specific time-variation	42
7.8	Practical implication assessment	42
7.9	Deliverables for Stage C	43
7.10	Decision point	45

8 Stage D: Robustness and Extensions	45
8.1 Objective and positioning	45
8.2 D.1: Tail behavior (quantile regression)	45
8.2.1 Motivation	45
8.2.2 Specification	46
8.2.3 Sector-specific tail behavior	46
8.2.4 Practical implications	46
8.2.5 Deliverables for D.1	47
8.3 D.2: Shock decomposition	47
8.3.1 Motivation	47
8.3.2 Factor construction	47
8.3.3 Multi-factor regression with Merton baseline	48
8.3.4 Diagnostic patterns	48
8.3.5 Connection to Stage 0 sector findings	48
8.3.6 Deliverables for D.2	49
8.4 D.3: Liquidity adjustment	49
8.4.1 Motivation	49
8.4.2 Liquidity-adjusted spread construction	49
8.4.3 Re-estimate Merton fit on default component	50
8.4.4 Diagnostic patterns	50
8.4.5 Sector-specific liquidity effects	50
8.4.6 Practical decision rule	50
8.4.7 Deliverables for D.3	51
8.5 Summary of Stage D	52
9 Stage A: Establish Cross-Sectional Variation	52
9.1 Objective	52
9.2 Specification A.1: Bucket-level betas	52
9.3 Specification A.2: Continuous characteristics	53
9.4 Establishing the stylized fact	53
9.5 Deliverables for Stage A	54
9.6 Decision point	55
10 Stage B: Does Merton Explain the Variation?	55
10.1 Objective	55
10.2 Specification B.1: Merton as offset (constrained)	55
10.3 Specification B.2: Decomposed components	56
10.4 Specification B.3: With sector adjustments	56
10.5 Benchmarking against Stage A	56
10.6 Specification B.4: Unrestricted for comparison	57
10.7 Theory vs reality table	57
10.8 Deliverables for Stage B	58
10.9 Decision point	59
11 Stage C: Does Static Merton Suffice or Do We Need Time-Variation?	60
11.1 Objective and prerequisite	60
11.2 Rolling window stability test	60
11.3 Stability of sector effects	60

11.4	Visual stability assessment	61
11.5	Conditional on instability: Macro driver analysis	61
11.6	Maturity-specific time-variation	62
11.7	Sector-specific time-variation	62
11.8	Practical implication assessment	62
11.9	Deliverables for Stage C	63
11.10	Decision point	64
12	Stage D: Robustness and Extensions	65
12.1	Objective and positioning	65
12.2	D.1: Tail behavior (quantile regression)	65
12.2.1	Motivation	65
12.2.2	Specification	65
12.2.3	Sector-specific tail behavior	66
12.2.4	Practical implications	66
12.2.5	Deliverables for D.1	66
12.3	D.2: Shock decomposition	67
12.3.1	Motivation	67
12.3.2	Factor construction	67
12.3.3	Multi-factor regression with Merton baseline	67
12.3.4	Diagnostic patterns	68
12.3.5	Connection to Stage 0 sector findings	68
12.3.6	Deliverables for D.2	68
12.4	D.3: Liquidity adjustment	69
12.4.1	Motivation	69
12.4.2	Liquidity-adjusted spread construction	69
12.4.3	Re-estimate Merton fit on default component	69
12.4.4	Diagnostic patterns	69
12.4.5	Sector-specific liquidity effects	70
12.4.6	Practical decision rule	70
12.4.7	Deliverables for D.3	70
12.5	Summary of Stage D	71
13	Stage E: Production Specification Selection	72
13.1	Objective	72
13.2	Decision framework: Hierarchical testing	72
13.3	Summary of specification hierarchy	75
13.4	Out-of-sample validation	76
13.4.1	Methodology	76
13.4.2	Performance metrics	76
13.4.3	Regime-specific performance	77
13.4.4	Sector-specific performance	77
13.5	Recommended approach	77
13.6	Deliverables for Stage E	78
14	Summary of Research Tasks	82
14.1	Resource requirements	87

15 Conclusion	88
15.1 Methodological contributions	88
15.2 Practical contributions	88
15.3 Academic contributions	89
15.4 Expected outcomes	89
15.5 Broader implications	90
15.6 Limitations and extensions	90
15.7 Final perspective	91

1 Objectives and High-Level Framework

1.1 Motivation and constraints

In practice, DTS is already embedded in risk systems, attribution frameworks, and portfolio construction for credit portfolios. Any enhanced modeling of spread dynamics must therefore:

- remain *DTS-centric*: spreads and returns should be modeled as functions of $DTS_{i,t}$ and index-level DTS factors;
- produce enhancements in the form of a *multiplicative adjustment* $\lambda_{i,t}$, rather than replacing DTS altogether;
- be implementable and interpretable for risk management, performance attribution, and relative-value work.

The empirical question is:

For a given macro environment, sector, rating, maturity, liquidity state, and idiosyncratic issuer profile, how do exogenous shocks—macro, sectoral, and issuer-specific—translate into spread movements, and how should these sensitivities be scaled relative to standard DTS?

1.2 Core modeling equation and DTS anchor

We work in discrete time at observation dates $t \in \{1, \dots, T\}$. For each bond i in universe U (IG or HY), denote:

- $OAS_{i,t}$: option-adjusted spread;
- $OASD_{i,t}$: OAS-based spread duration;
- $DTS_{i,t} = OAS_{i,t} \cdot OASD_{i,t}$: standard DTS;
- $OAS_t^{(U)}$: index-level OAS for universe U (IG or HY);
- $f_{DTS,t}^{(U)} = \frac{\Delta OAS_t^{(U)}}{OAS_{t-1}^{(U)}}$: index-level relative spread change, our DTS factor return.

We define bond-level relative spread changes as

$$y_{i,t} \equiv \frac{\Delta OAS_{i,t}}{OAS_{i,t-1}} = \frac{OAS_{i,t} - OAS_{i,t-1}}{OAS_{i,t-1}}. \quad (1)$$

The canonical DTS assumption for bond i in universe U is

$$y_{i,t} \approx f_{DTS,t}^{(U)} + \varepsilon_{i,t}, \quad (2)$$

up to idiosyncratic error terms. We *retain* the DTS factor-centric structure, but we generalize it to permit a state-dependent elasticity $\lambda_{i,t}$:

$$y_{i,t} \approx \lambda_{i,t} f_{DTS,t}^{(U)} + \varepsilon_{i,t}. \quad (3)$$

In excess-return space, letting $r_{i,t}^e$ denote the excess return of bond i over a duration-matched Treasury or swap hedge, we want to write

$$r_{i,t}^e \approx OAS_{i,t-1} \Delta t + DTS_{i,t} \lambda_{i,t} f_{DTS,t}^{(U)} + \text{idiosyncratic term}. \quad (4)$$

The research program is thus centered on specifying, estimating, and validating $\lambda_{i,t}$ as a function of observable bond characteristics and state variables.

1.3 Universe split: IG and HY

We can define the following two universes:

1. $U = \text{IG}$: the Bloomberg Barclays U.S. Corporate Investment Grade Index; and
2. $U = \text{HY}$: the Bloomberg Barclays U.S. High Yield 2% Issuer Cap Index.

It is not clear as to whether to pool IG and HY in estimation. To start with we pool IG and HY bond observations. The aim is to produce a unified λ specification for IG and HY, with potentially different parameterizations and state dependencies. Depending on findings we may have to split the universe into IG and HY bonds.

1.4 Integration with structural model theory

This research program explicitly incorporates predictions from structural credit models (specifically the Merton framework) as theoretical priors for $\lambda_{i,t}$. Following [Wuebben \[2025\]](#), structural credit theory establishes that:

1. **Cross-maturity proportionality fails severely in IG:** At spread levels below 300 bps, bonds with 1-year maturity have 4–6× higher percentage spread sensitivity than 10-year bonds from the same issuer.
2. **Maturity effects dominate credit quality effects:** In investment-grade markets, cross-maturity elasticity ratios of 400–500% dwarf same-maturity credit quality variations of 20–35%.
3. **Functional form predictions:** The Merton model implies specific functional forms for λ as functions of spread level s , maturity T , and the ratio $R/(sT)$ where R captures the fraction of debt value at risk.
4. **Regime-dependent patterns:** Proportionality failures are most severe in investment-grade (spreads <300 bps, maturity dispersion >3 years), moderate in high-yield (300–1000 bps), and paradoxically improve in distressed markets (>1000 bps).

These theoretical predictions provide powerful benchmarks for empirical estimation. Rather than searching blindly for λ patterns in data, we can:

- Test whether observed elasticity ratios match Merton predictions
- Use structural adjustment factors as starting values for iterative estimation
- Decompose empirical λ into theory-consistent and residual components
- Identify systematic deviations suggesting additional mechanisms (liquidity, sentiment, technical factors)

The research program therefore operates on two parallel tracks throughout:

Track 1: Theory-Guided Estimation

Start from Merton-implied $\lambda^{\text{Merton}}(s_i, T_i)$ and test whether empirical data support these predictions. Estimate deviations and assess their economic significance.

Track 2: Unrestricted Estimation

Allow data to determine λ coefficients without structural constraints. Compare to Track 1 to assess incremental explanatory power of unconstrained specifications.

1.5 Sequential research design

The stages proceed hierarchically with explicit decision points:

Stage 0: Raw Validation (Three-Pronged Approach)

Test Merton predictions directly from spread changes before any regression analysis using three complementary methodologies:

- **Bucket-level analysis:** Aggregate patterns across rating \times maturity \times sector buckets
- **Within-issuer analysis:** Pure maturity effects holding credit quality constant
- **Sector interaction analysis:** Formal inference about industry-specific deviations

Decision: Does theory provide adequate baseline? Where does it need augmentation?

Stage A: Establish Variation

Document that DTS betas differ across bonds. *Decision: If no variation, stop—standard DTS sufficient.*

Stage B: Explain Variation

Test whether Merton predictions match observed beta patterns. *Decision: Pure Merton / Calibrated / Need flexibility?*

Stage C: Test Stability

Assess whether static λ suffices or time-variation required. *Decision: Add macro state or keep static?*

Stage D: Robustness

Only after Stages A–C conclusive: test tail behavior, shock decomposition, liquidity. *Diagnose where/why theory fails.*

Stage E: Production Specification

Hierarchical selection: Standard DTS \rightarrow Pure Merton \rightarrow Calibrated \rightarrow Empirical (with sector adjustments where warranted). *Stop at simplest adequate model.*

Key principle: Each stage conditional on previous results. Don't add complexity until simple models demonstrably fail.

2 Data Construction and Sample Filters

2.1 Sample period

Primary sample: January 2013 – October 2025

Rationale:

- Excludes 2008–2009 crisis (extreme outlier, liquidity breakdown when Merton least reliable)
- 12+ years provides multiple credit cycles (2013 taper tantrum, 2015–16 energy crisis, 2020 COVID, 2022 rate shock)

Robustness sample: Use ICE data from January 2002 – December 2024

Test whether results hold when including crisis period (separate analysis, not primary).

2.2 Universe definition

For each universe $U \in \{\text{IG}, \text{HY}\}$:

- At each date t , the cross-section $\mathcal{I}_{U,t}$ consists of all bonds that are constituents of the relevant Bloomberg Barclays index on that date.
- We accept the full index methodology: new issues enter when eligible; bonds exit upon maturity, call, downgrade, or when they fail inclusion criteria.
- There is *no artificial restriction* to fixed-coupon non-call bullet bonds beyond what the index methodology already imposes.

Consequently, the panel (i, t) is naturally *unbalanced*: many bonds appear only for part of the sample; new issues are frequent, particularly in IG. We *embrace* this unbalanced nature and design the econometric work to handle it explicitly (see Section 2.6).

2.3 Expected sample size

Based on Bloomberg Barclays index historical composition:

	IG	HY
Avg bonds per month	5,500	1,800
Issuers with 2+ bonds outstanding	800	250
Issuer-weeks for within-issuer tests	180,000	45,000

Table 1: Expected sample size for primary sample (2013–2025)

This provides adequate power for all Stage 0 tests:

- Bucket-level analysis benefits from large samples per bucket (5,000–20,000 bond-weeks for major buckets)
- Within-issuer analysis has 225,000+ issuer-weeks for pooled estimation
- Sector interaction analysis pools millions of bond-week observations across sectors

2.4 Core variables

For each bond $i \in \mathcal{I}_{U,t}$ and date t :

- $OAS_{i,t}$ (in basis points);
- $OASD_{i,t}$ (in years);
- $DTS_{i,t} = OAS_{i,t} \cdot OASD_{i,t}$;
- clean price, accrued interest, yield-to-worst;
- time to maturity $M_{i,t}$ (years);
- index sector classification (e.g. Industrials, Financials, Utilities, etc.);

- credit rating, mapped to coarse buckets (e.g. AAA/AA, A, BBB for IG; BB, B, CCC for HY);
- issue size (amount outstanding), floating vs fixed, seniority;
- liquidity proxies where available: TRACE trading volume, bid-ask spread, number of trades, turnover.

For the index level (per universe U):

- $OAS_t^{(U)}$ and $OASD_t^{(U)}$, to define $DTS_t^{(U)}$ and $f_{DTS,t}^{(U)}$;
- sector- and rating-specific index OAS series $OAS_{s,t}^{(U)}$, $OAS_{r,t}^{(U)}$ for sector s and rating r (for Stage D decomposition).

2.5 Time frequency and return definitions

We start with *weekly* frequency to mitigate microstructure noise:

- Define a weekly observation grid t_1, t_2, \dots, t_T as e.g. Fridays or the last trading day of each week.
- Construct $OAS_{i,t}$ and $OAS_t^{(U)}$ on that grid by taking end-of-day values.

Define:

$$y_{i,t} \equiv \frac{\Delta OAS_{i,t}}{OAS_{i,t-1}} = \frac{OAS_{i,t} - OAS_{i,t-1}}{OAS_{i,t-1}}, \quad (5)$$

$$f_{DTS,t}^{(U)} \equiv \frac{\Delta OAS_t^{(U)}}{OAS_{t-1}^{(U)}} = \frac{OAS_t^{(U)} - OAS_{t-1}^{(U)}}{OAS_{t-1}^{(U)}}. \quad (6)$$

Robustness: Re-run key specifications (Stages A–B) at daily and monthly frequencies to assess sensitivity to horizon.

2.6 Panel structure and econometric approach

2.6.1 Unbalanced panel by design

The Bloomberg Barclays indices have substantial turnover:

- New issues enter monthly (IG: $\sim 50\text{--}100/\text{month}$, HY: $\sim 20\text{--}50/\text{month}$)
- Bonds mature, are called, or exit via downgrade/upgrade
- Average bond stays in IG index 6.5 years, HY index 4.2 years

We **embrace** this unbalanced structure rather than restricting to balanced sub-panels, which would:

- Eliminate 60%+ of observations
- Introduce survivorship bias (old bonds differ systematically from new)
- Miss new issue dynamics (relevant for practitioners)

2.7 Issuer identification methodology

Proper issuer identification is **critical** for the within-issuer analysis in Stage 0. We must correctly aggregate bonds to their Ultimate Parent to isolate pure maturity effects from cross-issuer heterogeneity.

2.7.1 Primary identification approach

- Use **Ultimate Parent ID** from Bloomberg or other data provider
- Match on **Seniority** (only compare senior unsecured to senior unsecured, subordinated to subordinated, etc.)
- Exclude subsidiaries with materially different credit profiles
- Manual validation for large issuers with complex corporate structures

2.7.2 Edge cases requiring special treatment

1. **Holding company vs operating subsidiary bonds:** Treat separately unless explicit guarantee exists. For example, separate bank holding company debt from bank operating company debt.
2. **Guaranteed debt:** Include in parent's bond set if guarantee is from Ultimate Parent and unconditional.
3. **Secured vs unsecured debt:** Analyze separately—different recovery rates imply different Merton elasticities even for same issuer.
4. **Acquisition/merger situations:** When issuer A acquires issuer B, bonds may trade differently during transition period. Flag issuer-weeks within 6 months of major corporate actions.
5. **Multi-national issuers:** Some issuers have bonds in multiple currencies or jurisdictions. Focus on USD-denominated bonds in primary analysis; cross-currency effects are a separate robustness check.

2.7.3 Validation procedure

For the 50 largest issuers by amount outstanding (covering approximately 40% of index market value):

1. Manual review of corporate structure
2. Verification that all bonds sharing Ultimate Parent ID have similar credit profiles
3. Documentation of any bonds excluded due to structural subordination or guarantee differences
4. Cross-check against rating agency issuer hierarchies

2.7.4 Sample filters for within-issuer analysis

For each issuer-week (i, t) , include in within-issuer analysis if:

1. $J_{i,t} \geq 2$ (at least two bonds outstanding from same Ultimate Parent + Seniority)
2. All bonds have valid OAS, maturity, and spread data
3. Maturity dispersion ≥ 2 years: $\max_j T_{ij} - \min_j T_{ij} \geq 2$
4. No bonds within 1 year of maturity (avoid pull-to-par distortions and front-end liquidity effects)
5. Spread changes $|\Delta s_{ij}/s_{ij}| < 200\%$ (exclude extreme outliers likely due to data errors or restructuring events)

3 Structural Model Priors for $\lambda_{i,t}$

Before proceeding to empirical estimation, we establish theoretical benchmarks from structural credit models that provide strong priors for the functional form and magnitude of $\lambda_{i,t}$. The theoretical framework developed in [Wuebbken \[2025\]](#) provides precise predictions for when and why proportional spread movements fail.

3.1 The Merton framework and spread elasticity

In the Merton (1974) structural model, credit spreads arise from the embedded put option in risky debt. The key quantity governing spread dynamics is the **elasticity** of the spread with respect to firm value:

$$\varepsilon_i \equiv \frac{\partial s_i}{\partial V} \cdot \frac{V}{s_i} \quad (7)$$

This measures the percentage change in spread for a 1% change in firm value. For bonds from the same issuer experiencing a common firm value shock $\Delta V/V$, percentage spread changes are:

$$\frac{\Delta s_i}{s_i} \approx \varepsilon_i \cdot \frac{\Delta V}{V} \quad (8)$$

Therefore, proportional spread movements require $\varepsilon_i = \varepsilon_j$ for all bonds i, j . The Merton model delivers an exact formula:

$$\varepsilon_i = -\frac{R_i}{T_i \cdot s_i} \quad (9)$$

where T_i is maturity, s_i is spread, and R_i is the ratio:

$$R_i = \frac{e^{x_i+rT_i} N(-d_{1,i})}{N(d_{2,i}) + e^{x_i+rT_i} N(-d_{1,i})} \quad (10)$$

with $x_i = \ln(V_i/D_i)$ the log-leverage ratio and $d_{1,i}, d_{2,i}$ the standard Black-Scholes parameters.

Economic interpretation: R_i measures the fraction of debt value coming from default states—the “value at risk.” For IG bonds, $R \approx 0.05–0.30$; for HY, $R \approx 0.30–0.80$; for distressed, $R \rightarrow 1$.

See [Wuebbken \[2025\]](#) for detailed derivations and calibrations of these relationships.

3.2 Adjustment factors from theory

Define the adjustment factor for bond i relative to a reference bond j as the ratio of elasticities:

$$\lambda_{i,j} \equiv \frac{\varepsilon_i}{\varepsilon_j} = \frac{R_i \cdot T_j \cdot s_j}{R_j \cdot T_i \cdot s_i} \quad (11)$$

For practical implementation, we decompose into **maturity adjustments** (same issuer, different T) and **credit quality adjustments** (different issuers, same T).

3.2.1 Maturity adjustment factors

For bonds from the same issuer with reference maturity $T_{\text{ref}} = 5$ years:

$$\lambda_T(T; s^*) = \frac{R(T, s^*) \cdot T_{\text{ref}} \cdot s(T_{\text{ref}}, s^*)}{R(T_{\text{ref}}, s^*) \cdot T \cdot s(T, s^*)} \quad (12)$$

where s^* is the calibrated spread at T_{ref} determining the issuer's leverage.

Key theoretical predictions:

1. At IG spread levels (50–300 bps), $\lambda_T(1y) \approx 2.4\text{--}3.6$ (short bonds much more sensitive)
2. At IG spread levels, $\lambda_T(10y) \approx 0.6\text{--}0.7$ (long bonds less sensitive)
3. Adjustment factors converge toward 1.0 as spreads widen to HY levels
4. The cross-maturity effect dominates: elasticity ratio 1y/10y ranges from 5.9 at 50 bps to 1.56 at 1000 bps

Table 2 summarizes theoretical predictions.

Table 2: Theoretical Maturity Adjustment Factors λ_T from Merton Model. Source: [Wuebbken \[2025\]](#).

Spread Level (bps)	Maturity Adjustment					$\lambda_T(T; 5y)$
	1y	3y	5y	7y	10y	
50 (AAA)	3.62	1.47	1.00	0.79	0.61	
100 (AA)	3.27	1.42	1.00	0.80	0.64	
200 (A-)	2.78	1.36	1.00	0.82	0.67	
300 (BBB-)	2.40	1.30	1.00	0.84	0.70	
500 (BB)	1.91	1.25	1.00	0.86	0.73	
1000 (B-)	1.26	1.12	1.00	0.91	0.81	
2000 (CCC)	1.08	1.05	1.00	0.97	0.93	

3.2.2 Credit quality adjustment factors

For bonds with same maturity $T = 5y$ but different spread levels, relative to reference $s_{\text{ref}} = 100$ bps:

$$\lambda_s(s; 100) = \frac{R(s) \cdot 100}{R(100) \cdot s} \quad (13)$$

Key theoretical predictions:

1. IG bonds safer than 100 bps have $\lambda_s > 1$ (higher sensitivity)
2. IG bonds riskier than 100 bps have $\lambda_s < 1$ (lower sensitivity)
3. Effect is moderate in IG: from 50 to 300 bps, λ_s varies from 1.15 to 0.75 (only 35% range)
4. Power-law approximation: $\lambda_s(s) \approx (s/100)^{-0.25}$ achieves $R^2 = 0.92$

Table 3 summarizes.

Table 3: Theoretical Credit Quality Adjustment Factors λ_s from Merton Model. Source: [Wuebben \[2025\]](#).

Spread (bps)	Exact Merton	Power Law $(s/100)^{-0.25}$
50	1.145	1.189
100	1.000	1.000
200	0.847	0.841
300	0.746	0.760
500	0.635	0.669
1000	0.468	0.562
2000	0.299	0.473

3.3 Regime-specific predictions

The Merton model predicts distinct behavior across five regimes [Wuebben \[2025\]](#):

Regime 1: IG with narrow maturity range ($s < 300$ bps, $\Delta T < 2y$): Standard DTS works reasonably. Maturity effects not applicable. Credit quality variation causes 20–35% deviation—acceptable.

Regime 2: IG with wide maturity range ($s < 300$ bps, $\Delta T > 3y$): Primary failure mode. Cross-maturity λ ratios of 3–6× create 300–500% deviations. This is where DTS models systematically fail.

Regime 3: HY with narrow maturity range ($300 < s < 1000$ bps, $\Delta T < 2y$): Cross-maturity effects reduced to 50–160% but still substantial. Same-maturity credit quality variation increases to 40–73%.

Regime 4: HY with wide maturity range ($300 < s < 1000$ bps, $\Delta T > 3y$): Both effects large. Comprehensive adjustments required.

Regime 5: Distressed ($s > 1000$ bps): Proportionality paradoxically improves. Both cross-maturity and same-maturity deviations decline.

3.4 Empirical testing strategy

The structural model priors suggest a specific empirical testing sequence:

Empirical Questions for Each Stage:

These questions test the theoretical predictions established in [Wuebben \[2025\]](#):

1. **Do raw data exhibit Merton-predicted patterns?** (Stage 0) Before any regression, compute elasticity ratios and compare to theory using three complementary approaches:
 - Bucket-level: Do aggregate patterns match?
 - Within-issuer: Does the mechanism work at issuer level?
 - Sector interactions: Are there systematic industry deviations?
2. **Is there cross-sectional variation to explain?** (Stage A) Document that betas differ. If not, standard DTS sufficient.
3. **Does Merton explain the variation?** (Stage B) Test whether theoretical λ matches empirical patterns.
4. **Is the relationship stable over time?** (Stage C) If Stage B works, test whether static λ suffices.
5. **Where does theory fail?** (Stage D) Diagnose failures: tails? liquidity? specific regimes?
6. **What production spec is adequate?** (Stage E) Hierarchical selection guided by theory, with sector adjustments where warranted.

4 Stage 0: Raw Validation Using Three Complementary Approaches

4.1 Objective and overview

Stage 0 provides assumption-free tests of Merton predictions before any complex regression analysis. We ask: *Do bonds exhibit sensitivities consistent with structural theory, and where do systematic deviations occur?*

To answer this comprehensively, we employ three complementary methodologies:

1. **Bucket-level analysis:** Groups bonds by observable characteristics (rating \times maturity \times sector) and estimates aggregate DTS sensitivities. Provides strong statistical power and tests whether Merton predictions hold on average across the market.
2. **Within-issuer analysis:** Examines bonds from the same issuer but different maturities, holding credit quality constant. Provides the cleanest test of Merton's cross-maturity predictions by isolating pure maturity effects.
3. **Sector interaction analysis:** Pools all bonds and formally tests whether sectors differ systematically in DTS sensitivity. Distinguishes real structural differences from sampling noise and identifies where sector-specific adjustments may be necessary.

Recommended workflow:

1. Run bucket-level analysis to establish aggregate patterns
2. Run within-issuer analysis to validate Merton predictions hold at issuer level

Table 4: Complementary Roles of Three Stage 0 Approaches

Approach	Strength	Limitation
Bucket-level	Large samples per bucket; strong statistical power; tests aggregate patterns	Mixes different issuers; cannot isolate pure maturity effect; descriptive sector comparisons only
Within-issuer	Cleanest Merton test; same credit quality; pure maturity variation	Smaller samples per issuer-week; requires multiple bonds per issuer; cannot test cross-issuer patterns
Sector interaction	Formal significance tests; distinguishes real from noise; identifies systematic deviations	Requires pooling across buckets; assumes common structure; less powerful per-sector

3. Run sector interaction tests to determine if sector-specific adjustments are necessary
4. Synthesize findings across all three approaches for Stage 0 conclusion

4.2 Bucket-Level Analysis

4.2.1 Methodology

Step 1: Define buckets

Define buckets k by:

- **Rating:** AAA/AA, A, BBB for IG; BB, B, CCC for HY
- **Maturity:** 1–2y, 2–3y, 3–5y, 5–7y, 7–10y, 10y+
- **Sector:** Industrial, Financial, Utility, Energy

For each bucket k , compute representative characteristics using bonds in that bucket:

- \bar{T}_k : median maturity
- \bar{s}_k : median spread
- $\lambda_k^{\text{Merton}} = \lambda_T(\bar{T}_k; 5y, \bar{s}_k) \times \lambda_s(\bar{s}_k; 100)$ using Tables 2–3

This creates $3 \times 6 \times 4 = 72$ buckets for IG and 72 for HY (some may be sparse or empty).

Step 2: Pooled regression by bucket

For each bucket k , estimate using all bond-week observations in that bucket:

$$y_{i,t} = \alpha^{(k)} + \beta^{(k)} \cdot f_{DTS,t}^{(U)} + \varepsilon_{i,t}^{(k)}, \quad i \in \text{bucket } k \quad (14)$$

where:

- $y_{i,t} = \Delta OAS_{i,t}/OAS_{i,t-1}$: percentage spread change for bond i in week t
- $f_{DTS,t}^{(U)} = \Delta OAS_t^{(U)}/OAS_{t-1}^{(U)}$: index-level percentage spread change

- $\beta^{(k)}$: empirical sensitivity for bonds in bucket k

Estimation: Pooled OLS within each bucket.

Standard errors: Cluster by week to allow arbitrary correlation across bonds within each week.

Sample size: Each major bucket (e.g., BBB 3–5y Industrials) has 5,000–20,000 bond-week observations over 2013–2025, providing strong statistical power.

Step 3: Compare empirical to theoretical

For each bucket k , we compare:

- $\hat{\beta}^{(k)}$: empirical DTS sensitivity from regression
- $\lambda_k^{\text{Merton}}$: theoretical prediction from Merton model

4.2.2 Statistical tests for bucket-level analysis

Test 1 (Level): Is $\hat{\beta}^{(k)} \approx \lambda_k^{\text{Merton}}$?

Use t -test: $H_0 : \beta^{(k)} = \lambda_k^{\text{Merton}}$ with clustered standard errors.

Test 2 (Cross-maturity pattern): Across maturity buckets (holding rating and sector fixed), do empirical betas exhibit predicted monotonicity?

Merton prediction: $\hat{\beta}^{(1-2y)} > \hat{\beta}^{(3-5y)} > \hat{\beta}^{(5-7y)} > \hat{\beta}^{(10y+)}$

Test: Spearman rank correlation between empirical $\hat{\beta}^{(k)}$ and theoretical $\lambda_k^{\text{Merton}}$ across maturity buckets. Expected strong positive correlation (short bonds more sensitive).

Test 3 (Regime pattern): Does the pattern hold strongly in IG but weaken in HY as theory predicts?

Split sample by spread level:

- IG spreads (<300 bps): Expect large cross-maturity dispersion
- HY spreads (300–1000 bps): Expect moderate dispersion
- Distressed (>1000 bps): Expect convergence toward $\beta \approx 1$

Test whether cross-maturity standard deviation of $\hat{\beta}^{(k)}$ declines as spreads widen (Regime 5 prediction).

Aggregated test across all buckets:

Pool all buckets and test whether average deviation from Merton is zero:

$$\text{Mean Deviation} = \frac{1}{K} \sum_{k=1}^K (\hat{\beta}^{(k)} - \lambda_k^{\text{Merton}}) \quad (15)$$

Test $H_0 : \text{Mean Deviation} = 0$ using bootstrap standard errors (resampling weeks).

Variation test:

Does most of the cross-bucket variation in $\hat{\beta}^{(k)}$ come from maturity differences as Merton predicts?

Compute:

- Total variation: $\text{Var}(\hat{\beta}^{(k)})$ across all buckets
- Within-rating variation: $\text{Var}(\hat{\beta}^{(k)} | \text{rating})$ holding rating constant

If within-rating variation is large and dominated by maturity, this confirms Merton's emphasis on maturity effects.

4.2.3 Deliverables for bucket-level analysis

- **Table 0.1:** Bucket-level results for key buckets
 - Rows: Rating \times Maturity combinations (e.g., BBB 1–2y, BBB 3–5y, BBB 7–10y)
 - Columns: $\hat{\beta}^{(k)}$ (empirical), $\lambda_k^{\text{Merton}}$ (theory), Ratio, t -stat for $H_0 : \beta = \lambda$, Sample size
 - Separate panels for IG and HY
 - Highlight cells where ratio outside [0.8, 1.2]
- **Table 0.2:** Cross-maturity pattern tests
 - For each rating class, show $\hat{\beta}^{(k)}$ across maturity buckets
 - Spearman correlation with theoretical $\lambda_k^{\text{Merton}}$
 - Test whether monotonicity holds (short > medium > long)
- **Figure 0.1:** Scatter plot: Empirical $\hat{\beta}^{(k)}$ (y-axis) vs Theoretical $\lambda_k^{\text{Merton}}$ (x-axis) for all buckets
 - 45-degree line for perfect agreement
 - Point size proportional to sample size
 - Color-code by spread level: IG (blue), HY (orange), Distressed (red)
 - Annotate outliers
- **Figure 0.2:** Cross-maturity patterns by rating
 - Separate panels for AAA/AA, A, BBB, BB, B
 - X-axis: Maturity (1y, 3y, 5y, 7y, 10y)
 - Y-axis: $\hat{\beta}$ (empirical, solid line with points) and λ^{Merton} (theoretical, dashed line)
 - Shows whether empirical pattern matches theoretical prediction
- **Figure 0.3:** Regime patterns
 - X-axis: Average spread level of bucket
 - Y-axis: Cross-maturity dispersion (std dev of $\hat{\beta}^{(k)}$ across maturity buckets)
 - Shows whether dispersion declines as spreads widen (convergence to Regime 5)

4.3 Within-Issuer Analysis

4.3.1 Theoretical motivation

The Merton (1974) structural model makes its strongest predictions about the *same issuer's* bonds at different maturities. For issuer i with bonds $j = 1, \dots, J_i$ outstanding, all bonds share:

- Same firm value V_i
- Same asset volatility σ_i
- Same leverage ratio $L_i = D_i/V_i$
- Same recovery rate R_i

The **only** difference is time to maturity T_{ij} .

When a shock hits firm value $\Delta V_i/V_i$, the percentage spread change for bond j is:

$$\frac{\Delta s_{ij}}{s_{ij}} = \varepsilon_{ij} \cdot \frac{\Delta V_i}{V_i} + \text{idiosyncratic noise} \quad (16)$$

where ε_{ij} is the elasticity from equation (9).

Key prediction: The ratio of elasticities for two bonds from the same issuer depends *only* on their maturities and spreads:

$$\frac{\varepsilon_{i1}}{\varepsilon_{i2}} = \frac{R_{i1} \cdot T_{i2} \cdot s_{i2}}{R_{i2} \cdot T_{i1} \cdot s_{i1}} \equiv \lambda_{i1,i2} \quad (17)$$

This is the **purest test** of Merton because all issuer-specific factors cancel out.

4.3.2 Empirical specification

Within-issuer, within-week fixed effects

For each issuer i in week t with $J_{i,t} \geq 2$ bonds outstanding, estimate:

$$\frac{\Delta s_{ij,t}}{s_{ij,t-1}} = \alpha_{i,t} + \beta \cdot \lambda_{ij,t}^{\text{Merton}} + \varepsilon_{ij,t} \quad (18)$$

where:

- $\alpha_{i,t}$: Issuer-week fixed effect (absorbs common shock $\Delta V_i/V_i$)
- β : Coefficient on Merton adjustment factor
- $\lambda_{ij,t}^{\text{Merton}}$: Predicted elasticity from Merton model for bond j of issuer i

Identification: Variation comes *entirely* from cross-maturity differences within the same issuer in the same week.

Theory prediction: $\beta = 1$ (Merton correctly predicts elasticity ratios)

Null hypothesis: $H_0 : \beta = 1$

4.3.3 Pooled estimation across issuer-weeks

After estimating equation (18) separately for each issuer-week with sufficient bonds, pool results using inverse-variance weighting:

$$\hat{\beta}_{\text{pooled}} = \frac{\sum_{i,t} w_{i,t} \hat{\beta}_{i,t}}{\sum_{i,t} w_{i,t}} \quad (19)$$

where $w_{i,t} = 1/\text{se}(\hat{\beta}_{i,t})^2$.

Pooled standard error:

$$\text{se}(\hat{\beta}_{\text{pooled}}) = \frac{1}{\sqrt{\sum_{i,t} w_{i,t}}} \quad (20)$$

Test statistic:

$$t = \frac{\hat{\beta}_{\text{pooled}} - 1}{\text{se}(\hat{\beta}_{\text{pooled}})} \sim t(df) \quad (21)$$

where df is the total number of issuer-weeks minus parameters.

4.3.4 Sample construction

Using the issuer identification methodology from Section 2.7, for each issuer-week (i, t) , include if:

1. $J_{i,t} \geq 2$ (at least two bonds outstanding from same Ultimate Parent + Seniority)
2. All bonds have valid OAS, maturity, and spread data
3. Maturity dispersion ≥ 2 years: $\max_j T_{ij} - \min_j T_{ij} \geq 2$
4. No bonds within 1 year of maturity (avoid pull-to-par distortions)
5. Spread changes $|\Delta s_{ij}/s_{ij}| < 200\%$ (exclude extreme outliers)

Expected sample size:

Universe	Issuers with 2+ bonds	Weeks	Issuer-weeks
Investment Grade	800	600	180,000
High Yield	250	600	45,000
Total	1,050	600	225,000

Table 5: Expected sample size for within-issuer analysis

This provides substantial statistical power despite smaller samples per issuer-week.

4.3.5 Concrete example: Illustrating within-issuer vs bucket-level

Suppose Ford Motor Company has these bonds on March 15, 2020 (during COVID shock):

Bond	Maturity	Spread	λ^{Merton}	$\Delta s/s$
Ford 2021	1y	150 bps	3.27	+60%
Ford 2025	5y	180 bps	1.00	+25%
Ford 2030	10y	200 bps	0.64	+18%

Table 6: Hypothetical Ford bonds during COVID shock

Within-issuer test:

$$\text{Elasticity ratio (1y/5y)} = \frac{60\%/150}{25\%/180} = \frac{0.40}{0.139} = 2.88 \quad (22)$$

$$\text{Merton prediction} = \frac{\lambda_{1y}^{\text{Merton}}}{\lambda_{5y}^{\text{Merton}}} = \frac{3.27}{1.00} = 3.27 \quad (23)$$

$$\text{Match} = \frac{2.88}{3.27} = 0.88 \quad (88\% \text{ of predicted}) \quad (24)$$

This is a **clean test**: same credit event (Ford's deterioration), only maturity differs.

Bucket-level test (what Section 4.2 does):

Compare all 1-year BBB bonds to all 5-year BBB bonds, mixing:

- Ford, GM, Boeing, AT&T, etc. (different credit stories)
- Different sectors (Auto, Aerospace, Telecom have different shock exposures)
- Different liquidity profiles
- Different idiosyncratic news

This is a **noisier test**: confounds maturity effect with issuer heterogeneity.

4.3.6 Diagnostic analyses

Heterogeneity by spread level

Merton predicts elasticity ratios converge toward 1.0 as spreads widen (Regime 5). Test:

$$\hat{\beta}_{i,t} = \gamma_0 + \gamma_s \log(\bar{s}_{i,t}) + u_{i,t} \quad (25)$$

where $\bar{s}_{i,t}$ is the average spread across issuer i 's bonds in week t .

Prediction: $\gamma_s > 0$ (higher spreads $\Rightarrow \beta$ closer to 1)

Heterogeneity by maturity dispersion

When an issuer has very wide maturity dispersion (e.g., 1y and 10y bonds), Merton predictions should be strongest. Test:

$$|\hat{\beta}_{i,t} - 1| = \delta_0 + \delta_{\Delta T} (\max_j T_{ij} - \min_j T_{ij}) + v_{i,t} \quad (26)$$

Prediction: $\delta_{\Delta T} < 0$ (wider dispersion \Rightarrow smaller deviation from theory)

Intuition: When issuer has both 1y and 10y bonds, the 4–6× elasticity difference is stark and harder to miss empirically.

Crisis vs normal periods

Does within-issuer proportionality break down during crises (liquidity evaporation, correlation breakdowns)?

Split sample:

- Normal: VIX < 20
- Stress: VIX $\in [20, 30]$
- Crisis: VIX > 30

Estimate $\hat{\beta}_{\text{pooled}}$ separately for each regime. Test:

$$H_0 : \beta_{\text{normal}} = \beta_{\text{crisis}} \quad (27)$$

Alternative hypothesis: $\beta_{\text{crisis}} < 1$ (Merton over-predicts front-end sensitivity in crises due to liquidity freezes)

4.3.7 Deliverables for within-issuer analysis

- **Table 0.3:** Within-issuer pooled results
 - Rows: Full sample, IG only, HY only, Normal (VIX<20), Crisis (VIX>30)
 - Columns: $\hat{\beta}$, SE, t -stat for $H_0 : \beta = 1$, p -value, N (issuer-weeks)
- **Table 0.4:** Diagnostic regressions
 - Spread level effect: $\hat{\gamma}_s$, SE, interpretation
 - Maturity dispersion effect: $\hat{\delta}_{\Delta T}$, SE, interpretation
- **Figure 0.4:** Distribution of issuer-week estimates
 - Histogram of $\hat{\beta}_{i,t}$ across all issuer-weeks
 - Vertical line at $\beta = 1$ (theory prediction)
 - Shaded region for pooled estimate ± 2 SE
 - Annotate median, IQR, fraction within [0.8, 1.2]
- **Figure 0.5:** Spread level diagnostic
 - Scatter plot: X-axis = average issuer spread $\bar{s}_{i,t}$, Y-axis = $\hat{\beta}_{i,t}$
 - Overlay: LOWESS smooth
 - Prediction: Convergence toward $\beta = 1$ as spreads widen
- **Figure 0.6:** Case studies for major issuers
 - For 5–10 major issuers (e.g., Apple, Ford, AT&T, JPMorgan)
 - Time series of $\hat{\beta}_{i,t}$ over sample period
 - Identify weeks where $\hat{\beta}_{i,t}$ deviates sharply (crisis episodes?)

4.4 Sector Interaction Analysis

4.4.1 Motivation: Descriptive vs inferential

The bucket-level approach creates **separate regressions** for each sector:

$$\text{Industrial: } y_{i,t} = \alpha^{(\text{Ind})} + \beta^{(\text{Ind})} f_{DTS,t} + \varepsilon_{i,t} \quad (28)$$

$$\text{Financial: } y_{i,t} = \alpha^{(\text{Fin})} + \beta^{(\text{Fin})} f_{DTS,t} + \varepsilon_{i,t} \quad (29)$$

$$\text{Utility: } y_{i,t} = \alpha^{(\text{Util})} + \beta^{(\text{Util})} f_{DTS,t} + \varepsilon_{i,t} \quad (30)$$

This yields **point estimates** such as $\hat{\beta}^{(\text{Ind})} = 1.02$, $\hat{\beta}^{(\text{Fin})} = 1.35$, $\hat{\beta}^{(\text{Util})} = 0.98$.

Descriptive observation: “Financials have higher β than Industrials.”

Unanswered question: *Is this difference statistically significant, or just sampling noise?*

The sector interaction specification answers this formally.

4.4.2 The sector interaction specification

Pool all sectors together and model differences explicitly:

$$y_{i,t} = \alpha + \beta_0[\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \sum_{s \in S} \beta_s \cdot \mathbb{1}_{\{i \in s\}} \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t} \quad (31)$$

where:

- β_0 : Base DTS sensitivity (reference sector = Industrial)
- β_s : **Additional** sensitivity for sector s relative to reference
- $\mathbb{1}_{\{i \in s\}}$: Indicator = 1 if bond i is in sector s
- $S = \{\text{Financial, Utility, Energy, ...}\}$: Set of non-reference sectors

Interpretation:

- Total sensitivity for Industrial bonds: $\beta_{\text{Ind}} = \beta_0$
- Total sensitivity for Financial bonds: $\beta_{\text{Fin}} = \beta_0 + \beta_{\text{Financial}}$
- Total sensitivity for Utility bonds: $\beta_{\text{Util}} = \beta_0 + \beta_{\text{Utility}}$

4.4.3 Statistical tests

Test 1: Do sectors differ at all?

Joint null hypothesis:

$$H_0 : \beta_{\text{Financial}} = \beta_{\text{Utility}} = \beta_{\text{Energy}} = \dots = 0 \quad (32)$$

Use Wald test (F-test):

$$F = \frac{(RSS_{\text{restricted}} - RSS_{\text{unrestricted}})/q}{RSS_{\text{unrestricted}}/df} \sim F(q, df) \quad (33)$$

where q is the number of sector dummies.

Interpretation:

- Reject ($p < 0.05$): Sectors have significantly different DTS sensitivities
- Fail to reject: Can use common β across sectors (parsimony)

Test 2: Does specific sector differ from reference?

For Financial sector:

$$H_0 : \beta_{\text{Financial}} = 0 \quad (34)$$

Use t -test:

$$t = \frac{\hat{\beta}_{\text{Financial}}}{\text{se}(\hat{\beta}_{\text{Financial}})} \sim t(df) \quad (35)$$

Test 3: Do two non-reference sectors differ from each other?

For Financial vs Utility:

$$H_0 : \beta_{\text{Financial}} - \beta_{\text{Utility}} = 0 \quad (36)$$

Use Wald test on linear combination:

$$t = \frac{\hat{\beta}_{\text{Financial}} - \hat{\beta}_{\text{Utility}}}{\sqrt{\text{Var}(\hat{\beta}_{\text{Financial}}) + \text{Var}(\hat{\beta}_{\text{Utility}}) - 2\text{Cov}(\hat{\beta}_{\text{Financial}}, \hat{\beta}_{\text{Utility}})}} \quad (37)$$

4.4.4 Economic rationale for sector differences

Why Financials might differ:

1. **Correlation risk:** Banks and insurance companies have correlated credit risk. When one bank struggles, contagion affects entire financial sector. Merton assumes firm-specific risk, but financials have systemic component.
Prediction: $\beta_{\text{Financial}} > 0$ (amplified sensitivity, especially during crises)
2. **Regulatory capital effects:** Mark-to-market accounting for trading books and regulatory capital requirements force deleveraging when spreads widen, creating mechanical amplification beyond fundamental credit risk.
3. **Liquidity provision role:** Financials are liquidity providers in credit markets. When their own credit deteriorates, market-wide liquidity dries up, creating feedback loops.
4. **Balance sheet opacity:** Complex derivatives exposures and off-balance-sheet vehicles create information asymmetry greater than for industrial firms.

Why Utilities might differ:

1. **Regulatory protection:** Rate regulation provides cash flow stability. Essential service status limits default risk. Counter-cyclical: demand stable even during recessions.
Prediction: $\beta_{\text{Utility}} < 0$ (dampened sensitivity to macro shocks)
2. **Rate sensitivity:** Long-duration assets (power plants, transmission infrastructure). Credit quality tied to regulatory environment, not economic cycles.
3. **Capital intensity:** High fixed costs, low variable costs. Leverage ratios higher than industrials, but leverage is accepted by regulators due to stable cash flows.

Why Energy might differ:

1. **Commodity price pass-through:** Energy bond spreads driven primarily by oil/gas prices. This is idiosyncratic to energy, not captured by broad DTS factor.
Prediction: β_{Energy} lower (idiosyncratic > systematic shocks)
2. **Boom-bust cycles:** Energy sector subject to extreme cyclical. Default rates spike during commodity downturns (2015–16, 2020).

4.4.5 Specification variants

Variant 1: Sector interactions with Merton scaling (preferred)

As in equation (31). Tests whether sectors differ in how much they deviate from Merton baseline.

Variant 2: Sector interactions without Merton scaling

$$y_{i,t} = \alpha + \beta_0 f_{DTS,t} + \sum_{s \in S} \beta_s \cdot \mathbb{1}_{\{i \in s\}} \cdot f_{DTS,t} + \varepsilon_{i,t} \quad (38)$$

Tests whether sectors differ in raw DTS sensitivity (ignoring Merton). Use if Stage 0 shows Merton doesn't work well.

Variant 3: Rating-specific sector interactions

Allow sector effects to differ by rating:

$$y_{i,t} = \alpha + \beta_0[\lambda_i f_{DTS,t}] + \sum_{s \in S} \sum_{r \in R} \beta_{s,r} \cdot \mathbb{1}_{\{i \in s, i \in r\}} \cdot [\lambda_i f_{DTS,t}] + \varepsilon_{i,t} \quad (39)$$

where $R = \{\text{AAA/AA, A, BBB, BB, B, CCC}\}$.

Rationale: Financial IG may differ from Financial HY in systematic ways (e.g., TBTF implicit guarantees for large bank IG bonds).

Cost: Many parameters ($|S| \times |R| \approx 4 \times 6 = 24$ interaction terms).

Decision rule: Only use if Variant 1 shows significant sector effects *and* you have theoretical reason to expect rating heterogeneity.

4.4.6 Implementation considerations

Clustering of standard errors

Even in pooled specification, use week clustering:

- All bonds (across all sectors) in week t experience common shocks
- Within-week correlation violates independence assumption
- Cluster by week (not by sector, not by bond) to get correct inference

Sample balance across sectors

Potential issue: Unbalanced samples can affect precision of interaction estimates.

Sector	N (bonds)	N (obs)	% of sample
Industrial	3,500	1,200,000	60%
Financial	1,200	400,000	20%
Utility	600	200,000	10%
Energy	400	150,000	8%
Other	300	50,000	2%
Total	6,000	2,000,000	100%

Table 7: Typical sector distribution in IG universe

Implications:

- Industrial (reference) has most observations $\Rightarrow \hat{\beta}_0$ most precise
- Financial has 1/3 the sample $\Rightarrow \hat{\beta}_{\text{Financial}}$ less precise (wider SE)
- Energy/Other have smallest samples \Rightarrow interaction estimates noisy

Robustness check: Re-estimate with balanced subsamples (downsample Industrial to match Financial). If results qualitatively similar, sample imbalance not driving findings.

Time-varying sector effects

Sector differences may be regime-dependent:

- Financials: Amplified sensitivity during financial crises (2008, 2020), normal otherwise

- Energy: Amplified sensitivity during oil price crashes (2015–16, 2020), normal otherwise
- Utilities: Stable across all regimes

Extended specification:

$$\beta_s \cdot \mathbb{1}_{\{i \in s\}} \rightarrow [\beta_{s,0} + \beta_{s,VIX} \cdot VIX_t] \cdot \mathbb{1}_{\{i \in s\}} \quad (40)$$

Test: Is $\beta_{Financial,VIX} > \beta_{Industrial,VIX}$? (Financials more stress-sensitive)

4.4.7 Deliverables for sector interaction analysis

- **Table 0.5:** Sector interaction estimates
 - Rows: Industrial (reference), Financial, Utility, Energy, Other
 - Columns: $\hat{\beta}_s$, SE, *t*-stat, *p*-value, Total β ($= \beta_0 + \beta_s$)
 - Joint test: *F*-statistic, *p*-value for all $\beta_s = 0$
- **Table 0.6:** Pairwise sector comparisons
 - Test Financial vs Utility, Financial vs Energy, etc.
 - Report difference, SE of difference, *p*-value
- **Figure 0.7:** Total β by sector
 - Bar chart with sectors on x-axis, total β on y-axis
 - Error bars: 95% confidence intervals
 - Horizontal line at $\beta = 1$ (Merton prediction)
 - Color-code: Blue for reference, Red for significant ($p < 0.05$), Gray for insignificant
- **Figure 0.8:** Sector deviations from reference
 - Bar chart with non-reference sectors on x-axis, β_s on y-axis
 - Error bars: 95% CI
 - Horizontal line at 0 (no difference from reference)
 - Annotate *p*-values
- **Figure 0.9:** Regime-dependent sector effects (if time-varying specification estimated)
 - Time series plot: X-axis = Date, Y-axis = Implied sector-specific β
 - Multiple lines: Industrial, Financial, Utility, Energy
 - Shade crisis periods (2020 COVID, 2022 rate shock)

4.5 Synthesis and Integration

4.5.1 Decision tree for Stage 0 conclusion

After completing all three analyses, synthesize findings using the following framework:

Stage 0 Synthesis Framework

Question 1: Does Merton work at bucket level?

From Section 4.2:

- If median ratio $\hat{\beta}^{(k)}/\lambda_k^{\text{Merton}} \in [0.8, 1.2]$ and $>80\%$ of buckets within $\pm 20\%$: **Yes**
- Otherwise: **Partially** or **No**

Question 2: Does Merton work within issuers?

From Section 4.3:

- If pooled $\hat{\beta} \in [0.9, 1.1]$ and p -value for $H_0 : \beta = 1 > 0.10$: **Yes**
- Otherwise: **No**

Question 3: Do sectors differ systematically?

From Section 4.4:

- If joint F -test $p < 0.05$ and ≥ 1 sector has $|\hat{\beta}_s| > 0.2$: **Yes**
- Otherwise: **No**

4.5.2 Interpretation matrix

4.5.3 Resolving apparent contradictions

Case 1: Bucket-level Merton fails, but within-issuer succeeds

Scenario:

- Bucket-level: Median ratio = 0.75, large dispersion
- Within-issuer: Pooled $\hat{\beta} = 0.98$, tight distribution

Diagnosis: Bucket approach confounds Merton's cross-maturity effect with cross-issuer heterogeneity within buckets. When we isolate pure within-issuer variation, Merton works.

Implication: Problem is not Merton theory but bucket construction. Need finer buckets (e.g., add spread level as dimension) or move to continuous characteristics (Stage A, Specification A.2).

Case 2: Both bucket and within-issuer show Merton works, but sectors differ

Scenario:

- Bucket-level: Median ratio = 1.05
- Within-issuer: Pooled $\hat{\beta} = 1.02$
- Sector interaction: $\hat{\beta}_{\text{Financial}} = 0.35$ ($p < 0.001$)

Diagnosis: Merton's cross-maturity predictions hold (both bucket and within-issuer validate), but there's a **level shift** for certain sectors. Financials move more than Merton predicts, but their cross-maturity *ratios* still obey Merton.

Implication: Use Merton cross-maturity adjustment λ_T , but add sector multiplier:

$$\lambda_i^{\text{prod}} = \lambda_T(T_i) \times \lambda_s(s_i) \times (1 + \hat{\beta}_{\text{sector}_i}) \quad (41)$$

Table 8: Stage 0 Synthesis: Interpretation of Combined Results

Bucket	Within-Issuer	Sectors	Conclusion and Recommendation
Yes	Yes	No	Merton works well across the board. Use pure Merton tables for production. Proceed to Stages A–C with high confidence in theoretical baseline.
Yes	Yes	Yes	Merton works but needs sector adjustments. Use $\lambda_i^{\text{prod}} = \lambda_i^{\text{Merton}} \times (1 + \hat{\beta}_{\text{sector}_i})$. Test sector effect stability in Stage C.
Yes	No	No	Merton works in aggregate but not within issuers. Cross-issuer heterogeneity exists beyond rating/maturity. Need issuer-specific calibration or finer bucket definitions.
Partial	Yes	No	Bucket approach too coarse; within-issuer Merton works. Use issuer-level analysis for production. Consider continuous characteristics (Stage A, Spec A.2).
Partial	Yes	Yes	Merton mechanism correct but bucket aggregation masks issuer and sector heterogeneity. Use within-issuer Merton with sector adjustments.
No	No	—	Merton fundamentally fails. Proceed with unrestricted empirical approach in Stages B–C. Investigate alternative mechanisms in Stage D (liquidity, jumps, correlations).

This preserves theoretical structure while accommodating sector-specific amplification.

Case 3: Within-issuer shows Merton works in IG but fails in HY

Scenario:

- Within-issuer IG: $\hat{\beta} = 1.02$ ($p = 0.65$)
- Within-issuer HY: $\hat{\beta} = 0.75$ ($p = 0.02$)

Diagnosis: Merton’s continuous diffusion assumption breaks down in HY. Jump-to-default risk, discrete credit events, and liquidity issues dominate. Front-end HY bonds don’t amplify as much as Merton predicts.

Implication: Use Merton adjustments for IG only. For HY, either:

- Use calibrated Merton with $\hat{\beta}_{\text{HY}} = 0.75$ as scaling factor
- Move to fully empirical approach for HY (Stage B, Specification B.3)

4.5.4 Complementary insights

Each analysis provides unique information that the others cannot:

- **Bucket-level:** Aggregate patterns across market. Most powerful for detecting overall success/failure of Merton. Establishes baseline for Stages A–C.
- **Within-issuer:** Validates that Merton’s *mechanism* is correct: same issuer, different maturities behave as predicted. Isolates pure maturity effect from other heterogeneity.
- **Sector interaction:** Identifies systematic deviations by industry. Reveals where Merton needs augmentation (e.g., Financial amplification) vs where it works universally.

Combined narrative for Stage 0 report (example):

Bucket-level analysis shows Merton predictions approximately hold in aggregate (median ratio 1.05), with stronger performance in IG than HY. Within-issuer analysis confirms Merton’s cross-maturity predictions work cleanly when holding credit quality constant, validating the theoretical mechanism. Sector interaction analysis reveals Financials have 33% amplified sensitivity beyond Merton baseline, likely due to correlation risk and regulatory capital effects; other sectors do not differ significantly.

Synthesis: Merton provides excellent starting point. Recommend *Calibrated Merton with sector adjustments* for production: Use theoretical λ_T and λ_s , add Financial sector multiplier of 1.33. This combines parsimony (theory-based) with empirical refinement (sector factor).

4.6 Combined Deliverables for Stage 0

4.6.1 Tables

1. **Table 0.1:** Bucket-level $\hat{\beta}^{(k)}$ for key rating \times maturity combinations
2. **Table 0.2:** Cross-maturity pattern tests and Spearman correlations
3. **Table 0.3:** Within-issuer pooled results by sample split
4. **Table 0.4:** Within-issuer diagnostic regressions
5. **Table 0.5:** Sector interaction estimates with joint test
6. **Table 0.6:** Pairwise sector comparisons
7. **Table 0.7:** Stage 0 synthesis summary (one-page overview of all three approaches)

4.6.2 Figures

1. **Figure 0.1:** Scatter of empirical vs theoretical β (bucket-level)
2. **Figure 0.2:** Cross-maturity patterns by rating
3. **Figure 0.3:** Regime convergence (dispersion vs spread level)
4. **Figure 0.4:** Distribution of within-issuer $\hat{\beta}_{i,t}$
5. **Figure 0.5:** Within-issuer $\hat{\beta}$ vs spread level
6. **Figure 0.6:** Case studies for major issuers

7. **Figure 0.7:** Total β by sector (bar chart)
8. **Figure 0.8:** Sector deviations from reference
9. **Figure 0.9:** Regime-dependent sector effects (if applicable)
10. **Figure 0.10:** Stage 0 synthesis visual (combining key findings from all three approaches)

4.6.3 Written summary

The Stage 0 written summary (5–7 pages) should address:

1. **Bucket-level findings** (1–2 pages):
 - Does Merton predict bucket-level sensitivities?
 - Is cross-maturity pattern correct?
 - Does pattern differ by spread level (regime convergence)?
 - Where do largest deviations occur?
2. **Within-issuer findings** (1–2 pages):
 - Does Merton work within issuer capital structures?
 - How much heterogeneity exists across issuers?
 - Does spread level matter (Regime 5 test)?
 - Do crises break proportionality?
3. **Sector interaction findings** (1–2 pages):
 - Do sectors differ significantly?
 - Which sectors differ and by how much?
 - Economic interpretation of sector patterns
 - Are differences stable or regime-dependent?
4. **Synthesis and recommendation** (1–2 pages):
 - Reconciliation across the three approaches
 - Overall assessment of Merton adequacy
 - Recommended baseline for Stages A–C
 - Specific adjustments needed (sectors, regimes, IG vs HY)

4.7 Decision Point

Based on Stage 0 results across all three approaches:

Path 1: Theory works well (all three approaches validate)

Conditions:

- Bucket-level median ratio $\in [0.9, 1.1]$
- Within-issuer pooled $\hat{\beta} \in [0.9, 1.1]$ with $p > 0.10$ for $H_0 : \beta = 1$
- Sector joint test $p > 0.10$ (no significant sector differences)

Action: Proceed to Stages A–C with theory-constrained specifications. Use λ^{Merton} as baseline. High confidence in theoretical foundation.

Path 2: Theory works with sector adjustments

Conditions:

- Bucket-level and within-issuer validate Merton
- Sector joint test $p < 0.05$ with ≥ 1 sector having $|\hat{\beta}_s| > 0.2$

Action: Adopt $\lambda_i^{\text{prod}} = \lambda_i^{\text{Merton}} \times (1 + \hat{\beta}_{\text{sector}_i})$. Proceed to Stage C to test stability of sector effects over time.

Path 3: Theory needs calibration

Conditions:

- Within-issuer $\hat{\beta}$ outside [0.9, 1.1] but patterns match (high R^2 , residuals unsystematic)
- Bucket-level shows systematic bias (all ratios > 1.2 or < 0.8)

Action: Adopt calibrated Merton: $\lambda^{\text{prod}} = \hat{\beta}_{\text{Merton}} \cdot \lambda^{\text{Merton}}$. Theory has right structure, wrong scale. Proceed to Stage C to test stability of calibration coefficient.

Path 4: Theory captures structure but misses details

Conditions:

- Moderate bucket-level fit (e.g., 60–80% of buckets within $\pm 20\%$)
- Within-issuer works for IG but not HY
- Some systematic sector or regime patterns

Action: Proceed to Stages A–C with both theory-guided and unrestricted tracks in parallel. Use Merton for IG, empirical for HY. Stage C will reveal if time-variation or regime-switching helps.

Path 5: Theory fundamentally fails

Conditions:

- Wrong patterns (e.g., long bonds more sensitive than short)
- Within-issuer $\hat{\beta}$ far from 1 with systematic deviations
- Bucket-level $R_{\text{Merton}}^2 < 0.5 \times R_{\text{buckets}}^2$

Action: Skip theory-constrained specifications. Proceed directly to Stage D (robustness) to diagnose *why* theory fails (liquidity? tails? specific shocks?). Then Stage E with unrestricted specification only.

Implication: Report that structural models don't provide adequate guidance for DTS adjustments in this market. Empirical approach necessary.

Key principle: The three-pronged Stage 0 approach provides robust evidence for choosing among these paths. No single analysis is sufficient—bucket-level can be confounded by issuer heterogeneity, within-issuer has smaller samples, and sector interactions require pooling assumptions.

Together, they provide a comprehensive diagnostic.

5 Stage A: Establish Cross-Sectional Variation

5.1 Objective

Stage A establishes the empirical fact that DTS betas differ across bonds *before* testing whether Merton explains why. This separation is critical:

- If no significant variation exists, stop here—standard DTS is adequate
- If variation exists but Merton doesn't explain it, we know theory fails
- If variation exists and Merton explains it, theory provides parsimonious structure

Relationship to Stage 0: Stage 0 tested whether Merton predictions hold using three complementary approaches. Stage A takes a step back and asks the more fundamental question: *Is there any systematic variation in DTS betas to explain?* This question is logically prior to testing whether theory explains the variation.

5.2 Specification A.1: Bucket-level betas

For each bucket k (defined by rating \times maturity \times sector), estimate:

$$y_{i,t} = \alpha^{(k)} + \beta^{(k)} f_{DTS,t}^{(U)} + \varepsilon_{i,t}^{(k)}, \quad i \in \text{bucket } k \quad (42)$$

Bucket definitions:

- Rating: AAA/AA, A, BBB for IG; BB, B, CCC for HY
- Maturity: 1–2y, 2–3y, 3–5y, 5–7y, 7–10y, 10y+
- Sector: Industrials, Financials, Utilities, Energy (can add more)

Creates $3 \times 6 \times 4 = 72$ buckets for IG, $3 \times 6 \times 4 = 72$ for HY (some will be sparse).

Estimation: Pooled OLS within each bucket, cluster standard errors by week.

Key outputs:

- Table A.1: $\hat{\beta}^{(k)}$ for all buckets with standard errors and t -statistics
- F -test for equality across buckets: $H_0 : \beta^{(1)} = \beta^{(2)} = \dots = \beta^{(K)}$
- **Critical decision:** If F -test fails to reject ($p > 0.10$), no significant variation—stop here, use standard DTS

5.3 Specification A.2: Continuous characteristics

Instead of discrete buckets, estimate how beta varies with continuous characteristics.

Two-step procedure:

Step 1: For each bond i , estimate bond-specific beta using rolling 2-year windows:

$$y_{i,t} = \alpha_i + \beta_i f_{DTS,t} + \varepsilon_{i,t} \quad (43)$$

Yields time-series of $\hat{\beta}_{i,\tau}$ for bond i at window midpoint τ .

Step 2: Cross-sectional regression of estimated betas on characteristics:

$$\hat{\beta}_{i,\tau} = \gamma_0 + \gamma_M M_{i,\tau} + \gamma_s s_{i,\tau} + \gamma_{M^2} M_{i,\tau}^2 + \gamma_{Ms} M_{i,\tau} \cdot s_{i,\tau} + u_{i,\tau} \quad (44)$$

where $M_{i,\tau}$ is maturity, $s_{i,\tau}$ is spread at window midpoint.

Standard errors: Bootstrap or cluster by bond (account for multiple windows per bond).

Key outputs:

- Coefficient estimates $\hat{\gamma}_M, \hat{\gamma}_s, \hat{\gamma}_{M^2}, \hat{\gamma}_{Ms}$
- R^2 of cross-sectional fit
- **Interpretation:** Do betas vary systematically with maturity and spread?

5.4 Establishing the stylized fact

Deliverable: Document that DTS betas are not constant. Show:

1. Statistical significance of variation (F -test from A.1, R^2 from A.2)
2. Economic significance: Range of $\hat{\beta}^{(k)}$ across buckets (e.g., min 0.6, max 1.8 implies $3\times$ variation)
3. Pattern: Short-maturity / low-spread bonds have higher betas (preliminary observation, not causal claim)

Comparison with Stage 0: Stage A results should be consistent with Stage 0 bucket-level analysis. The key difference is framing:

- Stage 0 asked: Do empirical betas match Merton predictions?
- Stage A asks: Do betas vary at all, regardless of whether they match theory?

If Stage 0 found Merton works well (Path 1), Stage A should find variation that aligns with Merton predictions. If Stage 0 found Merton partially works or fails, Stage A documents the variation that theory fails to fully explain.

5.5 Deliverables for Stage A

- **Table A.1:** Bucket-level $\hat{\beta}^{(k)}$ estimates
 - Rows: Maturity buckets
 - Columns: Rating buckets
 - Separate panels for IG and HY (and optionally by sector)
 - Include standard errors, t -statistics, sample size per bucket
- **Table A.2:** Tests of beta equality
 - F -test for H_0 : all $\beta^{(k)}$ equal, overall and by dimension:
 - * Across maturities (holding rating constant)
 - * Across ratings (holding maturity constant)
 - * Across sectors

- Report F -statistic, degrees of freedom, p -value
- **Decision rule:** If all $p > 0.10$, declare standard DTS adequate
- **Table A.3:** Continuous characteristic regression (Specification A.2)
 - Coefficients: $\hat{\gamma}_0, \hat{\gamma}_M, \hat{\gamma}_s, \hat{\gamma}_{M^2}, \hat{\gamma}_{Ms}$
 - Standard errors, t -statistics, R^2
 - Separate for IG and HY
- **Figure A.1:** Heatmap of $\hat{\beta}^{(k)}$ by maturity (x-axis) \times rating (y-axis). Color intensity represents beta magnitude. Separate panels for IG and HY.
- **Figure A.2:** Implied beta surface from Specification A.2: 3D plot or contour plot with maturity on x-axis, spread on y-axis, predicted $\hat{\beta}$ on z-axis/color.
- **Diagnostic summary (2 pages):**
 1. Is variation statistically significant? (F -test results)
 2. Is variation economically meaningful? (Range and IQR of $\hat{\beta}^{(k)}$)
 3. What characteristics drive variation? (Maturity vs spread vs sector)
 4. Does IG show more variation than HY? (Regime 2 prediction from Stage 0)
 5. **Recommendation:** Proceed to Stage B to test whether Merton explains patterns, or stop if no variation.

5.6 Decision point

- **If F -test $p < 0.01$ and $R^2 > 0.15$ in A.2:** Strong evidence of systematic variation. Proceed to Stage B with high confidence.
- **If F -test $0.01 < p < 0.10$:** Marginal variation. Proceed to Stage B but may find theory sufficient.
- **If F -test $p > 0.10$ and $R^2 < 0.05$:** No meaningful variation. **Stop here—standard DTS is adequate. Report this as primary finding.**

Consistency check with Stage 0: If Stage 0 found significant patterns (e.g., within-issuer $\hat{\beta} \neq 1$ or significant sector interactions), Stage A should find significant variation. If Stage A finds no variation but Stage 0 found patterns, revisit Stage 0 methodology—possible issue with bucket definitions or sample restrictions.

6 Stage B: Does Merton Explain the Variation?

6.1 Objective

Having established in Stage A that DTS betas vary across bonds, Stage B tests whether Merton's structural predictions explain this variation. This is the *core empirical test* of the theoretical framework.

Critical distinction: Stage A documented the *fact* of variation. Stage B tests whether *theory explains it*.

Relationship to Stage 0: Stage 0 provided preliminary evidence on whether Merton predictions hold. Stage B conducts more rigorous tests with formal model comparisons and benchmarking against the unrestricted variation documented in Stage A.

6.2 Specification B.1: Merton as offset (constrained)

Use Merton-predicted $\lambda_{i,t}^{\text{Merton}}$ as known adjustment factor:

$$y_{i,t} = \alpha + \beta_{\text{Merton}} \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t} \quad (45)$$

where $\lambda_{i,t}^{\text{Merton}} = \lambda_T(T_i; 5y, s_{i,t}) \times \lambda_s(s_{i,t}; 100\text{bps})$ using Tables 2–3.

Theory prediction: If Merton is exactly correct, $\beta_{\text{Merton}} = 1$.

Test: Wald test $H_0 : \beta = 1$ with clustered standard errors (week \times issuer).

Interpretation:

- $\hat{\beta}_{\text{Merton}} \in [0.9, 1.1]$: Merton predictions unbiased, theory works
- $\hat{\beta}_{\text{Merton}} \in [0.8, 1.2]$: Close enough for practical purposes
- $\hat{\beta}_{\text{Merton}} > 1.2$ or < 0.8 : Systematic bias, need calibration
- Low R^2 despite $\hat{\beta} \approx 1$: Merton captures mean but misses dispersion

6.3 Specification B.2: Decomposed components

Test maturity vs credit quality effects separately:

$$y_{i,t} = \alpha + \beta_T[\lambda_T(T_i; 5y, s_{i,t}) \cdot f_{DTS,t}] + \beta_s[\lambda_s(s_{i,t}; 100) \cdot f_{DTS,t}] + \varepsilon_{i,t} \quad (46)$$

Theory predictions:

- $\beta_T \approx 1$: Maturity adjustment works
- $\beta_s \approx 1$: Credit quality adjustment works
- If both hold: Merton decomposition empirically valid

Diagnostic patterns:

- $\beta_T \approx 1$ but $\beta_s \neq 1$: Maturity effects correct, quality effects need recalibration
- $\beta_T \neq 1$ but $\beta_s \approx 1$: Quality effects correct, maturity functional form wrong
- Both $\neq 1$: Need to reconsider entire Merton structure

6.4 Specification B.3: With sector adjustments

Building on Stage 0 sector interaction findings, test whether adding sector adjustments to Merton improves fit:

$$y_{i,t} = \alpha + \beta_0[\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \sum_{s \in S} \beta_s \cdot \mathbb{1}_{\{i \in s\}} \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t} \quad (47)$$

Comparison with Stage 0: This is identical to the sector interaction specification from Stage 0 (equation 31). In Stage B, we use it to assess incremental explanatory power:

- Does adding sector interactions improve R^2 significantly beyond pure Merton?
- Are sector coefficients stable when estimated jointly with Merton baseline?

6.5 Benchmarking against Stage A

Key comparison: Does theory-constrained Specification B.1 perform comparably to unrestricted bucket regressions from Stage A?

Metrics:

1. **R^2 comparison:**

- R_{buckets}^2 from Stage A (upper bound—fully flexible)
- R_{Merton}^2 from Specification B.1
- $R_{\text{Merton+Sector}}^2$ from Specification B.3
- If $R_{\text{Merton}}^2 > 0.9 \times R_{\text{buckets}}^2$: Theory captures 90%+ of explainable variation

2. **Bucket-level residuals:** For each bucket k , compute:

$$\text{Residual}_k = \hat{\beta}_{\text{Stage A}}^{(k)} - \lambda_k^{\text{Merton}} \quad (48)$$

If most residuals small (e.g., <0.2) and unsystematic, theory adequate.

3. **RMSE comparison:** Root mean squared error of spread change predictions:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i,t} (y_{i,t} - \hat{y}_{i,t})^2} \quad (49)$$

Compare Merton vs Merton+Sector vs bucket-based predictions.

6.6 Specification B.4: Unrestricted for comparison

Estimate fully flexible functional form:

$$\lambda_i = \beta_0 + \beta_M M_i + \beta_{M^2} M_i^2 + \beta_s s_i + \beta_{s^2} s_i^2 + \beta_{Ms} M_i \cdot s_i + \sum_{\text{rating}} \beta_r + \sum_{\text{sector}} \beta_{\text{sec}} \quad (50)$$

Then:

$$y_{i,t} = \alpha + [\hat{\lambda}_i \cdot f_{DTS,t}] + \varepsilon_{i,t} \quad (51)$$

Purpose: Assess incremental explanatory power beyond theory. If $R_{\text{unrestricted}}^2 \gg R_{\text{Merton}}^2$, theory misses important patterns.

6.7 Theory vs reality table

Critical deliverable: Direct comparison of empirical betas to Merton predictions from [Wuebben \[2025\]](#).

Decision criteria:

- **If 90%+ of ratios in [0.8, 1.2]:** Merton provides excellent baseline. Use pure Merton or calibrated version for production.
- **If systematic bias** (all ratios > 1.2 or < 0.8): Recalibrate with β_{Merton} from Specification B.1. Theory has right structure, wrong scale.
- **If high dispersion but no bias:** Heterogeneity beyond Merton dimensions. Proceed to unrestricted estimation, but Merton still useful as starting point.
- **If wrong patterns** (e.g., long bonds have higher sensitivity than short): Theory fundamentally fails. Investigate alternative mechanisms.

Table 9: Do Empirical Betas Match Merton Predictions? (Illustrative)

Bucket	$\hat{\beta}^{(k)}$ (Stage A)	$\lambda_k^{\text{Merton}}$ (Theory)	Ratio
1y, 50bps	3.45	3.62	0.95
1y, 100bps	3.15	3.27	0.96
1y, 200bps	2.85	2.78	1.03
3y, 50bps	1.52	1.47	1.03
3y, 100bps	1.38	1.42	0.97
3y, 200bps	1.29	1.36	0.95
10y, 50bps	0.58	0.61	0.95
10y, 100bps	0.61	0.64	0.95
10y, 200bps	0.65	0.67	0.97

6.8 Deliverables for Stage B

- **Table B.1:** Constrained Merton specifications

- Spec B.1: $\hat{\beta}_{\text{Merton}}$, standard error, Wald test p -value for $H_0 : \beta = 1$
- Spec B.2: $\hat{\beta}_T, \hat{\beta}_s$, standard errors, joint test p -value for $H_0 : (\beta_T, \beta_s) = (1, 1)$
- Spec B.3: $\hat{\beta}_0$, sector coefficients $\hat{\beta}_s$, joint test for sector effects
- Separate panels for IG and HY

- **Table B.2:** Model comparison

- Rows: Stage A buckets, Spec B.1 (Merton), Spec B.2 (decomposed), Spec B.3 (Merton+Sector), Spec B.4 (unrestricted)
- Columns: R^2 , RMSE, AIC, number of parameters
- ΔR^2 relative to Stage A buckets

- **Table B.3:** Theory vs Reality (as in Table 10)

- All maturity \times spread bucket combinations
- Empirical $\hat{\beta}^{(k)}$ from Stage A
- Theoretical $\lambda_k^{\text{Merton}}$
- Ratio and absolute deviation
- Highlight cells where $|\text{Ratio} - 1| > 0.25$

- **Figure B.1:** Scatter plot: Empirical $\hat{\beta}^{(k)}$ (y-axis) vs Theoretical $\lambda_k^{\text{Merton}}$ (x-axis) for all buckets. 45-degree line for perfect agreement. Color-code by:

- IG-narrow maturity (circles)
- IG-wide maturity (squares)
- HY-narrow (triangles)
- HY-wide (diamonds)
- Distressed (stars)

- **Figure B.2:** Residual analysis: $\hat{\beta}^{(k)} - \lambda_k^{\text{Merton}}$ by:
 - Panel A: By maturity (x-axis: 1y, 3y, 5y, 7y, 10y)
 - Panel B: By spread level (x-axis: 50, 100, 200, 300, 500, 1000 bps)
 - Panel C: By sector (x-axis: Industrials, Financials, Utilities, Energy)

Zero line indicates perfect Merton prediction. Look for systematic patterns.

- **Figure B.3:** Implied λ surface from Spec B.4 (unrestricted) vs Merton prediction:

- 3D surface plot or side-by-side contour plots
- X-axis: Maturity (1–10 years)
- Y-axis: Spread (50–1000 bps)
- Z-axis/color: Predicted λ
- Shows where unrestricted deviates from theory

- **Diagnostic summary (3–4 pages):**

1. **Does Merton work?** (Spec B.1 results: $\hat{\beta}_{\text{Merton}}$ and R^2)
2. **Which component drives fit?** (Spec B.2 results: maturity vs quality)
3. **Do sector adjustments help?** (Spec B.3 results: incremental R^2 from sectors)
4. **Where does theory succeed?** (Table B.3 analysis: which regimes have ratios near 1.0)
5. **Where does theory fail?** (Residual patterns from Figure B.2)
6. **Is unrestricted necessary?** (Table B.2 comparison: ΔR^2 and parameter efficiency)
7. **Practical recommendation:**
 - Use pure Merton tables (simplest)
 - Use calibrated Merton with $\hat{\beta}_{\text{Merton}}$ (simple + data-driven)
 - Use Merton with sector adjustments (theory + sector factors)
 - Need full unrestricted (complex but necessary)
8. **Consistency with Stage 0:** Do Stage B findings align with Stage 0 conclusions? If Stage 0 found significant sector effects, does Spec B.3 confirm their importance?

6.9 Decision point

Based on Stage B results:

Decision Tree for Stage C Entry:

Path 1: Theory works well

Condition: $\hat{\beta}_{\text{Merton}} \in [0.9, 1.1]$ and $R_{\text{Merton}}^2 > 0.85 \times R_{\text{buckets}}^2$

Action: Proceed to Stage C to test whether static λ^{Merton} suffices or time-variation needed.

High confidence in theoretical foundation.

Path 2: Theory works with sector adjustments

Condition: $\hat{\beta}_{\text{Merton}} \in [0.9, 1.1]$ but $R_{\text{Merton+Sector}}^2 > R_{\text{Merton}}^2 + 0.03$ with significant sector coefficients

Action: Adopt Merton with sector adjustments as baseline. Proceed to Stage C to test stability of both Merton coefficient and sector effects over time.

Path 3: Theory needs calibration

Condition: $\hat{\beta}_{\text{Merton}}$ outside $[0.9, 1.1]$ but patterns match (high R^2 , residuals unsystematic)

Action: Adopt calibrated Merton: $\lambda^{\text{prod}} = \hat{\beta}_{\text{Merton}} \cdot \lambda^{\text{Merton}}$. Proceed to Stage C to test stability of $\hat{\beta}_{\text{Merton}}$ over time.

Path 4: Theory captures structure but misses details

Condition: Moderate R^2_{Merton} (e.g., $0.6\text{--}0.8 \times$ buckets), some systematic residuals

Action: Proceed to Stage C with both theory-guided and unrestricted tracks in parallel. Stage C will reveal if time-variation helps or if static unrestricted better.

Path 5: Theory fundamentally fails

Condition: Wrong patterns (e.g., long bonds more sensitive than short), or $R^2_{\text{Merton}} < 0.5 \times R^2_{\text{buckets}}$

Action: Skip Stage C (no point testing time-variation of failed model). Proceed directly to Stage D (robustness) to diagnose *why* theory fails (liquidity? tails? specific shocks?). Then Stage E with unrestricted specification only.

Implication: Report that structural models don't provide adequate guidance for DTS adjustments in this market. Empirical approach necessary.

7 Stage C: Does Static Merton Suffice or Do We Need Time-Variation?

7.1 Objective and prerequisite

Prerequisite: Stage B showed that Merton $\lambda(s, T)$ explains cross-sectional variation (Paths 1–4 from Stage B decision tree).

Objective: Test whether the relationship between λ and (s, T) is stable over time, or whether macro state variables induce time-variation.

Key principle: Don't add time-variation until you've proven the simple static model fails.

Extended objective: If Stage 0 and Stage B found significant sector effects, also test whether sector adjustments are stable over time or regime-dependent.

7.2 Rolling window stability test

Divide sample into non-overlapping 1-year windows $w \in \{1, 2, \dots, W\}$. For each window:

$$y_{i,t} = \alpha_w + \beta_w \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t}, \quad t \in w \quad (52)$$

This yields time series of $\hat{\beta}_w$ with standard errors $\text{se}(\hat{\beta}_w)$.

Stability test: Chow test for structural break:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_W \quad (53)$$

Compute F -statistic comparing restricted (single β) vs unrestricted (separate β_w) models.

Decision rule:

- If $p > 0.10$: Static λ sufficient. **Stop Stage C here.** No need for time-varying adjustments. Report that Merton provides stable baseline.
- If $0.01 < p < 0.10$: Marginal instability. Proceed to investigate drivers but be skeptical of over-parameterization.
- If $p < 0.01$: Significant time-variation. Proceed to macro driver analysis.

7.3 Stability of sector effects

If Stage B adopted Merton with sector adjustments (Path 2), test whether sector coefficients are stable:

$$y_{i,t} = \alpha_w + \beta_{0,w}[\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \sum_{s \in S} \beta_{s,w} \cdot \mathbb{1}_{\{i \in s\}} \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t} \quad (54)$$

Tests:

1. Chow test for base coefficient: $H_0 : \beta_{0,1} = \beta_{0,2} = \dots = \beta_{0,W}$
2. Chow test for each sector: $H_0 : \beta_{s,1} = \beta_{s,2} = \dots = \beta_{s,W}$ for each s
3. Joint stability test: All coefficients stable across windows

Interpretation:

- If base β_0 stable but sector β_s vary: Sector effects are regime-dependent (e.g., Financials amplify more during crises)
- If base β_0 varies but sector β_s stable: Macro conditions affect all bonds similarly; sector differentials constant
- If both vary: Complex regime structure; may need VIX or spread-level interactions

7.4 Visual stability assessment

Time series plot: $\hat{\beta}_w$ over time with 95% confidence bands.

Interpretation:

- Confidence bands overlapping 1.0 throughout: Static Merton works, bands capture sampling variation
- Confidence bands tight but $\hat{\beta}_w$ drifts (e.g., 0.9 in 2010s, 1.1 in 2020s): Systematic shift, investigate macro drivers
- Wide swings during crises (2020, 2022) but stable otherwise: Regime-dependent but static in normal times

7.5 Conditional on instability: Macro driver analysis

Only if Chow test rejects, estimate second-stage regression:

$$\hat{\beta}_w = \delta_0 + \delta_{VIX} \cdot \overline{VIX}_w + \delta_{OAS} \cdot \log(\overline{OAS}_{index,w}) + \delta_r \cdot \bar{r}_{10y,w} + \eta_w \quad (55)$$

where \overline{X}_w denotes window-average of variable X .

Theory-based predictions Wuebben [2025]:

1. $\delta_{VIX} > 0$: High volatility amplifies sensitivity, especially for short maturities (flight-to-quality concentrates in long end, front end whipsaws more)
2. $\delta_{OAS} < 0$: Wide spreads reduce dispersion in λ (convergence to Regime 5 where all bonds near default, proportionality improves)

3. δ_r : Ambiguous. Higher rates increase discount effect (reduce duration), possibly dampening elasticities. Empirical question.

Economic significance threshold: Only declare time-variation meaningful if macro state changes λ by >20% over sample range.

Example: If $\delta_{VIX} = 0.01$ and VIX ranges from 10 to 40, the effect is $0.01 \times 30 = 0.30$, or 30% change in β . This is economically large.

If $\delta_{VIX} = 0.002$, effect is 6%—within noise, ignore.

7.6 Maturity-specific time-variation

Theory predicts time-variation should differ by maturity: short-maturity IG bonds most affected.

Test: Estimate rolling β_w separately for maturity buckets:

$$y_{i,t} = \alpha_{w,m} + \beta_{w,m} \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t}, \quad i \in \text{maturity bucket } m, t \in w \quad (56)$$

Then regress:

$$\hat{\beta}_{w,m} = \delta_{0,m} + \delta_{VIX,m} \cdot \overline{VIX}_w + \eta_{w,m} \quad (57)$$

Prediction: $\delta_{VIX,1y} > \delta_{VIX,5y} > \delta_{VIX,10y}$ (short bonds more regime-dependent).

If this pattern holds, supports theory-based intuition about crisis dynamics.

7.7 Sector-specific time-variation

Building on Stage 0 sector findings, test whether sector effects are regime-dependent:

$$\hat{\beta}_{s,w} = \delta_{s,0} + \delta_{s,VIX} \cdot \overline{VIX}_w + \eta_{s,w} \quad (58)$$

Predictions:

- $\delta_{\text{Financial},VIX} > 0$: Financial sector amplification increases during stress (correlation risk, liquidity spirals)
- $\delta_{\text{Utility},VIX} \approx 0$: Utility defensive characteristics stable across regimes
- $\delta_{\text{Energy},VIX}$: Ambiguous—energy may decouple from broad market during commodity-specific shocks

7.8 Practical implication assessment

Question: Even if time-variation is statistically significant, does it matter for portfolio management?

Scenario analysis: Compare static vs time-varying λ for:

1. **Risk model accuracy:** Does time-varying λ reduce tracking error in out-of-sample hedging?
2. **Crisis performance:** During 2020 COVID shock, did static λ severely misprice front-end IG?
3. **Operational complexity:** Time-varying λ requires daily macro state inputs and recalibration. Worth the cost?

Recommendation framework:

- If time-variation changes risk estimates by <10% except during rare crises: Use static λ , add crisis overlays manually
- If time-variation changes risk estimates by >20% routinely: Implement time-varying λ with macro state
- If time-variation important only for specific buckets (e.g., 1–2y IG): Use static for most bonds, time-varying for front end only
- If sector effects are regime-dependent: Use static sector adjustments with VIX-contingent overlays for Financials

7.9 Deliverables for Stage C

- **Table C.1:** Rolling window stability test
 - Rows: Time windows (2013–2014, 2014–2015, ..., 2024–2025)
 - Columns: $\hat{\beta}_w$, standard error, 95% CI, sample size
 - Separate panels for IG and HY
 - Chow test: F -statistic, p -value
- **Table C.2:** Sector coefficient stability (if applicable)
 - Rows: Windows
 - Columns: $\hat{\beta}_{0,w}$, $\hat{\beta}_{\text{Financial},w}$, $\hat{\beta}_{\text{Utility},w}$, $\hat{\beta}_{\text{Energy},w}$
 - Chow tests for each coefficient series
- **Table C.3:** Macro driver regression (conditional on instability)
 - Coefficients: $\hat{\delta}_{\text{VIX}}$, $\hat{\delta}_{\text{OAS}}$, $\hat{\delta}_r$
 - Standard errors, t -statistics, R^2
 - Economic significance: Effect of 1 SD change in each macro variable on β
 - Test predicted signs: $\delta_{\text{VIX}} > 0$, $\delta_{\text{OAS}} < 0$
- **Table C.4:** Maturity-specific time-variation
 - Rows: Maturity buckets (1–2y, 3–5y, 7–10y)
 - Columns: $\hat{\delta}_{\text{VIX},m}$, standard error, t -statistic
 - Test: Is $\delta_{\text{VIX},1y}$ significantly larger than $\delta_{\text{VIX},10y}$?
- **Table C.5:** Sector-specific time-variation (if applicable)
 - Rows: Sectors (Financial, Utility, Energy)
 - Columns: $\hat{\delta}_{s,\text{VIX}}$, standard error, t -statistic
 - Test: Is $\delta_{\text{Financial,VIX}} > 0$?
- **Figure C.1:** Time series of $\hat{\beta}_w$ for IG and HY
 - X-axis: Year (2013–2025)

- Y-axis: $\hat{\beta}_w$
- Point estimates with 95% confidence bands
- Horizontal line at $\beta = 1$ (theory prediction)
- Shade crisis periods (2020 COVID, 2022 rate shock)
- Interpretation: Does $\hat{\beta}_w$ spike during crises?
- **Figure C.2:** $\hat{\beta}_w$ vs macro state variables
 - Panel A: $\hat{\beta}_w$ (y-axis) vs \overline{VIX}_w (x-axis)
 - Panel B: $\hat{\beta}_w$ (y-axis) vs $\log(\overline{OAS}_w)$ (x-axis)
 - Scatter with OLS fit line
 - Color-code by time period (pre-2020, COVID, post-COVID)
 - Shows whether macro variables predict time-variation
- **Figure C.3:** Implied $\lambda_{i,t}$ for representative bonds over time
 - Three lines: 1-year BBB, 5-year BBB, 10-year BBB (all industrial)
 - Static λ (dashed) vs time-varying λ_t (solid)
 - Shows when and how much time-variation matters
- **Figure C.4:** Sector coefficient stability (if applicable)
 - Time series of $\hat{\beta}_{\text{Financial},w}$, $\hat{\beta}_{\text{Utility},w}$, $\hat{\beta}_{\text{Energy},w}$
 - Shows whether sector effects are stable or regime-dependent
- **Figure C.5:** Scenario analysis—crisis vs normal
 - Histogram of spread changes during normal periods ($VIX < 20$)
 - Histogram of spread changes during stress ($VIX > 30$)
 - Overlay: Static Merton prediction, time-varying prediction
 - Shows whether static model fails systematically in crises
- **Summary and recommendation (3–4 pages):**
 - 1. Is base relationship stable?** (Chow test results for β_w)
 - 2. Are sector effects stable?** (Chow test results for $\beta_{s,w}$)
 - 3. If unstable, what drives it?** (Macro driver analysis)
 - 4. Is instability economically meaningful?** (Effect size in % terms)
 - 5. Does theory-based intuition hold?** (VIX amplifies front-end, OAS compresses dispersion)
 - 6. Practical recommendation:**
 - Use static λ : Adequate for normal markets, simple implementation
 - Use time-varying λ : Necessary for crisis periods, worth complexity
 - Hybrid: Static baseline with crisis adjustments ($VIX > 30$)
 - For sectors: Static sector adjustments vs regime-dependent
 - 7. Implication for production:** If static suffices, Stage E will select among pure/calibrated Merton with or without sector adjustments. If time-varying needed, add macro state to production spec.

7.10 Decision point

- **If Chow test $p > 0.10$ for both base and sectors:** Static λ sufficient. Proceed to Stage D (robustness) with confidence in stable baseline.
- **If Chow test $p < 0.10$ for base but sectors stable:** Time-variation in overall level but not in sector differentials. Consider VIX overlay on base β_0 .
- **If Chow test $p > 0.10$ for base but sectors unstable:** Stable overall but regime-dependent sector effects. Use static base with VIX-contingent sector adjustments.
- **If Chow test $p < 0.01$ and effects $>20\%$ in crises:** Time-varying λ necessary. Incorporate macro state in Stage E production specification.

8 Stage D: Robustness and Extensions

8.1 Objective and positioning

Prerequisite: Stages A–C established whether/how Merton predictions hold for *mean* spread changes in standard conditions.

Objective: Test robustness across:

1. Tail events (quantile regression)
2. Shock types (systematic vs idiosyncratic decomposition)
3. Spread components (default vs liquidity)

Key framing: These are **secondary** tests. If Stages A–C show Merton fails, Stage D helps diagnose *why*. If Stages A–C validate Merton, Stage D confirms it's not just a mean effect.

Relationship to Stage 0: Stage D extends the robustness checks beyond what Stage 0's three-pronged approach covers. While Stage 0 tested Merton in aggregate (buckets), within issuers, and across sectors, Stage D examines distributional properties (tails), shock decomposition, and spread component separation.

8.2 D.1: Tail behavior (quantile regression)

8.2.1 Motivation

Merton model assumes normal shocks (geometric Brownian motion for firm value). If tails differ from mean, this suggests:

- Jump-to-default risk not captured by continuous diffusion
- Liquidity evaporation in stress (left tail)
- Asymmetric investor behavior (panic selling vs gradual buying)

8.2.2 Specification

For quantiles $\tau \in \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$, estimate:

$$Q_\tau(y_{i,t} | f_{DTS,t}, \lambda_i^{\text{Merton}}) = \alpha_\tau + \beta_\tau \cdot [\lambda_i^{\text{Merton}} \cdot f_{DTS,t}] \quad (59)$$

This models the τ -th conditional quantile of spread changes, allowing elasticity to differ across the distribution.

Merton prediction: $\beta_\tau \approx 1$ for all τ (all quantiles respect structural elasticities).

Diagnostics:

- If $\beta_{0.05} \gg \beta_{0.95}$: Left tail (spread widening) has amplified sensitivity. Consistent with jump-to-default or liquidity spirals.
- If $\beta_{0.95} > \beta_{0.05}$: Right tail (spread tightening) more sensitive. Suggests momentum/technical buying in rallies.
- If β_τ U-shaped (high at both tails): Both extreme moves behave differently than moderate moves—non-linearity in spread dynamics.

8.2.3 Sector-specific tail behavior

Building on Stage 0 sector findings, estimate quantile regressions separately by sector:

$$Q_\tau(y_{i,t} | f_{DTS,t}, \lambda_i^{\text{Merton}}, \text{sector}_i = s) = \alpha_{\tau,s} + \beta_{\tau,s} \cdot [\lambda_i^{\text{Merton}} \cdot f_{DTS,t}] \quad (60)$$

Hypothesis: If Stage 0 found Financials have amplified mean sensitivity, do they also have amplified tail sensitivity?

- $\beta_{0.05, \text{Financial}} > \beta_{0.05, \text{Industrial}}$: Financial left-tail risk exceeds mean amplification (correlation risk manifests in tails)
- $\beta_{0.05, \text{Utility}} \approx \beta_{0.50, \text{Utility}}$: Utility defensive characteristics hold in tails

8.2.4 Practical implications

Risk management:

- Use $\beta_{0.05}$ for Value-at-Risk (VaR) and Expected Shortfall (ES) calculations
- Use $\beta_{0.50}$ for expected return / attribution models
- If $\beta_{0.05} = 1.5 \times \beta_{0.50}$: Tail risk 50% larger than mean-based models predict

Stress testing:

- Standard Merton λ may underestimate losses in left-tail scenarios
- Adjust: $DTS_{i,t}^{*,\text{stress}} = \beta_{0.05} \cdot \lambda_i^{\text{Merton}} \cdot DTS_{i,t}$

8.2.5 Deliverables for D.1

- **Table D.1:** Quantile-specific β_τ estimates
 - Rows: $\tau \in \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$
 - Columns: $\hat{\beta}_\tau$, standard error, 95% CI
 - Separate panels for IG and HY
 - Test: $H_0 : \beta_{0.05} = \beta_{0.50}$ and $H_0 : \beta_{0.95} = \beta_{0.50}$
- **Table D.2:** Sector-specific tail behavior
 - Rows: Sectors (Industrial, Financial, Utility, Energy)
 - Columns: $\beta_{0.05}$, $\beta_{0.50}$, $\beta_{0.95}$, Ratio $\beta_{0.05}/\beta_{0.50}$
 - Shows which sectors have amplified tail risk
- **Figure D.1:** Plot of $\hat{\beta}_\tau$ across $\tau \in [0.05, 0.95]$
 - X-axis: Quantile τ
 - Y-axis: $\hat{\beta}_\tau$
 - Horizontal line at $\beta = 1$ (Merton prediction)
 - Confidence bands (bootstrap)
 - Interpretation: Flat line = Merton works across distribution. Upward/downward slope = asymmetry.
- **Figure D.2:** Sector-specific quantile plots
 - Separate lines for Industrial, Financial, Utility, Energy
 - Shows whether Financial tail amplification exceeds mean amplification
- **Interpretation note:** Are tail deviations concentrated in specific regimes (e.g., front-end IG)? If so, suggests these bonds have additional jump risk beyond Merton's continuous framework.

8.3 D.2: Shock decomposition

8.3.1 Motivation

Merton model treats all firm value shocks identically—whether macro, sector, or idiosyncratic, the elasticity λ_i should be the same because all operate through the firm's asset value.

Empirically test: Do different shock types exhibit different elasticities? If so, suggests mechanisms beyond firm fundamentals (e.g., liquidity contagion for sector shocks, information asymmetry for idiosyncratic shocks).

8.3.2 Factor construction

Decompose bond i 's spread change into orthogonal components:

$$y_{i,t} = \underbrace{f_{DTS,t}^{(G)}}_{\text{Global factor}} + \underbrace{f_{DTS,s(i),t}^{(S)}}_{\text{Sector factor}} + \underbrace{f_{DTS,i,t}^{(I)}}_{\text{Issuer-specific}} + \varepsilon_{i,t} \quad (61)$$

Estimation procedure:

Step 1: Global factor

$$f_{DTS,t}^{(G)} = \frac{\Delta OAS_t^{(U)}}{OAS_{t-1}^{(U)}} \quad (62)$$

Step 2: Sector factors (orthogonalized to global)

$$f_{DTS,s,t}^{(S)} = \frac{\Delta OAS_{s,t}^{(U)}}{OAS_{s,t-1}^{(U)}} - f_{DTS,t}^{(G)} \quad (63)$$

Step 3: Issuer-specific (residual)

$$f_{DTS,i,t}^{(I)} = y_{i,t} - f_{DTS,t}^{(G)} - f_{DTS,s(i),t}^{(S)} \quad (64)$$

8.3.3 Multi-factor regression with Merton baseline

Estimate:

$$y_{i,t} = \beta^{(G)}[\lambda_i^{\text{Merton}} \cdot f_{DTS,t}^{(G)}] + \beta^{(S)}[\lambda_i^{\text{Merton}} \cdot f_{DTS,s(i),t}^{(S)}] + \beta^{(I)}[\lambda_i^{\text{Merton}} \cdot f_{DTS,i,t}^{(I)}] + \varepsilon_{i,t} \quad (65)$$

Constrained specification: Impose $\beta^{(G)} = \beta^{(S)} = \beta^{(I)} = 1$, estimate single coefficient:

$$y_{i,t} = \beta \cdot \lambda_i^{\text{Merton}} \cdot [f_t^{(G)} + f_{s(i),t}^{(S)} + f_{i,t}^{(I)}] + \varepsilon_{i,t} \quad (66)$$

Merton prediction: $\beta^{(G)} \approx \beta^{(S)} \approx \beta^{(I)} \approx 1$ (all shocks respect structural elasticities).

8.3.4 Diagnostic patterns

- If $\beta^{(G)} \approx \beta^{(S)} \approx \beta^{(I)} \approx 1$: Merton applies uniformly—all shocks operate through firm value.
- If $\beta^{(S)} > \beta^{(G)}$: Sector shocks have amplified effects. Suggests contagion, correlation trading, or common liquidity factors beyond fundamentals.
- If $\beta^{(I)} \gg \beta^{(G)}$: Idiosyncratic news has exaggerated spread impact. Consistent with information asymmetry, adverse selection in trading.
- If $\beta^{(G)} < 1$ but $\beta^{(I)} > 1$: Bonds under-react to macro (diversified portfolios stabilize) but over-react to issuer-specific (concentrated positions, forced selling).

8.3.5 Connection to Stage 0 sector findings

If Stage 0 found Financials have amplified sensitivity, test whether this comes from:

- Global shocks: $\beta_{\text{Financial}}^{(G)} > \beta_{\text{Industrial}}^{(G)}$ (Financials more sensitive to macro)
- Sector shocks: $\beta_{\text{Financial}}^{(S)} > \beta_{\text{Industrial}}^{(S)}$ (correlation/contagion within Financial sector)
- Idiosyncratic shocks: $\beta_{\text{Financial}}^{(I)} > \beta_{\text{Industrial}}^{(I)}$ (information asymmetry in Financial credits)

This decomposition helps understand *why* sector effects exist, informing production model design.

8.3.6 Deliverables for D.2

- **Table D.3:** Variance decomposition
 - Rows: Rating \times maturity buckets
 - Columns: % variance from Global, Sector, Issuer-specific, Residual
 - Shows relative importance of each factor type
 - Separate for IG and HY (expect IG more global-driven, HY more issuer-specific)
- **Table D.4:** Shock-specific elasticities
 - Rows: $\beta^{(G)}$, $\beta^{(S)}$, $\beta^{(I)}$
 - Columns: Estimate, standard error, 95% CI
 - Test: $H_0 : \beta^{(G)} = \beta^{(S)} = \beta^{(I)} = 1$ (joint test)
 - Test: Pairwise $H_0 : \beta^{(G)} = \beta^{(S)}$, etc.
- **Table D.5:** Sector-specific shock decomposition
 - Rows: Sectors (Industrial, Financial, Utility, Energy)
 - Columns: $\beta_s^{(G)}$, $\beta_s^{(S)}$, $\beta_s^{(I)}$
 - Shows which shock types drive each sector's behavior
- **Figure D.3:** Bar chart of $\hat{\beta}$ by factor type
 - Three bars: Global, Sector, Issuer-specific
 - Error bars for 95% CI
 - Horizontal line at $\beta = 1$
 - Separate panels for IG and HY
- **Interpretation:** If all three ≈ 1 , Merton universally applicable. If sector/issuer deviate, need factor-specific adjustments in production models.

8.4 D.3: Liquidity adjustment

8.4.1 Motivation

OAS includes both default risk and liquidity premium. Merton model predicts elasticity for *default component only*. If liquidity shocks don't respect structural elasticities, need to decompose.

8.4.2 Liquidity-adjusted spread construction

Step 1: Estimate liquidity component cross-sectionally each period:

$$s_{i,t}^{\text{liq}} = \phi_0 + \phi_1 \text{BidAsk}_{i,t} + \phi_2 \log(\text{Size}_i) + \phi_3 \log(\text{Turnover}_{i,t}) + \phi_4 \text{Age}_{i,t} + \eta_{i,t} \quad (67)$$

Standard approach: Cross-sectional regression within each rating \times maturity cell to avoid confounding credit quality with liquidity.

Step 2: Define default component as residual:

$$OAS_{i,t}^{\text{def}} = OAS_{i,t} - \widehat{OAS}_{i,t}^{\text{liq}} \quad (68)$$

Step 3: Compute default-based spread changes:

$$y_{i,t}^{\text{def}} = \frac{\Delta OAS_{i,t}^{\text{def}}}{OAS_{i,t-1}^{\text{def}}} \quad (69)$$

8.4.3 Re-estimate Merton fit on default component

Run Stage B regression using $y_{i,t}^{\text{def}}$ as dependent variable:

$$y_{i,t}^{\text{def}} = \alpha + \beta_{\text{def}} \cdot [\lambda_i^{\text{Merton}} \cdot f_{DTS,t}^{\text{def}}] + \varepsilon_{i,t} \quad (70)$$

where $f_{DTS,t}^{\text{def}}$ is the index-level default spread factor (after removing liquidity).

Theory prediction: $\beta_{\text{def}} \approx 1$ (Merton works for default component) and $R_{\text{def}}^2 > R_{\text{total}}^2$ (less noise without liquidity shocks).

8.4.4 Diagnostic patterns

- If $\beta_{\text{def}} \approx 1$ and $\beta_{\text{total}} \approx 1$: Liquidity adjustment unnecessary—Merton works on total OAS.
- If $\beta_{\text{def}} \approx 1$ but $\beta_{\text{total}} < 1$: Liquidity shocks dampen spread movements. Total OAS underestimates default sensitivity.
- If $\beta_{\text{def}} < \beta_{\text{total}}$: Liquidity shocks amplify spread movements beyond Merton. Need separate liquidity beta.
- If improvement concentrated in illiquid bonds: Signals that liquidity noise matters for small size, low turnover bonds but not for liquid benchmarks.

8.4.5 Sector-specific liquidity effects

Building on Stage 0 findings, test whether liquidity adjustment differentially affects sectors:

- Financials: Often more liquid (active trading, market makers); liquidity adjustment may be less important
- Utilities: Often less liquid (buy-and-hold investors); liquidity adjustment may help
- Energy: Liquidity varies with oil price cycles; may need regime-dependent liquidity treatment

8.4.6 Practical decision rule

When to Use Liquidity Adjustment:

Liquid IG (BidAsk < 50bps, Size > \$1B):

- Use total OAS, liquidity adjustment negligible (<5bps typical)
- Liquidity R^2 low, not worth complexity

Illiquid IG (BidAsk > 100bps, Size < \$500M):

- Liquidity premium 20–50bps, material fraction of spread
- Consider decomposition for risk models

HY (BidAsk typically > 200bps):

- Decompose into OAS^{def} and OAS^{liq}
- Use λ^{def} from Merton for default component
- Add separate λ^{liq} empirically estimated (not theory-based)

8.4.7 Deliverables for D.3

- **Table D.6:** Liquidity model estimates
 - Cross-sectional regression: $\hat{\phi}_1$ (bid-ask), $\hat{\phi}_2$ (size), $\hat{\phi}_3$ (turnover), $\hat{\phi}_4$ (age)
 - R^2 by rating \times maturity bucket
 - Typical liquidity premium: Mean $\widehat{OAS}^{\text{liq}}$ by bucket
- **Table D.7:** Merton fit comparison
 - Rows: Total OAS, Default component only
 - Columns: $\hat{\beta}$, R^2 , RMSE
 - Test: Is β_{def} closer to 1 than β_{total} ?
 - Improvement: $\Delta R^2 = R_{\text{def}}^2 - R_{\text{total}}^2$
- **Table D.8:** Improvement by liquidity regime
 - Split sample by liquidity quartiles (bid-ask or turnover)
 - For each quartile: β_{total} , β_{def} , ΔR^2
 - Shows whether adjustment helps mainly for illiquid bonds
- **Table D.9:** Sector-specific liquidity effects
 - Rows: Sectors
 - Columns: β_{total} , β_{def} , ΔR^2
 - Shows which sectors benefit most from liquidity decomposition
- **Figure D.4:** Scatter plot: $\beta_{\text{def}} - \beta_{\text{total}}$ (y-axis) vs average bid-ask (x-axis)
 - Each point: a rating \times maturity bucket
 - Interpretation: Positive slope = illiquid bonds benefit more from decomposition
- **Interpretation and recommendation:**
 - For which bond types is liquidity adjustment material?
 - Does Merton work better on default component than total OAS?
 - Production implication: Use total OAS for liquid IG, decompose for HY?

8.5 Summary of Stage D

Stage D is **diagnostic**, not **decisional**. It answers:

1. **Where does Merton fail?** (Tails? Specific shock types? Liquidity-contaminated spreads?)
2. **How large are the failures?** (20% effect in tails? 50%? Negligible?)
3. **Are failures systematic?** (Always front-end IG? Always sector shocks? Always illiquid bonds?)
4. **Do failures relate to Stage 0 findings?** (Is Financial amplification driven by tails? sector shocks? liquidity?)

Use Stage D results in Stage E to decide:

- If tail effects large: Add quantile-specific λ for VaR/ES
- If sector shocks deviate: Consider sector-factor adjustments beyond static sector dummies
- If liquidity matters: Decompose OAS for HY, keep total for IG
- If Financial amplification comes from specific channel: Design targeted adjustment

But don't let Stage D derail the main program. If Stages A–C showed Merton works well on average (possibly with sector adjustments from Stage 0), Stage D refinements are enhancements, not requirements.

9 Stage A: Establish Cross-Sectional Variation

9.1 Objective

Stage A establishes the empirical fact that DTS betas differ across bonds *before* testing whether Merton explains why. This separation is critical:

- If no significant variation exists, stop here—standard DTS is adequate
- If variation exists but Merton doesn't explain it, we know theory fails
- If variation exists and Merton explains it, theory provides parsimonious structure

Relationship to Stage 0: Stage 0 tested whether Merton predictions hold using three complementary approaches. Stage A takes a step back and asks the more fundamental question: *Is there any systematic variation in DTS betas to explain?* This question is logically prior to testing whether theory explains the variation.

9.2 Specification A.1: Bucket-level betas

For each bucket k (defined by rating \times maturity \times sector), estimate:

$$y_{i,t} = \alpha^{(k)} + \beta^{(k)} f_{DTS,t}^{(U)} + \varepsilon_{i,t}^{(k)}, \quad i \in \text{bucket } k \quad (71)$$

Bucket definitions:

- Rating: AAA/AA, A, BBB for IG; BB, B, CCC for HY
- Maturity: 1–2y, 2–3y, 3–5y, 5–7y, 7–10y, 10y+
- Sector: Industrials, Financials, Utilities, Energy (can add more)

Creates $3 \times 6 \times 4 = 72$ buckets for IG, $3 \times 6 \times 4 = 72$ for HY (some will be sparse).

Estimation: Pooled OLS within each bucket, cluster standard errors by week.

Key outputs:

- Table A.1: $\hat{\beta}^{(k)}$ for all buckets with standard errors and t -statistics
- F -test for equality across buckets: $H_0 : \beta^{(1)} = \beta^{(2)} = \dots = \beta^{(K)}$
- **Critical decision:** If F -test fails to reject ($p > 0.10$), no significant variation—stop here, use standard DTS

9.3 Specification A.2: Continuous characteristics

Instead of discrete buckets, estimate how beta varies with continuous characteristics.

Two-step procedure:

Step 1: For each bond i , estimate bond-specific beta using rolling 2-year windows:

$$y_{i,t} = \alpha_i + \beta_i f_{DTS,t} + \varepsilon_{i,t} \quad (72)$$

Yields time-series of $\hat{\beta}_{i,\tau}$ for bond i at window midpoint τ .

Step 2: Cross-sectional regression of estimated betas on characteristics:

$$\hat{\beta}_{i,\tau} = \gamma_0 + \gamma_M M_{i,\tau} + \gamma_s s_{i,\tau} + \gamma_{M^2} M_{i,\tau}^2 + \gamma_{Ms} M_{i,\tau} \cdot s_{i,\tau} + u_{i,\tau} \quad (73)$$

where $M_{i,\tau}$ is maturity, $s_{i,\tau}$ is spread at window midpoint.

Standard errors: Bootstrap or cluster by bond (account for multiple windows per bond).

Key outputs:

- Coefficient estimates $\hat{\gamma}_M, \hat{\gamma}_s, \hat{\gamma}_{M^2}, \hat{\gamma}_{Ms}$
- R^2 of cross-sectional fit
- **Interpretation:** Do betas vary systematically with maturity and spread?

9.4 Establishing the stylized fact

Deliverable: Document that DTS betas are not constant. Show:

1. Statistical significance of variation (F -test from A.1, R^2 from A.2)
2. Economic significance: Range of $\hat{\beta}^{(k)}$ across buckets (e.g., min 0.6, max 1.8 implies 3× variation)
3. Pattern: Short-maturity / low-spread bonds have higher betas (preliminary observation, not causal claim)

Comparison with Stage 0: Stage A results should be consistent with Stage 0 bucket-level analysis. The key difference is framing:

- Stage 0 asked: Do empirical betas match Merton predictions?
- Stage A asks: Do betas vary at all, regardless of whether they match theory?

If Stage 0 found Merton works well (Path 1), Stage A should find variation that aligns with Merton predictions. If Stage 0 found Merton partially works or fails, Stage A documents the variation that theory fails to fully explain.

9.5 Deliverables for Stage A

- **Table A.1:** Bucket-level $\hat{\beta}^{(k)}$ estimates
 - Rows: Maturity buckets
 - Columns: Rating buckets
 - Separate panels for IG and HY (and optionally by sector)
 - Include standard errors, t -statistics, sample size per bucket
- **Table A.2:** Tests of beta equality
 - F -test for H_0 : all $\beta^{(k)}$ equal, overall and by dimension:
 - * Across maturities (holding rating constant)
 - * Across ratings (holding maturity constant)
 - * Across sectors
 - Report F -statistic, degrees of freedom, p -value
 - **Decision rule:** If all $p > 0.10$, declare standard DTS adequate
- **Table A.3:** Continuous characteristic regression (Specification A.2)
 - Coefficients: $\hat{\gamma}_0, \hat{\gamma}_M, \hat{\gamma}_s, \hat{\gamma}_{M^2}, \hat{\gamma}_{Ms}$
 - Standard errors, t -statistics, R^2
 - Separate for IG and HY
- **Figure A.1:** Heatmap of $\hat{\beta}^{(k)}$ by maturity (x-axis) \times rating (y-axis). Color intensity represents beta magnitude. Separate panels for IG and HY.
- **Figure A.2:** Implied beta surface from Specification A.2: 3D plot or contour plot with maturity on x-axis, spread on y-axis, predicted $\hat{\beta}$ on z-axis/color.
- **Diagnostic summary (2 pages):**
 1. Is variation statistically significant? (F -test results)
 2. Is variation economically meaningful? (Range and IQR of $\hat{\beta}^{(k)}$)
 3. What characteristics drive variation? (Maturity vs spread vs sector)
 4. Does IG show more variation than HY? (Regime 2 prediction from Stage 0)

5. Recommendation: Proceed to Stage B to test whether Merton explains patterns, or stop if no variation.

9.6 Decision point

- **If F -test $p < 0.01$ and $R^2 > 0.15$ in A.2:** Strong evidence of systematic variation. Proceed to Stage B with high confidence.
- **If F -test $0.01 < p < 0.10$:** Marginal variation. Proceed to Stage B but may find theory sufficient.
- **If F -test $p > 0.10$ and $R^2 < 0.05$:** No meaningful variation. **Stop here—standard DTS is adequate. Report this as primary finding.**

Consistency check with Stage 0: If Stage 0 found significant patterns (e.g., within-issuer $\hat{\beta} \neq 1$ or significant sector interactions), Stage A should find significant variation. If Stage A finds no variation but Stage 0 found patterns, revisit Stage 0 methodology—possible issue with bucket definitions or sample restrictions.

10 Stage B: Does Merton Explain the Variation?

10.1 Objective

Having established in Stage A that DTS betas vary across bonds, Stage B tests whether Merton's structural predictions explain this variation. This is the *core empirical test* of the theoretical framework.

Critical distinction: Stage A documented the *fact* of variation. Stage B tests whether *theory explains it*.

Relationship to Stage 0: Stage 0 provided preliminary evidence on whether Merton predictions hold. Stage B conducts more rigorous tests with formal model comparisons and benchmarking against the unrestricted variation documented in Stage A.

10.2 Specification B.1: Merton as offset (constrained)

Use Merton-predicted $\lambda_{i,t}^{\text{Merton}}$ as known adjustment factor:

$$y_{i,t} = \alpha + \beta_{\text{Merton}} \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t} \quad (74)$$

where $\lambda_{i,t}^{\text{Merton}} = \lambda_T(T_i; 5y, s_{i,t}) \times \lambda_s(s_{i,t}; 100\text{bps})$ using Tables 2–3.

Theory prediction: If Merton is exactly correct, $\beta_{\text{Merton}} = 1$.

Test: Wald test $H_0 : \beta = 1$ with clustered standard errors (week \times issuer).

Interpretation:

- $\hat{\beta}_{\text{Merton}} \in [0.9, 1.1]$: Merton predictions unbiased, theory works
- $\hat{\beta}_{\text{Merton}} \in [0.8, 1.2]$: Close enough for practical purposes
- $\hat{\beta}_{\text{Merton}} > 1.2$ or < 0.8 : Systematic bias, need calibration
- Low R^2 despite $\hat{\beta} \approx 1$: Merton captures mean but misses dispersion

10.3 Specification B.2: Decomposed components

Test maturity vs credit quality effects separately:

$$y_{i,t} = \alpha + \beta_T[\lambda_T(T_i; 5y, s_{i,t}) \cdot f_{DTS,t}] + \beta_s[\lambda_s(s_{i,t}; 100) \cdot f_{DTS,t}] + \varepsilon_{i,t} \quad (75)$$

Theory predictions:

- $\beta_T \approx 1$: Maturity adjustment works
- $\beta_s \approx 1$: Credit quality adjustment works
- If both hold: Merton decomposition empirically valid

Diagnostic patterns:

- $\beta_T \approx 1$ but $\beta_s \neq 1$: Maturity effects correct, quality effects need recalibration
- $\beta_T \neq 1$ but $\beta_s \approx 1$: Quality effects correct, maturity functional form wrong
- Both $\neq 1$: Need to reconsider entire Merton structure

10.4 Specification B.3: With sector adjustments

Building on Stage 0 sector interaction findings, test whether adding sector adjustments to Merton improves fit:

$$y_{i,t} = \alpha + \beta_0[\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \sum_{s \in S} \beta_s \cdot \mathbb{1}_{\{i \in s\}} \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t} \quad (76)$$

Comparison with Stage 0: This is identical to the sector interaction specification from Stage 0 (equation 31). In Stage B, we use it to assess incremental explanatory power:

- Does adding sector interactions improve R^2 significantly beyond pure Merton?
- Are sector coefficients stable when estimated jointly with Merton baseline?

10.5 Benchmarking against Stage A

Key comparison: Does theory-constrained Specification B.1 perform comparably to unrestricted bucket regressions from Stage A?

Metrics:

1. **R^2 comparison:**

- R^2_{buckets} from Stage A (upper bound—fully flexible)
- R^2_{Merton} from Specification B.1
- $R^2_{\text{Merton+Sector}}$ from Specification B.3
- If $R^2_{\text{Merton}} > 0.9 \times R^2_{\text{buckets}}$: Theory captures 90%+ of explainable variation

2. **Bucket-level residuals:** For each bucket k , compute:

$$\text{Residual}_k = \hat{\beta}_{\text{Stage A}}^{(k)} - \lambda_k^{\text{Merton}} \quad (77)$$

If most residuals small (e.g., < 0.2) and unsystematic, theory adequate.

3. **RMSE comparison:** Root mean squared error of spread change predictions:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i,t} (y_{i,t} - \hat{y}_{i,t})^2} \quad (78)$$

Compare Merton vs Merton+Sector vs bucket-based predictions.

10.6 Specification B.4: Unrestricted for comparison

Estimate fully flexible functional form:

$$\lambda_i = \beta_0 + \beta_M M_i + \beta_{M^2} M_i^2 + \beta_s s_i + \beta_{s^2} s_i^2 + \beta_{Ms} M_i \cdot s_i + \sum_{\text{rating}} \beta_r + \sum_{\text{sector}} \beta_{\text{sec}} \quad (79)$$

Then:

$$y_{i,t} = \alpha + [\hat{\lambda}_i \cdot f_{DTS,t}] + \varepsilon_{i,t} \quad (80)$$

Purpose: Assess incremental explanatory power beyond theory. If $R^2_{\text{unrestricted}} \gg R^2_{\text{Merton}}$, theory misses important patterns.

10.7 Theory vs reality table

Critical deliverable: Direct comparison of empirical betas to Merton predictions from [Wuebben \[2025\]](#).

Table 10: Do Empirical Betas Match Merton Predictions? (Illustrative)

Bucket	$\hat{\beta}^{(k)}$ (Stage A)	$\lambda_k^{\text{Merton}}$ (Theory)	Ratio
1y, 50bps	3.45	3.62	0.95
1y, 100bps	3.15	3.27	0.96
1y, 200bps	2.85	2.78	1.03
3y, 50bps	1.52	1.47	1.03
3y, 100bps	1.38	1.42	0.97
3y, 200bps	1.29	1.36	0.95
10y, 50bps	0.58	0.61	0.95
10y, 100bps	0.61	0.64	0.95
10y, 200bps	0.65	0.67	0.97

Decision criteria:

- **If 90%+ of ratios in [0.8, 1.2]:** Merton provides excellent baseline. Use pure Merton or calibrated version for production.
- **If systematic bias** (all ratios > 1.2 or < 0.8): Recalibrate with β_{Merton} from Specification B.1. Theory has right structure, wrong scale.
- **If high dispersion but no bias:** Heterogeneity beyond Merton dimensions. Proceed to unrestricted estimation, but Merton still useful as starting point.
- **If wrong patterns** (e.g., long bonds have higher sensitivity than short): Theory fundamentally fails. Investigate alternative mechanisms.

10.8 Deliverables for Stage B

- **Table B.1:** Constrained Merton specifications
 - Spec B.1: $\hat{\beta}_{\text{Merton}}$, standard error, Wald test p -value for $H_0 : \beta = 1$
 - Spec B.2: $\hat{\beta}_T, \hat{\beta}_s$, standard errors, joint test p -value for $H_0 : (\beta_T, \beta_s) = (1, 1)$
 - Spec B.3: $\hat{\beta}_0$, sector coefficients $\hat{\beta}_s$, joint test for sector effects
 - Separate panels for IG and HY
- **Table B.2:** Model comparison
 - Rows: Stage A buckets, Spec B.1 (Merton), Spec B.2 (decomposed), Spec B.3 (Merton+Sector), Spec B.4 (unrestricted)
 - Columns: R^2 , RMSE, AIC, number of parameters
 - ΔR^2 relative to Stage A buckets
- **Table B.3:** Theory vs Reality (as in Table 10)
 - All maturity \times spread bucket combinations
 - Empirical $\hat{\beta}^{(k)}$ from Stage A
 - Theoretical $\lambda_k^{\text{Merton}}$
 - Ratio and absolute deviation
 - Highlight cells where $|\text{Ratio} - 1| > 0.25$
- **Figure B.1:** Scatter plot: Empirical $\hat{\beta}^{(k)}$ (y-axis) vs Theoretical $\lambda_k^{\text{Merton}}$ (x-axis) for all buckets. 45-degree line for perfect agreement. Color-code by:
 - IG-narrow maturity (circles)
 - IG-wide maturity (squares)
 - HY-narrow (triangles)
 - HY-wide (diamonds)
 - Distressed (stars)
- **Figure B.2:** Residual analysis: $\hat{\beta}^{(k)} - \lambda_k^{\text{Merton}}$ by:
 - Panel A: By maturity (x-axis: 1y, 3y, 5y, 7y, 10y)
 - Panel B: By spread level (x-axis: 50, 100, 200, 300, 500, 1000 bps)
 - Panel C: By sector (x-axis: Industrials, Financials, Utilities, Energy)

Zero line indicates perfect Merton prediction. Look for systematic patterns.
- **Figure B.3:** Implied λ surface from Spec B.4 (unrestricted) vs Merton prediction:
 - 3D surface plot or side-by-side contour plots
 - X-axis: Maturity (1–10 years)
 - Y-axis: Spread (50–1000 bps)
 - Z-axis/color: Predicted λ

- Shows where unrestricted deviates from theory
- **Diagnostic summary (3–4 pages):**
 - Does Merton work?** (Spec B.1 results: $\hat{\beta}_{\text{Merton}}$ and R^2)
 - Which component drives fit?** (Spec B.2 results: maturity vs quality)
 - Do sector adjustments help?** (Spec B.3 results: incremental R^2 from sectors)
 - Where does theory succeed?** (Table B.3 analysis: which regimes have ratios near 1.0)
 - Where does theory fail?** (Residual patterns from Figure B.2)
 - Is unrestricted necessary?** (Table B.2 comparison: ΔR^2 and parameter efficiency)
 - Practical recommendation:**
 - Use pure Merton tables (simplest)
 - Use calibrated Merton with $\hat{\beta}_{\text{Merton}}$ (simple + data-driven)
 - Use Merton with sector adjustments (theory + sector factors)
 - Need full unrestricted (complex but necessary)
 - Consistency with Stage 0:** Do Stage B findings align with Stage 0 conclusions? If Stage 0 found significant sector effects, does Spec B.3 confirm their importance?

10.9 Decision point

Based on Stage B results:

Decision Tree for Stage C Entry:

Path 1: Theory works well

Condition: $\hat{\beta}_{\text{Merton}} \in [0.9, 1.1]$ and $R^2_{\text{Merton}} > 0.85 \times R^2_{\text{buckets}}$

Action: Proceed to Stage C to test whether static λ^{Merton} suffices or time-variation needed.

High confidence in theoretical foundation.

Path 2: Theory works with sector adjustments

Condition: $\hat{\beta}_{\text{Merton}} \in [0.9, 1.1]$ but $R^2_{\text{Merton}+\text{Sector}} > R^2_{\text{Merton}} + 0.03$ with significant sector coefficients

Action: Adopt Merton with sector adjustments as baseline. Proceed to Stage C to test stability of both Merton coefficient and sector effects over time.

Path 3: Theory needs calibration

Condition: $\hat{\beta}_{\text{Merton}}$ outside $[0.9, 1.1]$ but patterns match (high R^2 , residuals unsystematic)

Action: Adopt calibrated Merton: $\lambda^{\text{prod}} = \hat{\beta}_{\text{Merton}} \cdot \lambda^{\text{Merton}}$. Proceed to Stage C to test stability of $\hat{\beta}_{\text{Merton}}$ over time.

Path 4: Theory captures structure but misses details

Condition: Moderate R^2_{Merton} (e.g., $0.6\text{--}0.8 \times \text{buckets}$), some systematic residuals

Action: Proceed to Stage C with both theory-guided and unrestricted tracks in parallel. Stage C will reveal if time-variation helps or if static unrestricted better.

Path 5: Theory fundamentally fails

Condition: Wrong patterns (e.g., long bonds more sensitive than short), or $R^2_{\text{Merton}} < 0.5 \times R^2_{\text{buckets}}$

Action: Skip Stage C (no point testing time-variation of failed model). Proceed directly to Stage D (robustness) to diagnose *why* theory fails (liquidity? tails? specific shocks?). Then Stage E with unrestricted specification only.

Implication: Report that structural models don't provide adequate guidance for DTS adjustments in this market. Empirical approach necessary.

11 Stage C: Does Static Merton Suffice or Do We Need Time-Variation?

11.1 Objective and prerequisite

Prerequisite: Stage B showed that Merton $\lambda(s, T)$ explains cross-sectional variation (Paths 1–4 from Stage B decision tree).

Objective: Test whether the relationship between λ and (s, T) is stable over time, or whether macro state variables induce time-variation.

Key principle: Don't add time-variation until you've proven the simple static model fails.

Extended objective: If Stage 0 and Stage B found significant sector effects, also test whether sector adjustments are stable over time or regime-dependent.

11.2 Rolling window stability test

Divide sample into non-overlapping 1-year windows $w \in \{1, 2, \dots, W\}$. For each window:

$$y_{i,t} = \alpha_w + \beta_w \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t}, \quad t \in w \quad (81)$$

This yields time series of $\hat{\beta}_w$ with standard errors $\text{se}(\hat{\beta}_w)$.

Stability test: Chow test for structural break:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_W \quad (82)$$

Compute F -statistic comparing restricted (single β) vs unrestricted (separate β_w) models.

Decision rule:

- **If $p > 0.10$:** Static λ sufficient. **Stop Stage C here.** No need for time-varying adjustments. Report that Merton provides stable baseline.
- **If $0.01 < p < 0.10$:** Marginal instability. Proceed to investigate drivers but be skeptical of over-parameterization.
- **If $p < 0.01$:** Significant time-variation. Proceed to macro driver analysis.

11.3 Stability of sector effects

If Stage B adopted Merton with sector adjustments (Path 2), test whether sector coefficients are stable:

$$y_{i,t} = \alpha_w + \beta_{0,w}[\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \sum_{s \in S} \beta_{s,w} \cdot \mathbb{1}_{\{i \in s\}} \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t} \quad (83)$$

Tests:

1. Chow test for base coefficient: $H_0 : \beta_{0,1} = \beta_{0,2} = \dots = \beta_{0,W}$
2. Chow test for each sector: $H_0 : \beta_{s,1} = \beta_{s,2} = \dots = \beta_{s,W}$ for each s

3. Joint stability test: All coefficients stable across windows

Interpretation:

- If base β_0 stable but sector β_s vary: Sector effects are regime-dependent (e.g., Financials amplify more during crises)
- If base β_0 varies but sector β_s stable: Macro conditions affect all bonds similarly; sector differentials constant
- If both vary: Complex regime structure; may need VIX or spread-level interactions

11.4 Visual stability assessment

Time series plot: $\hat{\beta}_w$ over time with 95% confidence bands.

Interpretation:

- Confidence bands overlapping 1.0 throughout: Static Merton works, bands capture sampling variation
- Confidence bands tight but $\hat{\beta}_w$ drifts (e.g., 0.9 in 2010s, 1.1 in 2020s): Systematic shift, investigate macro drivers
- Wide swings during crises (2020, 2022) but stable otherwise: Regime-dependent but static in normal times

11.5 Conditional on instability: Macro driver analysis

Only if Chow test rejects, estimate second-stage regression:

$$\hat{\beta}_w = \delta_0 + \delta_{VIX} \cdot \overline{VIX}_w + \delta_{OAS} \cdot \log(\overline{OAS}_{index,w}) + \delta_r \cdot \bar{r}_{10y,w} + \eta_w \quad (84)$$

where \overline{X}_w denotes window-average of variable X .

Theory-based predictions Wuebben [2025]:

1. $\delta_{VIX} > 0$: High volatility amplifies sensitivity, especially for short maturities (flight-to-quality concentrates in long end, front end whipsaws more)
2. $\delta_{OAS} < 0$: Wide spreads reduce dispersion in λ (convergence to Regime 5 where all bonds near default, proportionality improves)
3. δ_r : Ambiguous. Higher rates increase discount effect (reduce duration), possibly dampening elasticities. Empirical question.

Economic significance threshold: Only declare time-variation meaningful if macro state changes λ by >20% over sample range.

Example: If $\delta_{VIX} = 0.01$ and VIX ranges from 10 to 40, the effect is $0.01 \times 30 = 0.30$, or 30% change in β . This is economically large.

If $\delta_{VIX} = 0.002$, effect is 6%—within noise, ignore.

11.6 Maturity-specific time-variation

Theory predicts time-variation should differ by maturity: short-maturity IG bonds most affected.

Test: Estimate rolling β_w separately for maturity buckets:

$$y_{i,t} = \alpha_{w,m} + \beta_{w,m} \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t}, \quad i \in \text{maturity bucket } m, t \in w \quad (85)$$

Then regress:

$$\hat{\beta}_{w,m} = \delta_{0,m} + \delta_{\text{VIX},m} \cdot \overline{\text{VIX}}_w + \eta_{w,m} \quad (86)$$

Prediction: $\delta_{\text{VIX},1y} > \delta_{\text{VIX},5y} > \delta_{\text{VIX},10y}$ (short bonds more regime-dependent).

If this pattern holds, supports theory-based intuition about crisis dynamics.

11.7 Sector-specific time-variation

Building on Stage 0 sector findings, test whether sector effects are regime-dependent:

$$\hat{\beta}_{s,w} = \delta_{s,0} + \delta_{s,\text{VIX}} \cdot \overline{\text{VIX}}_w + \eta_{s,w} \quad (87)$$

Predictions:

- $\delta_{\text{Financial,VIX}} > 0$: Financial sector amplification increases during stress (correlation risk, liquidity spirals)
- $\delta_{\text{Utility,VIX}} \approx 0$: Utility defensive characteristics stable across regimes
- $\delta_{\text{Energy,VIX}}$: Ambiguous—energy may decouple from broad market during commodity-specific shocks

11.8 Practical implication assessment

Question: Even if time-variation is statistically significant, does it matter for portfolio management?

Scenario analysis: Compare static vs time-varying λ for:

1. **Risk model accuracy:** Does time-varying λ reduce tracking error in out-of-sample hedging?
2. **Crisis performance:** During 2020 COVID shock, did static λ severely misprice front-end IG?
3. **Operational complexity:** Time-varying λ requires daily macro state inputs and recalibration. Worth the cost?

Recommendation framework:

- If time-variation changes risk estimates by $<10\%$ except during rare crises: Use static λ , add crisis overlays manually
- If time-variation changes risk estimates by $>20\%$ routinely: Implement time-varying λ with macro state
- If time-variation important only for specific buckets (e.g., 1–2y IG): Use static for most bonds, time-varying for front end only
- If sector effects are regime-dependent: Use static sector adjustments with VIX-contingent overlays for Financials

11.9 Deliverables for Stage C

- **Table C.1:** Rolling window stability test
 - Rows: Time windows (2013–2014, 2014–2015, ..., 2024–2025)
 - Columns: $\hat{\beta}_w$, standard error, 95% CI, sample size
 - Separate panels for IG and HY
 - Chow test: F -statistic, p -value
- **Table C.2:** Sector coefficient stability (if applicable)
 - Rows: Windows
 - Columns: $\hat{\beta}_{0,w}$, $\hat{\beta}_{\text{Financial},w}$, $\hat{\beta}_{\text{Utility},w}$, $\hat{\beta}_{\text{Energy},w}$
 - Chow tests for each coefficient series
- **Table C.3:** Macro driver regression (conditional on instability)
 - Coefficients: $\hat{\delta}_{\text{VIX}}$, $\hat{\delta}_{\text{OAS}}$, $\hat{\delta}_r$
 - Standard errors, t -statistics, R^2
 - Economic significance: Effect of 1 SD change in each macro variable on β
 - Test predicted signs: $\delta_{\text{VIX}} > 0$, $\delta_{\text{OAS}} < 0$
- **Table C.4:** Maturity-specific time-variation
 - Rows: Maturity buckets (1–2y, 3–5y, 7–10y)
 - Columns: $\hat{\delta}_{\text{VIX},m}$, standard error, t -statistic
 - Test: Is $\delta_{\text{VIX},1y}$ significantly larger than $\delta_{\text{VIX},10y}$?
- **Table C.5:** Sector-specific time-variation (if applicable)
 - Rows: Sectors (Financial, Utility, Energy)
 - Columns: $\hat{\delta}_{s,\text{VIX}}$, standard error, t -statistic
 - Test: Is $\delta_{\text{Financial,VIX}} > 0$?
- **Figure C.1:** Time series of $\hat{\beta}_w$ for IG and HY
 - X-axis: Year (2013–2025)
 - Y-axis: $\hat{\beta}_w$
 - Point estimates with 95% confidence bands
 - Horizontal line at $\beta = 1$ (theory prediction)
 - Shade crisis periods (2020 COVID, 2022 rate shock)
 - Interpretation: Does $\hat{\beta}_w$ spike during crises?
- **Figure C.2:** $\hat{\beta}_w$ vs macro state variables
 - Panel A: $\hat{\beta}_w$ (y-axis) vs $\overline{\text{VIX}}_w$ (x-axis)
 - Panel B: $\hat{\beta}_w$ (y-axis) vs $\log(\overline{\text{OAS}}_w)$ (x-axis)
 - Scatter with OLS fit line

- Color-code by time period (pre-2020, COVID, post-COVID)
- Shows whether macro variables predict time-variation
- **Figure C.3:** Implied $\lambda_{i,t}$ for representative bonds over time
 - Three lines: 1-year BBB, 5-year BBB, 10-year BBB (all industrial)
 - Static λ (dashed) vs time-varying λ_t (solid)
 - Shows when and how much time-variation matters
- **Figure C.4:** Sector coefficient stability (if applicable)
 - Time series of $\hat{\beta}_{\text{Financial},w}$, $\hat{\beta}_{\text{Utility},w}$, $\hat{\beta}_{\text{Energy},w}$
 - Shows whether sector effects are stable or regime-dependent
- **Figure C.5:** Scenario analysis—crisis vs normal
 - Histogram of spread changes during normal periods ($VIX < 20$)
 - Histogram of spread changes during stress ($VIX > 30$)
 - Overlay: Static Merton prediction, time-varying prediction
 - Shows whether static model fails systematically in crises
- **Summary and recommendation (3–4 pages):**
 1. **Is base relationship stable?** (Chow test results for β_w)
 2. **Are sector effects stable?** (Chow test results for $\beta_{s,w}$)
 3. **If unstable, what drives it?** (Macro driver analysis)
 4. **Is instability economically meaningful?** (Effect size in % terms)
 5. **Does theory-based intuition hold?** (VIX amplifies front-end, OAS compresses dispersion)
 6. **Practical recommendation:**
 - Use static λ : Adequate for normal markets, simple implementation
 - Use time-varying λ : Necessary for crisis periods, worth complexity
 - Hybrid: Static baseline with crisis adjustments ($VIX > 30$)
 - For sectors: Static sector adjustments vs regime-dependent
 7. **Implication for production:** If static suffices, Stage E will select among pure/calibrated Merton with or without sector adjustments. If time-varying needed, add macro state to production spec.

11.10 Decision point

- **If Chow test $p > 0.10$ for both base and sectors:** Static λ sufficient. Proceed to Stage D (robustness) with confidence in stable baseline.
- **If Chow test $p < 0.10$ for base but sectors stable:** Time-variation in overall level but not in sector differentials. Consider VIX overlay on base β_0 .
- **If Chow test $p > 0.10$ for base but sectors unstable:** Stable overall but regime-dependent sector effects. Use static base with VIX-contingent sector adjustments.
- **If Chow test $p < 0.01$ and effects >20% in crises:** Time-varying λ necessary. Incorporate macro state in Stage E production specification.

12 Stage D: Robustness and Extensions

12.1 Objective and positioning

Prerequisite: Stages A–C established whether/how Merton predictions hold for *mean* spread changes in standard conditions.

Objective: Test robustness across:

1. Tail events (quantile regression)
2. Shock types (systematic vs idiosyncratic decomposition)
3. Spread components (default vs liquidity)

Key framing: These are **secondary** tests. If Stages A–C show Merton fails, Stage D helps diagnose *why*. If Stages A–C validate Merton, Stage D confirms it's not just a mean effect.

Relationship to Stage 0: Stage D extends the robustness checks beyond what Stage 0's three-pronged approach covers. While Stage 0 tested Merton in aggregate (buckets), within issuers, and across sectors, Stage D examines distributional properties (tails), shock decomposition, and spread component separation.

12.2 D.1: Tail behavior (quantile regression)

12.2.1 Motivation

Merton model assumes normal shocks (geometric Brownian motion for firm value). If tails differ from mean, this suggests:

- Jump-to-default risk not captured by continuous diffusion
- Liquidity evaporation in stress (left tail)
- Asymmetric investor behavior (panic selling vs gradual buying)

12.2.2 Specification

For quantiles $\tau \in \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$, estimate:

$$Q_\tau(y_{i,t} | f_{DTS,t}, \lambda_i^{\text{Merton}}) = \alpha_\tau + \beta_\tau \cdot [\lambda_i^{\text{Merton}} \cdot f_{DTS,t}] \quad (88)$$

This models the τ -th conditional quantile of spread changes, allowing elasticity to differ across the distribution.

Merton prediction: $\beta_\tau \approx 1$ for all τ (all quantiles respect structural elasticities).

Diagnostics:

- If $\beta_{0.05} \gg \beta_{0.95}$: Left tail (spread widening) has amplified sensitivity. Consistent with jump-to-default or liquidity spirals.
- If $\beta_{0.95} > \beta_{0.05}$: Right tail (spread tightening) more sensitive. Suggests momentum/technical buying in rallies.
- If β_τ U-shaped (high at both tails): Both extreme moves behave differently than moderate moves—non-linearity in spread dynamics.

12.2.3 Sector-specific tail behavior

Building on Stage 0 sector findings, estimate quantile regressions separately by sector:

$$Q_\tau(y_{i,t} \mid f_{DTS,t}, \lambda_i^{\text{Merton}}, \text{sector}_i = s) = \alpha_{\tau,s} + \beta_{\tau,s} \cdot [\lambda_i^{\text{Merton}} \cdot f_{DTS,t}] \quad (89)$$

Hypothesis: If Stage 0 found Financials have amplified mean sensitivity, do they also have amplified tail sensitivity?

- $\beta_{0.05, \text{Financial}} > \beta_{0.05, \text{Industrial}}$: Financial left-tail risk exceeds mean amplification (correlation risk manifests in tails)
- $\beta_{0.05, \text{Utility}} \approx \beta_{0.50, \text{Utility}}$: Utility defensive characteristics hold in tails

12.2.4 Practical implications

Risk management:

- Use $\beta_{0.05}$ for Value-at-Risk (VaR) and Expected Shortfall (ES) calculations
- Use $\beta_{0.50}$ for expected return / attribution models
- If $\beta_{0.05} = 1.5 \times \beta_{0.50}$: Tail risk 50% larger than mean-based models predict

Stress testing:

- Standard Merton λ may underestimate losses in left-tail scenarios
- Adjust: $DTS_{i,t}^{*, \text{stress}} = \beta_{0.05} \cdot \lambda_i^{\text{Merton}} \cdot DTS_{i,t}$

12.2.5 Deliverables for D.1

- **Table D.1:** Quantile-specific β_τ estimates
 - Rows: $\tau \in \{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$
 - Columns: $\hat{\beta}_\tau$, standard error, 95% CI
 - Separate panels for IG and HY
 - Test: $H_0 : \beta_{0.05} = \beta_{0.50}$ and $H_0 : \beta_{0.95} = \beta_{0.50}$
- **Table D.2:** Sector-specific tail behavior
 - Rows: Sectors (Industrial, Financial, Utility, Energy)
 - Columns: $\beta_{0.05}$, $\beta_{0.50}$, $\beta_{0.95}$, Ratio $\beta_{0.05}/\beta_{0.50}$
 - Shows which sectors have amplified tail risk
- **Figure D.1:** Plot of $\hat{\beta}_\tau$ across $\tau \in [0.05, 0.95]$
 - X-axis: Quantile τ
 - Y-axis: $\hat{\beta}_\tau$
 - Horizontal line at $\beta = 1$ (Merton prediction)
 - Confidence bands (bootstrap)

- Interpretation: Flat line = Merton works across distribution. Upward/downward slope = asymmetry.
- **Figure D.2:** Sector-specific quantile plots
 - Separate lines for Industrial, Financial, Utility, Energy
 - Shows whether Financial tail amplification exceeds mean amplification
- **Interpretation note:** Are tail deviations concentrated in specific regimes (e.g., front-end IG)? If so, suggests these bonds have additional jump risk beyond Merton's continuous framework.

12.3 D.2: Shock decomposition

12.3.1 Motivation

Merton model treats all firm value shocks identically—whether macro, sector, or idiosyncratic, the elasticity λ_i should be the same because all operate through the firm's asset value.

Empirically test: Do different shock types exhibit different elasticities? If so, suggests mechanisms beyond firm fundamentals (e.g., liquidity contagion for sector shocks, information asymmetry for idiosyncratic shocks).

12.3.2 Factor construction

Decompose bond i 's spread change into orthogonal components:

$$y_{i,t} = \underbrace{f_{DTS,t}^{(G)}}_{\text{Global factor}} + \underbrace{f_{DTS,s(i),t}^{(S)}}_{\text{Sector factor}} + \underbrace{f_{DTS,i,t}^{(I)}}_{\text{Issuer-specific}} + \varepsilon_{i,t} \quad (90)$$

Estimation procedure:

Step 1: Global factor

$$f_{DTS,t}^{(G)} = \frac{\Delta OAS_t^{(U)}}{OAS_{t-1}^{(U)}} \quad (91)$$

Step 2: Sector factors (orthogonalized to global)

$$f_{DTS,s,t}^{(S)} = \frac{\Delta OAS_{s,t}^{(U)}}{OAS_{s,t-1}^{(U)}} - f_{DTS,t}^{(G)} \quad (92)$$

Step 3: Issuer-specific (residual)

$$f_{DTS,i,t}^{(I)} = y_{i,t} - f_{DTS,t}^{(G)} - f_{DTS,s(i),t}^{(S)} \quad (93)$$

12.3.3 Multi-factor regression with Merton baseline

Estimate:

$$y_{i,t} = \beta^{(G)}[\lambda_i^{\text{Merton}} \cdot f_{DTS,t}^{(G)}] + \beta^{(S)}[\lambda_i^{\text{Merton}} \cdot f_{DTS,s(i),t}^{(S)}] + \beta^{(I)}[\lambda_i^{\text{Merton}} \cdot f_{DTS,i,t}^{(I)}] + \varepsilon_{i,t} \quad (94)$$

Constrained specification: Impose $\beta^{(G)} = \beta^{(S)} = \beta^{(I)} = 1$, estimate single coefficient:

$$y_{i,t} = \beta \cdot \lambda_i^{\text{Merton}} \cdot [f_t^{(G)} + f_{s(i),t}^{(S)} + f_{i,t}^{(I)}] + \varepsilon_{i,t} \quad (95)$$

Merton prediction: $\beta^{(G)} \approx \beta^{(S)} \approx \beta^{(I)} \approx 1$ (all shocks respect structural elasticities).

12.3.4 Diagnostic patterns

- **If** $\beta^{(G)} \approx \beta^{(S)} \approx \beta^{(I)} \approx 1$: Merton applies uniformly—all shocks operate through firm value.
- **If** $\beta^{(S)} > \beta^{(G)}$: Sector shocks have amplified effects. Suggests contagion, correlation trading, or common liquidity factors beyond fundamentals.
- **If** $\beta^{(I)} \gg \beta^{(G)}$: Idiosyncratic news has exaggerated spread impact. Consistent with information asymmetry, adverse selection in trading.
- **If** $\beta^{(G)} < 1$ **but** $\beta^{(I)} > 1$: Bonds under-react to macro (diversified portfolios stabilize) but over-react to issuer-specific (concentrated positions, forced selling).

12.3.5 Connection to Stage 0 sector findings

If Stage 0 found Financials have amplified sensitivity, test whether this comes from:

- Global shocks: $\beta_{\text{Financial}}^{(G)} > \beta_{\text{Industrial}}^{(G)}$ (Financials more sensitive to macro)
- Sector shocks: $\beta_{\text{Financial}}^{(S)} > \beta_{\text{Industrial}}^{(S)}$ (correlation/contagion within Financial sector)
- Idiosyncratic shocks: $\beta_{\text{Financial}}^{(I)} > \beta_{\text{Industrial}}^{(I)}$ (information asymmetry in Financial credits)

This decomposition helps understand *why* sector effects exist, informing production model design.

12.3.6 Deliverables for D.2

- **Table D.3:** Variance decomposition
 - Rows: Rating \times maturity buckets
 - Columns: % variance from Global, Sector, Issuer-specific, Residual
 - Shows relative importance of each factor type
 - Separate for IG and HY (expect IG more global-driven, HY more issuer-specific)
- **Table D.4:** Shock-specific elasticities
 - Rows: $\beta^{(G)}, \beta^{(S)}, \beta^{(I)}$
 - Columns: Estimate, standard error, 95% CI
 - Test: $H_0 : \beta^{(G)} = \beta^{(S)} = \beta^{(I)} = 1$ (joint test)
 - Test: Pairwise $H_0 : \beta^{(G)} = \beta^{(S)}$, etc.
- **Table D.5:** Sector-specific shock decomposition
 - Rows: Sectors (Industrial, Financial, Utility, Energy)
 - Columns: $\beta_s^{(G)}, \beta_s^{(S)}, \beta_s^{(I)}$
 - Shows which shock types drive each sector's behavior
- **Figure D.3:** Bar chart of $\hat{\beta}$ by factor type
 - Three bars: Global, Sector, Issuer-specific

- Error bars for 95% CI
- Horizontal line at $\beta = 1$
- Separate panels for IG and HY
- **Interpretation:** If all three ≈ 1 , Merton universally applicable. If sector/issuer deviate, need factor-specific adjustments in production models.

12.4 D.3: Liquidity adjustment

12.4.1 Motivation

OAS includes both default risk and liquidity premium. Merton model predicts elasticity for *default component only*. If liquidity shocks don't respect structural elasticities, need to decompose.

12.4.2 Liquidity-adjusted spread construction

Step 1: Estimate liquidity component cross-sectionally each period:

$$s_{i,t}^{\text{liq}} = \phi_0 + \phi_1 \text{BidAsk}_{i,t} + \phi_2 \log(\text{Size}_i) + \phi_3 \log(\text{Turnover}_{i,t}) + \phi_4 \text{Age}_{i,t} + \eta_{i,t} \quad (96)$$

Standard approach: Cross-sectional regression within each rating \times maturity cell to avoid confounding credit quality with liquidity.

Step 2: Define default component as residual:

$$OAS_{i,t}^{\text{def}} = OAS_{i,t} - \widehat{OAS}_{i,t}^{\text{liq}} \quad (97)$$

Step 3: Compute default-based spread changes:

$$y_{i,t}^{\text{def}} = \frac{\Delta OAS_{i,t}^{\text{def}}}{OAS_{i,t-1}^{\text{def}}} \quad (98)$$

12.4.3 Re-estimate Merton fit on default component

Run Stage B regression using $y_{i,t}^{\text{def}}$ as dependent variable:

$$y_{i,t}^{\text{def}} = \alpha + \beta_{\text{def}} \cdot [\lambda_i^{\text{Merton}} \cdot f_{DTS,t}^{\text{def}}] + \varepsilon_{i,t} \quad (99)$$

where $f_{DTS,t}^{\text{def}}$ is the index-level default spread factor (after removing liquidity).

Theory prediction: $\beta_{\text{def}} \approx 1$ (Merton works for default component) and $R_{\text{def}}^2 > R_{\text{total}}^2$ (less noise without liquidity shocks).

12.4.4 Diagnostic patterns

- **If $\beta_{\text{def}} \approx 1$ and $\beta_{\text{total}} \approx 1$:** Liquidity adjustment unnecessary—Merton works on total OAS.
- **If $\beta_{\text{def}} \approx 1$ but $\beta_{\text{total}} < 1$:** Liquidity shocks dampen spread movements. Total OAS underestimates default sensitivity.
- **If $\beta_{\text{def}} < \beta_{\text{total}}$:** Liquidity shocks amplify spread movements beyond Merton. Need separate liquidity beta.
- **If improvement concentrated in illiquid bonds:** Signals that liquidity noise matters for small size, low turnover bonds but not for liquid benchmarks.

12.4.5 Sector-specific liquidity effects

Building on Stage 0 findings, test whether liquidity adjustment differentially affects sectors:

- Financials: Often more liquid (active trading, market makers); liquidity adjustment may be less important
- Utilities: Often less liquid (buy-and-hold investors); liquidity adjustment may help
- Energy: Liquidity varies with oil price cycles; may need regime-dependent liquidity treatment

12.4.6 Practical decision rule

When to Use Liquidity Adjustment:

Liquid IG (BidAsk < 50bps, Size > \$1B):

- Use total OAS, liquidity adjustment negligible (<5bps typical)
- Liquidity R^2 low, not worth complexity

Illiiquid IG (BidAsk > 100bps, Size < \$500M):

- Liquidity premium 20–50bps, material fraction of spread
- Consider decomposition for risk models

HY (BidAsk typically > 200bps):

- Decompose into OAS^{def} and OAS^{liq}
- Use λ^{def} from Merton for default component
- Add separate λ^{liq} empirically estimated (not theory-based)

12.4.7 Deliverables for D.3

- **Table D.6:** Liquidity model estimates

- Cross-sectional regression: $\hat{\phi}_1$ (bid-ask), $\hat{\phi}_2$ (size), $\hat{\phi}_3$ (turnover), $\hat{\phi}_4$ (age)
- R^2 by rating \times maturity bucket
- Typical liquidity premium: Mean $\widehat{OAS}^{\text{liq}}$ by bucket

- **Table D.7:** Merton fit comparison

- Rows: Total OAS, Default component only
- Columns: $\hat{\beta}$, R^2 , RMSE
- Test: Is β_{def} closer to 1 than β_{total} ?
- Improvement: $\Delta R^2 = R_{\text{def}}^2 - R_{\text{total}}^2$

- **Table D.8:** Improvement by liquidity regime

- Split sample by liquidity quartiles (bid-ask or turnover)
- For each quartile: β_{total} , β_{def} , ΔR^2

- Shows whether adjustment helps mainly for illiquid bonds
- **Table D.9:** Sector-specific liquidity effects
 - Rows: Sectors
 - Columns: β_{total} , β_{def} , ΔR^2
 - Shows which sectors benefit most from liquidity decomposition
- **Figure D.4:** Scatter plot: $\beta_{\text{def}} - \beta_{\text{total}}$ (y-axis) vs average bid-ask (x-axis)
 - Each point: a rating \times maturity bucket
 - Interpretation: Positive slope = illiquid bonds benefit more from decomposition
- **Interpretation and recommendation:**
 - For which bond types is liquidity adjustment material?
 - Does Merton work better on default component than total OAS?
 - Production implication: Use total OAS for liquid IG, decompose for HY?

12.5 Summary of Stage D

Stage D is **diagnostic**, not **decisional**. It answers:

1. **Where does Merton fail?** (Tails? Specific shock types? Liquidity-contaminated spreads?)
2. **How large are the failures?** (20% effect in tails? 50%? Negligible?)
3. **Are failures systematic?** (Always front-end IG? Always sector shocks? Always illiquid bonds?)
4. **Do failures relate to Stage 0 findings?** (Is Financial amplification driven by tails? sector shocks? liquidity?)

Use Stage D results in Stage E to decide:

- If tail effects large: Add quantile-specific λ for VaR/ES
- If sector shocks deviate: Consider sector-factor adjustments beyond static sector dummies
- If liquidity matters: Decompose OAS for HY, keep total for IG
- If Financial amplification comes from specific channel: Design targeted adjustment

But don't let Stage D derail the main program. If Stages A–C showed Merton works well on average (possibly with sector adjustments from Stage 0), Stage D refinements are enhancements, not requirements.

13 Stage E: Production Specification Selection

13.1 Objective

Stage E selects the parsimonious production model that balances theoretical coherence, empirical fit, and implementation cost.

Key principle: Use *hierarchical testing* guided by theory. Stop at the simplest adequate model. Don't over-engineer.

Philosophical stance: Theory provides a strong prior. Only deviate when data strongly reject it. The burden of proof is on the more complex model.

Integration of Stage 0 findings: The three-pronged Stage 0 analysis (bucket-level, within-issuer, sector interaction) provides critical input for production specification. Sector adjustments identified in Stage 0 and validated in Stages B–C should be incorporated where statistically and economically significant.

13.2 Decision framework: Hierarchical testing

Level 1: Is standard DTS adequate?

Test: Stage A, F -test for $H_0 : \beta^{(1)} = \beta^{(2)} = \dots = \beta^{(K)} = 1$

Decision rule:

- If fail to reject ($p > 0.10$): **Use standard DTS.** No adjustments needed. Done.
- If reject: Proceed to Level 2

Production spec: $y_{i,t} \approx f_{DTS,t}$

Parameters: 0

Implementation: Trivial (already in all systems)

Level 2: Does pure Merton suffice?

Test: Stage B Specification B.1, $H_0 : \beta_{\text{Merton}} = 1$

Decision rule:

- If $\hat{\beta}_{\text{Merton}} \in [0.9, 1.1]$ **and** $R_{\text{Merton}}^2 > 0.9 \times R_{\text{buckets}}^2$ **and** Stage 0 sector interactions not significant:

Use Pure Merton (Tables 2–3 from [Wuebbgen \[2025\]](#)). Done.

- If Stage 0 found significant sector effects: Proceed to Level 2a
- If systematic bias ($\hat{\beta}_{\text{Merton}}$ outside $[0.9, 1.1]$): Proceed to Level 3
- If poor fit ($R_{\text{Merton}}^2 < 0.7 \times R_{\text{buckets}}^2$): Proceed to Level 4

Production spec:

$$\lambda_i^{\text{prod}} = \lambda_T(T_i; 5y, s_i) \times \lambda_s(s_i; 100) \quad (100)$$

Parameters: 0 (lookup tables from [Wuebbgen \[2025\]](#))

Implementation: Read current s_i, T_i , look up λ from tables or use closed-form $(s/100)^{-0.25}$

Level 2a: Pure Merton with sector adjustments

Condition: Stage 0 found significant sector effects (joint F -test $p < 0.05$, ≥ 1 sector with $|\hat{\beta}_s| > 0.2$), confirmed in Stage B Specification B.3

Test: Stage B Specification B.3, joint test for sector coefficients

Decision rule:

- If $\hat{\beta}_0 \in [0.9, 1.1]$ and sector effects stable in Stage C: **Use Pure Merton with Sector Adjustments.** Done.
- If sector effects regime-dependent (Stage C): Proceed to Level 2b
- If $\hat{\beta}_0$ outside $[0.9, 1.1]$: Proceed to Level 3a

Production spec:

$$\lambda_i^{\text{prod}} = \lambda_T(T_i; 5y, s_i) \times \lambda_s(s_i; 100) \times (1 + \hat{\beta}_{\text{sector}_i}) \quad (101)$$

Parameters: $|S| - 1$ (sector dummies, typically 3–4)

Implementation: Read $s_i, T_i, \text{sector}_i$, apply Merton lookup, multiply by sector factor

Level 2b: Pure Merton with regime-dependent sector adjustments

Condition: Stage C found sector effects vary with VIX or spread level

Production spec:

$$\lambda_{i,t}^{\text{prod}} = \lambda_i^{\text{Merton}} \times (1 + \hat{\beta}_{s,0} + \hat{\beta}_{s,\text{VIX}} \cdot \text{VIX}_t) \quad (102)$$

where $s = \text{sector}_i$.

Parameters: $2 \times (|S| - 1)$ (base and VIX coefficient per sector)

Implementation: Requires daily VIX feed, sector classification, lookup tables

Level 3: Calibrated Merton (without sectors)

Condition: Merton has right functional form but wrong scale, no significant sector effects

Production spec:

$$\lambda_i^{\text{prod}} = c_0 \cdot \lambda_T(T_i; 5y, s_i) \times \lambda_s(s_i; 100)^{c_s} \quad (103)$$

where:

- c_0 : overall scaling factor (estimated from Stage B: $\hat{c}_0 = \hat{\beta}_{\text{Merton}}$)
- c_s : adjust power-law exponent (theory suggests -0.25 , data may differ)

Estimation:

$$y_{i,t} = \alpha + \beta_0[\lambda_T(T_i) \cdot f_{DTS,t}] + \beta_s[\lambda_s(s_i)^c \cdot f_{DTS,t}] + \varepsilon_{i,t} \quad (104)$$

Grid search over $c \in [-0.5, 0]$ to maximize R^2 , then calibrate β_0, β_s .

Decision rule:

- If $\hat{c}_0 \in [0.8, 1.2]$ and $\hat{c}_s \in [-0.35, -0.15]$: Theory approximately correct. Use calibrated version. Done.
- If outside these ranges: Proceed to Level 4

Parameters: 2 (c_0, c_s)

Implementation: Requires one-time calibration using sample data, then fixed multipliers

Level 3a: Calibrated Merton with sector adjustments

Condition: Need both calibration (Level 3) and sector adjustments (Level 2a)

Production spec:

$$\lambda_i^{\text{prod}} = c_0 \cdot \lambda_T(T_i)^{c_T} \times \lambda_s(s_i)^{c_s} \times (1 + \hat{\beta}_{\text{sector}_i}) \quad (105)$$

Parameters: 2–3 (calibration) + 3–4 (sectors) = 5–7 total

Implementation: Moderate complexity; annual recalibration recommended

Level 4: Full empirical (without sectors)

Condition: Calibrated Merton inadequate, need unrestricted functional form, no significant sector effects

Production spec:

$$\lambda_i^{\text{prod}} = \exp(\beta_0 + \beta_M \log M_i + \beta_s \log s_i + \beta_{M^2} (\log M_i)^2 + \beta_{Ms} \log M_i \cdot \log s_i) \quad (106)$$

Estimate via Stage B Specification B.4 (unrestricted).

Decision rule:

- If $R^2_{\text{unrestricted}} - R^2_{\text{calibrated}} > 0.05$: Empirical spec justified. Done.
- Otherwise: Stay at Level 3 (principle of parsimony)

Parameters: 5–8 (maturity terms, spread terms, interactions)

Implementation: Requires regression estimation, periodic recalibration (annually)

Level 4a: Full empirical with sector adjustments

Condition: Need unrestricted functional form plus sector effects

Production spec:

$$\lambda_i^{\text{prod}} = \exp(\beta_0 + \beta_M \log M_i + \beta_s \log s_i + \beta_{M^2} (\log M_i)^2 + \beta_{Ms} \log M_i \cdot \log s_i + \sum_{\text{sector}} \beta_{\text{sec}}) \quad (107)$$

Parameters: 8–12 (maturity terms, spread terms, interactions, sector dummies)

Implementation: Complex; requires annual recalibration, sector classification

Level 5: Time-varying (optional add-on to any level)

Condition: Stage C showed significant instability ($p < 0.01$ in Chow test) **and** macro state changes λ by >30% during crises

Production spec:

$$\lambda_{i,t}^{\text{prod}} = \lambda_i^{\text{base}} \times \exp(\gamma_{\text{VIX}} \cdot \text{VIX}_t + \gamma_{\text{OAS}} \cdot \log(\text{OAS}_{\text{index},t})) \quad (108)$$

where λ_i^{base} comes from Level 2, 2a, 3, 3a, 4, or 4a.

Decision rule:

- If improvement in crisis periods substantial (RMSE reduction >20%) **and** operational complexity acceptable:

Production spec = Time-varying with base + 2 macro parameters

- Otherwise: Use static λ_i^{base} with manual crisis overlays

Parameters: Base parameters + 2 (macro state coefficients)

Implementation: Requires daily macro data feeds, dynamic recalculation

13.3 Summary of specification hierarchy

Table 11: Model Comparison: Parsimony vs Performance

Specification	Params	R^2	Impl.	Use When
Standard DTS	0	R_0^2	Trivial	No cross-sectional variation
Pure Merton	0	R_M^2	Simple	Theory unbiased, no sector effects
Pure Merton + Sectors	3–4	R_{M+S}^2	Simple	Theory unbiased, stable sector effects
Calibrated Merton	2	$R_M^2 + \delta_1$	Moderate	Theory right structure, needs scaling
Calibrated + Sectors	5–7	$R_{M+S}^2 + \delta_1$	Moderate	Calibration needed, stable sector effects
Empirical	5–8	$R_M^2 + \delta_2$	Complex	Theory inadequate, no sector effects
Empirical + Sectors	8–12	$R_{M+S}^2 + \delta_2$	Complex	Theory inadequate, sector effects present
Any + Time-varying	+2	$+\delta_3$	Very complex	Significant instability, large crisis effects

Expected incremental R^2 gains:

- Pure Merton vs Standard DTS: +5–15% (if Regime 2 prevalent)
- Adding sector adjustments: +2–5% (if Stage 0 found significant effects)
- Calibrated vs Pure Merton: +1–3% (scaling correction)
- Empirical vs Calibrated: +2–5% (if sectors or interactions matter beyond theory)
- Time-varying vs Static: +1–2% overall, +10–20% in crises

Implementation cost ranking:

1. Standard DTS: Zero (already implemented)

2. Pure Merton: Low (lookup tables, one-time coding)
3. Pure Merton + Sectors: Low (add sector classification)
4. Calibrated Merton: Low-moderate (one-time calibration, quarterly review)
5. Calibrated + Sectors: Moderate (calibration plus sector tracking)
6. Empirical: Moderate-high (annual recalibration, more inputs)
7. Empirical + Sectors: High (many parameters, sector classification)
8. Time-varying: High (daily macro feeds, operational burden)

13.4 Out-of-sample validation

For each candidate specification at its decision level, conduct rolling-window out-of-sample test.

13.4.1 Methodology

Window structure:

- Training window: 3 years (rolling)
- Test window: 1 year (out-of-sample)
- Roll forward by 1 year, repeat

Procedure for each window:

1. Estimate specification parameters using training data
2. Generate predictions $\hat{y}_{i,t}^{\text{model}} = \hat{\lambda}_{i,t-1} \cdot f_{DTS,t}$ for test window
3. Compute performance metrics
4. Roll window forward

13.4.2 Performance metrics

Metric 1: Forecast accuracy

$$\text{RMSE}_{\text{OOS}} = \sqrt{\frac{1}{N_{\text{test}}} \sum_{i,t \in \text{test}} (y_{i,t} - \hat{y}_{i,t})^2} \quad (109)$$

Compare across specifications. Lower is better.

Metric 2: Out-of-sample R^2

$$R_{\text{OOS}}^2 = 1 - \frac{\sum (y_{i,t} - \hat{y}_{i,t})^2}{\sum (y_{i,t} - \bar{y})^2} \quad (110)$$

Can be negative if model performs worse than mean. Compare to in-sample R^2 to assess overfitting.

Metric 3: Hedge effectiveness

Construct hedges based on model predictions:

- Portfolio: Long 1–2y BBB bonds
- Hedge: Short $h_t = \lambda_{1y}^{\text{model}} / \lambda_{5y}^{\text{model}}$ units of 5y BBB bonds

Compute tracking error of hedged portfolio. Lower tracking error = better model.

Metric 4: Sharpe ratio of mispricing signal

Define mispricing:

$$\text{Mispricing}_{i,t} = y_{i,t} - \hat{y}_{i,t}^{\text{model}} \quad (111)$$

If bond under-reacted last period (positive mispricing), predict mean-reversion. Trade on signal:

$$\text{Position}_{i,t+1} = -\text{sign}(\text{Mispricing}_{i,t}) \quad (112)$$

Better model \Rightarrow stronger mean-reversion \Rightarrow higher Sharpe on position.

13.4.3 Regime-specific performance

Evaluate each specification separately during:

1. **Normal periods:** VIX < 20
2. **Stress periods:** VIX $\in [20, 30]$
3. **Crisis periods:** VIX > 30

Key questions:

- Does time-varying specification outperform static mainly in crises? If so, operational complexity may not be worth 98% of the time?
- Do sector adjustments help uniformly or mainly in specific regimes?
- Does Financial sector amplification (if present) manifest primarily during stress?

13.4.4 Sector-specific performance

If Stage 0 found significant sector effects, evaluate performance by sector:

- Compare RMSE for Industrial, Financial, Utility, Energy separately
- Test whether sector adjustments improve performance *within* each sector
- Identify if certain sectors drive overall improvement (e.g., sector adjustments help mainly for Financials)

13.5 Recommended approach

Decision Protocol:

Step 1: Review Stage 0 findings

- Did within-issuer analysis validate Merton mechanism?
- Did sector interaction analysis find significant sector effects?
- Which sectors differ and by how much?

Step 2: Execute hierarchical testing (Levels 1–5)

- If Stage 0 found no sector effects: Test Levels $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$
- If Stage 0 found sector effects: Test Levels $1 \rightarrow 2a \rightarrow 3a \rightarrow 4a$
- At each level, conduct stopping criterion test

Step 3: At the level where you stop, conduct out-of-sample validation against:

- Previous level (simpler)
- Next level (more complex)

Step 4: Apply parsimony rule

- If next level improves out-of-sample RMSE by $<5\%$: Stick with current level
- If next level improves out-of-sample RMSE by $>10\%$: Adopt more complex model
- If improvement 5–10%: Consider implementation costs

Step 5: Document production specification with:

- Parameter values (if any)
- Sector adjustments (if any)
- Implementation pseudo-code
- Recalibration frequency
- Performance benchmarks by regime and sector

Guiding principle: Occam's Razor—prefer simplest model with adequate fit. A 2-parameter model with $R^2 = 0.75$ beats a 20-parameter model with $R^2 = 0.78$.

13.6 Deliverables for Stage E

- **Table E.1:** Hierarchical test results
 - Rows: Levels 1, 2, 2a, 3, 3a, 4, 4a, 5
 - Columns: Test statistic, p -value, Decision (PASS / Proceed to next level / N/A)
 - Mark the stopping level
 - Note which path taken (with or without sectors)
- **Table E.2:** Model comparison (all candidate specs)
 - Rows: Standard DTS, Pure Merton, Pure Merton + Sectors, Calibrated, Calibrated + Sectors, Empirical, Empirical + Sectors, Time-varying variants
 - Columns: Parameters, In-sample R^2 , OOS R^2 , OOS RMSE, Hedge tracking error, Sharpe ratio
 - Highlight recommended specification
- **Table E.3:** Performance by regime

- Panel A: Normal periods ($VIX < 20$)
- Panel B: Stress periods ($VIX 20\text{--}30$)
- Panel C: Crisis periods ($VIX > 30$)
- For each panel: OOS R^2 and RMSE for each specification
- Shows where complexity pays off

- **Table E.4:** Performance by sector (if sector adjustments used)

- Rows: Industrial, Financial, Utility, Energy
- Columns: RMSE (no sector adj), RMSE (with sector adj), Improvement
- Shows which sectors drive benefit of sector adjustments

- **Table E.5:** Recommended production specification

- Specification name and level
- Parameter estimates with standard errors
- Sector adjustment factors (if applicable)
- Implementation formula
- Recalibration frequency
- Expected R^2 improvement over baseline
- Complexity rating (Low / Moderate / High)

- **Figure E.1:** Out-of-sample R^2 over rolling windows

- X-axis: Test window start date
- Y-axis: OOS R^2
- Multiple lines: Key specifications (Standard DTS, Pure Merton, Merton + Sectors, Calibrated, etc.)
- Shade crisis periods
- Shows stability and relative performance over time

- **Figure E.2:** Forecast error distribution

- Histogram of $(y_{i,t} - \hat{y}_{i,t})$ for recommended spec
- Overlay: Normal distribution with same mean/variance
- Q-Q plot in corner panel
- Check for unmodeled fat tails or asymmetry

- **Figure E.3:** Scatter: Predicted vs actual spread changes

- X-axis: $\hat{y}_{i,t}$ (model prediction)
- Y-axis: $y_{i,t}$ (actual)
- 45-degree line
- Color-code by regime (IG-narrow, IG-wide, HY, distressed)
- Shows where model works best/worst

- **Figure E.4:** Sector-specific performance (if applicable)
 - Bar chart: RMSE by sector, with and without sector adjustments
 - Shows incremental benefit of sector factors
- **Implementation blueprint (5–7 pages):**

1. **Algorithmic steps:**

- Input data required (bond OAS, maturity, rating, sector; macro state if time-varying)
- Step-by-step calculation of λ_i or $\lambda_{i,t}$
- Application of sector adjustments (if applicable)
- Computation of adjusted DTS: $DTS_{i,t}^* = \lambda_i \cdot DTS_{i,t}$

2. **Pseudo-code / function definitions:**

```

function compute_lambda(OAS, Maturity, Sector):
    # Pure Merton + Sectors example
    lambda_T = lookup_maturity_adjustment(Maturity, OAS)
    lambda_s = (OAS / 100) ^ (-0.25)
    lambda_base = lambda_T * lambda_s

    # Apply sector adjustment (from Stage 0/B)
    sector_factor = get_sector_factor(Sector)
    # e.g., Financial: 1.33, Utility: 1.0, Energy: 1.0

    lambda = lambda_base * sector_factor
    return lambda

function lookup_maturity_adjustment(T, s):
    # Bilinear interpolation on Table 3.1
    # Returns lambda_T(T; 5y, s)
    ...

function get_sector_factor(Sector):
    # From Stage 0 sector interaction estimates
    factors = {
        'Industrial': 1.00,  # reference
        'Financial': 1.33,   # beta_Financial = 0.33
        'Utility': 0.98,     # beta_Utility = -0.02
        'Energy': 1.15       # beta_Energy = 0.15
    }
    return factors.get(Sector, 1.00)

```

3. **Lookup tables:** Full tables if using Pure Merton (or reference to theory paper), plus sector adjustment factors

4. **Recalibration protocol:**

- Frequency: Quarterly review for Pure Merton (check if spreads/maturities still in table range); Annual for Calibrated/Empirical; Semi-annual for sector factors

- Procedure: Re-estimate using trailing 3-year window, compare to current parameters
- Trigger for update: If new estimates differ by >20%, adopt new parameters
- Sector factor review: Re-run Stage 0 sector interaction test annually
- Documentation: Maintain version history of parameter changes

5. Edge case handling:

- Very short maturity (< 6 months): Use 6-month λ as floor, don't extrapolate
- Distressed spreads (> 2000 bps): Cap λ at 1.2 (proportionality improves in extreme stress)
- Missing sector: Use aggregate cross-sector average (no sector adjustment)
- New issues (< 3 months old): Apply new issue adjustment γ_{new} if Stage B found significant effect
- Sector reclassification: Use new sector's adjustment factor immediately

6. Integration with existing systems:

- Risk models: Replace $DTS_{i,t}$ with $DTS_{i,t}^* = \lambda_i \cdot DTS_{i,t}$ in portfolio risk calculations
- Attribution: Decompose factor return into $f_{DTS,t}$, cross-sectional λ_i effect, and sector effect
- Relative value: Flag bonds with large deviations from λ -adjusted fair value
- Sector allocation: Account for differential DTS sensitivity when sizing sector bets

7. Performance monitoring:

- Track out-of-sample R^2 monthly, by sector
- Alert if R^2 drops below threshold (e.g., < 50% of historical average)
- Compare to benchmark (Standard DTS) quarterly
- Monitor sector factor stability (flag if sector $\hat{\beta}_s$ changes by >0.1)

• Comparative performance analysis (3–4 pages):

1. Executive summary: How much does recommended spec improve over baseline DTS?

- RMSE reduction: e.g., 25% lower forecast errors

2. Which regimes see largest gains?

- IG with wide maturity range: 40% RMSE reduction (Regime 2 targeted)

3. Economic value: Translate statistical improvements to portfolio management gains

– Example 1: Hedging efficiency

“Portfolio manager long \$100M of 1-year BBB industrials, wants to hedge with 5-year BBB.

Standard DTS: Hedge ratio = 1.0, residual tracking error = 120 bps/year

Merton-adjusted: Hedge ratio = 3.2, residual tracking error = 40 bps/year

Value: 80 bps lower tracking error, equivalent to \$800K lower unexpected P&L volatility”

- **Example 2: Relative value signals**
 “Identifying rich/cheap bonds within issuer capital structure.
 Standard DTS: Assumes all bonds move proportionally, misses 300–500% cross-maturity differences.
 Merton-adjusted: Properly scales for maturity, identifies mispricings averaging 15 bps.
 Trade: Long cheap (e.g. 10y), short rich (e.g. 1y), earn 30 bps as convergence occurs over 6 months.”
- **Example 3: Sector allocation with DTS adjustment**
 “Allocating risk budget across sectors.
 Standard DTS: Treats Financial and Industrial DTS identically.
 With sector adjustment: Recognizes Financial has 33% amplified sensitivity.
 Result: For same DTS exposure, reduce Financial allocation by 25% to equalize true risk.
 Avoids unintended concentration in high-sensitivity sector.”
- **Example 4: Portfolio construction**
 “Constructing credit barbell (short + long maturity, avoid intermediate).
 Standard DTS: Treats all durations equally, over-concentrates DTS risk in short end.
 Merton-adjusted: Recognizes front-end 3–4× more sensitive, rebalances to equalize risk contribution.
 Result: More stable returns, 25% reduction in unexpected drawdowns during spread volatility spikes.”

4. Limitations and caveats:

- Out-of-sample performance may degrade if regime shifts outside historical experience
- Liquidity crises (2008-style) may break Merton assumptions
- Model requires clean data (OAS, maturity, sector, liquidity proxies)—quality control critical
- Parameter drift possible—requires monitoring and periodic recalibration
- Sector factors estimated on historical data may not capture future sector dynamics

5. Sensitivity to implementation choices:

- Bucket definitions: Finer buckets improve fit but reduce sample size per bucket
- Frequency: Daily more noisy, monthly smoother but lags; weekly optimal balance
- Clustering: Week clustering adequate, bond clustering more conservative but similar results
- Sector classification: Results robust to minor reclassifications; major changes (e.g., splitting Financials into Banks/Insurance) may require re-estimation

14 Summary of Research Tasks

This section provides an actionable checklist for executing the research program.

Phase 0: Data and Universe Construction

1. Extract full histories of bond-level OAS, OASD, ratings, maturities, sectors, liquidity proxies for all constituents of Bloomberg Barclays U.S. Corporate IG and HY indices (2013–2025)
2. Construct weekly observation grid (Friday close or last trading day)

3. Apply bond-level filters: seniority, maturity ≥ 1 y, liquidity ≥ 5 trades/month, price 30–150% par
4. **Construct issuer identifiers** using Ultimate Parent + Seniority matching (Section 2.7)
 - Validate for top 50 issuers by amount outstanding
 - Document edge cases (holding company vs subsidiary, secured vs unsecured)
 - Cross-check against rating agency issuer hierarchies
5. Compute $y_{i,t} = \Delta OAS_{i,t}/OAS_{i,t-1}$ and $f_{DTS,t}^{(U)} = \Delta OAS_t^{(U)}/OAS_{t-1}^{(U)}$
6. Document sample size: bonds per week, issuers with 2+ bonds, total issuer-weeks, bonds per sector
7. **Deliverable:** Sample construction memo (3–5 pages) + summary statistics table

Phase 0.5: Structural Model Setup

8. Implement Merton model functions: $\lambda_T(T; s)$ and $\lambda_s(s)$ using formulas from theory paper
9. Generate lookup tables (Tables 2–3) if not already available
10. For each bond-date, compute $\lambda_{i,t}^{\text{Merton}} = \lambda_T(T_i; 5y, s_i) \times \lambda_s(s_i; 100)$
11. Validate: Spot-check 10–20 bonds to ensure λ^{Merton} values sensible (e.g., 1y IG $\lambda \approx 3$, 10y IG $\lambda \approx 0.6$)
12. **Deliverable:** Merton calculation functions (code) + validation memo

Phase 1: Stage 0 — Bucket-Level Analysis

13. Define buckets: Rating (AAA/AA, A, BBB for IG; BB, B, CCC for HY) \times Maturity (1–2y, 2–3y, 3–5y, 5–7y, 7–10y, 10y+) \times Sector (Industrial, Financial, Utility, Energy)
14. For each bucket k , estimate: $y_{i,t} = \alpha^{(k)} + \beta^{(k)} f_{DTS,t} + \varepsilon_{i,t}^{(k)}$
15. Compute $\lambda_k^{\text{Merton}}$ for each bucket using median characteristics
16. Statistical tests: t -test for $H_0 : \beta^{(k)} = \lambda_k^{\text{Merton}}$, Spearman correlation (maturity vs λ), regime pattern test
17. **Deliverables:** Tables 0.1–0.2, Figures 0.1–0.3

Phase 1.5: Stage 0 — Within-Issuer Analysis

18. Identify all issuer-weeks with ≥ 2 bonds outstanding (same Ultimate Parent + Seniority, maturity dispersion ≥ 2 years)

19. For each issuer-week, estimate within-issuer regression:

$$\frac{\Delta s_{ij,t}}{s_{ij,t-1}} = \alpha_{i,t} + \beta_{i,t} \cdot \lambda_{ij,t}^{\text{Merton}} + \epsilon_{ij,t}$$

20. Pool across issuer-weeks using inverse-variance weighting
 21. Diagnostic tests: Spread level effect, maturity dispersion effect, crisis vs normal comparison
 22. **Deliverables:** Tables 0.3–0.4, Figures 0.4–0.6, case studies for major issuers

Phase 1.6: Stage 0 — Sector Interaction Analysis

23. Estimate sector interaction specification:

$$y_{i,t} = \alpha + \beta_0 [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \sum_{s \in S} \beta_s \cdot \mathbb{1}_{\{i \in s\}} \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t}$$

24. Statistical tests: Joint F -test for all $\beta_s = 0$, individual t -tests, pairwise comparisons
 25. If significant sector effects found, estimate regime-dependent specification with VIX interactions
 26. **Deliverables:** Tables 0.5–0.6, Figures 0.7–0.9

Phase 1.7: Stage 0 — Synthesis

27. Apply synthesis framework (Section 4.5): Answer Questions 1–3
 28. Resolve any apparent contradictions between approaches
 29. Determine path for subsequent stages (Paths 1–5 from Section 4.7)
 30. **Deliverables:** Table 0.7 (synthesis summary), Figure 0.10, written summary (5–7 pages)
 31. **Decision point:** Does Merton provide adequate baseline? Are sector adjustments needed?

Phase 2: Stage A — Establish Variation

32. Estimate bucket-level regressions (Spec A.1): $y_{i,t} = \alpha^{(k)} + \beta^{(k)} f_{DTS,t} + \varepsilon_{i,t}^{(k)}$
 33. F -test for equality: $H_0 : \beta^{(1)} = \dots = \beta^{(K)}$
 34. Two-step estimation (Spec A.2): Rolling window bond-specific betas, then cross-sectional regression on $(M_i, s_i, M_i^2, M_i \cdot s_i)$
 35. **Deliverables:** Tables A.1–A.3, Figures A.1–A.2, diagnostic summary (2 pages)
 36. **Decision point:** Is variation significant? (If F -test $p > 0.10$, stop—standard DTS adequate)

Phase 3: Stage B — Test Theory

37. Estimate constrained Merton (Spec B.1): $y_{i,t} = \alpha + \beta_{\text{Merton}} \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t}$
38. Test: Wald $H_0 : \beta_{\text{Merton}} = 1$
39. Estimate decomposed Merton (Spec B.2): Separate β_T and β_s
40. Estimate Merton with sector adjustments (Spec B.3): Include sector interaction terms from Stage 0
41. Estimate unrestricted (Spec B.4): Flexible $\lambda(M_i, s_i, \text{sector}_i)$
42. Construct Theory vs Reality table: Empirical $\hat{\beta}^{(k)}$ vs Theoretical $\lambda_k^{\text{Merton}}$ for all buckets
43. Compare R^2 : Does $R_{\text{Merton}}^2 > 0.9 \times R_{\text{buckets}}^2$? Does adding sectors help?
44. **Deliverables:** Tables B.1–B.3, Figures B.1–B.3, diagnostic summary (3–4 pages)
45. **Decision point:** Pure Merton / Merton + Sectors / Calibrated / Unrestricted? Follow decision tree from Section 10

Phase 4: Stage C — Test Stability

46. Divide sample into non-overlapping 1-year windows
47. For each window, estimate: $y_{i,t} = \alpha_w + \beta_w \cdot [\lambda_{i,t}^{\text{Merton}} \cdot f_{DTS,t}] + \varepsilon_{i,t}$
48. Chow test: $H_0 : \beta_1 = \beta_2 = \dots = \beta_W$
49. If sector adjustments used, test stability of sector coefficients across windows
50. **If stable** ($p > 0.10$): Stop Stage C, use static λ (with or without sector adjustments)
51. **If unstable:** Estimate macro driver regression: $\hat{\beta}_w = \delta_0 + \delta_{\text{VIX}} \overline{\text{VIX}}_w + \delta_{\text{OAS}} \log(\overline{\text{OAS}}_w) + \eta_w$
52. Assess economic significance: Does macro state change λ by $>20\%$?
53. Maturity-specific and sector-specific time-variation tests
54. **Deliverables:** Tables C.1–C.5, Figures C.1–C.5, summary (3–4 pages)
55. **Decision point:** Static sufficient / Time-varying necessary? Follow thresholds from Section 11

Phase 5: Stage D — Robustness

56. Quantile regression (D.1): Estimate $Q_\tau(y_{i,t}) = \alpha_\tau + \beta_\tau \cdot [\lambda_i^{\text{Merton}} \cdot f_{DTS,t}]$ for $\tau \in \{0.05, 0.50, 0.95\}$
57. Sector-specific tail behavior: Estimate quantile regressions by sector
58. Test: Is $\beta_{0.05}$ significantly different from $\beta_{0.50}$? Do sectors differ in tail behavior?

59. Shock decomposition (D.2): Construct orthogonalized factors $f_t^{(G)}, f_{s,t}^{(S)}, f_{i,t}^{(I)}$
60. Estimate: $y_{i,t} = \beta^{(G)}[\lambda_i^{\text{Merton}} f_t^{(G)}] + \beta^{(S)}[\lambda_i^{\text{Merton}} f_{s,t}^{(S)}] + \beta^{(I)}[\lambda_i^{\text{Merton}} f_{i,t}^{(I)}] + \varepsilon_{i,t}$
61. Sector-specific shock decomposition: Which shock types drive sector effects?
62. Test: $H_0 : \beta^{(G)} = \beta^{(S)} = \beta^{(I)}$
63. Liquidity adjustment (D.3): Estimate $s_{i,t}^{\text{liq}}$ cross-sectionally, define $OAS_{i,t}^{\text{def}} = OAS_{i,t} - \widehat{OAS}_{i,t}^{\text{liq}}$
64. Re-estimate Stage B using $y_{i,t}^{\text{def}}$, compare β_{def} to β_{total}
65. Sector-specific liquidity effects
66. **Deliverables:** Tables D.1–D.9, Figures D.1–D.4, interpretation notes
67. **Use in Stage E:** Identify where/why Merton fails, inform production spec refinements
- Phase 6: Stage E — Production Specification**
68. Review Stage 0 findings on sector effects
69. Execute hierarchical testing: Levels 1–5 (with or without sector track) from Section 13
70. At each level, conduct stopping criterion test
71. Identify stopping level (e.g., Level 2a = Pure Merton + Sectors)
72. Rolling-window out-of-sample validation: 3-year training, 1-year test, roll forward
73. Compute performance metrics: OOS R^2 , RMSE, hedge tracking error, Sharpe ratio of mis-pricing signal
74. Regime-specific and sector-specific performance evaluation
75. Compare recommended spec to previous and next levels
76. If next level improves RMSE by <5%, stick with current (parsimony); if >10%, adopt next level
77. Document production specification: Formula, parameters, sector factors, implementation, recalibration protocol, edge cases
78. Create implementation blueprint: Pseudo-code, lookup tables, sector adjustment factors, system integration guide
79. Comparative performance analysis: Economic value examples (hedging, relative value, sector allocation, portfolio construction)
80. **Deliverables:** Tables E.1–E.5, Figures E.1–E.4, implementation blueprint (5–7 pages), comparative analysis (3–4 pages), final recommendation template

Phase 7: Final Report and Presentation

81. Compile all stage deliverables into comprehensive report (60–90 pages)
82. Executive summary (3–5 pages): Key findings, recommended spec, expected improvements, sector adjustment factors
83. Technical appendix: Detailed methodology, robustness checks, sensitivity analyses
84. Create presentation deck for stakeholders (20–30 slides):
 - Motivation: Why DTS adjustments needed
 - Theory: Merton predictions in plain language
 - Stage 0 findings: Three-pronged validation results, sector effects
 - Results: What we found (stage by stage)
 - Recommendation: Production spec with performance comparison
 - Implementation: Roadmap and timeline
85. Supplementary materials: Code repository, data dictionary, replication instructions

14.1 Resource requirements

Analysis	Computation Time	Analyst Time
Phase 0: Data construction	4–8 hours	3–5 days
Phase 0.5: Merton setup	1–2 hours	1–2 days
Phase 1: Bucket-level	2–4 hours	2–3 days
Phase 1.5: Within-issuer	4–8 hours	2–3 days
Phase 1.6: Sector interaction	1–2 hours	1–2 days
Phase 1.7: Synthesis	—	2–3 days
Phase 2: Stage A	2–4 hours	2–3 days
Phase 3: Stage B	2–4 hours	3–4 days
Phase 4: Stage C	2–4 hours	2–3 days
Phase 5: Stage D	4–8 hours	3–4 days
Phase 6: Stage E	4–8 hours	4–5 days
Phase 7: Final report	—	5–7 days
Total	26–52 hours	30–44 days

Table 12: Estimated resource requirements for complete research program

Notes:

- Computation time assumes modern hardware (16+ cores, 64GB RAM) and efficient code
- Within-issuer analysis most computationally intensive (225,000+ separate regressions)
- Analyst time includes data preparation, code development, quality checks, interpretation, and documentation

- Parallelization can reduce wall-clock time significantly
- Total timeline: 6–10 weeks with dedicated analyst

15 Conclusion

This research program provides a comprehensive framework for enhancing the Duration-Times-Spread (DTS) model through theory-guided empirical estimation. The key innovations relative to standard approaches include:

15.1 Methodological contributions

1. **Three-pronged Stage 0 validation:** Rather than relying on a single approach, we test Merton predictions using bucket-level analysis (aggregate patterns), within-issuer analysis (pure maturity effects), and sector interaction analysis (formal inference on industry differences). This triangulation provides robust evidence for subsequent stages.
2. **Sequential testing with explicit decision points:** Rather than running all analyses in parallel, we proceed hierarchically with clear stopping rules. This prevents over-engineering and respects the principle that simpler models are preferable when adequate.
3. **Theory as baseline, not constraint:** Merton predictions provide strong priors that sharpen empirical tests and reduce parameter space, but we systematically evaluate when theory fails and data require more flexibility.
4. **Separation of existence and explanation:** Stage A establishes *that* DTS betas vary before Stage B tests *whether* theory explains it. This clarifies where standard DTS fails versus where our adjustments add value.
5. **Formal sector effect testing:** Moving beyond descriptive sector comparisons to formal statistical inference about whether Financial, Utility, and Energy bonds systematically differ in DTS sensitivity. This determines whether production models require sector-specific adjustments.
6. **Hierarchical model selection with sector track:** The Level 1–5 framework guided by theory now includes explicit paths for incorporating sector adjustments (Levels 2a, 3a, 4a). We stop at the simplest adequate model rather than maximizing in-sample fit.

15.2 Practical contributions

1. **Implementable specifications:** Every stage produces actionable outputs—lookup tables, regression coefficients, sector adjustment factors, implementation pseudo-code—not just statistical tests.
2. **Economic significance thresholds:** We don’t just test statistical significance but also whether effects are large enough to matter for portfolio management (e.g., 20% threshold for time-variation, 0.2 threshold for sector effects).
3. **Sector-aware production specifications:** When Stage 0 identifies significant sector effects (e.g., Financial amplification), these are carried through to production with explicit adjustment factors and monitoring protocols.

4. **Production readiness:** Stage E delivers a complete implementation blueprint with recalibration protocols, edge case handling, sector classification requirements, and performance monitoring.
5. **Regime awareness:** Rather than assuming one-size-fits-all, we explicitly acknowledge that DTS adjustments are most critical in Regime 2 (IG with wide maturity range) and that sector effects may be regime-dependent.

15.3 Academic contributions

1. **First comprehensive empirical test with proper identification:** This program provides the first market-wide test of structural model spread dynamics using modern bond data (2013–2025) with proper within-issuer identification of pure maturity effects.
2. **Theory-data dialogue:** Rather than pure theory (no empirics) or pure empirics (ignoring theory), we create a dialogue where theory guides specification and data discipline theory.
3. **Quantification of regime and sector effects:** We move beyond qualitative statements about when proportional spread movements fail to precise estimates of elasticity ratios by spread level, maturity, and sector.
4. **Decomposition of failures:** Stage D systematically diagnoses *why* theory fails (tails? liquidity? specific shocks? specific sectors?) rather than just documenting that it does.
5. **Sector mechanism identification:** The combination of Stage 0 sector interactions with Stage D shock decomposition reveals *why* sector effects exist (correlation risk? liquidity? information asymmetry?), informing both academic understanding and production model design.

15.4 Expected outcomes

Based on prior research and the theoretical framework, we anticipate:

1. **Stage 0 will validate Merton for IG cross-maturity effects:** The structural model's prediction [Wuebben \[2025\]](#) that 1-year IG bonds have 3–4× higher sensitivity than 10-year bonds should be confirmed by both bucket-level and within-issuer tests.
2. **Stage 0 will identify Financial sector amplification:** We expect $\hat{\beta}_{\text{Financial}} > 0$ (likely 0.2–0.4), indicating Financials have 20–40% higher DTS sensitivity than Industrials, driven by correlation risk and regulatory capital effects.
3. **Stage A will show significant variation:** *F*-tests will reject beta equality with $p < 0.01$, motivating the need for adjustments.
4. **Stage B will show Merton explains most IG variation:** For investment-grade with wide maturity dispersion, $R_{\text{Merton}}^2 > 0.85 \times R_{\text{buckets}}^2$ likely. Adding sector adjustments will provide incremental improvement (R^2 gain of 2–5%).
5. **Stage C will find static relationships largely sufficient:** Time-variation likely marginal except in rare crises. Sector effects likely stable in normal periods but may amplify during stress (particularly for Financials).

6. **Stage D will identify Financial amplification sources:** Quantile regressions and shock decomposition will reveal whether Financial sector effects come from tail events, sector-wide contagion, or baseline amplification.
7. **Stage E will recommend Calibrated Merton with Sector Adjustments:** Most likely outcome is Level 2a or 3a, providing 25–35% RMSE improvement over Standard DTS with moderate implementation complexity.

15.5 Broader implications

This research program has implications beyond corporate credit DTS:

1. **Methodology template:** The three-pronged validation framework (aggregate, within-unit, interaction effects) applies to other markets where structural models exist and heterogeneity may be sector-specific.
2. **Theory-practice integration:** Demonstrates how academic models can inform practitioner tools when subjected to rigorous empirical validation rather than blind application.
3. **Risk model enhancement:** Improved DTS specifications with sector adjustments translate directly to better portfolio risk estimates, hedge ratios, and stress test scenarios.
4. **Alpha generation:** Properly scaled relative-value signals (rich/cheap within issuer capital structures, accounting for sector effects) can generate trading alpha when market prices deviate from theory-adjusted fair values.
5. **Sector allocation:** Understanding differential DTS sensitivity across sectors enables more accurate risk budgeting and avoids unintended concentration in high-sensitivity sectors.

15.6 Limitations and extensions

Several important limitations should be acknowledged:

1. **Merton is approximate:** Even if empirically validated, the structural model rests on simplifying assumptions (log-normal firm value, constant parameters, no jumps). Stage D robustness checks partially address this.
2. **Parameter instability:** Coefficients estimated on 2013–2025 data may not hold if credit cycle or market structure changes. Sector effects particularly may evolve with regulatory changes or market structure shifts.
3. **Data quality dependence:** Results critically depend on accurate OAS, maturity, sector classification, and liquidity data. Sector reclassifications require monitoring.
4. **Sample period limitations:** 2013–2025 includes only one true crisis (COVID 2020). Performance in 2008-style liquidity breakdown uncertain. Sector effects during such crises may differ.
5. **Sector classification granularity:** Current framework uses 4 broad sectors. Finer distinctions (e.g., Banks vs Insurance, Exploration vs Integrated Energy) may reveal additional heterogeneity.

Future extensions could include:

1. **International markets:** Apply framework to EUR, GBP, emerging market credit. Test whether Merton predictions and sector effects are universal or US-specific.
2. **Finer sector specialization:** Develop sub-sector adjustments for Financials (Banks, Insurance, REITs), Energy (Exploration, Integrated, Midstream), etc.
3. **Integration with equity signals:** Combine with equity volatility, CDS spreads, structural model equity-based default probabilities. Test whether sector effects in bonds align with equity market patterns.
4. **Machine learning augmentation:** Use ML to model λ as non-linear function of characteristics, but benchmark against Merton baseline and sector-adjusted specifications to ensure interpretability.
5. **Transaction costs and implementation:** Assess whether improved DTS adjustments (including sector factors) survive real-world trading costs, market impact, and operational constraints.
6. **Dynamic sector effects:** Model how sector adjustment factors evolve with regulatory changes, market structure shifts, and macroeconomic conditions.

15.7 Final perspective

The Duration-Times-Spread framework is elegantly simple and widely adopted, but its assumption of proportional spread movements breaks down systematically in investment-grade markets with maturity dispersion and exhibits sector-specific patterns that standard implementations ignore. This research program demonstrates that structural credit theory [Wuebben \[2025\]](#), particularly the Merton model, provides the key to understanding *when* and *why* proportionality fails, while rigorous empirical analysis reveals *where* additional adjustments (notably for sectors like Financials) are necessary.

By integrating theoretical priors with a three-pronged empirical validation approach, we deliver not just statistical improvements but economically meaningful enhancements that respect both the wisdom of markets and the insights of financial theory. The hierarchical testing framework ensures we adopt only the complexity that data justify—whether that means pure Merton tables, sector-adjusted specifications, or more flexible empirical models—resulting in production specifications that are simultaneously theoretically grounded, empirically validated, and practically implementable.

The ultimate deliverable is not merely a set of regression coefficients but a complete framework for understanding credit spread dynamics—one that portfolio managers can use with confidence for hedging and relative value, risk managers can rely on for accurate measurements across sectors, and researchers can build upon for future innovations.

References

- Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, 29(2), 449–470.
- Black, F., & Cox, J. C. (1976). Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance*, 31(2), 351–367.
- Leland, H. E. (1994). Corporate debt value, bond covenants, and optimal capital structure. *Journal of Finance*, 49(4), 1213–1252.

- Ben Dor, A., Dynkin, L., Hyman, J., Houweling, P., Van Leeuwen, E., & Penninga, O. (2010). DTS (Duration Times Spread). *Barclays Capital Quantitative Portfolio Strategy Research*.
- Wooldridge, J. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.
- Koenker, R. (2005). *Quantile Regression*. Cambridge University Press.
- Campbell, J. Y., & Taksler, G. B. (2003). Equity volatility and corporate bond yields. *Journal of Finance*, 58(6), 2321–2350.
- Chen, L., Lesmond, D. A., & Wei, J. (2007). Corporate yield spreads and bond liquidity. *Journal of Finance*, 62(1), 119–149.
- Collin-Dufresne, P., Goldstein, R. S., & Martin, J. S. (2001). The determinants of credit spread changes. *Journal of Finance*, 56(6), 2177–2207.
- Wuebben, B. J. (2025). When do credit spreads move proportionally? A structural model analysis of the Merton framework. *Working Paper*, AllianceBernstein.